Massimo Rudan

Physics of Semiconductor Devices

Second Edition

Springer

Physics of Semiconductor Devices

Massimo Rudan

Physics of Semiconductor Devices

Second Edition

Springer

Massimo Rudan
DEI
University of Bologna
Bologna, Italy

ISBN 978-3-319-63153-0 ISBN 978-3-319-63154-7 (eBook)
DOI 10.1007/978-3-319-63154-7

Library of Congress Control Number: 2017947902

Printed on acid-free paper

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

To Rossella and Morgana

Foreword to the First Edition

We are delighted to write the foreword of this comprehensive book by Massimo Rudan, entitled *Physics of Semiconductor Devices*, a treatise that goes to the heart of the physics involved in the study of electron devices. While the literature on this topic is very wide and diversified, it typically covers only part of the whole subject: either semiconductor physics or device theory. This book bridges the gap between these two aspects of the discipline, as it thoroughly covers both of them and even addresses process technology.

Why do we need a book on physics written by an engineer for engineers? Are not physics books written by physicists good enough? The answer is rooted in the difference of motivation of science as opposed to engineering. The scientist wants to know *why* and investigates the cause or the underlying fundamental laws from observed effects. The engineer wants to know *what for* and pursues a purpose, like building a device, based on such effects. The engineer tries to arrange physical objects in such a way that a desired and foreseeable action is achieved as can be expected by virtue of the laws of physics. The engineer goes from the known causes to the desired effects. In this sense, the engineer's objective is the inverse of the scientist's endeavor.

Is there common ground between physics and engineering? Yes, and that is why physicists and engineers should be friends. Obviously, knowledge of the appropriate part of physics (but not all physics) is indispensable for the engineer to build something useful. That is what this book aims at in the area of electronic devices. On the other hand, physicists build sophisticated laboratory equipment (sometimes with the help of engineers) for the purpose of investigating new physical phenomena. The builders of such unique tools do not have to worry about mass manufacturing in contrast to engineering, where low-cost production and reliability are major concerns, yet another difference of motivation.

The complexity of the physical background required for a deep understanding of device behavior makes it difficult to pursue a *deductive* teaching methodology, where every new concept is justified on fundamental physical principles. The limited number of lecture hours, and the need to comply with the program and objectives of engineering courses, is often such that many teachers, as well as textbooks, tend to

ignore such a background and to use an *inductive* approach instead. So doing, some fundamental concepts are delivered to students intuitively, in the form of physical models with some degree of simplification and no proof behind them. While the inductive approach may be useful for practical purposes, it bears the drawback of hiding assumptions and limitations of the model, thereby leading to an insufficient depth of understanding and lack of criticism on the validity of the obtained results.

This book is fully functional to a *deductive* teaching approach and, as such, represents a very useful tool not just for students but for researchers and scientists as well. The reader can thus find within the same volume both a rigorous justification for the semiclassical and quantum-mechanical models adopted in practice to clarify the device operation, such as the most popular drift-diffusion transport model, and detailed descriptions of the operating principles of the most important semiconductor devices. The evolution of microelectronics toward nanoscale structures makes it imperative to master quantum mechanics and solid-state physics in order to understand the operation of miniaturized devices. This book is therefore timely and meets a need felt by many students as well as teachers.

Roughly speaking, this volume addresses four main areas: (i) a review of analytical mechanics and electromagnetism, (ii) an introduction to statistical and quantum mechanics, (iii) a treatment of solid-state physics and carrier transport in semiconductors, and (iv) a description of the basic device concepts and their operating principles. The Appendices highlight mathematical concepts and proofs of theorems cited somewhere throughout the volume. For example, Chap. 6, where the Boltzmann transport equation is derived and discussed; Chap. 17, where the fundamental concepts of crystalline solid-state physics are introduced; and Chap. 18, with the description of the basic properties of electrons and holes in semiconductors under equilibrium conditions, highlight the rigorous, yet elegant, approach to the description of delicate physical concepts, such as the Bloch theorem, the band structure of crystalline solids in the first Brillouin zone, and the phonon vibrational spectra. The author's deep mathematical background shows up in several demonstrations, where the use of sophisticated mathematical tools leads to simple, albeit rigorous, proofs of the concepts addressed in the treatment. An important specificity of this book is also that simplifying assumptions are widely discussed and their impact on final results is typically highlighted.

In conclusion, *Physics of Semiconductor Devices* by Massimo Rudan represents a remarkable piece of work that is likely to provide a valuable learning tool to electrical engineering students willing to deepen their understanding of this fascinating field. It will also be a useful consultation tool for researchers, scientists, and engineers involved in the fundamental aspects of material science and device design.

Bologna, Italy — Giorgio Baccarani
Zürich, Switzerland — Henry Baltes
September 2014

Preface to the Second Edition

The second edition of this book has been enriched mostly in the parts devoted to semiconductors and devices (Chaps. 20 through 25). The analysis of the net recombination rate in polycrystalline materials, relevant for, e.g., thin-film transistors, is carried out in Chap. 20. The derivation of the characteristics of the bipolar transistor, photodiode, and solar cell has been added to Chap. 21, along with some considerations about the scaling limits of the bipolar architecture. The new material of Chap. 22 features the analysis of the *n*-substrate MOS capacitor, MOS photo-capacitor, and *p*-channel MOSFET, followed by the description of the CMOS architecture. A more general theory of the MOSFET, including the diffusive contribution to the channel current, is given; a brief outline about the scaling rules of semiconductor devices and the design steps of integrated circuits completes the new material of this chapter. Chapter 23 has been extended substantially by including the description of ion implantation. As for the measurement techniques, Chap. 25 has been enriched with the theory of the Hall effect in the nonideal case and with the resistivity measurement based on the van der Pauw method.

More material has sparsely been added to other chapters: the derivation of the Schrödinger equation as the Euler equation of a constrained functional is extended to more than one dimension in Chap. 1; the theory of the Green function in the top half plane, instrumental to the van der Pauw measurement scheme, has been added to Chap. 4; a full derivation of the field generated by a point-like charge is now included in Chap. 5. The demonstration of the Boltzmann H-theorem and a discussion about the Kac-ring paradox have been added to Chap. 6. In the chapters devoted to quantum mechanics, the discussion of Chap. 13 about the properties of the one-dimensional Schrödinger equation is more extended, and the treatment of the time-dependent perturbation of Chap. 14 is now embracing the harmonic perturbation of a general form and the spatially periodic case; the latter analysis provides in turn the selection rules that are exploited in the discussion of the electron-phonon interaction, added to Chap. 17. The transport model worked out in Chap. 19 is used in the complements of the chapter to derive the Onsager relations.

The Appendices have also been enriched with new material, with the aim of completing a number of topics tackled in the book's chapters; several sections about

matrix analysis have been added to Appendix A, followed by other sections showing the application to the numerical solution of the semiconductor equations in one dimension. Appendix B has been enriched with the theories of Lagrange multipliers, conformal mapping, and contraction mappings; finally, the approximations to the Fermi integrals, the illustration of the Bernoulli numbers and polynomials, the calculation of a special integral using Cauchy's residue theorem, the treatment of the random walk problem, and the central limit theorem have been added to Appendix C.

About 50 problems, 50 figures, and 3 tables have been added to those of the first edition and about 20 figures reworked; several new figures are in color. A subject index featuring about 2,600 entries has been inserted at the end of the volume.

Bologna, Italy Massimo Rudan
June 2017

Preface to the First Edition

This volume originates from the lectures on solid-state electronics and microelectronics that I have been giving since 1978 at the School of Engineering of the University of Bologna. Its scope is to provide the reader with a book that, starting from the elementary principles of classical mechanics and electromagnetism, introduces the concepts of quantum mechanics and solid-state theory and describes the basic physics of semiconductors including the hierarchy of transport models, ending up with the standard mathematical model of semiconductor devices and the analysis of the behavior of basic devices. The ambition of the work has been to write a book, self-contained as far as possible, that would be useful for both students and researchers; to this purpose, a strong effort has been made to elucidate physical concepts, mathematical derivations, and approximation levels, without being verbose.

The book is divided into eight parts. Part I deals with analytical mechanics and electromagnetism; purposedly, the material is not given in the form of a resumé: quantum mechanics and solid-state theory concepts are so richly intertwined with the classical ones that presenting the latter in an abridged form may make the reading unwieldy and the connections more difficult to establish. Part II provides the introductory concepts of statistical mechanics and quantum mechanics, followed by the description of the general methods of quantum mechanics. The problem of bridging the classical concepts with the quantum ones is first tackled using the historical perspective, covering the years from 1900 to 1926. The type of statistical description necessary for describing the experiments, and the connection with the limiting case of the same experiments involving massive bodies, is related to the properties of the doubly stochastic matrices. Part III illustrates a number of applications of the Schrödinger equation: elementary cases, solutions by factorization, and time-dependent perturbation theory. Part IV analyzes the properties of systems of particles, with special attention to those made of identical particles, and the methods for separating the equations. The concepts above are applied in Part V to the analysis of periodic structures, with emphasis to crystals of the cubic type and to silicon in particular, which, since the late 1960s, has been and still is the most important material for the fabrication of integrated circuits. Part VI illustrates

the single-electron dynamics in a periodic structure and derives the semiclassical Boltzmann transport equation; from the latter, the hydrodynamic and drift-diffusion models of semiconductor devices are obtained using the moments expansion. The drift-diffusion model is used in Part VII to work out analytically the electrical characteristics for the basic devices of the bipolar and MOS type. Finally, Part VIII presents a collection of items which, although important per se, are not in the book's mainstream: some of the fabrication-process steps of integrated circuits (thermal diffusion, thermal oxidation, layer deposition, epitaxy) and methods for measuring the semiconductor parameters.

In the preparation of the book, I have been helped by many colleagues. I wish to thank, in particular, Giorgio Baccarani, Carlo Jacoboni, and Rossella Brunetti, who gave me important suggestions about the matter's distribution in the book, read the manuscript, and, with their observations, helped me to clarify and improve the text; I wish also to thank, for reading the manuscript and giving me their comments, Giovanni Betti Beneventi, Fabrizio Buscemi, Gaetano D'Emma, Antonio Gnudi, Elena Gnani, Enrico Piccinini, Susanna Reggiani, and Paolo Spadini.

Last, but not least, I wish to thank the students, undergraduate, graduate, and postdocs, who for decades have accompanied my teaching and research activity with stimulating curiosity. Many comments, exercises, and complements of this book are the direct result of questions and comments that came from them.

Bologna, Italy
September 2014

Massimo Rudan

Contents

Part VIII Miscellany

About the Author

Massimo Rudan received a degree in electrical engineering in 1973 and a degree in physics in 1976, both from the University of Bologna, Italy. His research interests are in the field of physics of carrier transport and numerical analysis of semiconductor devices. In 1986, he was a visiting scientist, on a one-year assignment, at the IBM Thomas J. Watson Research Center at Yorktown Heights, NY, studying the discretization techniques for the higher-order moments of the Boltzmann transport equation.

From 1979, he has been teaching annual courses in the Faculty of Engineering of the University of Bologna, firstly as lecturer and then as associate professor. In 1990, he became full professor of microelectronics at the University of Bologna. An IEEE fellow (2008) and life fellow (2014), M.R. is author, coauthor, or editor of 12 books and has authored or coauthored about 250 technical papers published in major journals or presented in international conferences, in the fields of electronics, solid-state physics, and solid-state sensors.

M.R. has coordinated several research projects funded by the European Commission, international companies and foundations, the National Council of Research, and the National Ministry of University and Research. In 2001, he was one of the founders of the Advanced Research Center for Electronic Systems (ARCES) of the University of Bologna.

Acronyms

AM	Air mass. Length of the optical path, through Earth's atmosphere, of light rays coming from an extraterrestrial source. The AM value at sea level, when the source is at the zenith, is set equal to unity and used as normalization parameter.
BJT	Bipolar junction transistor. A transistor whose operation is obtained by a suitable arrangement of two *p*-*n* junctions. The term "bipolar" is used because both electrons and holes are involved in the device functioning.
BTE	Boltzmann transport equation. The equation expressing the continuity of the distribution function in the phase space.
CVD	Chemical vapor deposition. A deposition process in which the material to be deposited is the product of a chemical reaction that takes place on the surface of the substrate or in its vicinity.
DD	Drift diffusion. The term indicates a transport model for semiconductors made, for each energy band, of the balance equations for the carrier number and average velocity. Such equations contain the electric field and the magnetic induction; as a consequence, their solution must be calculated consistently with that of the Maxwell equations. Compare with the HD model.
DRAM	Dynamic random access memory. A type of random-access, solid-state memory where each bit is stored in a separate capacitor. The charged/discharged states of the latter correspond to the logic states of the memory bit. Due to the leakage of the circuits connected to the capacitor, the stored datum tends to fade unless it is periodically refreshed (whence the designation "dynamic"). DRAM memories are volatile, namely, they lose the data when the power supply is removed.
HD	Hydro dynamic. The term indicates a transport model for semiconductors made, for each energy band, of the balance equation for the carrier number, average velocity, average kinetic energy, and average flux of the kinetic energy. Such equations contain the electric field and the magnetic induction; as a consequence, their solution must be calculated

	consistently with that of the Maxwell equations. Compare with the DD model.
IC	Integrated circuit. Also called *chip* or *microchip*. An assembly of electronic circuits on the same plate of a semiconductor material. The idea was proposed in the early 1950s and demonstrated in 1958; it provided an enormous improvement, both in cost and performance, with respect to the manual assembly of circuits using discrete components.
IGFET	Insulated gate field effect transistor. A device architecture demonstrated in the early 1930s. Its first implementation (1960) using a thermally oxidized silicon layer gave rise to the MOSFET architecture.
LASER	Light amplification by stimulated emission of radiation. A device made of a material having a high probability of radiative emission, in which the electron concentration of high-energy states is artificially kept high by an external power source. The material produces a coherent emission, and a suitable feedback makes it to oscillate; this results in a nearly monochromatic light.
LOCOS	Local oxidation. The technological process consisting in depositing and patterning a layer of silicon nitride over the areas where the substrate's oxidation must be prevented.
MBE	Molecular beam epitaxy. A low-temperature epitaxial process based on evaporation.
MIS	Metal insulator semiconductor. Structure made of the superposition of a metal contact, an insulator, and a semiconductor.
MOS	Metal oxide semiconductor. Structure made of the superposition of a metal contact, an oxide that acts as an insulator, and a semiconductor.
MOSFET	Metal oxide semiconductor, field effect transistor. A transistor whose active region is an MOS structure. In last-generation devices, the insulator may be deposited instead of being obtained by oxidizing the semiconductor underneath. The MOSFET has been for decades, and still is, the fundamental device of the integrated-circuit architecture.
PCM	Phase change memory. A solid-state memory whose logic states are associated with a high-resistance or low-resistance condition of the material; such conditions are in turn associated to the material's phases, amorphous or crystalline, respectively. The material is forced to switch from one phase to the other by an externally applied electric signal.
PDE	Partial differential equation.
PV	Photo voltaic. The adjective refers to physical processes that convert the energy of a radiation into electric energy (e.g., *photovoltaic effect*) or to devices where such processes occur (e.g., *photovoltaic cell*).
PVD	Physical vapor deposition. A deposition process in which the material to be deposited does not react chemically with other substances.
SGOI	Silicon germanium on insulator. A technology analogous to SOI. SGOI increases the speed of the transistors by straining the material under the gate, thus making the electron mobility higher.

SOI Silicon on insulator. A technology introduced in 1998 for semiconductor manufacturing, in which the standard silicon substrate is replaced with a layered structure of the silicon-insulator-silicon type. SOI reduces the parasitic capacitances and the short-channel effect in MOS transistors.

SOS Silicon on sapphire. A technological process that consists in growing a thin layer of silicon on a wafer made of sapphire (Al_2O_3).

SS Subthreshold slope. The inverse of $\mathrm{d}\log_{10}(I)/\mathrm{d}V_G$ in a device where a conductive channel is controlled by a gate; I is the channel current, V_G the gate voltage. A smaller SS corresponds to a faster transition of the device from the *on* to the *off* state or vice versa.

TFT Thin film transistor. A type of field-effect transistor fabricated by depositing the active semiconductor layer, the dielectric layer, and the metallic contacts over a nonconducting substrate. An important application of TFTs is in liquid-crystal displays; for this reason, a typical type of nonconducting substrate of TFTs is glass.

List of Figures

List of Tables

Part I
A Review of Analytical Mechanics and Electromagnetism

Chapter 1
Analytical Mechanics

1.1 Introduction

The differential equations that govern the dynamic problems can be derived from variational principles, upon which the theories of Euler and Lagrange, and those of Hamilton and Jacobi, are based. Apart from being the source of the greatest intellectual enjoyment, these theories have the definite advantage of generality. Their concepts can in fact be extended to cases where the Newtonian equation of dynamics does not apply. Among such cases there are the equations governing the electromagnetic field and those related to the quantum description of the particles' motion.

The invariance property of the Lagrange equations with respect to a change of coordinates gives origin to the concept of generalized coordinates and conjugate momenta; in turn, the introduction of the Hamilton equations provides a picture in which the intrinsic symmetry of the roles of coordinates and momenta becomes apparent. Basing on the concept of conjugate coordinates, the time evolution of a particle or of a system of particles is described in the phase space instead of the coordinate space. This chapter and the next one illustrate the basic principles of Analytical Mechanics. Their purpose is to introduce a number of concepts that are not only useful *per se*, but also constitute a basis for the concepts of Quantum Mechanics that are introduced in later chapters. The third chapter devoted to Analytical Mechanics describes a number of important examples that will be applied to later developments illustrated in the book.

As the velocity of particles within a semiconductor device is small with respect to that of light, the nonrelativistic form of the mechanical laws is sufficient for the purposes of this book. The relativistic form is used only in a few paragraphs belonging to the chapter devoted to examples, to the purpose of describing a specific type of collision between particles. This chapter starts with the description of the Lagrangian function and the Lagrange equations that are derived as a consequence of the variational calculus, followed by the derivation of the Hamiltonian function and Hamilton equations. Next, the Hamilton-Jacobi equation is derived after

M. Rudan, *Physics of Semiconductor Devices*, DOI 10.1007/978-3-319-63154-7_1

discussing the time-energy conjugacy. The chapter continues with the definition of the Poisson brackets and the derivation of some properties of theirs, and concludes with the description of the phase space and state space.

1.2 Variational Calculus

Consider a real function $w(\xi)$ defined in the interval $\xi \in [a, b]$ and differentiable in its interior at least twice. The first two derivatives will be indicated with $\dot{w}$ e $\ddot{w}$. Now, define the integral

$$G[w] = \int_a^b g(w, \dot{w}, \xi)\, \mathrm{d}\xi\,, \tag{1.1}$$

where the form of the function $g(w, \dot{w}, \xi)$ is prescribed. If (1.1) is calculated for any function w fulfilling the requisites stated above, with a and b fixed, the result is some real number G whose value depends on the choice of w. By this procedure, (1.1) establishes a correspondence $G[w]$ between a set of functions and a set of numbers. Such a correspondence is called *functional*.

It is interesting to extend to the case of functionals some concepts and procedures that apply to functions proper; among these the concept of extremum is important. In fact, one defines the *extremum function* of a functional by a method similar to that used for defining the extremum point of a function: some w is an extremum of G if a variation $\mathrm{d}w$ in (1.1) produces a variation $\mathrm{d}G$ that is infinitesimal of an order higher than that of $\mathrm{d}w$. The procedure by which the extremum functions are calculated is called *variational calculus*.

To proceed it is necessary to define the variation $\mathrm{d}w$. For this one lets $\delta w = \alpha\, \eta$, with $\eta(\xi)$ an arbitrary function defined in $[a, b]$ and differentiable in its interior, and α a real parameter. The function δw thus defined is the finite variation of w. The sum $w + \delta w$ tends to w in the limit $\alpha \to 0$. As a consequence, such a limit provides the infinitesimal variation $\mathrm{d}w$. For simplicity it is convenient to restrict the choice of η to the case $\eta(a) = \eta(b) = 0$, so that $w + \delta w$ coincides with w at the integration boundaries for any value of α. Now, replacing w with $w + \alpha\, \eta$ in (1.1) makes G a function of α, whose derivative is

$$\frac{\mathrm{d}G}{\mathrm{d}\alpha} = \int_a^b \left[\frac{\partial g}{\partial(w + \alpha\eta)}\, \eta + \frac{\partial g}{\partial(\dot{w} + \alpha\dot{\eta})}\, \dot{\eta} \right] \mathrm{d}\xi\,. \tag{1.2}$$

According to the definition given above, if w is an extremum function of G then it must be $\lim_{\alpha\to 0} \mathrm{d}G/\mathrm{d}\alpha = 0$; in this case, in fact, the first-order term in the power expansion of G with respect to α vanishes, and the variation of G becomes second order in α or higher. In conclusion, one proceeds by imposing that the right-hand side of (1.2) vanishes for $\alpha = 0$. Then, integrating by parts the second term in brackets yields

$$\int_a^b \frac{\partial g}{\partial w}\,\eta\,\mathrm{d}\xi + \left[\frac{\partial g}{\partial \dot{w}}\,\eta\right]_a^b = \int_a^b \left(\frac{\mathrm{d}}{\mathrm{d}\xi}\frac{\partial g}{\partial \dot{w}}\right)\eta\,\mathrm{d}\xi \tag{1.3}$$

where, in turn, the integrated part vanishes because $\eta(a) = \eta(b) = 0$. This makes the two integrals in (1.3) equal to each other. On the other hand, such an equality must hold for any choice of η due to the arbitrariness of the latter. It follows that the integrands must be equal to each other, namely,

$$\frac{\mathrm{d}}{\mathrm{d}\xi}\frac{\partial g}{\partial \dot{w}} = \frac{\partial g}{\partial w}\,. \tag{1.4}$$

The relation (1.4) thus found is a second-order differential equation in the unknown w, whose explicit form is easily calculated:

$$\frac{\partial^2 g}{\partial \dot{w}^2}\,\ddot{w} + \frac{\partial^2 g}{\partial w \partial \dot{w}}\,\dot{w} + \frac{\partial^2 g}{\partial \xi \partial \dot{w}} = \frac{\partial g}{\partial w}\,. \tag{1.5}$$

Its solution provides the extremum function w sought. To actually find a solution one must associate with (1.4) suitable boundary conditions, e.g., $w(a) = w_a$, $\dot{w}(a) = \dot{w}_a$, or $w(a) = w_a$, $w(b) = w_b$, and so on. As g does not contain $\ddot{w}$, (1.4) is linear with respect to $\ddot{w}$. It is also worth noting that, consistently with what happens in the case of functions proper, the above calculation does not provide in itself any information about w being a minimum or maximum of G. Such an information must be sought through additional calculations.

The analysis above is easily extended to the case where g depends on several functions w_1, w_2 ,... and the corresponding derivatives. Introducing the vectors $\mathbf{w}(\xi) = (w_1, w_2, \ldots, w_n)$, $\dot{\mathbf{w}}(\xi) = (\dot{w}_1, \dot{w}_2, \ldots, \dot{w}_n)$ one finds that the set of n extremum functions $w_i(\xi)$ of functional

$$G[\mathbf{w}] = \int_a^b g(\mathbf{w}, \dot{\mathbf{w}}, \xi)\,\mathrm{d}\xi \tag{1.6}$$

is the solution of the set of differential equations

$$\frac{\mathrm{d}}{\mathrm{d}\xi}\frac{\partial g}{\partial \dot{w}_i} = \frac{\partial g}{\partial w_i}\,, \qquad i = 1, \ldots, n\,, \tag{1.7}$$

supplemented with suitable boundary conditions. Equations (1.7) are called *Euler equations* of the functional G.

Each Eq. (1.7) is homogeneous with respect to the derivatives of g and does not contain g itself. As a consequence, the differential equations (1.7) are invariant when g is replaced with $A\,g + B$, where $A, B \neq 0$ are constants. As the boundary conditions of w_i are not affected by that, the solutions w_i are invariant under the transformation. Moreover, it can be shown that the solutions are invariant under a

more general transformation. In fact, consider an arbitrary function $h = h(\mathbf{w}, \xi)$ and let $g' = g + \mathrm{d}h/\mathrm{d}\xi$, thus transforming (1.6) into

$$G'[\mathbf{w}] = A \int_a^b g(\mathbf{w}, \dot{\mathbf{w}}, \xi)\, \mathrm{d}\xi + h(\mathbf{w}_b, \xi_b) - h(\mathbf{w}_a, \xi_a)\,. \tag{1.8}$$

When each w_i is replaced with $w_i + \mathrm{d}w_i$, the terms involving h do not vary because the variations vanish at the boundaries of the integration domain. Thus, the variation of G' equals that of the integral, namely, it is of a higher order than $\mathrm{d}w_i$. In conclusion, the extremum functions of G are also extremum functions of G'. This means that the solutions $w_i(\xi)$ are invariant under addition to g of the total derivative of an arbitrary function that depends on $\mathbf{w}$ and ξ only. This reasoning does not apply if h depends also on the derivatives $\dot{\mathbf{w}}$, because in general the derivatives of the variations do not vanish at the boundaries.

1.3 Lagrangian Function

In many cases the solution of a physical problem is achieved by solving a set of second-order differential equations of the form $\ddot{w}_i = \ddot{w}_i(\mathbf{w}, \dot{\mathbf{w}}, \xi)$. For instance, for nonrelativistic velocities the law of motion of a particle of constant mass m is Newton's law $\mathbf{F} = m\,\mathbf{a}$ which, in a Cartesian frame, takes the form

$$m\,\ddot{x}_i = F_i(\mathbf{r}, \dot{\mathbf{r}}, t)\,, \qquad i = 1, 2, 3\,. \tag{1.9}$$

In (1.9), $\mathbf{r}(t) = (x_1, x_2, x_3)$ is the particle's position vector[1] at t. In the following the particle's velocity will be indicated with $\mathbf{u} = \dot{\mathbf{r}}$.

Equations (1.9) and (1.7) have the same form, as is easily found by observing that t is the analogue of ξ and x_i is that of w_i. As a consequence, one may argue that (1.9) could be deduced as Euler equations of a suitable functional. This problem is in fact the inverse of that solved in Sect. 1.2: there, the starting point is the function g, whose derivatives provide the coefficients of the differential equations (1.7); here, the coefficients of the differential equation are given, while the function g is to be found. For the inverse problem the existence of a solution is not guaranteed in general; if a solution exists, finding the function g may be complicate because the process requires an integration. In other terms, the direct problem involves only the somewhat "mechanical" process of calculating derivatives, whereas the inverse problem involves the integration which is, so to speak, an art.

When dealing with the dynamics of a particle or of a system of particles, the function g, if it exists, is called *Lagrangian function* and is indicated with L. The

[1] The units in (1.9) are: $[m] = \mathrm{kg}$, $[\mathbf{r}] = \mathrm{m}$, $[\dot{\mathbf{r}}] = \mathrm{m\ s^{-1}}$, $[\ddot{x}_i] = \mathrm{m\ s^{-2}}$, $[F_i] = \mathrm{N}$, where "N" stands for Newton.

equations corresponding to (1.7) are called *Lagrange equations*. The expression of the Lagrangian function depends on the form of the force F_i in (1.9). Some examples are given in the following. It is important to note that by "system of particles" it is meant a collection of particles that interact with each other. If there were no interactions it would be possible to tackle the dynamics of each particle separately; in other terms, each particle would constitute a system in itself, described by a smaller number of degrees of freedom.

1.3.1 Force Deriving from a Potential Energy

Consider the case of a force deriving from a potential energy, namely $\mathbf{F} = -\operatorname{grad} V$ with $V = V(\mathbf{r}, t)$, so that (1.9) becomes

$$m\,\dot{u}_i = -\frac{\partial V}{\partial x_i}\,. \tag{1.10}$$

Using the replacements $\xi \leftarrow t$, $w_i \leftarrow x_i$, $g \leftarrow L$ and equating (1.7) and (1.10) side by side yield

$$\frac{\partial L}{\partial x_i} = -\frac{\partial V}{\partial x_i}\,, \qquad \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial u_i} = \frac{\mathrm{d}}{\mathrm{d}t}(m\,u_i)\,, \qquad i = 1, 2, 3\,. \tag{1.11}$$

The first of (1.11) shows that the sum $T = L + V$ does not depend on the coordinates x_i. Inserting $L = T - V$ into the second of (1.11) and taking $i = 1$ show that the difference $\Phi = \partial T/\partial u_1 - m u_1$ does not depend on time either, so it depends on the u_i components at most. Integrating Φ with respect to u_1 yields $T = m u_1^2/2 + T_1(u_2, u_3, t)$, with T_1 yet undetermined. Differentiating this expression of T with respect to u_2 and comparing it with the second of (1.11) specified for $i = 2$ yield $T = m\,(u_1^2 + u_2^2)/2 + T_2(u_3, t)$, with T_2 undetermined. Repeating the procedure for $i = 3$ finally provides $T = m\,(u_1^2 + u_2^2 + u_3^2)/2 + T_0(t)$, with T_0 an undetermined function of time only. The latter, in turn, can be viewed as the time derivative of another function h. Remembering the invariance property discussed at the end of Sect. 1.2 with reference to (1.8), one lets $T_0 = 0$. In conclusion, indicating with u the modulus of $\mathbf{u}$ it is $T = m u^2/2$, and the Lagrangian function reads

$$L = \frac{1}{2}\,m u^2 - V\,. \tag{1.12}$$

The derivation of (1.12) may appear lengthy. However, the procedure is useful because it is applicable to forces of a more complicate form.

1.3.2 Electromagnetic Force

Consider a charged particle subjected to an electromagnetic field and let m and e be its mass and charge, respectively. The particle's velocity $\mathbf{u}$ is assumed to be nonrelativistic. The electromagnetic field acts on the particle with the *Lorentz force* (Sect. 4.11) [2]

$$\mathbf{F} = e\,(\mathbf{E} + \mathbf{u} \wedge \mathbf{B})\,, \tag{1.13}$$

where the electric field $\mathbf{E}$ and the magnetic induction $\mathbf{B}$ are in turn expressed through the scalar potential $\varphi = \varphi(\mathbf{r}, t)$ and the vector potential $\mathbf{A} = \mathbf{A}(\mathbf{r}, t)$ as (Sect. 4.4)

$$\mathbf{E} = -\operatorname{grad}\varphi - \frac{\partial \mathbf{A}}{\partial t}\,, \qquad \mathbf{B} = \operatorname{rot}\mathbf{A}\,. \tag{1.14}$$

Letting $i = 1$ in (1.9) one finds from (1.13) $m\dot{u}_1 = e\,(E_1 + u_2 B_3 - u_3 B_2)$. Using for E_1, B_3, B_2 the expressions extracted from (1.14) yields

$$m\,\dot{u}_1 + e\left(\frac{\partial A_1}{\partial t} + u_2\,\frac{\partial A_1}{\partial x_2} + u_3\,\frac{\partial A_1}{\partial x_3}\right) = e\left(-\frac{\partial \varphi}{\partial x_1} + u_2\,\frac{\partial A_2}{\partial x_1} + u_3\,\frac{\partial A_3}{\partial x_1}\right). \tag{1.15}$$

Now, using $u_i = \dot{x}_i$ transforms the term in parentheses at the left-hand side of (1.15) into $\mathrm{d}A_1/\mathrm{d}t - u_1\,\partial A_1/\partial x_1$, which gives (1.15) the more compact form

$$\frac{\mathrm{d}}{\mathrm{d}t}\,(m\,u_1 + e\,A_1) = \frac{\partial}{\partial x_1}\,(e\,\mathbf{u}\cdot\mathbf{A} - e\,\varphi)\,. \tag{1.16}$$

Similar expressions are found for $i = 2, 3$. Comparing with (1.7) in the same manner as in Sect. 1.3.1 yields

$$\frac{\partial L}{\partial x_i} = \frac{\partial}{\partial x_i}\,(e\,\mathbf{u}\cdot\mathbf{A} - e\,\varphi)\,, \qquad \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial u_i} = \frac{\mathrm{d}}{\mathrm{d}t}\,(m\,u_i + e\,A_i)\,, \qquad i = 1, 2, 3\,. \tag{1.17}$$

Note that (1.17) reduce to (1.11) when $\mathbf{A} = 0$, with $e\,\varphi = V$. The first of (1.17) shows that the sum $T = L + e\,\varphi - e\,\mathbf{u}\cdot\mathbf{A}$ does not depend on the coordinates x_i. Inserting $L = T - e\,\varphi + e\,\mathbf{u}\cdot\mathbf{A}$ into the second of (1.17) transforms the latter into $\mathrm{d}(\partial T/\partial u_i)/\mathrm{d}t = \mathrm{d}(m\,u_i)/\mathrm{d}t$. Like in Sect. 1.3.2 the procedure eventually yields $T = mu^2/2$. In conclusion, the Lagrangian function of a particle subjected to the Lorentz force (1.13) is

[2] The units in (1.13) are: $[\mathbf{F}] = \mathrm{N}$, $[e] = \mathrm{C}$, $[\mathbf{E}] = \mathrm{V\,m^{-1}}$, $[\mathbf{u}] = \mathrm{m\,s^{-1}}$, $[\mathbf{B}] = \mathrm{V\,s\,m^{-2}} = \mathrm{Wb\,m^{-2}} = \mathrm{T}$, where "N," "C," "V," "Wb," and "T" stand for Newton, Coulomb, Volt, Weber, and Tesla, respectively. The coefficients in (1.13) differ from those of [10] because of the different units adopted there. In turn, the units in (1.14) are: $[\varphi] = \mathrm{V}$, $[\mathbf{A}] = \mathrm{V\,s\,m^{-1}} = \mathrm{Wb\,m^{-1}}$.

$$L = \frac{1}{2} m u^2 - e\varphi + e\,\mathbf{u}\cdot\mathbf{A}\,. \tag{1.18}$$

It is shown in Sect. 4.4 that the **E** and **B** fields are invariant under the *gauge transformation*

$$\varphi \leftarrow \varphi - \frac{\partial h}{\partial t}\,, \qquad \mathbf{A} \leftarrow \mathbf{A} + \operatorname{grad} h\,, \tag{1.19}$$

where $h(\mathbf{r}, t)$ is an arbitrary function. Using (1.19) in (1.18) transforms the terms containing the potentials as

$$-e\varphi + e\,\mathbf{u}\cdot\mathbf{A} \leftarrow -e\varphi + e\,\mathbf{u}\cdot\mathbf{A} + e\,\frac{\mathrm{d}h}{\mathrm{d}t}\,, \tag{1.20}$$

namely, the transformed Lagrangian function differs from the original one by the total derivative of an arbitrary function that depends on position and time only. As a consequence, the solutions $x_i(t)$ are invariant under the gauge transformation (1.19). This is easily understood by observing that the invariance of the **E** and **B** fields makes the Lorentz force (1.13) invariant as well. As a consequence, the particle's dynamics is not influenced by the gauge transformation.

1.3.3 Work

The elementary work exerted by a force **F** acting on a particle of mass m during the time $\mathrm{d}t$ is $\mathbf{F}\cdot\mathrm{d}\mathbf{r}$, where **r** is the particle's position at t in a Cartesian frame and $\mathrm{d}\mathbf{r} = \mathbf{u}\,\mathrm{d}t$ the elementary displacement. Let $P = \mathbf{r}(t = a)$, $Q = \mathbf{r}(t = b)$ be the boundaries of the particle's trajectory. The work exerted from P to Q is found by integrating $\mathbf{F}\cdot\mathrm{d}\mathbf{r}$ over the trajectory, namely,

$$\int_P^Q \mathbf{F}\cdot\,\mathrm{d}\mathbf{r} = m\int_a^b \dot{\mathbf{u}}\cdot\mathbf{u}\,\mathrm{d}t = \frac{1}{2}m\int_a^b \frac{\mathrm{d}u^2}{\mathrm{d}t}\,\mathrm{d}t = T(b) - T(a)\,, \tag{1.21}$$

where the relation $T = mu^2/2$ has been used. The exerted work is then equal to the variation of T, which is the same quantity that appears in (1.12, 1.18) and is called *kinetic energy* of the particle. If a system having n degrees of freedom is considered instead of a single particle, the work exerted by the forces is defined as the sum of terms of the form (1.21). As a consequence, the kinetic energy of the system is the sum of the kinetic energies of the individual particles. The expression of the system's kinetic energy in Cartesian coordinates is

$$T = \sum_{i=1}^{n} \frac{1}{2}\, m_i\, u_i^2 = \sum_{i=1}^{n} \frac{1}{2}\, m_i\, \dot{x}_i^2\,, \tag{1.22}$$

that is, a positive-definite quadratic form in the velocities. The masses in (1.22) take the same value when they are referred to the same particle. When other types of coordinates are used, the kinetic energy is still a second-degree function of the velocities; however, the function's coefficients may depend on the coordinates (an example is given in Sect. 2.8).

When a force deriving from a potential energy $V = V(\mathbf{r}, t)$ is considered, like that of Sect. 1.3.1, the integrand of (1.21) becomes $-\operatorname{grad} V \cdot \mathrm{d}\mathbf{r}$. To calculate the integral it is necessary to account for the explicit dependence of V on t by using mutually consistent values of $\mathbf{r}$ and t; in other terms, the integral in (1.21) can actually be calculated only after determining the function $\mathbf{r}(t)$. An exception occurs when V has no explicit dependence on time; in this case one finds

$$\int_P^Q \mathbf{F} \cdot \mathrm{d}\mathbf{r} = -\int_P^Q \operatorname{grad} V \cdot \mathrm{d}\mathbf{r} = -\int_P^Q \mathrm{d}V = V(P) - V(Q)\,, \tag{1.23}$$

namely, to calculate the integral it suffices to know the boundaries of the trajectory. Moreover, when $V = V(\mathbf{r})$ the Lagrangian function (1.12) does not depend explicitly on time either. It is shown in Sect. 1.6 that in this case also the sum $T + V$ of the kinetic and potential energies is independent of time. A dynamic property that does not depend on time is called *constant of motion*. A force field that makes $T + V$ a constant of motion is called *conservative*.

When a force of the form $\mathbf{F} = e\,(\mathbf{E} + \mathbf{u} \wedge \mathbf{B})$ is considered, like that of Sect. 1.3.2, the scalar multiplication by $\mathrm{d}\mathbf{r} = \mathbf{u}\,\mathrm{d}t$ shows that the second term of the force does not contribute to the work because $\mathbf{u} \wedge \mathbf{B} \cdot \mathbf{u} = 0$ (Sect. A.7). Remembering the first of (1.14), the integral corresponding to that of (1.23) reads

$$\int_P^Q \mathbf{F} \cdot \mathrm{d}\mathbf{r} = -e \int_P^Q \left(\operatorname{grad} \varphi + \frac{\partial \mathbf{A}}{\partial t} \right) \cdot \mathrm{d}\mathbf{r}\,. \tag{1.24}$$

If the electromagnetic field is independent of time, the calculation is the same as in (1.23) and the exerted work is $e\varphi(P) - e\varphi(Q)$.

1.3.4 Hamilton Principle—Synchronous Trajectories

From the analysis of Sect. 1.2 it follows that the solutions $x_i(t)$ of the motion equations (1.9) are the extremum functions of the functional

$$S[\mathbf{r}] = \int_a^b L(\mathbf{r}, \dot{\mathbf{r}}, t)\,\mathrm{d}t\,. \tag{1.25}$$

On the other hand, $\mathbf{r}(t)$ describes the particle's trajectory. The latter is also called *natural trajectory* to distinguish it from the $\mathbf{r} + \delta\mathbf{r}$ trajectories that are obtained through a variation. In summary, the natural trajectory of the particle is the extremum function of (1.25). This statement is called *Hamilton principle*.

The integration boundaries in (1.25) determine a time interval $b-a$ that measures the motion's duration between the initial and final position of the particle, $\mathbf{r}(a)$ and $\mathbf{r}(b)$ respectively. The duration is the same also for the $\mathbf{r}+\delta\mathbf{r}$ trajectories. In fact, remembering the derivation of Sect. 1.2, the variation $\delta\mathbf{r}$ vanishes at the integration boundaries, so that any trajectory obtained through a variation has the same initial and final positions as the natural one at the same instants a and b. Moreover, any position $\mathbf{r}+\delta\mathbf{r}$ between $\mathbf{r}(a)$ and $\mathbf{r}(b)$ is considered at the same instant as the position $\mathbf{r}$ of the natural trajectory. For this reason the $\mathbf{r}+\delta\mathbf{r}$ trajectories of the functional (1.25) are called *synchronous*.

1.4 Generalized Coordinates

The extremum functions are calculated as shown in Sect. 1.3 also when a system of N particles, instead a single particle, is considered. The starting point is still (1.9), where a new index is introduced to distinguish the masses. The number of coordinates that describe the motion of all particles in the system is not necessary equal to $3N$; in fact, a number of constraints may exist that limit the relative positions of the particles. As a consequence, letting $n \leq 3N$ indicate the number of degrees of freedom of the system, the set $x_1(t), \ldots, x_n(t)$ suffices to determine the positions of the particles at time t.

Depending on the problem in hand it may be more convenient to use a new set of coordinates $q_1(t), \ldots, q_n(t)$ instead of the Cartesian set $x_1(t), \ldots, x_n(t)$. For this, it is necessary to extend the calculation of the extremum functions to the case where the new set is used. Let the relation between the old and new coordinates be

$$\begin{cases} q_1 = q_1(x_1, \ldots, x_n, t) \\ \quad \vdots \\ q_n = q_n(x_1, \ldots, x_n, t) \end{cases} \qquad \begin{cases} x_1 = x_1(q_1, \ldots, q_n, t) \\ \quad \vdots \\ x_n = x_n(q_1, \ldots, q_n, t) \end{cases} \tag{1.26}$$

The coordinates q_i, whose units are not necessarily a length, are called *generalized coordinates*. Their time derivatives $\dot{q}_i = \mathrm{d}q_i/\mathrm{d}t$ are called *generalized velocities*. The explicit dependence on time in (1.26) is present if a relative motion of the two frames exists: e.g., the relations $q_1 = x_1 - v_0 t$, $q_2 = x_2$, $q_3 = x_3$ transform the $x_1x_2x_3$ set into the $q_1q_2q_3$ set that moves with respect to the former one with the velocity v_0 along the first axis.

Differentiating q_i twice with respect to time and using the first of (1.26) provide a relation of the form $\ddot{q}_i = \ddot{q}_i(x_1, \dot{x}_1, \ddot{x}_1, \ldots, x_n, \dot{x}_n, \ddot{x}_n, t)$. The above, after eliminating the second derivatives $\ddot{x}_1, \ldots, \ddot{x}_n$ through (1.9), becomes $\ddot{q}_i = \ddot{q}_i(x_1, \dot{x}_1, \ldots, x_n, \dot{x}_n, t)$. Finally, replacing $x_1, \dot{x}_1, \ldots, x_n, \dot{x}_n$ extracted from the second of (1.26) yields

$$\ddot{q}_i = \ddot{q}_i(\mathbf{q}, \dot{\mathbf{q}}, t), \qquad i = 1, \ldots, n, \tag{1.27}$$

where $\mathbf{q}$ indicates the set $q_1, \ldots, q_n$, and $\dot{\mathbf{q}}$ indicates the corresponding derivatives. Equations (1.27) have the same form as (1.9), hence they must be deducible as the extremum functions of a functional. To show this, one starts from (1.25) by writing the Lagrangian function in the new set of coordinates. A rather lengthy calculation based on the chain-differentiation rule yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}_i} = \frac{\partial L}{\partial q_i}, \qquad i = 1, \ldots, n, \tag{1.28}$$

that is, the Lagrange equations written in the q_i coordinates. Specifically, (1.28) turns out to be the Lagrange equations of the functional

$$S[\mathbf{q}] = \int_a^b L(\mathbf{q}, \dot{\mathbf{q}}, t)\, \mathrm{d}t. \tag{1.29}$$

This result is very important because it shows that the Lagrange equations are invariant under a change of coordinates of the type (1.26).

The solution of (1.28) provides the time evolution of the coordinates q_i describing the particles' motion. As (1.28) are n second-order equations, to determine their solution it is necessary to specify at $t = a$ the values of the n functions q_i and of the correspondent derivatives $\dot{q}_i$, namely, a total of $2n$ constants. The function $p_i = \partial L/\partial \dot{q}_i$ is called *generalized momentum* or *conjugate momentum* of q_i. From this definition and from (1.28) it follows

$$p_i = \frac{\partial L}{\partial \dot{q}_i}, \qquad \dot{p}_i = \frac{\partial L}{\partial q_i}, \qquad i = 1, \ldots, n. \tag{1.30}$$

The derivative $\dot{p}_i$ is called *generalized force*. Due to the definitions (1.30), the generalized momentum and force depend on the same coordinates as the Lagrangian function, namely, $p_i = p_i(\mathbf{q}, \dot{\mathbf{q}}, t)$, $\dot{p}_i = \dot{p}_i(\mathbf{q}, \dot{\mathbf{q}}, t)$.

1.5 Hamiltonian Function

From (1.30) one derives the following expression of the total derivative with respect to time of the Lagrangian function:

$$\frac{\mathrm{d}L}{\mathrm{d}t} = \frac{\partial L}{\partial t} + \sum_{i=1}^{n}\left(\frac{\partial L}{\partial q_i}\dot{q}_i + \frac{\partial L}{\partial \dot{q}_i}\ddot{q}_i\right) = \frac{\partial L}{\partial t} + \sum_{i=1}^{n}\left(\dot{p}_i\,\dot{q}_i + p_i\,\ddot{q}_i\right). \tag{1.31}$$

The quantity in parentheses in (1.31) is the time derivative of $p_i\,\dot{q}_i$, so that

$$\frac{\partial L}{\partial t} = -\frac{\mathrm{d}H}{\mathrm{d}t}, \qquad H = \sum_{i=1}^{n} p_i\,\dot{q}_i - L. \tag{1.32}$$

The quantity H defined by (1.32) is called *Hamiltonian function*. Remembering the derivation of the Lagrangian function one observes that L, H, and $p_i\dot{q}_i$ have the units of an energy. In turn, the product energy $\times$ time is called *action*. In particular, the functional (1.29) is called *action integral* [56, Chap. 8]. From the above observation it follows that $q_i\, p_i$ has the units of an action in all coordinate sets.

By way of example one takes the single-particle Lagrangian functions (1.12) and (1.18), where the Cartesian coordinates are used. The momentum conjugate to x_i is given, respectively, by

$$L = \frac{1}{2}\, mu^2 - V \to p_i = m\, u_i\,, \quad L = \frac{1}{2}\, mu^2 - e\varphi + e\mathbf{u}\cdot\mathbf{A} \to p_i = mu_i + eA_i\,. \tag{1.33}$$

The expression of H is found from (1.32) after introducing the vector $\mathbf{p} = (p_1, p_2, p_3)$ and indicating its modulus with p. For the case $L = mu^2/2 - V$ one finds

$$H = \frac{1}{2}\, mu^2 + V = \frac{1}{2m}\, p^2 + V\,, \tag{1.34}$$

while the case $L = mu^2/2 - e\varphi + e\mathbf{u}\cdot\mathbf{A}$ yields

$$H = \frac{1}{2}\, mu^2 + e\varphi = \frac{1}{2m}\, |\mathbf{p} - e\mathbf{A}|^2 + e\varphi\,. \tag{1.35}$$

Still using the Cartesian coordinates, (1.34) is readily extended to the case of a system of particles having n degrees of freedom. The force acting on the ith degree of freedom at time t is given by a generalization of (1.10),

$$m_i\dot{u}_i = -\frac{\partial V}{\partial x_i}\,, \tag{1.36}$$

where the time derivative is calculated at t and the $x_1, \ldots, x_n$ coordinates appearing in V are calculated at t as well. For the sake of simplicity the coordinate index i is also used to distinguish the masses in (1.36). It is implied that the same value of m_i must be applied to the indices associated with the same particle. The Lagrangian function is calculated in the same manner as in Sect. 1.3 and reads

$$L = \sum_{i=1}^{n} \frac{1}{2}\, m_i u_i^2 - V(\mathbf{r}, t)\,, \qquad p_i = m_i u_i\,, \tag{1.37}$$

whence

$$H = \sum_{i=1}^{n} \frac{1}{2}\, m_i u_i^2 + V(\mathbf{r}, t) = \sum_{i=1}^{n} \frac{1}{2m_i}\, p_i^2 + V(\mathbf{r}, t)\,. \tag{1.38}$$

Comparing the two forms of H shown in (1.38), one notes that the second of (1.37) is exploited to express the Hamiltonian function in terms of the **r**, **p** sets instead of the **r**, **u** sets. This procedure is generalized in Sect. 1.6.

1.6 Hamilton Equations

As the Lagrangian function depends on $\mathbf{q}$, $\dot{\mathbf{q}}$, and t, the generalized momentum p_i defined by (1.30) depends on the same variables at most. It is useful to consider also the inverse relations, where the generalized velocities $\dot{q}_i$ are expressed in terms of $\mathbf{q}$, $\mathbf{p}$, and t. The two sets of relations are

$$\begin{cases} p_1 = p_1(q_1, \dot{q}_1, \ldots, q_n, \dot{q}_n, t) \\ \quad \vdots \\ p_n = p_n(q_1, \dot{q}_1, \ldots, q_n, \dot{q}_n, t) \end{cases} \qquad \begin{cases} \dot{q}_1 = \dot{q}_1(q_1, p_1, \ldots, q_n, p_n, t) \\ \quad \vdots \\ \dot{q}_n = \dot{q}_n(q_1, p_1, \ldots, q_n, p_n, t) \end{cases} \tag{1.39}$$

A simple example is given by the two cases of (1.33). Letting $q_i = x_i$, $\dot{q}_i = u_i$, the first case gives (1.39) the form $p_i = m\,\dot{q}_i$ and $\dot{q}_i = p_i/m$, while the second one gives (1.39) the form $p_i = m\,\dot{q}_i + e\,A_i(\mathbf{q}, t)$ and $\dot{q}_i = [p_i - e\,A_i(\mathbf{q}, t)]/m$.

Introducing the second of (1.39) into the definition (1.32) of the Hamiltonian function expresses the latter in terms of $\mathbf{q}$, $\mathbf{p}$, and t. The derivatives of the Hamiltonian function with respect to the new variables q_i, p_i are very significant. In fact, for any index r one finds

$$\frac{\partial H}{\partial q_r} = \sum_{i=1}^{n} p_i \frac{\partial \dot{q}_i}{\partial q_r} - \left(\frac{\partial L}{\partial q_r} + \sum_{i=1}^{n} \frac{\partial L}{\partial \dot{q}_i} \frac{\partial \dot{q}_i}{\partial q_r} \right) = -\frac{\partial L}{\partial q_r} = -\dot{p}_r \,. \tag{1.40}$$

The two sums in (1.40) cancel each other thanks to the first of (1.30), while the last equality is due to the second of (1.30). The derivative with respect to p_r is found by the same token,

$$\frac{\partial H}{\partial p_r} = \left(\dot{q}_r + \sum_{i=1}^{n} p_i \frac{\partial \dot{q}_i}{\partial p_r} \right) - \sum_{i=1}^{n} \frac{\partial L}{\partial \dot{q}_i} \frac{\partial \dot{q}_i}{\partial p_r} = \dot{q}_r \,. \tag{1.41}$$

The results of (1.40, 1.41) are condensed in the *Hamilton equations*

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \,, \qquad \dot{p}_i = -\frac{\partial H}{\partial q_i} \,, \qquad i = 1, \ldots, n \,, \tag{1.42}$$

that provide a set of $2n$ differential equations of the first order in the $2n$ independent unknowns $q_1, \ldots, q_n, p_1, \ldots, p_n$. It is important to note that from (1.42) one readily derives the following:

$$\frac{\partial \dot{q}_i}{\partial q_i} + \frac{\partial \dot{p}_i}{\partial p_i} = \frac{\partial^2 H}{\partial q_i \partial p_i} - \frac{\partial^2 H}{\partial p_i \partial q_i} = 0\,. \tag{1.43}$$

The Hamilton equations (1.42) provide the time evolution of the generalized coordinates q_i; as a consequence, they are equivalent to the Lagrange equations (1.28). Another way of obtaining the Hamilton equations is to derive them as the extremum equations of a suitable functional. This is shown in Sect. 1.7.

In contrast to the Lagrange equations (1.28), that are n second-order differential equations, the Hamilton equations (1.42) are $2n$, first-order differential equations. To determine the solution of the latter it is necessary to prescribe the values of the $2n$ unknowns $q_1, \ldots, q_n, p_1, \ldots, p_n$ at the initial time $t = a$, that is, $2n$ constants. Therefore, the number of constants to be prescribed is the same as for the Lagrange equations. The independent functions $q_1, \ldots, q_n, p_1, \ldots, p_n$ are called *canonical coordinates*. For each index i the functions q_i, p_i are called *conjugate coordinates*. Thanks to (1.42) the total derivative of H reads

$$\frac{\mathrm{d}H}{\mathrm{d}t} = \frac{\partial H}{\partial t} + \sum_{i=1}^{n} \left(\frac{\partial H}{\partial p_i} \dot{p}_i + \frac{\partial H}{\partial q_i} \dot{q}_i \right) = \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}\,, \tag{1.44}$$

where the last equality derives from the first of (1.32). If the Lagrangian function does not depend explicitly on time it follows $\mathrm{d}H/\mathrm{d}t = 0$, namely, H is a constant of motion. Its value is fixed by the values of the canonical coordinates at the initial time $t = a$. From (1.44) it also follows that $\mathrm{d}H/\mathrm{d}t = 0$ is equivalent to $\partial H/\partial t = 0$. In other terms, the Hamiltonian function is a constant of motion if it does not depend explicitly on time, and vice versa.

If the Lagrangian function does not depend on one of the coordinates, say, q_r, the latter is called *cyclic* or *ignorable*. From the second of (1.30) it follows that if q_r is cyclic, its conjugate momentum p_r is a constant of motion. Moreover, due to the second of (1.42) it is $\partial H/\partial q_r = 0$, namely, the Hamiltonian function does not depend on q_r either.

1.7 Time–Energy Conjugacy—Hamilton-Jacobi Equation

Equations (1.42) can also be derived as the extremum equations of a functional. To show this it suffices to replace the Lagrangian function taken from the second of (1.32) into the functional (1.29), thus yielding

$$S = \int_a^b \left(\sum_{i=1}^{n} p_i \dot{q}_i - H \right) \mathrm{d}t\,. \tag{1.45}$$

Using in (1.45) the expressions of the generalized velocities given by the second of (1.39), the integrand becomes a function of q_i, p_i, and t. Then, the extremum

equations are found by introducing the variations in the coordinates, that become $q_i + \alpha_i \eta_i$. Like in the case of (1.1) it is assumed that η_i vanishes at a and b. Similarly, the conjugate momenta become $p_i + \beta_i \zeta_i$. Differentiating (1.45) with respect to α_i or β_i yields, respectively,

$$\frac{\partial S}{\partial \alpha_i} = \int_a^b \left[(p_i + \beta_i \zeta_i)\, \dot{\eta}_i - \frac{\partial H}{\partial (q_i + \alpha_i \eta_i)} \eta_i \right] \mathrm{d}t \,, \tag{1.46}$$

$$\frac{\partial S}{\partial \beta_i} = \int_a^b \left[(\dot{q}_i + \alpha_i \dot{\eta}_i)\, \zeta_i - \frac{\partial H}{\partial (p_i + \beta_i \zeta_i)} \zeta_i \right] \mathrm{d}t \,. \tag{1.47}$$

Letting $\alpha_1 = \ldots = \beta_n = 0$ in (1.46, 1.47), integrating by parts the term containing $\dot{\eta}_i$, and using the condition $\eta_i(a) = \eta_i(b) = 0$ provide

$$\left(\frac{\partial S}{\partial \alpha_i} \right)_0 = - \int_a^b \left(\dot{p}_i + \frac{\partial H}{\partial q_i} \right) \eta_i \, \mathrm{d}t \,, \quad \left(\frac{\partial S}{\partial \beta_i} \right)_0 = \int_a^b \left(\dot{q}_i - \frac{\partial H}{\partial p_i} \right) \zeta_i \, \mathrm{d}t \,. \tag{1.48}$$

As in Sect. 1.2 the equations for the extremum functions are found by letting $(\partial S / \partial \alpha_i)_0 = 0$, $(\partial S / \partial \beta_i)_0 = 0$. Such equations coincide with (1.42). It is worth observing that as no integration by part is necessary for obtaining the second of (1.48), the derivation of (1.42) does not require any prescription for the boundary conditions of ζ_i. On the other hand, considering that in the Hamilton equations q_i and p_i are independent variables, one can add the prescription $\zeta_i(a) = \zeta_i(b) = 0$. Although the latter is not necessary here, it becomes useful in the treatment of the canonical transformations, as shown in Sect. 2.2.

In the coordinate transformations discussed so far, time was left unchanged. This aspect is not essential: in fact, within the coordinate transformation one can replace t with another parameter that depends single-valuedly on t. This parameter, say, $\theta(t)$, is equally suitable for describing the evolution of the particles' system; proceeding in this way transforms (1.45) into

$$S = \int_{\theta(a)}^{\theta(b)} \left(\sum_{i=1}^{n} p_i \dot{q}_i \frac{\mathrm{d}t}{\mathrm{d}\theta} - H \frac{\mathrm{d}t}{\mathrm{d}\theta} \right) \mathrm{d}\theta = \int_{\theta(a)}^{\theta(b)} \left(\sum_{i=1}^{n} p_i q_i' - H t' \right) \mathrm{d}\theta \,, \tag{1.49}$$

where the primes indicate the derivatives with respect to θ. Now, letting $q_{n+1} = t$, $p_{n+1} = -H$, the Lagrangian function is recast in the more compact form $L = \sum_{i=1}^{n+1} p_i q_i'$. Remembering the definition (1.30) of the conjugate momenta, it follows that the latter becomes $p_i = \partial L / \partial q_i'$. In conclusion, the negative Hamiltonian function is the momentum conjugate to θ. This result is not due to any particular choice of the relation $\theta(t)$, hence it holds also for the identical transformation $\theta = t$; in other terms, $-H$ is the momentum conjugate to t.

If the upper limit b in the action integral S (Eq. (1.29)) is considered as a variable, the Lagrangian is found to be the total time derivative of S. Letting $b \leftarrow t$, from the (1.45) form of S one derives its total differential

$$\mathrm{d}S = \sum_{i=1}^{n} p_i \,\mathrm{d}q_i - H\,\mathrm{d}t\,. \tag{1.50}$$

As a consequence it is $p_i = \partial S/\partial q_i$, $H = -\partial S/\partial t$. Remembering that H depends on the generalized coordinates and momenta, and on time, one may abridge the above findings into the relation

$$\frac{\partial S}{\partial t} + H\left(q_1, \ldots, q_n\,, \frac{\partial S}{\partial q_1}, \ldots, \frac{\partial S}{\partial q_n}\,, t\right) = 0\,, \qquad p_i = \frac{\partial S}{\partial q_i}\,, \tag{1.51}$$

that is, a partial-differential equation in the unknown function S. The former is called *Hamilton-Jacobi equation*, while the latter in this context is called *Hamilton's principal function*. As (1.51) is a first-order equation in the $n + 1$ variables $q_1, \ldots, q_n$, t, the solution S contains $n + 1$ integration constants. One of them is an additive constant on S, as is easily found by observing that (1.51) contains the derivatives of S, not S itself. For this reason the additive constant is irrelevant and can be set to zero, so that the integration constants reduce to n. It is shown in Sect. 2.2 that (1.51) provides the time evolution of the generalized coordinates q_i. As a consequence it is equivalent to the Lagrange equations (1.28) and to the Hamilton equations (1.42) for describing the system's dynamics.

1.8 Poisson Brackets

Let ϱ, σ be arbitrary functions of the canonical coordinates, differentiable with respect to the latter. The *Poisson bracket* of ϱ and σ is defined as the function[3]

$$[\varrho, \sigma] = \sum_{i=1}^{n} \left(\frac{\partial \varrho}{\partial q_i} \frac{\partial \sigma}{\partial p_i} - \frac{\partial \varrho}{\partial p_i} \frac{\partial \sigma}{\partial q_i} \right)\,. \tag{1.52}$$

From (1.52) it follows $[\varrho, \sigma] = -[\sigma, \varrho]$, $[\varrho, \varrho] = 0$. Also, due to (1.42) it is

$$\frac{\mathrm{d}\varrho}{\mathrm{d}t} = \frac{\partial \varrho}{\partial t} + [\varrho, H]\,. \tag{1.53}$$

Letting $\varrho = H$ shows that (1.44) is a special case of (1.53). If ϱ is a constant of motion, then

[3]The definition and symbol (1.52) of the Poisson bracket conform to those of [56, Sect. 9-5]. In [84, Sect. 42], instead, the definition has the opposite sign and the symbol $\{\varrho, \sigma\}$ is used. In [136, Sect. 11] the definition is the same as that adopted here, while the symbol $\{\varrho, \sigma\}$ is used.

$$\frac{\partial \varrho}{\partial t} + [\varrho, H] = 0\,. \tag{1.54}$$

If ϱ does not depend explicitly on time, (1.53) yields

$$\frac{\mathrm{d}\varrho}{\mathrm{d}t} = [\varrho, H] \tag{1.55}$$

where, in turn, the right-hand side is equal to zero if ϱ is a constant of motion, while it is different from zero in the other case. Special cases of the Poisson bracket are

$$\big[q_i, q_j\big] = 0\,, \qquad \big[p_i, p_j\big] = 0\,, \qquad \big[q_i, p_j\big] = \delta_{ij}\,, \tag{1.56}$$

with δ_{ij} the Kronecker symbol (A.18). Other interesting expressions are found by introducing the $2\,n$-dimensional vectors $\mathbf{s}$, $\mathbf{e}$ defined as

$$\mathbf{s} = \begin{bmatrix} q_1 \\ \vdots \\ q_n \\ p_1 \\ \vdots \\ p_n \end{bmatrix}\,, \qquad \mathbf{e} = \begin{bmatrix} \partial H/\partial p_1 \\ \vdots \\ \partial H/\partial p_n \\ -\partial H/\partial q_1 \\ \vdots \\ -\partial H/\partial q_n \end{bmatrix}\,. \tag{1.57}$$

Using the definitions in (1.57) one finds

$$\dot{\mathbf{s}} = \mathbf{e}\,, \qquad \mathrm{div}_{\mathbf{s}}\,\dot{\mathbf{s}} = \sum_{i=1}^{n} \left(\frac{\partial \dot{q}_i}{\partial q_i} + \frac{\partial \dot{p}_i}{\partial p_i} \right) = 0\,, \tag{1.58}$$

the first of which expresses the Hamilton equations (1.42) in vector form, while the second one derives from (1.43). The symbol div_s indicates the divergence with respect to all the variables that form vector $\mathbf{s}$ (Sect. A.3). Now, taking an arbitrary function ϱ like that used in (1.53) and calculating the divergence of the product $\varrho\,\dot{\mathbf{s}}$ yields, thanks to (1.58) and to (A.16, A.12),

$$\mathrm{div}_{\mathbf{s}}(\varrho\,\dot{\mathbf{s}}) = \varrho\,\mathrm{div}_{\mathbf{s}}\,\dot{\mathbf{s}} + \dot{\mathbf{s}} \cdot \mathrm{grad}_{\mathbf{s}}\varrho = \sum_{i=1}^{n} \left(\frac{\partial \varrho}{\partial q_i}\,\dot{q}_i + \frac{\partial \varrho}{\partial p_i}\,\dot{p}_i \right) = [\varrho, H]\,. \tag{1.59}$$

1.9 Phase Space and State Space

Given a system of particles having n degrees of freedom it is often convenient to describe its dynamic properties by means of a geometrical picture. To this purpose one introduces a $2n$-dimensional space whose coordinates are $q_1, \ldots, q_n, p_1, \ldots, p_n$.

This space has no definite metrical structure; one simply assumes that q_i and p_i are plotted as Cartesian coordinates of an Euclidean space [82, Chap. 6-5]. Following Gibbs, the space thus defined is often called *phase space*. However, it is convenient to better specify the terminology by using that of Ehrenfest, in which the term γ-*space* is used for this space (the citations of Gibbs and Ehrenfest are in [136, Sect. 17]). At some instant t the whole set of canonical coordinates $q_1, \ldots, q_n$, $p_1, \ldots, p_n$ corresponds to a point of the γ-space. Such a point is called *phase point*. In turn, the *state* of a mechanical system at some instant t is defined as the set of its canonical coordinates at that instant. It follows that the phase point represents the dynamic state of the system at t. As time evolves, the phase points representing the state at different instants provide a curve of the γ-space called *phase trajectory*.

A generalization of the γ-space is obtained by adding the time t as a $(2n + 1)$th coordinate. The $(2n + 1)$-dimensional space thus obtained is called *state space* [82, Chap. 6-5]. The curve of the state space describing the system's dynamics is called *state trajectory*. Consider two systems governed by the same Hamiltonian function and differing only in the initial conditions. The latter are represented by two distinct points of the $2n$-dimensional section of the state space corresponding to $t = 0$. The subsequent time evolution of the two systems provides two state trajectories that never cross each other. In fact, if a crossing occurred at, say, $t = \bar{t}$, the canonical coordinates of the two Hamiltonian functions would be identical there, thus making the initial conditions of the subsequent motion identical as well. As a consequence, the two state trajectories would coincide for $t \geq \bar{t}$. However, the same reasoning holds when considering the motion backward in time ($t \leq \bar{t}$). Thus, the two trajectories should coincide at all times, thus contradicting the hypothesis that the initial conditions at $t = 0$ were different.

A similar reasoning about the crossing of trajectories is possible in the γ-space. The conclusion is that the phase trajectories do not cross each other if the Hamiltonian function does not depend explicitly on time. Instead, they may cross each other if the Hamiltonian function depends on time; the crossing, however, occurs at different times (in other terms, the set of canonical coordinates of the first system calculated at $t = t_1$ may coincide with the set of canonical coordinates of the second system calculated at $t = t_2$ only in the case $t_2 \neq t_1$).

Despite of the larger number of dimensions, the adoption of the state space is convenient for the geometrical representation of the system's dynamics, because a trajectory is uniquely specified by the initial point and no crossing of trajectories occurs. With the provision stated above, this applies also to the γ-space. In contrast, consider a geometrical picture of the Lagrangian type, in which the generalized coordinates $q_1, \ldots, q_n$ only are used. The latter may be considered as the Cartesian coordinates of an n-dimensional Euclidean space called *configuration space*. To specify a trajectory in such a space it is necessary to prescribe the position $q_1, \ldots, q_n$ and velocity $\dot{q}_1, \ldots, \dot{q}_n$ of the system at $t = 0$. If one considers two or more systems differing only in the initial conditions, the motion of each system could start from every point of the configuration space and in every direction. As a consequence, it would be impossible to obtain an ordered picture of the trajectories, which will inevitably cross each other.

As mentioned above, the γ-space for a system having n degrees of freedom is a $2n$-dimensional space whose coordinates are $q_1, \ldots, q_n, p_1, \ldots, p_n$. It is sometimes convenient to use a different type of phase space whose dimension is twice the number of degrees of freedom possessed by each of the system's particles. To specify this issue, consider the case of a system made of N point-like particles, with no constraints. In this case each particle (say, the jth one) has 3 degrees of freedom and its dynamic state at the time t is determined by the 6 canonical coordinates $\bar{q}_{1j}, \bar{q}_{2j}, \bar{q}_{3j}, \bar{p}_{1j}, \bar{p}_{2j}, \bar{p}_{3j}$. Together, the latter identify a point X_j of a 6-dimensional phase space called μ-*space*.[4] At the time t the system as a whole is represented in the μ-space by the set of N points $X_1, \ldots, X_N$.

1.10 Complements

1.10.1 Higher-Order Variational Calculus

The variational calculus described in Sect. 1.2 can be extended to cases where the function g in (1.1) depends on derivatives of a higher order than the first. Consider for instance the functional

$$G[w] = \int_a^b g(w, \dot{w}, \ddot{w}, \xi)\, \mathrm{d}\xi \,. \tag{1.60}$$

Following the procedure of Sect. 1.2 and assuming that the derivative $\dot{\eta}$ vanishes at a and b yield the following differential equation for the extremum functions of (1.60):

$$-\frac{\mathrm{d}^2}{\mathrm{d}\xi^2}\frac{\partial g}{\partial \ddot{w}} + \frac{\mathrm{d}}{\mathrm{d}\xi}\frac{\partial g}{\partial \dot{w}} = \frac{\partial g}{\partial w} \,. \tag{1.61}$$

1.10.2 Lagrangian Invariance and Gauge Invariance

It is shown in Sect. 1.2 that the extremum functions $w_i(\xi)$ are invariant under addition to g of the total derivative of an arbitrary function h that depends on $\mathbf{w}$ and ξ only (refer to Eq. (1.8)). Then, it is mentioned in Sect. 1.3.2 that the $\mathbf{E}$ and $\mathbf{B}$ fields are invariant under the gauge transformation (1.19), where $h(\mathbf{r}, t)$ is an arbitrary function. These two properties have in fact the same origin, namely, the description based upon a Lagrangian function. In fact, as shown in Sect. 4.2, a Lagrangian description is possible also in the case of a system having a continuous distribution of the degrees of freedom like, for instance, the electromagnetic field.

[4] The letter "μ" stands for "molecule," whereas the letter "γ" in the term "γ-space" stands for "gas."

1.10.3 Variational Calculus with Constraints

In several problems it is required that the function w, introduced in Sect. 1.2 as the extremum function of functional (1.1), be able to fulfill one or more constraints. By way of example consider the constraint

$$G_0 = \int_a^b g_0(w, \dot{w}, \xi)\, \mathrm{d}\xi\,, \tag{1.62}$$

where the function g_0 and the number G_0 are prescribed. A typical case where (1.62) occurs is that of finding the maximum area bounded by a perimeter of given length (*Dido's problem*). For this reason, extremum problems having a constraint like (1.62) are called *isoperimetric* even when they have no relation with geometry [143, Par. 4-1].

To tackle the problem one extends the definition of the variation of w by letting $\delta w = \alpha_1 \eta_1 + \alpha_2 \eta_2$, where $\eta_1(\xi)$, $\eta_2(\xi)$ are arbitrary functions that are differentiable in the interior of $[a, b]$ and fulfill the conditions $\eta_1(a) = \eta_1(b) = 0$, $\eta_2(a) = \eta_2(b) = 0$. If w is an extremum function of G that fulfills (1.62), replacing w with $w + \delta w$ transforms (1.1, 1.62) to a pair of functions of the α_1, α_2 parameters, namely,

$$G = G(\alpha_1, \alpha_2)\,, \qquad G_0(\alpha_1, \alpha_2) = G_0(0, 0) = \text{const.} \tag{1.63}$$

The first of (1.63) has an extremum at $\alpha_1 = \alpha_2 = 0$, while the second one establishes a relation between α_1 and α_2. The problem is thus reduced to that of calculating a constrained extremum, and is solved by the method of the *Lagrange multipliers*, described in Sect. B.6. For this, one considers the function $G_\lambda = G(\alpha_1, \alpha_2) + \lambda\, G_0(\alpha_1, \alpha_2)$, with λ an indeterminate parameter, and calculates the free extremum of G_λ by letting

$$\left(\frac{\partial G_\lambda}{\partial \alpha_1}\right)_0 = 0\,, \qquad \left(\frac{\partial G_\lambda}{\partial \alpha_2}\right)_0 = 0\,, \tag{1.64}$$

where index 0 stands for $\alpha_1 = \alpha_2 = 0$. The rest of the calculation is the same as in Sect. 1.2; the two relations (1.64) turn out to be equivalent to each other and provide the same Euler equation. More specifically, from the definition of G and G_0 as integrals of g and g_0 one finds that the Euler equation of this case is obtained from that of Sect. 1.2 by replacing g with $g_\lambda = g + \lambda g_0$:

$$\frac{\mathrm{d}}{\mathrm{d}\xi}\frac{\partial g_\lambda}{\partial \dot{w}} = \frac{\partial g_\lambda}{\partial w}\,. \tag{1.65}$$

As (1.65) is a second-order equation, its solution w contains two integration constants. The λ multiplier is an additional indeterminate constant. The three constants are found from the constraint (1.62) and from the two relations provided by the boundary or initial conditions of w.

1.10.4 An Interesting Example of Extremum Equation

Consider the Hamilton-Jacobi equation (1.51) for a single particle of mass m. Using the Cartesian coordinates and a Hamiltonian function of the form

$$H = \frac{p^2}{2m} + V(x_1, x_2, x_3, t)\,, \quad p^2 = p_1^2 + p_2^2 + p_3^2\,, \tag{1.66}$$

the Hamilton-Jacobi equation reads

$$\frac{\partial S}{\partial t} + \frac{|\operatorname{grad} S|^2}{2m} + V(x_1, x_2, x_3, t) = 0\,, \qquad p_i = \frac{\partial S}{\partial q_i}\,. \tag{1.67}$$

If V is independent of time, then $H = E$ and the separation $S = W - E\,t$ (Sect. 2.4) yields $\partial S/\partial t = -E$, $\operatorname{grad} S = \operatorname{grad} W = \mathbf{p}$. It follows

$$\frac{|\operatorname{grad} W|^2}{2m} + V(x_1, x_2, x_3) = E\,. \tag{1.68}$$

Both Hamilton's principal (S) and characteristic (W) functions have the dimensions of an action and are defined apart from an additive constant. Also, the form of $|\operatorname{grad} W|$ is uniquely defined by that of $V - E$. In turn, E is prescribed by the initial conditions of the particle's motion.

Consider now the case where $E \geq V$ within a closed domain Ω whose boundary is $\partial\Omega$. As grad W is real, the motion of the particle is confined within Ω, and grad W vanishes at the boundary $\partial\Omega$. The Hamilton-Jacobi equation for W (1.68) is recast in a different form by introducing an auxiliary function w such that

$$w = w_0 \exp(W/\mu)\,, \tag{1.69}$$

with μ a constant having the dimensions of an action. The other constant w_0 is used for prescribing the dimensions of w. Apart from this, the choice of w_0 is arbitrary due to the arbitrariness of the additive constant of W. Taking the gradient of (1.69) yields $\mu\ \operatorname{grad} w = w\ \operatorname{grad} W$, with $w \neq 0$ due to the definition. As grad W vanishes at the boundary, grad w vanishes there as well. As a consequence, w is constant over the boundary. Inserting (1.69) into (1.68) yields

$$\frac{\mu^2}{2m}\frac{|\operatorname{grad} w|^2}{w^2} + V(x_1, x_2, x_3) = E\,, \tag{1.70}$$

which determines $|\operatorname{grad} w/w|$ as a function of $V - E$. Rearranging the above and observing that $\operatorname{div}(w \operatorname{grad} w) = w\nabla^2 w + |\operatorname{grad} w|^2$ (Sect. A.1) provides

$$\frac{\mu^2}{2m}\left[\operatorname{div}(w \operatorname{grad} w) - w\nabla^2 w\right] + (V - E)\,w^2 = 0\,. \tag{1.71}$$

Integrating (1.71) over Ω and remembering that grad w vanishes at the boundary,

$$\int_{\Omega} w \left[-\frac{\mu^2}{2m} \nabla^2 w + (V - E)\, w \right] \mathrm{d}\Omega = 0\,. \tag{1.72}$$

The term in brackets of (1.72) does not necessarily vanish. In fact, the form of w is such that only the integral as a whole vanishes. On the other hand, by imposing that the term in brackets vanishes, and replacing μ with the reduced Planck constant $\hbar$, yield

$$-\frac{\hbar^2}{2m} \nabla^2 w + (V - E)\, w = 0\,, \tag{1.73}$$

namely, the Schrödinger equation independent of time (7.44). This result shows that the Schrödinger equation derives from a stronger constraint than that prescribed by the Hamilton-Jacobi equation.

An immediate consequence of replacing the integral relation (1.72) with the differential equation (1.73) is that the domain of w is not limited any more by the condition $E \geq V$, but may extend to infinity. Another consequence is that if the boundary conditions are such that w vanishes over the boundary (which, as said above, may also be placed at infinity), then (1.73) is solvable only for specific values of E, that form its spectrum of eigenvalues. Moreover it can be demonstrated, basing on the form of the Schrödinger equation, that the condition $E \geq V_{\min}$ must be fulfilled (Sect. 8.2.3).

It is interesting to note another relation between the Schrödinger and the Hamilton-Jacobi equations. For the sake of simplicity one takes the one-dimensional case of the Hamilton-Jacobi equation expressed in terms of w (1.70):

$$\frac{\mu^2}{2m} (w')^2 + V(x)\, w^2 = E\, w^2\,, \tag{1.74}$$

where the prime indicates the derivative with respect to x. The left-hand side of the equation may be considered the generating function $g = g(w, w', x)$ of a functional G, defined over an interval of the x axis that may extend to infinity:

$$G[w] = \int_a^b \left[\frac{\mu^2}{2m} (w')^2 + V\, w^2 \right] \mathrm{d}x\,. \tag{1.75}$$

One then seeks the extremum function w of G that fulfills the constraint

$$G_0[w] = \int_a^b w^2 \, \mathrm{d}x = 1\,. \tag{1.76}$$

The problem is solved by the method of 1.10.3, namely, by letting $g_0 = w^2$, $g_E = g - E\, g_0$, and applying the Euler equation to g_E:

$$\frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial g_E}{\partial w'} = \frac{\mathrm{d}}{\mathrm{d}x}\frac{\mu^2}{m}w' = \frac{\mu^2}{m}w'', \qquad \frac{\partial g_E}{\partial w} = 2\,(V-E)\,w\,, \tag{1.77}$$

showing that the Schrödinger equation is actually the Euler equation of the functional G subjected to the constraint G_0, with the eigenvalue E provided by the Lagrange multiplier. This result holds also in the higher-dimensional cases (Prob. 1.6), and is in fact the method originally used by Schrödinger to determine the time-independent equation [118, Eqs. (23,24)].

1.10.5 Constant-Energy Surfaces

Consider the γ-space for a system having n degrees of freedom (Sect. 1.9). If the system is conservative, the relation $H(q_1,\ldots,q_n,\,p_1,\ldots,p_n) = E$ introduces a constraint among the canonical coordinates. Due to this, at each instant of time the latter must belong to the $(2n-1)$-dimensional surface $H = E$ of the phase space, that is called *constant-energy surface*. As E is prescribed by the initial conditions, the phase point of a conservative system always belongs to the same constant-energy surface.

For a system having one degree of freedom the relation describing the constant-energy surface reduces to $H(q,p) = E$ that describes a curve in the $q\,p$ plane. The corresponding state trajectory is a curve of the three-dimensional $q\,p\,t$ space.

Problems

1.1 In the xy plane find the geodesic $y = y(x)$ through the points $A \equiv (a, y_a)$, $B \equiv (b, y_b)$, $A \neq B$.

1.2 Given the Hamiltonian function $H = p^2/(2\,m) + (c/2)\,x^2$, $m, c > 0$ (that describes the *linear harmonic oscillator*, Sect. 3.3), find the constant-energy curves of the $x\,p$ plane corresponding to different values of the total energy E.

1.3 Given the Hamiltonian function of the harmonic oscillator of the general form $H = p^2/(2\,m) + (c/s)\,|x|^s$, $m, c, s > 0$, find the constant-energy curves of the $x\,p$ plane corresponding to a fixed total energy E and to different values of parameter s.

1.4 Draw the state trajectory of a linear harmonic oscillator.

1.5 With reference to the harmonic oscillator of the general form used in Prob. 1.3, draw the state trajectory in the $s \to \infty$ limit. Hint: from the analysis of Prob. 3.1 it turns out that the problem reduces to describing the motion of a particle within a square well.

1.6 Extend to the three-dimensional case the derivation of the Schrödinger equation (1.77) from a Hamilton-Jacobi equation subjected to a constraint. Hint: use the results of Sect. 4.2.

Chapter 2
Coordinate Transformations and Invariance Properties

2.1 Introduction

An important generalization of the subject of coordinate transformation is that of canonical transformation, which leads to the concept of generating function and, through it, to the definition of the principal function and characteristic function of Hamilton. The principal function is found to coincide with the solution of the Hamilton-Jacobi equation introduced in the previous chapter, thus showing the equivalence of the approach based on the variational principles with that based on the canonical transformations. Connected with the Hamiltonian formalism is also the important concept of separability. Still based on Hamilton's principal function is the concept of phase velocity applied to a mechanical system, that brings about an analogy with the electromagnetic theory. The aspects mentioned above give another indication about the generality that is reached by the formalism of Analytical Mechanics illustrated in this part of the book.

Another fundamental issue is that of the invariance properties of the mechanical systems. It is shown that, basing only on the observation of symmetries possessed by the Lagrangian function or other functions connected to it, one derives the existence of invariance properties of the system. Among these are the constants of motion, namely, the dynamic properties that are constant in time and are therefore known from the motion's initial condition.

Of special relevance among the constants of motion are the total energy, the total momentum, and the total angular momentum of the system. The conservation of the total energy is related to the uniformity of time, that of the total momentum is related to the uniformity of space, and, finally, that of the total angular momentum is related to the isotropy of space. Besides the theoretical interest connected to it, the knowledge of a constant of motion is important also for practical purposes: by introducing a known constraint among the canonical coordinates, it is of use in the separation procedure.

M. Rudan, *Physics of Semiconductor Devices*, DOI 10.1007/978-3-319-63154-7_2

This chapter completes the illustration of the basic principles of Analytical Mechanics started in the previous one. The purpose of the two chapters is to introduce a number of concepts that are not only useful *per se*, but also constitute a basis for the concepts of Quantum Mechanics that are introduced in later chapters. The first subject is that of the canonical transformations, followed by the definition and properties of the Hamilton characteristic function and of the phase velocity. Then, the invariance properties that derive from the symmetries of the Lagrangian function are discussed. The chapter continues with a short description of the Maupertuis principle and of the expression of the angular momentum in spherical coordinates. The last paragraphs deal with the linear motion and the action-angle variables.

2.2 Canonical Transformations

Section 1.4 introduced the generalized coordinates $q_1, \ldots, q_n$ that are defined by the first of (1.26) starting from a set of Cartesian coordinates $x_1, \ldots, x_n$. From this, one defines the generalized velocities $\dot{q}_i$ and, from the second of (1.26), calculates the Lagrangian function in the new variables $L(\mathbf{q}, \dot{\mathbf{q}}, t)$. The conjugate momenta p_i are derived from the latter using the first of (1.30) and, finally, the new Hamiltonian function is determined from the second of (1.32). From this, the Hamilton equations (1.42) in the new coordinates are deduced. The process depicted here is a series of replacement, elimination, and differentiation steps.

Relations like (1.26), that transform a set of coordinates into another one, are called *point transformations*. It has been observed in Sect. 1.6 that the canonical coordinates $q_1, \ldots, q_n, p_1, \ldots, p_n$ are mutually independent. It follows that the point transformations are not the most general coordinate transformations, because they act on the $q_1, \ldots, q_n$ only. The most general transformations act simultaneously on the generalized coordinate and momenta, hence they have the form

$$\begin{cases} \tilde{q}_1 = \tilde{q}_1(q_1, \ldots, q_n, p_1, \ldots, p_n, t) \\ \quad \vdots \\ \tilde{q}_n = \tilde{q}_n(q_1, \ldots, q_n, p_1, \ldots, p_n, t) \\ \tilde{p}_1 = \tilde{p}_1(q_1, \ldots, q_n, p_1, \ldots, p_n, t) \\ \quad \vdots \\ \tilde{p}_n = \tilde{p}_n(q_1, \ldots, q_n, p_1, \ldots, p_n, t) \end{cases} \tag{2.1}$$

where q_i, p_i indicate the old canonical coordinates and $\tilde{q}_i, \tilde{p}_i$ indicate the new ones. If H is the Hamiltonian function in the old coordinates, introducing into H the inverse transformations of (2.1) yields a function $\tilde{H}$ that depends on the new coordinates and on time.

For an arbitrary choice of the form of (2.1) it is not possible in general to deduce from $\tilde{H}$ the Hamilton equations in the new coordinates. For this it is in fact necessary to limit the choice of the transformation (2.1) to the cases where the resulting $\tilde{H}$ is a Hamiltonian function proper, namely, it is such that

$$\frac{\mathrm{d}\tilde{q}_i}{\mathrm{d}t} = \frac{\partial \tilde{H}}{\partial \tilde{p}_i}\,, \qquad \frac{\mathrm{d}\tilde{p}_i}{\mathrm{d}t} = -\frac{\partial \tilde{H}}{\partial \tilde{q}_i}\,, \qquad i = 1, \ldots, n\,. \tag{2.2}$$

are fulfilled. The transformations (2.1) that make (2.2) to hold are called *canonical transformations*. The procedure by which the Hamilton equations in the old coordinates are found has been illustrated in Sect. 1.6 and is based on the derivation of the extremum equation of the action integral (1.45). To obtain (2.2) the same calculation must be repeated based on the action integral defined in the new coordinates. It follows that for two sets of coordinates q_i, p_i and $\tilde{q}_i, \tilde{p}_i$ connected by a canonical transformation, the following must hold simultaneously:

$$S = \int_a^b \left(\sum_{i=1}^n p_i \frac{\mathrm{d}q_i}{\mathrm{d}t} - H \right) \mathrm{d}t\,, \qquad \tilde{S} = \int_a^b \left(\sum_{i=1}^n \tilde{p}_i \frac{\mathrm{d}\tilde{q}_i}{\mathrm{d}t} - \tilde{H} \right) \mathrm{d}t\,. \tag{2.3}$$

The difference between the two integrals in (2.3) can be set equal to an arbitrary constant because the calculation uses only the variations of S or $\tilde{S}$. As the limits of the two integrals are fixed to the same values a and b, the constant can in turn be written as the integral between a and b of the total time derivative of an arbitrary function K. In this way the relation between the two integrands in (2.3) reads

$$\sum_{i=1}^n p_i \frac{\mathrm{d}q_i}{\mathrm{d}t} - H = \sum_{i=1}^n \tilde{p}_i \frac{\mathrm{d}\tilde{q}_i}{\mathrm{d}t} - \tilde{H} + \frac{\mathrm{d}K}{\mathrm{d}t}\,. \tag{2.4}$$

It is worth reminding that in the derivation of the Hamilton equations in Sect. 1.7 it is assumed that all variations of generalized coordinates and momenta vanish at the integration limits. Here this applies to both the old and new sets of coordinates. As a consequence, K can be made to depend on all $4\,n$ coordinates q_i, p_i, $\tilde{q}_i$, $\tilde{p}_i$, and on time t. Due to the $2\,n$ relations (2.1) that define the canonical transformation, only $2\,n$ coordinates are independent. As a consequence K can be made to depend on $2\,n$ coordinates chosen among the $4\,n$ available ones, and on time. The most interesting cases are those where K has one of the following forms [56, Sect. 9-1]:

$$K_1 = K_1(\mathbf{q}, \tilde{\mathbf{q}}, t)\,, \quad K_2 = K_2(\mathbf{q}, \tilde{\mathbf{p}}, t)\,, \quad K_3 = K_3(\mathbf{p}, \tilde{\mathbf{q}}, t)\,, \quad K_4 = K_4(\mathbf{p}, \tilde{\mathbf{p}}, t)\,. \tag{2.5}$$

By way of example, select the first form: replacing K_1 into (2.4), calculating $\mathrm{d}K_1/\mathrm{d}t$, and multiplying both sides by $\mathrm{d}t$ yield

$$\sum_{i=1}^{n}\left(\frac{\partial K_1}{\partial q_i}-p_i\right)\mathrm{d}q_i+\sum_{i=1}^{n}\left(\frac{\partial K_1}{\partial \tilde{q}_i}+\tilde{p}_i\right)\mathrm{d}\tilde{q}_i+\left(\frac{\partial K_1}{\partial t}+H-\tilde{H}\right)\mathrm{d}t=0\,, \tag{2.6}$$

where the left-hand side is a total differential in the independent variables q_i, $\tilde{q}_i$, and t. To fulfill (2.6) the parentheses must vanish independently from each other, whence

$$p_i=\frac{\partial K_1}{\partial q_i}\,,\qquad \tilde{p}_i=-\frac{\partial K_1}{\partial \tilde{q}_i}\,,\qquad \tilde{H}=H+\frac{\partial K_1}{\partial t}\,,\qquad i=1,\ldots,n\,. \tag{2.7}$$

As K_1 is prescribed, the first two equations in (2.7) provide $2\,n$ relations involving the $4\,n$ coordinates $q_i, p_i, \tilde{q}_i, \tilde{p}_i$, that constitute the canonical transformations sought. Using the latter one expresses the right-hand side of the third of (2.7) in terms of $\tilde{q}_i, \tilde{p}_i$, thus yielding the new Hamiltonian function $\tilde{H}(\tilde{\mathbf{q}},\tilde{\mathbf{p}},t)$. The procedure is the same for the other functions listed in (2.5), that can all be defined starting from K_1. In fact, letting

$$K_2(\mathbf{q},\tilde{\mathbf{p}},t)=K_1(\mathbf{q},\tilde{\mathbf{q}},t)+\sum_{i=1}^{n}\tilde{p}_i\tilde{q}_i\,, \tag{2.8}$$

and applying the same procedure used to determine (2.7) yield

$$p_i=\frac{\partial K_2}{\partial q_i}\,,\qquad \tilde{q}_i=\frac{\partial K_2}{\partial \tilde{p}_i}\,,\qquad \tilde{H}=H+\frac{\partial K_2}{\partial t}\,,\qquad i=1,\ldots,n\,. \tag{2.9}$$

In (2.8) the independent variables are $q_i, \tilde{p}_i$, so that the coordinates $\tilde{q}_i$ are expressed through them. Similarly, when K_3 is used one lets

$$K_3(\mathbf{p},\tilde{\mathbf{q}},t)=K_1(\mathbf{q},\tilde{\mathbf{q}},t)-\sum_{i=1}^{n}p_iq_i\,, \tag{2.10}$$

to find

$$q_i=-\frac{\partial K_3}{\partial p_i}\,,\qquad \tilde{p}_i=-\frac{\partial K_3}{\partial \tilde{q}_i}\,,\qquad \tilde{H}=H+\frac{\partial K_3}{\partial t}\,,\qquad i=1,\ldots,n\,. \tag{2.11}$$

Finally, in the case of K_4 one lets

$$K_4(\mathbf{p},\tilde{\mathbf{p}},t)=K_1(\mathbf{q},\tilde{\mathbf{q}},t)+\sum_{i=1}^{n}\tilde{p}_i\tilde{q}_i-\sum_{i=1}^{n}p_iq_i\,, \tag{2.12}$$

whence

$$q_i = -\frac{\partial K_4}{\partial p_i}\,, \qquad \tilde{q}_i = \frac{\partial K_4}{\partial \tilde{p}_i}\,, \qquad \tilde{H} = H + \frac{\partial K_4}{\partial t}\,, \qquad i = 1, \ldots, n\,. \tag{2.13}$$

Regardless of the choice of K, the relation between the old and new Hamiltonian function is always of the form $\tilde{H} = H + \partial K/\partial t$. As the canonical transformation is completely determined when K is prescribed, the latter is called *generating function of the canonical transformation*. Two interesting examples are those produced by the generating functions $F_1 = \sum_{i=1}^{n} q_i \tilde{q}_i$ and $F_2 = \sum_{i=1}^{n} q_i \tilde{p}_i$. Applying (2.7) to F_1 yields $\tilde{q}_i = p_i$ and $\tilde{p}_i = -q_i$. As a consequence, the effect of the transformation generated by F_1 is that of exchanging the roles of the generalized coordinates and momenta. This result shows that the distinction between coordinates and momenta is not fundamental, namely, these two groups of variables globally constitute a set of $2n$ independent coordinates. Applying (2.9) to F_2 provides the identical transformation $\tilde{q}_i = q_i$, $\tilde{p}_i = p_i$. A generalization of this example is found using $F_2 = \sum_{i=1}^{n} z_i(\mathbf{q}, t)\tilde{p}_i$, where z_i are arbitrary functions. The new coordinates are in this case $\tilde{q}_i = z_i(\mathbf{q}, t)$ which, as indicated at the beginning of this section, are point transformations. This example shows that all point transformations are canonical.

2.3 An Application of the Canonical Transformation

The discussion of Sect. 2.2 has shown that a canonical transformation based on an arbitrary generating function K brings a Hamiltonian function $H(\mathbf{q}, \mathbf{p}, t)$ into a new one $\tilde{H}(\tilde{\mathbf{q}}, \tilde{\mathbf{p}}, t)$. One may then exploit the arbitrariness of K to obtain the form of $\tilde{H}$ that is most convenient for solving the problem in hand. For instance, remembering the definition of cyclic coordinate given in Sect. 1.6, one may seek a transformation such that the new canonical coordinates $\tilde{q}_i$, $\tilde{p}_i$ are all cyclic. In this case, thanks to (2.2), it is $\mathrm{d}\tilde{q}_i/\mathrm{d}t = \partial\tilde{H}/\partial\tilde{p}_i = 0$, $\mathrm{d}\tilde{p}_i/\mathrm{d}t = -\partial\tilde{H}/\partial\tilde{q}_i = 0$, namely, each new canonical coordinate is a constant of motion.

The simplest way to obtain this result is to set the new Hamiltonian function equal to zero. Remembering from Sect. 2.2 that in every canonical transformation the relation between the old and new Hamiltonian function is $\tilde{H} = H + \partial K/\partial t$, one finds in this case the relation $\partial K/\partial t + H = 0$. To proceed it is convenient to choose a generating function of the $K_2 = K_2(\mathbf{q}, \tilde{\mathbf{p}}, t)$ type in which, as noted above, the new momenta $\tilde{p}_i$ are constants of motion. Given that the aim is to obtain the relation $\tilde{H} = 0$, the generating function of this problem is the particular function of the coordinates q_i, $\tilde{p}_i$, and t, that fulfills the equation $\partial K_2/\partial t + H = 0$. In other terms, the generating function becomes the problem's unknown. A comparison with (1.51) shows that the equation to be solved is that of Hamilton-Jacobi, and that K_2 coincides with Hamilton's principal function S.

As mentioned in Sect. 1.7, (1.51) is a first-order, partial differential equation in the unknown S and in the $n+1$ variables $q_1, \ldots, q_n, t$. As one of the $n+1$ integration constants can be set to zero, the actual integration constants are n. This seems contradictory because the Hamilton-Jacobi equation is expected to be equivalent to those of Hamilton or Lagrange for the description of the system's motion. As a consequence, the number of constants involved should be $2\,n$. The contradiction is easily removed by observing that n more constants appear in S, to be identified with the new momenta $\tilde{p}_1, \ldots, \tilde{p}_n$: remembering the canonical transformations (2.9) to be used in connection with the generating function of the K_2 type one finds

$$p_i = \frac{\partial S}{\partial q_i}\,, \qquad \tilde{q}_i = \frac{\partial S}{\partial \tilde{p}_i}\,, \qquad i = 1, \ldots, n\,. \tag{2.14}$$

Calculating the first of (2.14) at the initial time $t = a$ yields a set of n algebraic equations in the n unknowns $\tilde{p}_1, \ldots, \tilde{p}_n$. In fact, at $t = a$ the old canonical coordinates q_i, p_i are known because they are the problem's initial conditions. The solution of such algebraic equations yields the first set of motion's constants $\tilde{p}_1, \ldots, \tilde{p}_n$. Then, one considers the second of (2.14) at $t = a$, whose right-hand sides, at this point of the procedure, are known. As a consequence, the second of (2.14) yields the new generalized coordinates $\tilde{q}_1, \ldots, \tilde{q}_n$, that are the second set of motion's constants.

It is worth observing that the procedure depicted above provides also the time evolution of the old canonical coordinates. In fact, after all constants have been calculated, Eq. (2.14) forms $2\,n$ relations in the $2\,n + 1$ variables $q_1, \ldots, q_n, p_1, \ldots, p_n, t$. From them one extracts the relations $q_1 = q_1(t), \ldots, p_n = p_n(t)$. This shows that the Hamilton-Jacobi picture is equivalent to those based on the Hamilton or Lagrange equations for the solution of the mechanical problem.

2.4 Separation—Hamilton's Characteristic Function

The Hamilton-Jacobi equation (1.51) can be recast in a more symmetric form by letting $q_{n+1} = t$ and incorporating $\partial S/\partial t = \partial S/\partial q_{n+1}$ into the other term:

$$C\left(q_1, \ldots, q_{n+1}\,, \frac{\partial S}{\partial q_1}, \ldots, \frac{\partial S}{\partial q_{n+1}}\right) = 0\,. \tag{2.15}$$

Solving (2.15) becomes simpler if one of the coordinates, say q_i, and the corresponding momentum $p_i = \partial S/\partial q_i$ appear in (2.15) only through a relation $c_i = c_i(q_i, \partial S/\partial q_i)$ that does not contain any other coordinate, nor derivatives with respect to them, nor time. In this case q_i is called *separable coordinate* and the solution of (2.15) can be written as $S = S_i + W_i$, where S_i depends only on q_i and W_i depends on the other coordinates and time [84, Sect. 48]. Replacing this expression of S into (2.15) and extracting c_i yield a relation of the form $C_i = c_i$ with

$$C_i = C_i\left(q_1,\ldots,q_{i-1},q_{i+1},\ldots,q_{n+1}\,,\frac{\partial W_i}{\partial q_1},\ldots,\frac{\partial W_i}{\partial q_{i-1}},\frac{\partial W_i}{\partial q_{i+1}},\ldots,\frac{\partial W_i}{\partial q_{n+1}}\right),$$

$$c_i = c_i\left(q_i,\frac{\partial S_i}{\partial q_i}\right). \tag{2.16}$$

The equality $C_i = c_i$ must hold for any value of the coordinates. As this is possible only if the two sides are constant, $C_i = c_i$ separates and yields the pair

$$c_i\left(q_i,\frac{\partial S_i}{\partial q_i}\right) = c_{i0}\,, \tag{2.17}$$

$$C\left(q_1,\ldots,q_{i-1},q_{i+1},\ldots,q_{n+1}\,,\frac{\partial W_i}{\partial q_1},\ldots,\frac{\partial W_i}{\partial q_{i-1}},\frac{\partial W_i}{\partial q_{i+1}},\ldots,\frac{\partial W_i}{\partial q_{n+1}},c_{i0}\right) = 0\,,$$

where C does not contain q_i nor the corresponding derivative. The solution of the first of (2.17) provides $S_i(q_i)$. The latter contains two constants, namely c_{i0} and the integration constant. As noted earlier, the latter can be set to zero because an additive constant on S is irrelevant. In conclusion, c_{i0} is the only constant that remains after this step. In turn, the solution of the second of (2.17), which is an n-variable differential equation, contains n more constant, one of which is additive and can be disregarded. It follows that the total number of integration constants in the set (2.17) is still n.

If all coordinates are separable one has $S = \sum_{i=1}^{n} S_i(q_i)$ and the problem is solved by n individual integrations (an example is given in Sect. 3.10). In this case one says that the Hamilton-Jacobi equation is *completely separable*. A special case of separable coordinate is that of the cyclic coordinate. If q_i is cyclic, in fact, (2.17) reduces to $\partial S_i/\partial q_i = c_{i0}$, whence $S_i = c_{i0}q_i$ and $S = c_{i0}q_i + W_i$. If the cyclic coordinate is $q_{n+1} = t$, the above becomes

$$\frac{\partial S_{n+1}}{\partial t} = -E\,, \qquad S_{n+1} = -E\,t\,, \tag{2.18}$$

where the symbol E is used for the constant $c_{n+1,0}$. It is worth noting that the units of E are always those of an energy regardless of the choice of the generalized coordinates q_i. Comparing (2.18) with (1.51) yields $H = E =$ cost, consistently with the hypothesis that H does not depend on t. Using the symbol W instead of W_{n+1} provides the pair

$$H\left(q_1,\ldots,q_n\,,\frac{\partial W}{\partial q_1},\ldots,\frac{\partial W}{\partial q_n}\right) = E\,, \qquad S = W - Et\,, \tag{2.19}$$

that holds when H is a constant of motion. The first of (2.19) is a differential equation in the generalized coordinates only, called *time-independent Hamilton-Jacobi equation*. The unknown function W is called *Hamilton's characteristic function*.

2.5 Phase Velocity

The dynamics of a mechanical system can be obtained from Hamilton's principal function $S(\mathbf{q}, \tilde{\mathbf{p}}, t)$ as shown in Sect. 2.3. After S has been determined it is possible to build up an interesting geometrical construction, that is shown below. The indication of the constants $\tilde{p}_i$ is omitted for the sake of conciseness.

To begin, fix the time t and let $S(\mathbf{q}, t) = S_0$, where S_0 is some constant. This relation describes a surface belonging to the configuration space $q_1, \ldots, q_n$. Now change the time by $\mathrm{d}t$: the corresponding variation in S is obtained from (1.50) and reads $\mathrm{d}S = \sum_{i=1}^{n} p_i \mathrm{d}q_i - H\mathrm{d}t = \mathbf{p} \cdot \mathrm{d}\mathbf{q} - H\mathrm{d}t$. In this relation each component $p_i = \partial S / \partial q_i$ is calculated in terms of the coordinates $q_1, \ldots, q_n$ at the instant t, hence the vector $\mathbf{p} = \mathrm{grad}_{\mathbf{q}} S$ is a function of $\mathbf{q}$ calculated at that instant. If $\mathbf{q}$ belongs to the surface $S = S_0$, then $\mathbf{p}$ is normal to the surface at $\mathbf{q}$. Now let $S' = S + \mathrm{d}S = S_0$, where S_0 is the same constant as before. The relation $S'(\mathbf{q}, t) = S_0$ provides the new surface into which $S = S_0$ evolves in the interval $\mathrm{d}t$. As both $S = S_0$, $S + \mathrm{d}S = S_0$ hold, it must be $\mathrm{d}S = 0$, namely, $\mathbf{p} \cdot \mathrm{d}\mathbf{q} = H\mathrm{d}t$.

When H has no explicit dependence on t, thanks to (2.19) the relation $\mathbf{p} \cdot \mathrm{d}\mathbf{q} = H\mathrm{d}t$ becomes $\mathbf{p} \cdot \mathrm{d}\mathbf{q} = E\ \mathrm{d}t$, with $\mathbf{p} = \mathrm{grad}_{\mathbf{q}} W$. In this case, letting φ be the angle between the vectors $\mathbf{p}$ and $\mathrm{d}\mathbf{q}$ (whose moduli are indicated with p, $\mathrm{d}q$), and excluding the points where $\mathbf{p} = 0$, one obtains

$$\cos\varphi \, \frac{\mathrm{d}q}{\mathrm{d}t} = \frac{E}{p} \,. \tag{2.20}$$

The product $\cos\varphi\ \mathrm{d}q$ in (2.20) is the projection of $\mathrm{d}\mathbf{q}$ over the direction of $\mathbf{p}$, hence it provides the variation of $\mathbf{q}$ in the direction normal to the surface $S = S_0$. When the Cartesian coordinates are used, the product $\cos\varphi\ \mathrm{d}q$ is a length and the left-hand side of (2.20) is a velocity that provides the displacement of the point $\mathbf{q}$ during the time interval $\mathrm{d}t$ and in the direction normal to the surface $S = S_0$.

As shown above, the vector $\mathbf{p}$ is normal to the $S = S_0$ surface at each point of the latter. Consider for simplicity the case of a single particle of mass m in the conservative case, and use the Cartesian coordinates; from $\mathbf{p} = m\,\dot{\mathbf{q}}$ one finds that at each instant the surface $S = S_0$ is normal to the particle's trajectory. This makes the surface $S = S_0$ the analogue of the constant-phase surface found in the wave theory (e.g., Sect. 5.9). For this reason, $\cos\varphi\ \mathrm{d}q/\mathrm{d}t$ is called *phase velocity*.

Due to its definition, the phase velocity depends on the position $\mathbf{q}$ and is the velocity with which each point of the $S = S_0$ surface moves. It is worth adding that the phase velocity does not coincide with the actual velocity of any of the system's particles. To show this it suffices to consider the single particle's case with $\mathbf{p} = m\,\dot{\mathbf{q}}$: from (2.20) one finds that the modulus of the phase velocity is in fact inversely proportional to that of the particle.

2.6 Invariance Properties

A number of dynamic properties of the system of particles under consideration can be inferred directly from the form of the Lagrangian or Hamiltonian function, without the need of solving the motion's equations. An example is the conservation of the momentum conjugate to a cyclic coordinate (Sect. 1.6). Other properties are discussed below.

2.6.1 Time Reversal

It is found by inspection that the expression (1.22) of the system's kinetic energy in Cartesian coordinates is invariant when t is replaced with $-t$ (time reversal). A rather lengthy calculation based on the chain-differentiation rule shows that this property still holds after a coordinate transformation.

In some cases the whole Lagrangian function is invariant under time reversal. This makes the Lagrange equations (1.28) invariant as well. Assume that (1.28) are solved starting from a given initial condition at $t = a$ to the final instant $t = b$. Then, replace t with $t' = -t$ and solve the Lagrange equations again, using $q_i(t = b)$ and $-\dot{q}_i(t = b)$ as initial conditions. Letting $q_i' = \mathrm{d}q_i/\mathrm{d}t' = -\dot{q}_i$, (1.28) become

$$\frac{\mathrm{d}}{\mathrm{d}(-t')}\frac{\partial L}{\partial(-q_i')} = \frac{\mathrm{d}}{\mathrm{d}t'}\frac{\partial L}{\partial q_i'} = \frac{\partial L}{\partial q_i}\,, \qquad i = 1, \ldots, n \tag{2.21}$$

where, due to the hypothesis of invariance, the Lagrangian function is the same as that used to describe the motion from $t = a$ to $t = b$. It follows that the trajectories of the second motion are equal to those of the first one. Moreover, the initial velocities of the second motion, calculated at $t' = -b$, are opposite to those of the first motion at the same point. Due to the arbitrariness of b, at each point of a trajectory the velocity described by the time t' is opposite to that described by t. A motion having this property is called *reversible*.

Taking the examples of Sect. 1.3 and remembering the form (1.12, 1.18) of the corresponding Lagrangian functions one finds that, in the first example, the motion is reversible if the potential energy V is invariant under time reversal, namely, $V(-t) = V(t)$, while in the second example the motion is reversible if $\varphi(-t) = \varphi(t)$ and $\mathbf{A}(-t) = -\mathbf{A}(t)$.

2.6.2 Translation of Time

Consider the case where the Hamiltonian function is invariant with respect to translations of the origin of time. The invariance holds also for an infinitesimal

translation $\mathrm{d}t$, hence it is $\mathrm{d}H/\mathrm{d}t = 0$. In other terms H is a constant of motion. When this happens, as illustrated in Sect. 1.6, the Lagrangian and Hamiltonian functions have no explicit dependence on time, and vice versa.

2.6.3 Translation of the Coordinates

Another interesting case occurs when the Lagrangian function is invariant with respect to translations of the coordinates' origin. By way of example consider an N-particle system with no constraints, whence $n = 3N$, and use the Cartesian coordinates x_{js}. Here the first index is associated with the particles and the second one with the axes. Then, choose an infinitesimal translation $\mathrm{d}h_1$ in the direction of the first axis and, similarly, infinitesimal translations $\mathrm{d}h_2$ and $\mathrm{d}h_3$ in the other two directions. Thus, each coordinate $x_{j1}, j = 1, \ldots, N$ within the Lagrangian function is replaced by $x_{j1} + \mathrm{d}h_1$, and so on. The translation invariance then yields

$$\mathrm{d}L = \mathrm{d}h_1 \sum_{j=1}^{N} \frac{\partial L}{\partial x_{j1}} + \mathrm{d}h_2 \sum_{j=1}^{N} \frac{\partial L}{\partial x_{j2}} + \mathrm{d}h_3 \sum_{j=1}^{N} \frac{\partial L}{\partial x_{j3}} = 0\,. \tag{2.22}$$

Each sum in (2.22) vanishes independently of the others due to the arbitrariness of the translations. Taking the sum multiplying $\mathrm{d}h_1$ and using (1.28) yield

$$\sum_{j=1}^{N} \frac{\partial L}{\partial x_{j1}} = \sum_{j=1}^{N} \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{x}_{j1}} = \frac{\mathrm{d}}{\mathrm{d}t} \sum_{j=1}^{N} p_{j1} = \frac{\mathrm{d}P_1}{\mathrm{d}t} = 0\,, \tag{2.23}$$

where $P_1 = \sum_{j=1}^{N} p_{j1}$ is the first component of the *total momentum*

$$\mathbf{P} = \sum_{j=1}^{N} \mathbf{p}_j \tag{2.24}$$

of the system of particles. The other two components are treated in the same manner. In conclusion, if the Lagrangian function is invariant with respect to translations of the coordinates' origin, then the total momentum of the system is a constant of motion.

The above reasoning applies independently to each axis. As a consequence, if the Lagrangian function is such that the sum $\sum_{j=1}^{N} \partial L/\partial x_{j1}$ vanishes, while the analogous sums associated with the other two axes do not vanish, then P_1 is a constant of motion, while P_2, P_3 are not.

An important example of a Lagrangian function, that is invariant with respect to translations of the coordinates' origin, is found when the force $\mathbf{F}_i$ acting on the ith

particle derives from a potential energy V that depends only on the relative distances $r_{jk} = |\mathbf{r}_j - \mathbf{r}_k|$ among the particles, $k \neq j$. An example is given in Sect. 3.7.

2.6.4 Rotation of the Coordinates

Consider the case where the Lagrangian function is invariant with respect to rotations of the coordinates around an axis that crosses the origin. Like in Sect. 2.6.3 a system of N particles with no constraints is assumed, and the Cartesian coordinates are used. Let π be the plane that contains the origin and is normal to the rotation axis. It is convenient to use on π a polar reference (Sect. B.2) in which the rotation is defined over π by the angle φ. In turn, let ϑ_j be the angle between the rotation axis and the position vector $\mathbf{r}_j = (x_{j1}, x_{j2}, x_{j3})$ of the jth particle. The meaning of the angles is the same as in Fig. B.1, where the axes x, y define plane π; then, ϑ_j and $\mathbf{r}_j$ are represented by ϑ and $\mathbf{r}$ of the figure, respectively. If an infinitesimal rotation $\mathrm{d}\varphi$ is considered, the position vector $\mathbf{r}_j$ undergoes a variation $\mathrm{d}\mathbf{r}_j$ parallel to π and of magnitude $|\mathrm{d}\mathbf{r}_j| = r_j \sin\vartheta_j \, \mathrm{d}\varphi$. To specify the direction of $\mathrm{d}\mathbf{r}_j$ one takes the unit vector $\mathbf{a}$ of the rotation axis and associates with the rotation the vector $\mathbf{a}\,\mathrm{d}\varphi$ such that

$$\mathrm{d}\mathbf{r}_j = \mathbf{a}\,\mathrm{d}\varphi \wedge \mathbf{r}_j \,. \tag{2.25}$$

The corresponding variations $\mathrm{d}\dot{\mathbf{r}}_j$ of the velocities are found by differentiating (2.25) with respect to time. The variation of the Lagrangian function is

$$\mathrm{d}L = \sum_{j=1}^{N} \sum_{s=1}^{3} \left(\frac{\partial L}{\partial x_{js}} \, \mathrm{d}x_{js} + \frac{\partial L}{\partial \dot{x}_{js}} \, \mathrm{d}\dot{x}_{js} \right) , \tag{2.26}$$

where the variations of the components are found from (2.25) and read $\mathrm{d}x_{js} = \mathrm{d}\varphi \left(\mathbf{a} \wedge \mathbf{r}_j\right)_s$, $\mathrm{d}\dot{x}_{js} = \mathrm{d}\varphi \left(\mathbf{a} \wedge \dot{\mathbf{r}}_j\right)_s$. Replacing the latter in (2.26) and using (1.28) yield

$$\mathrm{d}L = \mathrm{d}\varphi \sum_{j=1}^{N} \left(\mathbf{a} \wedge \mathbf{r}_j \cdot \dot{\mathbf{p}}_j + \mathbf{a} \wedge \dot{\mathbf{r}}_j \cdot \mathbf{p}_j \right) . \tag{2.27}$$

Due to the rotational invariance, (2.27) vanishes for any $\mathrm{d}\varphi$. Letting the sum vanish after exchanging the scalar and vector products, and remembering that $\mathbf{a}$ is constant, one finds

$$\mathbf{a} \cdot \sum_{j=1}^{N} \left(\mathbf{r}_j \wedge \dot{\mathbf{p}}_j + \dot{\mathbf{r}}_j \wedge \mathbf{p}_j \right) = \mathbf{a} \cdot \frac{\mathrm{d}}{\mathrm{d}t} \sum_{j=1}^{N} \mathbf{r}_j \wedge \mathbf{p}_j = \mathbf{a} \cdot \frac{\mathrm{d}\mathbf{M}}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t} \left(\mathbf{M} \cdot \mathbf{a} \right) = 0 \,, \tag{2.28}$$

where

$$\mathbf{M} = \sum_{j=1}^{N} \mathbf{r}_j \wedge \mathbf{p}_j \tag{2.29}$$

is the *total angular momentum* of the system of particles. In conclusion, if the Lagrangian function is invariant with respect to rotations of the coordinates around an axis that crosses the origin, then the projection of the system's total angular momentum $\mathbf{M}$ over the rotation axis is a constant of motion.

2.7 Maupertuis Principle

Besides the Hamilton principle described in Sect. 1.3.4, other variational principles exist. Among them is the *Maupertuis*, or *least action*, principle, that applies to a particle subjected to conservative forces. Let $V = V(x_1, x_2, x_3)$ be the potential energy and $E = \text{const}$ the total energy, and let A and B indicate the two points of the (x_1, x_2, x_3) space that limit the trajectory of the particle. The Maupertuis principle states that the natural trajectory between A and B is the one that minimizes the functional

$$G = \int_{AB} \sqrt{E - V}\, \mathrm{d}s\,, \qquad \mathrm{d}s^2 = \mathrm{d}x_1^2 + \mathrm{d}x_2^2 + \mathrm{d}x_3^2\,, \tag{2.30}$$

where the integral is carried out along the trajectory. The form of (2.30) explains the term "least action": in fact, the relation $p^2/(2m) = m\,u^2/2 = E - V$ shows that the integrand $\sqrt{E - V}$ is proportional to the particle's momentum p; as a multiplicative constant is irrelevant for calculating the extremum functions, the minimization of G is equivalent to that of the action $\int_{AB} p\, \mathrm{d}s$.

To calculate the extremum functions of (2.30) it is convenient to parametrize the coordinates in the form $x_i = x_i(\xi)$, where the parameter ξ takes the same limiting values, say, $\xi = a$ at A and $\xi = b$ at B, for the natural and all the virtual trajectories. Letting $\dot{x}_i = \mathrm{d}x_i/\mathrm{d}\xi$ one finds $(\mathrm{d}s/\mathrm{d}\xi)^2 = \dot{x}_1^2 + \dot{x}_2^2 + \dot{x}_3^2$ which, remembering (1.7), yields the extremum condition for (2.30):

$$\delta \int_a^b \theta\, \mathrm{d}\xi = 0\,, \qquad \theta = \sqrt{E - V}\, \sqrt{\dot{x}_1^2 + \dot{x}_2^2 + \dot{x}_3^2}\,, \qquad \frac{\mathrm{d}}{\mathrm{d}\xi} \frac{\partial \theta}{\partial \dot{x}_i} = \frac{\partial \theta}{\partial x_i}\,. \tag{2.31}$$

The following relations are useful to work out the last of (2.31): $\mathrm{d}s/\mathrm{d}t = u = \sqrt{(2/m)(E - V)}$, $\mathrm{d}x_i = \dot{x}_i\, \mathrm{d}\xi$, $\mathrm{d}x_i/\mathrm{d}t = u_i$. One finds

$$\frac{\partial \theta}{\partial \dot{x}_i} = \sqrt{E - V}\, \frac{\dot{x}_i}{\mathrm{d}s/\mathrm{d}\xi} = \sqrt{\frac{m}{2}}\, u\, \frac{\dot{x}_i\, \mathrm{d}\xi}{\mathrm{d}s} = \sqrt{\frac{m}{2}}\, \frac{\mathrm{d}x_i}{\mathrm{d}t} = \sqrt{\frac{m}{2}}\, u_i\,, \tag{2.32}$$

$$\frac{\partial\theta}{\partial x_i} = \frac{\mathrm{d}s}{\mathrm{d}\xi}\,\frac{-\partial V/\partial x_i}{2\sqrt{E-V}} = \frac{\mathrm{d}s}{\mathrm{d}\xi}\,\frac{F_i}{2\sqrt{m/2}\,u} = \frac{\mathrm{d}t}{\mathrm{d}\xi}\,\frac{F_i}{\sqrt{2m}}\,, \tag{2.33}$$

with F_i the ith component of the force. The last of (2.31) then yields

$$\frac{\mathrm{d}}{\mathrm{d}\xi}\left(\sqrt{\frac{m}{2}}\,u_i\right) = \frac{\mathrm{d}t}{\mathrm{d}\xi}\,\frac{F_i}{\sqrt{2m}}\,, \qquad F_i = m\,\frac{\mathrm{d}u_i}{\mathrm{d}\xi}\,\frac{\mathrm{d}\xi}{\mathrm{d}t} = m\,\frac{\mathrm{d}u_i}{\mathrm{d}t}\,. \tag{2.34}$$

In conclusion, the equation that provides the extremum condition for functional G is equivalent to Newton's second law $\mathbf{F} = m\,\mathbf{a}$.

2.8 Spherical Coordinates—Angular Momentum

Consider a single particle of mass m and use the transformation from the Cartesian (x, y, z) to the spherical (r, ϑ, φ) coordinates shown in Sect. B.1. The kinetic energy is given by (B.7), namely,

$$T = \frac{m}{2}\,(\dot{r}^2 + r^2\dot{\vartheta}^2 + r^2\dot{\varphi}^2\sin^2\vartheta)\,. \tag{2.35}$$

If the force acting onto the particle is derivable from a potential energy $V = V(x, y, z, t)$, the Lagrangian function in the spherical reference is $L = T - V(r, \vartheta, \varphi, t)$, where T is given by (2.35). The momenta conjugate to the spherical coordinates are

$$\begin{cases} p_r &= \partial L/\partial\dot{r} &= m\dot{r} \\ p_\vartheta &= \partial L/\partial\dot{\vartheta} &= mr^2\,\dot{\vartheta} \\ p_\varphi &= \partial L/\partial\dot{\varphi} &= mr^2\,\dot{\varphi}\,\sin^2\vartheta \end{cases} \tag{2.36}$$

Using (2.36), the kinetic energy is recast as

$$T = \frac{1}{2m}\left(p_r^2 + \frac{p_\vartheta^2}{r^2} + \frac{p_\varphi^2}{r^2\,\sin^2\vartheta}\right). \tag{2.37}$$

The components of the momentum $\mathbf{p}$ derived from the Lagrangian function written in the Cartesian coordinates are $m\dot{x}$, $m\dot{y}$, $m\dot{z}$. It follows that the components of the angular momentum $\mathbf{M} = \mathbf{r} \wedge \mathbf{p}$ written in the Cartesian and spherical references are

$$\begin{cases} M_x = m\,(y\,\dot{z} - z\,\dot{y}) &= -mr^2\,(\dot{\vartheta}\,\sin\varphi + \dot{\varphi}\,\sin\vartheta\,\cos\vartheta\,\cos\varphi) \\ M_y = m\,(z\,\dot{x} - x\,\dot{z}) &= mr^2\,(\dot{\vartheta}\,\cos\varphi - \dot{\varphi}\,\sin\vartheta\,\cos\vartheta\,\sin\varphi) \\ M_z = m\,(x\,\dot{y} - y\,\dot{x}) &= mr^2\,\dot{\varphi}\,\sin^2\vartheta \end{cases} \tag{2.38}$$

The square modulus of the angular momentum in spherical coordinates reads

$$M^2 = m^2 r^4 \left(\dot{\vartheta}^2 + \dot{\varphi}^2 \sin^2\vartheta\right) = p_\vartheta^2 + \frac{p_\varphi^2}{\sin^2\vartheta}\,, \tag{2.39}$$

where the last equality is due to (2.36). From (2.37, 2.39) one finds

$$T = \frac{1}{2m}\left(p_r^2 + \frac{M^2}{r^2}\right). \tag{2.40}$$

If M is a constant of motion, (2.40) shows that the kinetic energy depends on r and $\dot{r}$ only. Comparing (2.36) with (2.38) one also notices that

$$p_\varphi = M_z\,, \tag{2.41}$$

namely, the component along the z axis of the angular momentum turns out to be the momentum conjugate to the φ coordinate. The latter describes the rotations along the same axis. In contrast, the other two components of $\mathbf{M}$ are not conjugate momenta. This result is due to the asymmetry of the relations (B.1) that connect the Cartesian to the spherical coordinates, and does not ascribe any privileged role to the z axis. In fact, by exchanging the Cartesian axes one makes p_φ to coincide with M_x or M_y.

Another example refers to a particle of mass m and charge e subjected to an electromagnetic field. Remembering (1.33) one has $L = (1/2)\,m\,u^2 - e\,U + e\,\mathbf{u}\cdot\mathbf{A}$, where the scalar potential is indicated with U to avoid confusion with the φ coordinate. It follows

$$L = \frac{1}{2}\,m\left(\dot{r}^2 + r^2\,\dot{\vartheta}^2 + r^2\,\dot{\varphi}^2\,\sin^2\vartheta\right) - e\,U + e\,\mathbf{u}\cdot\mathbf{A}\,, \tag{2.42}$$

where the components of $\mathbf{u} = \dot{\mathbf{r}}$ are given by (B.6), and U, A depend on the coordinates and time. Indicating the components of $\mathbf{A}$ with A_x, A_y, A_z, the momenta read

$$\begin{cases} p_r = \partial L/\partial\dot{r} = m\,\dot{r} + e\,A_x\,\sin\vartheta\,\cos\varphi + e\,A_y\,\sin\vartheta\,\sin\varphi + e\,A_z\,\cos\vartheta\,\cos\varphi \\ p_\vartheta = \partial L/\partial\dot{\vartheta} = m\,r^2\,\dot{\vartheta} + e\,A_x\,r\,\cos\vartheta\,\cos\varphi + e\,A_y\,r\,\cos\vartheta\,\sin\varphi - e\,A_z\,r\,\sin\vartheta \\ p_\varphi = \partial L/\partial\dot{\varphi} = m\,r^2\,\dot{\varphi}\,\sin^2\vartheta - e\,A_x\,r\,\sin\vartheta\,\sin\varphi + e\,A_y\,r\,\sin\vartheta\,\cos\varphi \end{cases} \tag{2.43}$$

Thanks to (B.1, B.6), the third of (2.43) can be written as

$$p_\varphi = m\,(x\,\dot{y} - y\,\dot{x}) + e\,(x\,A_y - y\,A_x) = x\,(m\,\dot{y} + e\,A_y) - y\,(m\,\dot{x} + e\,A_x)\,, \tag{2.44}$$

that coincides with the component of the angular momentum $\mathbf{M} = \mathbf{r}\wedge\mathbf{p} = \mathbf{r}\wedge(m\,\mathbf{u} + e\,\mathbf{A})$ along the z axis. This result shows that (2.41) holds also when the force acting onto the particle derives from an electromagnetic field.

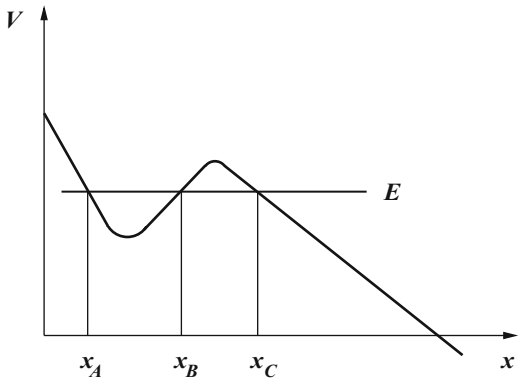


Fig. 2.1 Example of potential energy discussed in Sect. 2.9

2.9 Linear Motion

The *linear motion* is the motion of a system having only one degree of freedom. Using the Cartesian coordinate x, and assuming the case where the force acting onto the particle derives from a potential energy $V(x)$, gives the Hamiltonian function (1.32) the form $H = p^2/(2m) + V(x)$. As shown in Sect. 1.6, a Hamiltonian function of this type is a constant of motion whence, remembering that here it is $p = m\dot{x}$,

$$\frac{1}{2} m\dot{x}^2 + V(x) = E = \text{const.} \tag{2.45}$$

The constant E is called *total energy*. Its value is given by the initial conditions $x_0 = x(t = a)$, $\dot{x}_0 = \dot{x}(t = a)$. As the kinetic energy $m\dot{x}^2/2$ cannot be negative, the motion is possible only in the intervals of the x axis such that $V(x) \leq E$. In particular, the velocity $\dot{x}$ vanishes at the points where $V = E$. Instead, the intervals where $V > E$ cannot be reached by the particle. Equation (2.45) is separable and provides a relation of the form $t = t(x)$,

$$t = a \pm \sqrt{\frac{m}{2}} \int_{x_0}^{x} \frac{\mathrm{d}\xi}{\sqrt{E - V(\xi)}} \,. \tag{2.46}$$

By way of example consider a situation like that shown in Fig. 2.1, where it is assumed that to the right of x_C the potential energy V keeps decreasing as $x \to \infty$. If the initial position of the particle is $x_0 = x_C$, there the velocity vanishes and the particle is subjected to a positive force $F = -\mathrm{d}V/\mathrm{d}x > 0$. As a consequence, the particle's motion will always be oriented to the right starting from x_C. Such a motion is called *unlimited*. If the initial position is $x_0 > x_C$ and the initial velocity is negative, the particle moves to the left until it reaches the position x_C, where it bounces back. The subsequent motion is the same as described above.

A different situation arises when the initial position of the particle belongs to an interval limited by two zeros of the function $E - V(x)$ like, e.g., x_A and x_B in Fig. 2.1. The motion is confined between x_A and x_B and, for this reason, is called *limited*. The particle bounces back and forth endlessly under the effect of a force that does not depend on time. As a consequence, the time necessary to complete a cycle $x_A \to x_B \to x_A$ is the same for each cycle. In other terms, the motion is periodic in time. Also, from (2.46) it is found by inspection that the time spent by the particle in the $x_A \to x_B$ part of the cycle is the same as that spent in the $x_B \to x_A$ part. The period of the oscillation is then found to be

$$T = 2\sqrt{\frac{m}{2}} \int_{x_A}^{x_B} \frac{\mathrm{d}x}{\sqrt{E - V(x)}} \,. \tag{2.47}$$

Note that the period depends on the total energy E. However, there are exceptions, as the example of Sect. 3.3 shows.

2.10 Action-Angle Variables

Consider a linear, conservative motion of constant energy E (Sect. 2.9) and let q, p be two canonical coordinates describing it. The following hold

$$H(q,p) = E \,, \qquad p = p(q,E) \,. \tag{2.48}$$

The second of (2.48) is derived from the first one by solving for the momentum, and provides the phase trajectory (Sect. 1.9) starting from the initial conditions q_0, p_0 of the motion. As shown below, in several mechanical systems of interest the phase trajectory has some special characteristic that is worth examining.

Consider, first, the situation where the phase trajectory is closed: in this case, after a time T has elapsed from $t = 0$, the canonical coordinates take again the values q_0, p_0. As a consequence, for $t > T$ the motion repeats itself, and so on. It follows that both q and p are periodic functions of time with the same period T. As discussed in Sect. 2.9, this type of periodic motion, in which both q and p are bounded, is typically found when the initial position q_0 lies between two zeros of $E - V$, and is of the oscillatory type. It is also indicated with the astronomical term *libration*.

A second important situation occurs when p is a periodic function of q. In this type of motion q is unbounded. However, when q increases by a period the configuration of the mechanical system remains practically unchanged. In fact, in this type of motion the canonical coordinate q is always an angle of rotation: the motion is still periodic and is referred to as *rotation*. Note that the same mechanical

system may give rise to a libration or a rotation, depending on the motion's initial conditions: a typical example is that of the simple pendulum where q is identified with the angle of deflection [56, Chap. 10.6]. The *action variable* is defined as

$$J(E) = \oint p(q, E)\, \mathrm{d}q\,, \tag{2.49}$$

where the integral is carried out over a complete period of libration or rotation, depending on the case under investigation. The name given to J stems from the fact that, as mentioned in Sect. 1.5, the product $q\,p$ has the units of an action in all coordinate sets. The action variable is a constant of motion because it depends on E only. Inverting $J(E)$ yields $H = H(J)$. Now one applies a canonical transformation generated by a Hamilton characteristic function of the form $W = W(q, J)$. Remembering the procedure depicted in Sect. 2.4, W is the solution of

$$H\left(q, \frac{\partial W}{\partial q}\right) = E\,. \tag{2.50}$$

Applying (2.14) one finds the generalized coordinate $w = \partial W/\partial J$, called *angle variable*, conjugate to J. The pair J, w constitutes the set of canonical coordinates called *action-angle variables*. Finally, the Hamilton equations in the new coordinates read

$$\dot{w} = \frac{\partial H}{\partial J} = \text{const}\,, \qquad \dot{J} = -\frac{\partial H}{\partial w} = 0\,. \tag{2.51}$$

The time evolution of the action-angle variables is then $w = \dot{w}\,t + w_0$, $J = \text{const}$. From the first of (2.51) it also follows that the units of $\dot{w}$ are those of a frequency. The usefulness of the action-angle variables becomes apparent when one calculates the change Δw over a complete libration or rotation cycle of q:

$$\Delta w = \oint \mathrm{d}w = \oint \frac{\partial w}{\partial q}\,\mathrm{d}q = \oint \frac{\partial^2 W}{\partial q\,\partial J}\,\mathrm{d}q = \frac{\mathrm{d}}{\mathrm{d}J}\oint \frac{\partial W}{\partial q}\,\mathrm{d}q = 1\,, \tag{2.52}$$

where the last equality derives from combining $p = \partial W/\partial q$ with (2.49). On the other hand, if T is the time necessary for completing a cycle of q, then it is $\Delta w = w(T) - w(0) = \dot{w}\,T$, whence $\dot{w} = 1/T$. Thus, the frequency $\nu = \dot{w}$ is that associated with the periodic motion of q. In conclusion, the action-angle variables provide a straightforward method to determine the frequency of a periodic motion without the need of solving the motion equation. The method is applicable also to conservative systems having more than one degree of freedom, provided there exists at least one set of coordinates in which the Hamilton-Jacobi equation is completely separable [56, Chap. 10.7].

2.11 Complements

2.11.1 *Infinitesimal Canonical Transformations*

Consider a system with n degrees of freedom whose Hamiltonian function is $H(q_1, \ldots, q_n, p_1, \ldots, p_n, t)$. Remembering (1.42) one finds that the canonical coordinates at $t + \mathrm{d}t$ are expressed, in terms of the same coordinates at t, by the relations

$$q_i + \mathrm{d}q_i = q_i + \frac{\partial H}{\partial p_i}\,\mathrm{d}t\,, \qquad p_i + \mathrm{d}p_i = p_i - \frac{\partial H}{\partial q_i}\,\mathrm{d}t\,. \tag{2.53}$$

Letting $\tilde{q}_i = q_i + \mathrm{d}q_i$, $\tilde{p}_i = p_i + \mathrm{d}p_i$ gives (2.53) the same form as (2.1), namely, that of a coordinate transformation. It is interesting to check whether such a transformation is canonical. For this, one notes that the transformation (2.53) differs by infinitesimal quantities from the identical transformation $\tilde{q}_i = q_i$, $\tilde{p}_i = p_i$; as a consequence one expects the generating function of (2.53), if it exists, to differ by an infinitesimal function from $F_2 = \sum_{i=1}^{n} q_i \tilde{p}_i$ which, as shown in Sect. 2.2, generates the identical transformation. One then lets

$$K_2 = \sum_{i=1}^{n} q_i \tilde{p}_i + \epsilon\, G(\mathbf{q}, \tilde{\mathbf{p}}, t)\,, \tag{2.54}$$

where ϵ is an infinitesimal quantity. From the first two equations in (2.9) it follows

$$p_i = \frac{\partial K_2}{\partial q_i} = \tilde{p}_i + \epsilon\,\frac{\partial G}{\partial q_i}\,, \qquad \tilde{q}_i = \frac{\partial K_2}{\partial \tilde{p}_i} = q_i + \epsilon\,\frac{\partial G}{\partial \tilde{p}_i}\,. \tag{2.55}$$

In the last term of (2.55) one may replace $\tilde{p}_i$ with p_i on account of the fact that the difference between $\epsilon\,\partial G/\partial \tilde{p}_i$ and $\epsilon\,\partial G/\partial p_i$ is infinitesimal of a higher order. Then, letting $\epsilon = \mathrm{d}t$, $G(\mathbf{q}, \mathbf{p}, t) = H(\mathbf{q}, \mathbf{p}, t)$, and $K_2 = \sum_{i=1}^{n} q_i \tilde{p}_i + H(\mathbf{q}, \mathbf{p}, t)\,\mathrm{d}t$, makes (2.55) identical to (2.53). Note that this replacement transforms the third of (2.9) into $\tilde{H} = H + (\partial H/\partial t)\,\mathrm{d}t$, as should be (compare with (1.44)).

The above reasoning shows that the $H\,\mathrm{d}t$ term in (2.54) generates a canonical transformation that produces the variations of the canonical coordinates in the time interval $\mathrm{d}t$. Such a transformation is called *infinitesimal canonical transformation*. On the other hand, as the application of more than one canonical transformation is still canonical, the evolution of the coordinates q_i, p_i during a finite interval of time can be thought of as produced by a sequence of infinitesimal canonical transformations generated by the Hamiltonian function. In other terms, the Hamiltonian function generates the motion of the system.

2.11.2 Constants of Motion

It has been shown in Sect. 2.11.1 that a sequence of infinitesimal canonical transformation generated by the Hamiltonian function determines the time evolution of the canonical coordinates q_i, p_i. If such a sequence starts with the initial conditions q_{i0}, p_{i0}, at some later time t the transformation equations (2.53) take the form

$$q_i = q_i(\mathbf{q}_0, \mathbf{p}_0, t)\,, \qquad p_i = p_i(\mathbf{q}_0, \mathbf{p}_0, t)\,. \tag{2.56}$$

The relations (2.56) are nothing else than the solution of the mechanical problem; in fact, they express the canonical coordinates at time t, given the initial conditions. From another viewpoint, they show that the solution of the problem contains $2\,n$ constants. They are not necessarily constants of motion, in fact, their values at $t > 0$ are in general different from those at $t = 0$. If the system has extra properties (like, e.g., the invariance properties discussed in Sect. 2.6), it also has one or more constants of motion. The latter keep the value that they possessed at $t = 0$, so they are expressible as combinations of the canonical coordinates at $t = 0$; by way of example, the total energy E of a single particle subjected to a conservative force reads $E = (p_{01}^2 + p_{02}^2 + p_{03}^2)/(2\,m) + V(x_{10}, x_{20}, x_{30})$.

For a system having n degrees of freedom the total number of independent combinations of the initial conditions cannot exceed the number of the initial conditions themselves. As a consequence, for such a system the maximum number of independent constants of motion is $2\,n$.

Problem

2.1 Given the Hamiltonian function $H = p^2/(2\,m) + (c/2)\,x^2$, $m, c > 0$ (that describes the *linear harmonic oscillator*, Sect. 3.3), find the oscillation frequency ν using the action-angle variables.

Chapter 3
Applications of the Concepts of Analytical Mechanics

3.1 Introduction

This chapter provides a number of important examples of application of the principles of Analytical Mechanics. The examples are chosen with reference to the applications to Quantum Mechanics shown in later chapters. The first sections treat the problems of the square well, linear harmonic oscillator, and central motion. The subsequent sections deal with the two-particle interaction: first, the description of the collision is given, along with the calculation of the energy exchange involved in it, with no reference to the form of the potential energy; this is followed by the treatment of the collision when the potential energy is of the repulsive-Coulomb type. The chapter continues with the treatment of a system of strongly bound particles: the diagonalization of its Hamiltonian function shows that the motion of the particles is a superposition of harmonic oscillations. Finally, the motion of a particle subjected to a periodic potential energy is analyzed, including the case where a weak perturbation is superimposed to the periodic part. A number of complements are also given, that include the treatment of the collision with a potential energy of the attractive-Coulomb type, and that of the collision of two relativistic particles.

3.2 Particle in a Square Well

As a first example of linear motion consider the case of a potential energy V of the form shown in Fig. 3.1. Such a potential energy is called *square well* and is to be understood as the limiting case of a potential energy well whose sides have a finite, though very large, slope. It follows that the force $F = -\mathrm{d}V/\mathrm{d}x$ is everywhere equal to zero with the exception of the two points $-x_M$ and $+x_M$, where it tends to $+\infty$

M. Rudan, *Physics of Semiconductor Devices*, DOI 10.1007/978-3-319-63154-7_3

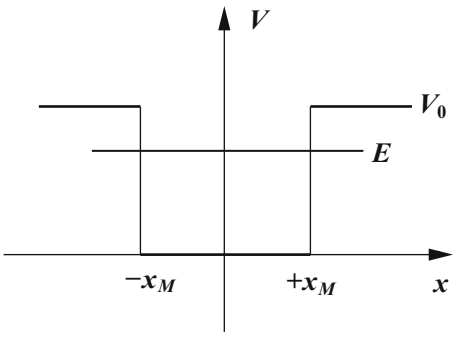


Fig. 3.1 The example of the square well analyzed in Sect. 3.2. Only the case $0 \leq E \leq V_0$ is shown

and $-\infty$, respectively. From the discussion of Sect. 2.9 it follows that the case $E < 0$ is forbidden. The motion of the particle is finite for $0 \leq E \leq V_0$, while it is infinite for $E > V_0$.

Considering the $0 \leq E \leq V_0$ case first, the motion is confined within the well and the velocity of the particle is constant in the interval $-x_M < x < +x_M$, where the Hamiltonian function yields $m\dot{x}^2/2 = E$. If the particle's motion is oriented to the right, the velocity is $\dot{x} = \sqrt{2E/m}$. When the particle reaches the position x_M its velocity reverses instantly to become $\dot{x} = -\sqrt{2E/m}$. The motion continues at a constant velocity until the particle reaches the position $-x_M$ where it reverses again, and so on. As the spatial interval corresponding to a full cycle is $4\,x_M$, the oscillation period is $T = \sqrt{8\,m/E}\,x_M$.

To treat the $E > V_0$ case assume that the particle is initially at a position $x < -x_M$ with a motion oriented to the right. The Hamiltonian function outside the well yields $m\dot{x}^2/2 + V_0 = E$. The constant velocity is $\dot{x} = \sqrt{2\,(E - V_0)/m}$ until the particle reaches the position $-x_M$. There the velocity increases abruptly to $\dot{x} = \sqrt{2E/m}$ and keeps this value until the particle reaches the other edge of the well, $+x_M$. There, the velocity decreases abruptly back to the initial value $\dot{x} = \sqrt{2\,(E - V_0)/m}$, and the particle continues its motion at a constant velocity in the positive direction.

3.3 Linear Harmonic Oscillator

An important example of linear motion is found when the force derives from a potential energy of the form $V = c\,x^2/2$, with $c > 0$. The force acting on the particle turns out to be $F = -\mathrm{d}V/\mathrm{d}x = -c\,x$, namely, it is linear with respect to x, vanishes at $x = 0$, and has a modulus that increases as the particle departs from the origin. Also, due to the positiveness of c, the force is always directed towards the origin. A force of this type is also called *linear elastic force*, and c is called *elastic constant* (Fig. 3.2).

From the discussion of Sect. 2.9 it follows that the case $E < 0$ is forbidden. The motion of the particle is always finite because for any $E \geq 0$ it is confined between

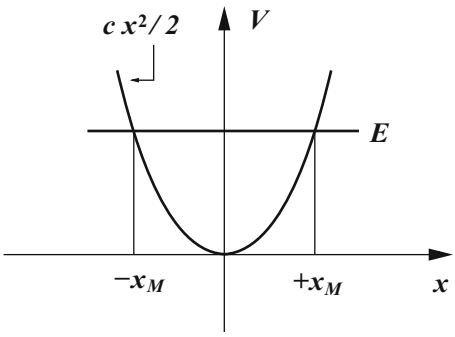


Fig. 3.2 The example of the linear harmonic oscillator analyzed in Sect. 3.3

the two zeros $x_M = \pm\sqrt{2\,E/c}$ of the equation $V = E$. The Hamiltonian function reads

$$H = \frac{1}{2\,m}\,p^2 + \frac{1}{2}\,c\,x^2 = E = \text{const}, \tag{3.1}$$

yielding the motion's equation $\dot{p} = m\,\ddot{x} = -\partial H/\partial x = -c\,x$ whose solution is

$$x(t) = x_M\,\cos(\omega\,t + \alpha_0)\,, \quad \dot{x}(t) = -\omega\,x_M\,\sin(\omega\,t + \alpha_0)\,, \quad \omega = \sqrt{c/m}\,. \tag{3.2}$$

Due to the form of (3.2), a particle whose motion is derived from the Hamiltonian function (3.1) is called *linear harmonic oscillator*. The maximum elongation $x_M > 0$ and the initial phase α_0 are readily expressed in terms of the initial conditions $x_0 = x(t = 0)$, $\dot{x}_0 = \dot{x}(t = 0)$. In fact, letting $t = 0$ in (3.2) yields $x_M^2 = x_0^2 + \dot{x}_0^2/\omega^2$ and $\tan\alpha_0 = -\dot{x}_0/(\omega\,x_0)$. The total energy in terms of the initial conditions reads $E = m\,\dot{x}_0^2/2 + c\,x_0^2/2$ and, finally, the oscillation's period is $T = 2\,\pi/\omega$. Note that T depends on the two parameters m, c appearing in the Hamiltonian function, but not on the total energy (in other terms, for a given pair m, c the oscillation period does not depend on the initial conditions). As mentioned in Sect. 2.9, this is an exceptional case.

3.4 Central Motion

Consider the case of a particle of mass m acted upon by a force that derives from a potential energy of the form $V = V(r)$, where $r = \sqrt{x^2 + y^2 + z^2}$ is the modulus of the position vector $\mathbf{r}$ of the particle. The force

$$\mathbf{F} = -\,\text{grad}\,V = -\frac{\mathrm{d}V}{\mathrm{d}r}\,\frac{\mathbf{r}}{r}\,, \tag{3.3}$$

depends on r only and is oriented along $\mathbf{r}$. For this reason it is called *central force*. In turn, the point whence the force originates (in this case, the origin of the reference) is called *center of force*. The corresponding Lagrangian function

$$L = \frac{1}{2} m\dot{r}^2 - V(r)\,, \qquad \mathbf{p} = m\,\dot{\mathbf{r}}\,, \qquad \dot{r} = |\dot{\mathbf{r}}| \tag{3.4}$$

turns out to be invariant under any rotation (Sect. 2.6.4). Remembering (2.28, 2.29), this type of invariance means that the projection of the angular momentum $\mathbf{M}$ onto any direction is conserved. It follows that for a particle acted upon by a central force, the vector $\mathbf{M}$ itself is conserved. On the other hand, it is $\mathbf{M} = \mathbf{r} \wedge m\,\dot{\mathbf{r}}$, so the constant angular momentum is fixed by the initial conditions of the motion.

As $\mathbf{M}$ is normal to the plane defined by $\mathbf{r}$ and $\dot{\mathbf{r}}$, the trajectory of the particle lies always on such a plane. It is then useful to select the Cartesian reference by aligning, e.g., the z axis with $\mathbf{M}$. In this way, the trajectory belongs to the x, y plane and two coordinates eventually suffice to describe the motion. Turning to the spherical coordinates (B.1) and using (2.40) yields the Hamiltonian function

$$H = \frac{1}{2m}\left(p_r^2 + \frac{M^2}{r^2}\right) + V(r) = \frac{p_r^2}{2m} + V_e(r)\,, \quad V_e = V + \frac{M^2}{2mr^2}\,, \tag{3.5}$$

with $M^2 = p_\vartheta^2 + p_\varphi^2/\sin^2\vartheta$, $\mathbf{M} = \text{const}$, and $p_r = m\dot{r}$, $p_\vartheta = mr^2\dot{\vartheta}$, $p_\varphi = M_z = mr^2\dot{\varphi}\,\sin\vartheta$. However, the z axis has been aligned with $\mathbf{M}$, which is equivalent to letting $\vartheta = \pi/2$. It turns out $M_z = M$, and $p_r = m\dot{r}$, $p_\vartheta = 0$, $p_\varphi = M = mr^2\dot{\varphi}$, so that

$$H = \frac{1}{2m}\left(p_r^2 + \frac{p_\varphi^2}{r^2}\right) + V(r) = \frac{p_r^2}{2m} + V_e(r)\,, \quad V_e = V + \frac{p_\varphi^2}{2mr^2}\,. \tag{3.6}$$

As the total energy is conserved it is $H = E$, where E is known from the initial conditions. The intervals of r where the motion can actually occur are those in which $E \geq V_e$. Letting $r_0 = r(t = 0)$, the time evolution of the radial part is found from $p_r^2 = m^2(\mathrm{d}r/\mathrm{d}t)^2 = 2\,m\,(E - V_e)$, namely

$$t(r) = \pm\sqrt{\frac{m}{2}}\int_{r_0}^{r} \frac{\mathrm{d}\xi}{\sqrt{E - V_e(\xi)}}\,. \tag{3.7}$$

From $p_\varphi = mr^2\dot{\varphi} = \text{const}$ it follows that φ depends monotonically on time, and also that $\mathrm{d}t = (mr^2/p_\varphi)\,\mathrm{d}\varphi$. Combining the latter with (3.7) written in differential form, $\mathrm{d}t = \pm\sqrt{m/2}\;[E - V_e(r)]^{-1/2}\;\mathrm{d}r$, yields the equation for the trajectory,

$$\varphi(r) = \varphi_0 \pm \frac{p_\varphi}{\sqrt{2m}}\int_{r_0}^{r} \frac{\mathrm{d}\xi}{\xi^2\,\sqrt{E - V_e(\xi)}}\,, \tag{3.8}$$

with $\varphi_0 = \varphi(t = 0)$. Finally, elimination of r from $t(r)$ and $\varphi(r)$ provides the time evolution of φ. It is convenient to let the initial time $t = 0$ correspond to an extremum of the possible values of r. In this way the sign of t and $\varphi - \varphi_0$ changes at $r = r_0$. By this choice the trajectory is symmetric with respect to the line drawn from the origin to the point of coordinates r_0, φ_0, and the evolution of the particle's motion over each half of the trajectory is symmetric with respect to time.

3.5 Two-Particle Collision

Consider a system made of two particles whose masses are m_1, m_2. The system is isolated, namely, the particles are not subjected to any forces apart those due to the mutual interaction. As a consequence, the Lagrangian function is invariant under coordinate translations, and the Hamiltonian function is invariant under time translations. Thus, as shown in Sect. 2.6, the total momentum and total energy of the system are conserved.

The type of motion that is considered is such that the distance between the particles is initially so large as to make the interaction negligible. The interaction becomes significant when the particles come closer to each other; when they move apart, the interaction becomes negligible again. This type of interaction is called *collision*. The values of the dynamic quantities that hold when the interaction is negligible are indicated as *asymptotic values*. The labels a and b will be used to mark the asymptotic values before and after the interaction, respectively.

It is worth specifying that it is assumed that the collision does not change the internal state of the particles (for this reason it is more appropriately termed *elastic collision* [84, Sect. 17]). When the distance is sufficiently large, the particles can be considered as isolated: they move at a constant velocity and the total energy of the system is purely kinetic, $E_a = T_a$ and $E_b = T_b$. On the other hand, due to the invariance under time translation the total energy of the system is conserved, $E_b = E_a$. In conclusion it is $T_b = T_a$, namely, in an elastic collision the asymptotic kinetic energy of the system is conserved.

An analysis of the collision based only on the asymptotic values is incomplete because it does not take into account the details of the interaction between the two particles. However it provides a number of useful results, so it is worth pursuing. Letting $\mathbf{r}_1$ and $\mathbf{r}_2$ be the positions of the particles in a reference O, the position of the center of mass and the relative position of the particles are

$$\mathbf{R} = \frac{m_1\,\mathbf{r}_1 + m_2\,\mathbf{r}_2}{m_1 + m_2}\,, \qquad \mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2\,. \tag{3.9}$$

The corresponding velocities are $\mathbf{v}_1 = \dot{\mathbf{r}}_1$, $\mathbf{v}_2 = \dot{\mathbf{r}}_2$, and $\mathbf{v} = \dot{\mathbf{r}}$. The relations between the velocities are obtained by differentiating (3.9) with respect to time. Solving for $\mathbf{v}_1$, $\mathbf{v}_2$ yields

$$\mathbf{v}_1 = \dot{\mathbf{R}} + \frac{m_2}{m_1 + m_2}\,\mathbf{v}\,, \qquad \mathbf{v}_2 = \dot{\mathbf{R}} - \frac{m_1}{m_1 + m_2}\,\mathbf{v}\,. \tag{3.10}$$

Letting $\dot{R} = |\dot{\mathbf{R}}|$, the system's kinetic energy before the interaction is

$$T_a = \frac{1}{2}\,m_1\,v_{1a}^2 + \frac{1}{2}\,m_2\,v_{2a}^2 = \frac{1}{2}\,(m_1 + m_2)\,\dot{R}_a^2 + \frac{1}{2}\,m\,v_a^2\,, \tag{3.11}$$

where $m = m_1\,m_2/(m_1 + m_2)$ is called *reduced mass*. The expression of the kinetic energy after the interaction is obtained from (3.11) by replacing a with b.

The total momentum before the collision is $\mathbf{P}_a = m_1\,\mathbf{v}_{1a} + m_2\,\mathbf{v}_{2a} = (m_1 + m_2)\,\dot{\mathbf{R}}_a$. The conservation of $\mathbf{P}$ due to the invariance under coordinate translations yields $\mathbf{P}_b = \mathbf{P}_a$, whence $\dot{\mathbf{R}}_b = \dot{\mathbf{R}}_a$. Using (3.11) in combination with the conservation rules $\dot{\mathbf{R}}_b = \dot{\mathbf{R}}_a$ and $T_b = T_a$ yields $v_b = v_a$, namely, the asymptotic modulus of the relative velocity is conserved.

The analysis is now repeated in a new reference B in which the particles' positions are defined as

$$\mathbf{s}_1 = \mathbf{r}_1 - \mathbf{R} = \frac{m_2}{m_1 + m_2}\,(\mathbf{r}_1 - \mathbf{r}_2)\,, \quad \mathbf{s}_2 = \mathbf{r}_2 - \mathbf{R} = \frac{m_1}{m_1 + m_2}\,(\mathbf{r}_2 - \mathbf{r}_1)\,. \tag{3.12}$$

By construction, the origin of B coincides with the system's center of mass. The relative position in B is the same as in O, in fact

$$\mathbf{s} = \mathbf{s}_1 - \mathbf{s}_2 = \mathbf{r}_1 - \mathbf{r}_2 = \mathbf{r}\,. \tag{3.13}$$

From (3.12, 3.13) one finds

$$m_1\,\mathbf{s}_1 = -m_2\,\mathbf{s}_2\,, \qquad \mathbf{s}_1 = \frac{m_2}{m_1 + m_2}\,\mathbf{s}\,, \qquad \mathbf{s}_2 = -\frac{m_1}{m_1 + m_2}\,\mathbf{s}\,. \tag{3.14}$$

The velocities in reference B are $\mathbf{u}_1 = \dot{\mathbf{s}}_1$, $\mathbf{u}_2 = \dot{\mathbf{s}}_2$, and $\mathbf{u} = \dot{\mathbf{s}}$. The relations among the latter are found by differentiating (3.12, 3.13) and read

$$\mathbf{u}_1 = \mathbf{v}_1 - \dot{\mathbf{R}}\,, \qquad \mathbf{u}_2 = \mathbf{v}_2 - \dot{\mathbf{R}}\,, \qquad \mathbf{u} = \mathbf{v}\,, \tag{3.15}$$

$$\mathbf{u}_1 = \frac{m_2}{m_1 + m_2}\,\mathbf{u}\,, \qquad \mathbf{u}_2 = -\frac{m_1}{m_1 + m_2}\,\mathbf{u}\,, \tag{3.16}$$

which in turn yield

$$\mathbf{v}_1 = \dot{\mathbf{R}} + \frac{m_2}{m_1 + m_2}\,\mathbf{u}\,, \qquad \mathbf{v}_2 = \dot{\mathbf{R}} - \frac{m_1}{m_1 + m_2}\,\mathbf{u}\,. \tag{3.17}$$

Thanks to (3.16) the system's kinetic energy before and after the interaction, in reference B, is

$$K_a = \frac{1}{2} m_1 u_{1a}^2 + \frac{1}{2} m_2 u_{2a}^2 = \frac{1}{2} m u_a^2 , \qquad K_b = \frac{1}{2} m u_b^2 . \tag{3.18}$$

The conservation of the kinetic energy, $K_b = K_a$, yields $u_b = u_a$. Using the third of (3.15) then yields

$$u_b = u_a = v_b = v_a , \tag{3.19}$$

that is, the asymptotic modulus of the relative velocity is conserved and has the same value in the two references. Moreover, (3.16) shows that it is also $u_{1b} = u_{1a}$ and $u_{2b} = u_{2a}$, namely, in reference B the asymptotic kinetic energy is conserved for each particle separately.

3.6 Energy Exchange in the Two-Particle Collision

To complete the asymptotic analysis of the two-particle collision it is useful to choose for O a reference such that $\mathbf{v}_{2a} = 0$. In this case (3.10) yields $\mathbf{v}_a = \mathbf{v}_{1a}$, whence the total momentum reads $(m_1 + m_2)\,\dot{\mathbf{R}}_a = m_1 \mathbf{v}_{1a} = m_1 \mathbf{v}_a$. Remembering that $\dot{\mathbf{R}}_b = \dot{\mathbf{R}}_a$ one finds $\dot{\mathbf{R}}_b = m_1\,\mathbf{v}_a/(m_1 + m_2)$. Using the latter relation in the second of (3.17) specified with the b label yields, after multiplying both sides by m_2,

$$m_2\,\mathbf{v}_{2b} = m\,\mathbf{v}_a - m\,\mathbf{u}_b . \tag{3.20}$$

The triangle formed by the vectors $m_2\mathbf{v}_{2b}$, $m\mathbf{v}_a$, and $m\mathbf{u}_b$ is isosceles because $\mathbf{v}_a$ and $\mathbf{u}_b$ have the same modulus (Fig. 3.3). Letting χ, θ be the angle between $m\,\mathbf{v}_a$ and $m\,\mathbf{u}_b$ and, respectively, the common value of the other two angles, a scalar multiplication of (3.20) by $\mathbf{v}_a$ yields $m_2\,v_{2b}\cos\theta = m\,v_a - m\,u_b\cos\chi = m\,v_a\,(1 - \cos\chi)$. Using $2\theta + \chi = \pi$ and $v_a = v_{1a}$ transforms the latter into

$$m_2 v_{2b} = m v_{1a}\,\frac{1 - \cos\chi}{\cos[(\pi - \chi)/2]} = 2 m v_{1a}\,\sin(\chi/2) . \tag{3.21}$$

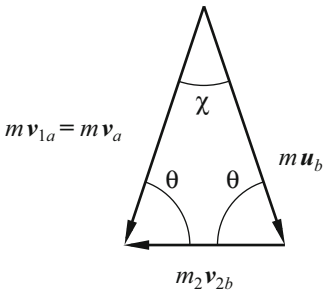


Fig. 3.3 Graphic representation of the vector relation (3.20)

This relation allows one to calculate, in reference O where the particle of mass m_2 is initially at rest, the modulus of the final velocity of this particle in terms of the initial velocity of the other particle and of angle χ. As only v_{1a} is prescribed, the two quantities v_{2b} and χ cannot be determined separately. The reason for this (already mentioned in Sect. 3.5) is that (3.21) is deduced using the motion's asymptotic conditions without considering the interaction's details. In fact, the calculation is based only on the momentum and total-energy conservation and on the hypothesis that the collision is elastic. From (3.21) one derives the relation between the kinetic energies $T_{1a} = (1/2)\, m_1 v_{1a}^2$ and $T_{2b} = (1/2)\, m_2 v_{2b}^2$,

$$T_{2b}\,(\chi) = \frac{4\, m_1\, m_2}{(m_1 + m_2)^2}\, T_{1a}\, \sin^2(\chi/2)\,. \tag{3.22}$$

As in reference O the particle of mass m_2 is initially at rest, T_{2b} is the variation of the kinetic energy of this particle due to the collision, expressed in terms of χ. The maximum variation is $T_{2b}(\chi = \pm\pi)$. The conservation relation for the kinetic energy $T_{1b} + T_{2b} = T_{1a}$ coupled with (3.22) yields the kinetic energy of the particle of mass m_1 after the collision,

$$T_{1b} = T_{1a} - T_{2b} = \left[1 - \frac{4\, m_1\, m_2}{(m_1 + m_2)^2}\, \sin^2(\chi/2)\right] T_{1a}\,. \tag{3.23}$$

Expressing T_{1a} and T_{1a} in (3.23) in terms of the corresponding velocities yields the modulus of the final velocity of the particle of mass m_1 as a function of its initial velocity and of angle χ,

$$v_{1b} = [(m_1^2 + m_2^2 + 2m_1 m_2 \cos\chi)^{1/2}/(m_1 + m_2)]\, v_{1a}\,. \tag{3.24}$$

Although expressions (3.21–3.24) are compact, the use of angle χ is inconvenient. It is preferable to use the angle, say ψ, between vectors $\mathbf{v}_{1b}$ and $\mathbf{v}_{1a} = \mathbf{v}_a$ that belong to the same reference O (Fig. 3.7).[1] A scalar multiplication by $\mathbf{v}_{1a}$ of the conservation relation for momentum, $m_1\mathbf{v}_{1b} = m_1\mathbf{v}_{1a} - m_2\mathbf{v}_{2b}$, followed by the replacement of the expressions of v_{2b} and v_{1b} extracted from (3.21, 3.24), eventually yields $\cos\psi = (m_1 + m_2\cos\chi)/(m_1^2 + m_2^2 + 2m_1 m_2\cos\chi)^{1/2}$. Squaring both sides of the latter provides the relation between χ and ψ,

$$\tan\psi = \frac{\sin\chi}{m_1/m_2 + \cos\chi}\,. \tag{3.25}$$

[1] Note that the angle γ between two momenta $\mathbf{p}' = m'\mathbf{v}'$ and $\mathbf{p}'' = m''\mathbf{v}''$ is the same as that between the corresponding velocities because the masses cancel out in the angle's definition $\gamma = \arccos[\mathbf{p}' \cdot \mathbf{p}''/(p'p'')]$. In contrast, a relation like (3.20) involving a triad of vectors holds for the momenta but not (with the exception of the trivial cases of equal masses) for the corresponding velocities.

Using (3.25) one expresses (3.21–3.24) in terms of the deflection ψ (in reference O) of the particle of mass m_1. If $m_1 > m_2$, then $\psi < \pi/2$, while it is $\psi = \chi/2$ if $m_1 = m_2$. When $m_1 < m_2$ and $\chi = \arccos(-m_1/m_2)$, then $\psi = \pi/2$; finally, if $m_1 \ll m_2$ it is $\psi \simeq \chi$ and, from (3.21–3.24), it follows $v_{2b} \simeq 0$, $T_{2b} \simeq 0$, $v_{1b} \simeq v_{1a}$, $T_{1b} \simeq T_{1a}$. In other terms, when $m_1 \ll m_2$ the particle of mass m_2 remains at rest; the other particle is deflected, but its kinetic energy is left unchanged.

In reference O, the angle between the final velocity of the particle of mass m_2, initially at rest, and the initial velocity of the other particle has been defined above as $\theta = (\pi - \chi)/2$. Replacing the latter in (3.25) provides the relation between ψ and θ,

$$\tan\psi = \frac{\sin(2\theta)}{m_1/m_2 - \cos(2\theta)} \,. \tag{3.26}$$

3.7 Central Motion in the Two-Particle Interaction

Consider an isolated two-particle system where the force acting on each particle derives from a potential energy $V = V(\mathbf{r}_1, \mathbf{r}_2)$. Using the symbols defined in Sect. 3.5 yields the Lagrangian function $L = m_1 v_1^2/2 + m_2 v_2^2/2 - V(\mathbf{r}_1, \mathbf{r}_2)$. Now assume that the potential energy V depends on the position of the two particles only through the modulus of their distance, $r = |\mathbf{r}_1 - \mathbf{r}_2|$. In this case it is convenient to use the coordinates and velocities relative to the center of mass, (3.12–3.17), to find

$$L = \frac{1}{2}(m_1 + m_2)\,\dot{R}^2 + \frac{1}{2}\,m\,\dot{s}^2 - V(s)\,, \qquad \dot{s} = |\dot{\mathbf{s}}| = |\mathbf{u}|\,. \tag{3.27}$$

As discussed in Sects. 2.6.3 and 3.5 the total momentum is conserved, whence $\dot{R}$ is constant. Another way of proving this property is noting that the components of $\mathbf{R}$ are cyclic (Sect. 1.6). The first term at the right-hand side of (3.27), being a constant, does not influence the subsequent calculations. The remaining terms, in turn, are identical to those of (3.4). This shows that when in a two-particle system the potential energy depends only on the relative distance, adopting suitable coordinates makes the problem identical to that of the central motion. One can then exploit the results of Sect. 3.4. Once the time evolution of $\mathbf{s}$ is found, the description of the motion of the individual particles is recovered from (3.12) to (3.17), where the constant $\dot{\mathbf{R}}$ is determined by the initial conditions.

The total energy of the two-particle system is conserved and, in the new reference, it reads

$$\frac{1}{2}\,m\,\dot{s}^2 + V(s) = E_B\,, \qquad E_B = E - \frac{1}{2}(m_1 + m_2)\,\dot{R}^2\,. \tag{3.28}$$

The total angular momentum is constant as well. Starting from the original reference and using (3.12–3.17) yields

$$\mathbf{M} = \mathbf{r}_1 \wedge m_1 \mathbf{v}_1 + \mathbf{r}_2 \wedge m_2 \mathbf{v}_2 = (m_1 + m_2)\, \mathbf{R} \wedge \dot{\mathbf{R}} + m\, \mathbf{s} \wedge \mathbf{u}\,. \tag{3.29}$$

The constancy of $\dot{\mathbf{R}}$ yields $\mathbf{R} = \mathbf{R}_0 + \dot{\mathbf{R}}\, t$, with $\mathbf{R}_0$ the initial value of $\mathbf{R}$, whence $(\mathbf{R}_0 + \dot{\mathbf{R}}\, t) \wedge \dot{\mathbf{R}} = \mathbf{R}_0 \wedge \dot{\mathbf{R}}$. Thus, the first term at the right-hand side of (3.29) is constant, which makes $\mathbf{M}_B = m\, \mathbf{s} \wedge \mathbf{u}$ a constant as well. The latter vector is parallel to $\mathbf{M}$ because the motion is confined to a fixed plane (Sect. 3.4). Then, aligning the z axis with $\mathbf{M}$, turning to polar coordinates over the x, y plane ($s_x = s \cos\varphi$, $s_y = s \sin\varphi$), and using (3.8), one finds

$$\varphi(s) = \varphi_0 \pm \frac{M_B}{\sqrt{2m}} \int_{s_0}^{s} \frac{\mathrm{d}\xi}{\xi^2 \sqrt{E_B - V_e(\xi)}}\,, \tag{3.30}$$

with $V_e(s) = V(s) + M_B^2/(2ms^2)$. It is important to note that the factor M_B in (3.30) is the scalar coefficient of $\mathbf{M}_B = M_B \mathbf{k}$, with $\mathbf{k}$ the unit vector of the z axis. As a consequence, M_B may have a sign. As observed in Sect. 2.9, the admissible values of s are those belonging to the interval such that $E_B \geq V_e(s)$. If two or more disjoint intervals exist that have this property, the actual interval of the motion is determined by the initial conditions. The motion is limited or unlimited, depending on the extent of this interval.

The analysis cannot be pursued further unless the form of the potential energy V is specified. This is done in Sect. 3.8 with reference to the Coulomb case.

3.8 Coulomb Field

An important example is that of a potential energy of the form $V \propto 1/r$ that occurs for the gravitational and for the electrostatic force. In the latter case the term *Coulomb potential energy* is used for V, that reads

$$V(s) = \frac{\kappa\, Z_1\, Z_2\, q^2}{4\pi\, \varepsilon_0\, s}\,, \qquad s > 0\,, \tag{3.31}$$

with $q > 0$ the elementary electric charge, $Z_1\, q$ and $Z_2\, q$ the absolute value of the net charge of the first and second particle, respectively, ε_0 the vacuum permittivity, and, finally, $\kappa = 1\ (-1)$ in the repulsive (attractive) case. The form of V fixes the additive constant of the energy so that $V(\infty) = 0$. The repulsive case only is considered here, whence V_e is strictly positive and $E_B \geq V_e > 0$. Defining the lengths

$$\lambda = \frac{Z_1\, Z_2\, q^2}{8\pi\, \varepsilon_0\, E_B} > 0\,, \qquad \mu = \frac{M_B}{\sqrt{2m\, E_B}} \tag{3.32}$$

yields $V_e/E_B = 2\lambda/s + \mu^2/s^2$. The zeros of $E_B - V_e = E_B\,(s^2 - 2\,\lambda\,s - \mu^2)/s^2$ are

$$s_A = \lambda - \sqrt{\lambda^2 + \mu^2}\,, \qquad s_B = \lambda + \sqrt{\lambda^2 + \mu^2}\,, \tag{3.33}$$

where s_A is negative and must be discarded as s is strictly positive. The only acceptable zero is then $s_B \geq 2\lambda > 0$ that corresponds to the position where the radial velocity $\dot{s} = \pm\sqrt{2\,(E_B - V)/m}$ reverses, and must therefore be identified with s_0 (Sect. 3.4). The definitions (3.32, 3.33) are now inserted into the expression (3.30) of the particle's trajectory. To calculate the integral it is convenient to use a new variable w such that $(s_0 - \lambda)/(\xi - \lambda) = (\mu^2 - s_0^2\,w^2)/(\mu^2 + s_0^2\,w^2)$. The range of w corresponding to $s_0 \leq \xi \leq s$ is

$$0 \leq w \leq \frac{|\mu|}{s_0}\sqrt{\frac{s - s_0}{s + s_0 - 2\lambda}}\,, \qquad s \geq s_0 \geq 2\lambda > 0\,. \tag{3.34}$$

From (3.30) the trajectory in the s, φ reference is thus found to be

$$\varphi(s) = \varphi_0 \pm 2\,\arctan\left(\frac{\mu}{s_0}\sqrt{\frac{s - s_0}{s + s_0 - 2\lambda}}\right). \tag{3.35}$$

Next, the trajectory in the Cartesian reference s_x, s_y is found by replacing (3.35) into $s_x = s\,\cos\varphi$, $s_y = s\,\sin\varphi$ and eliminating s from the pair $s_x(s)$, $s_y(s)$ thus found. A graphic example is given in Fig. 3.4. It is worth observing that in the derivation of (3.35) the factor $|\mu|$ appears twice, in such a way as to compensate for the sign of μ. The result then holds irrespective of the actual sign of $M_B = \sqrt{2m\,E_B}\,\mu$. It still holds for $M_B = 0$, that yields $\varphi(s) = \varphi_0$; such a case corresponds to a straight line crossing the origin of the s, φ reference: along this line the modulus s of the relative position decreases until it reaches s_0, then it increases from this point on.

When $M_B \neq 0$ the angles corresponding to the asymptotic conditions of the motion are found by letting $s \to \infty$, namely, $\varphi_a = \varphi_0 - 2\,\arctan(\mu/s_0)$ and $\varphi_b = \varphi_0 + 2\,\arctan(\mu/s_0)$. The total deflection is then $\varphi_b - \varphi_a$ which, in each curve of Fig. 3.4, is the angle between the two asymptotic directions. Now one combines the definition of angle χ given in Sect. 3.6 with the equality $\mathbf{u} = \mathbf{v}$ taken from the last of (3.15); with the aid of Fig. 3.5 one finds

$$\chi = \pi - (\varphi_b - \varphi_a) = \pi - 4\,\arctan\left(\frac{\mu}{s_0}\right). \tag{3.36}$$

The definitions (3.32, 3.33) show that (3.36) eventually provides the relation $\chi = \chi(E_B, M_B)$. In contrast with the approach of Sect. 3.6, where the asymptotic conditions only were considered, here the analysis has been brought to the end by considering a specific type of interaction.

When μ ranges from $-\infty$ to $+\infty$ at a fixed E_B, the definitions (3.32, 3.33) make the ratio $\mu/s_0 = \mu/s_A$ to range from -1 to $+1$. If $\mu/s_0 = 1$ (-1), then $\chi = 0$ (2π), namely, no deflection between $\mathbf{u}_a$ and $\mathbf{u}_b$ occurs. If $\mu/s_0 = 0$, then $\chi = \pi$, namely,

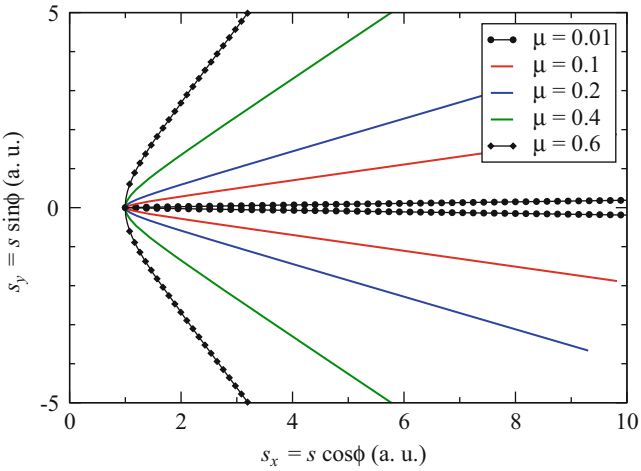


Fig. 3.4 Graphic representation of the trajectory (3.35) for different values of the angular momentum. The curves have been obtained by setting the parameters' values to $s_0 = 1$, $\varphi_0 = 0$, $\lambda = 0.5$, and $\mu = 0.01, \ldots, 0.6$ (the units are arbitrary)

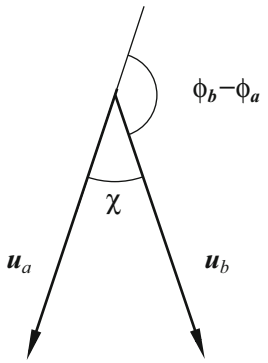


Fig. 3.5 Graphic representation of (3.36)

the motion's direction reverses at s_0 as noted above. From $\mathbf{u} = \mathbf{v}$ one finds that χ is also the angle between $\mathbf{v}_a = \mathbf{v}_{1a} - \mathbf{v}_{2a}$ and $\mathbf{v}_b = \mathbf{v}_{1b} - \mathbf{v}_{2b}$.

3.9 System of Particles Near an Equilibrium Point

Consider a system of N particles, not necessarily identical to each other, subjected to conservative forces. The mass and instantaneous position of the jth particle are indicated with m_j and $\mathbf{R}_j = (X_{j1}, X_{j2}, X_{j3})$, respectively. It is assumed that there are no constraints, so that the number of degrees of freedom of the system is $3N$. The Hamiltonian function reads

$$H_a = T_a + V_a = \sum_{j=1}^{N} \frac{P_j^2}{2\,m_j} + V_a(X_{11}, X_{12}, \ldots)\,, \tag{3.37}$$

with $P_j^2 = m_j^2\,(\dot{X}_{j1}^2 + \dot{X}_{j2}^2 + \dot{X}_{j3}^2)$. The force acting on the jth particle along the kth axis is $F_{jk} = -\partial V_a/\partial X_{jk}$. The Hamilton equations (Sect. 1.6) read

$$\dot{X}_{jk} = \frac{\partial H_a}{\partial P_{jk}} = \frac{P_{jk}}{m_j}\,, \qquad \dot{P}_{jk} = -\frac{\partial H_a}{\partial X_{jk}} = -\frac{\partial V_a}{\partial X_{jk}} = F_{jk}\,. \tag{3.38}$$

They show that the relation $F_{jk} = m_j\,\ddot{X}_{jk}$, which yields the dynamics of the jth particle along the kth axis, involves the positions of all particles in the system due to the coupling of the latter.

Define the $3N$-dimensional vector $\mathbf{R} = (X_{11}, \ldots, X_{N3})$ that describes the instantaneous position of the system in the configuration space, and let $\mathbf{R}_0$ be a position where the potential energy V_a has a minimum, namely, $(\partial V_a/\partial X_{jk})_{\mathbf{R}_0} = 0$ for all j, k. Such a position is called *equilibrium point* of the system. To proceed, assume that the instantaneous displacement $\mathbf{R} - \mathbf{R}_0$ with respect to the equilibrium point is small. In this case one approximates V with a second-order Taylor expansion around $\mathbf{R}_0$. To simplify the notation new symbols are adopted, namely, $s_1 = X_{11}$, $s_2 = X_{12}, \ldots, s_{3j+k-3} = X_{jk}, \ldots$, and $h_n = s_n - s_{n0}$, with $n = 1, 2, \ldots, 3N$ and s_{n0} the equilibrium position. Remembering that the first derivatives of V_a vanish at $\mathbf{R}_0$ one finds

$$V_a \simeq V_{a0} + \frac{1}{2} \sum_{k=1}^{3N} h_k \sum_{n=1}^{3N} c_{kn}\, h_n\,, \qquad c_{kn} = \left(\frac{\partial^2 V_a}{\partial h_k\, \partial h_n} \right)_{\mathbf{R}_0}\,. \tag{3.39}$$

In (3.39) it is $V_{a0} = V_a(\mathbf{R}_0)$, and the terms c_{kn} are called *elastic coefficients*. As the approximate form of the potential energy is quadratic in the displacements, each component of the force is a linear combination of the latter,

$$F_r = -\frac{\partial V_a}{\partial s_r} = -\frac{\partial V_a}{\partial h_r} = -\sum_{n=1}^{3N} c_{rn}\, h_n\,, \qquad r = 3j + k - 3\,. \tag{3.40}$$

To recast the kinetic energy in terms of the new symbols it is necessary to indicate the masses with μ_n, $n = 1, \ldots, 3N$, where $\mu_{3j-2} = \mu_{3j-1} = \mu_{3j} = m_j$, $j = 1, \ldots, N$. Observing that $\dot{X}_{jk} = \dot{s}_{3j+k-3} = \dot{h}_{3j+k-3}$, one finds a quadratic form in the derivatives of the displacements,

$$T_a = \frac{1}{2} \sum_{j=1}^{N} \sum_{k=1}^{3} \mu_{3j+k-3} \dot{X}_{jk}^2 = \frac{1}{2} \sum_{j=1}^{N} \sum_{k=1}^{3} \mu_{3j+k-3} \dot{h}_{3j+k-3}^2 = \frac{1}{2} \sum_{n=1}^{3N} \mu_n \dot{h}_n^2\,. \tag{3.41}$$

The relations obtained so far are readily recast in matrix form. First, one defines the *mass matrix* $\mathbf{M}$ as the real, $3N \times 3N$ diagonal matrix whose entries are $[\mathbf{M}]_{kn} = \mu_n \, \delta_{kn} > 0$, with δ_{kn} the Kronecker symbol (A.18). By construction, the mass matrix is symmetric and positive definite; the entries of its inverse are $[\mathbf{M}^{-1}]_{kn} = \delta_{kn}/\mu_n$. Then, one defines the *elastic matrix* $\mathbf{C}$ as the real, $3N \times 3N$ matrix whose entries are $[\mathbf{C}]_{kn} = c_{kn}$. The entries of the elastic matrix are the second derivatives of the potential energy V_a; as the order of the derivation is irrelevant, the matrix is symmetric. Also, the derivatives are calculated in a minimum of V_a; from the first of (3.39) it follows that the quadratic form at the right-hand side equals $V_a - V_{a0}$ which, by construction, is positive. It follows that the elastic matrix is positive definite, namely, for any choice of the displacements (excluding the case where all displacements are zero) the quadratic form generated by the matrix is positive. Finally, let $\mathbf{h}$ be the column vector of entries $h_1, h_2, \ldots$, and $\mathbf{h}^T$ its transpose. Combining (3.37, 3.39, 3.41) expresses the Hamiltonian function in terms of the sum of two quadratic forms,

$$H_a - V_{a0} = \frac{1}{2} \dot{\mathbf{h}}^T \mathbf{M} \dot{\mathbf{h}} + \frac{1}{2} \mathbf{h}^T \mathbf{C} \, \mathbf{h} \, . \tag{3.42}$$

3.10 Diagonalization of the Hamiltonian Function

Thanks to the properties of the matrices $\mathbf{M}$ and $\mathbf{C}$, the right-hand side of (3.42) can be set in diagonal form. To this purpose one considers the eigenvalue equation

$$\mathbf{C} \, \mathbf{g}_\sigma = \lambda_\sigma \, \mathbf{M} \, \mathbf{g}_\sigma \, , \qquad \sigma = 1, \ldots, 3N \, , \tag{3.43}$$

where the eigenvalues λ_σ are real because $\mathbf{C}$ and $\mathbf{M}$ are real and symmetric. As all coefficients of (3.43) are real, the eigenvectors $\mathbf{g}_\sigma$ are real as well. Also, due to the positive definiteness of $\mathbf{C}$ and $\mathbf{M}$, the eigenvalues are positive and the eigenvectors are linearly independent. They can also be selected in order to fulfill the property of being orthonormal with respect to $\mathbf{M}$, namely, $\mathbf{g}_\sigma^T \, \mathbf{M} \, \mathbf{g}_\tau = \delta_{\sigma\tau}$.

Each of the $3N$ eigenvectors $\mathbf{g}_\sigma$ has $3N$ entries. Thus, the set of eigenvectors can be arranged to form a $3N \times 3N$ real matrix $\mathbf{G}$, whose σth column is the σth eigenvector. The inverse matrix $\mathbf{G}^{-1}$ exists because, by construction, the columns of $\mathbf{G}$ are linearly independent. The orthonormality relation between the eigenvectors can now be expressed in matrix form as[2]

$$\mathbf{G}^T \, \mathbf{M} \, \mathbf{G} = \mathbf{I} \, , \tag{3.44}$$

[2]The procedure illustrated here is an application to the case of real matrices of the unitary transformation illustrated in Sect. A.11.2.

with $\mathbf{I}$ the identity matrix. Equation (3.44) is the basic ingredient for the diagonalization of (3.42). From it one preliminarily derives four more relations,

$$\mathbf{G}^T\,\mathbf{M} = \mathbf{G}^{-1}\,, \qquad \mathbf{G}\,\mathbf{G}^T\,\mathbf{M} = \mathbf{I}\,, \qquad \mathbf{M}\,\mathbf{G} = \left(\mathbf{G}^T\right)^{-1}\,, \qquad \mathbf{M}\,\mathbf{G}\,\mathbf{G}^T = \mathbf{I}\,. \tag{3.45}$$

The first of (3.45) is obtained by right multiplying (3.44) by $\mathbf{G}^{-1}$ and using $\mathbf{G}\,\mathbf{G}^{-1} = \mathbf{I}$. Left multiplying by $\mathbf{G}$ the first of (3.45) yields the second one. The third of (3.45) is obtained by left multiplying (3.44) by $(\mathbf{G}^T)^{-1}$. Finally, right multiplying by $\mathbf{G}^T$ the third of (3.45) yields the fourth one. To complete the transformation of the equations into a matrix form one defines the eigenvalue matrix $\mathbf{L}$ as the real, $3N \times 3N$ diagonal matrix whose entries are $[\mathbf{L}]_{\sigma\tau} = \lambda_\tau\,\delta_{\sigma\tau} > 0$. The set of $3N$ eigenvalue equations (3.43) then takes one of the two equivalent forms

$$\mathbf{C}\,\mathbf{G} = \mathbf{M}\,\mathbf{G}\,\mathbf{L}\,, \qquad \mathbf{G}^T\,\mathbf{C}\,\mathbf{G} = \mathbf{L}\,. \tag{3.46}$$

The first of (3.46) is the analogue of (3.43), while the second form is obtained from the first one by left multiplying by $\mathbf{G}^T$ and using (3.44). The diagonalization of (3.42) is now accomplished by inserting the second and fourth of (3.45) into the potential-energy term of (3.42) to obtain

$$\mathbf{h}^T\,\mathbf{C}\,\mathbf{h} = \mathbf{h}^T\,(\mathbf{M}\,\mathbf{G}\,\mathbf{G}^T)\,\mathbf{C}\,(\mathbf{G}\,\mathbf{G}^T\,\mathbf{M})\,\mathbf{h} = (\mathbf{h}^T\,\mathbf{M}\,\mathbf{G})\,(\mathbf{G}^T\,\mathbf{C}\,\mathbf{G})\,(\mathbf{G}^T\,\mathbf{M}\,\mathbf{h})\,, \tag{3.47}$$

where the associative law has been used. At the right-hand side of (3.47), the term in the central parenthesis is replaced with $\mathbf{L}$ due to the second of (3.46). The term in the last parenthesis is a column vector for which the short-hand notation $\mathbf{b} = \mathbf{G}^T\,\mathbf{M}\,\mathbf{h}$ is introduced. Note that $\mathbf{b}$ depends on time because $\mathbf{h}$ does. The first of (3.45) shows that $\mathbf{h} = \mathbf{G}\,\mathbf{b}$, whence $\mathbf{h}^T = \mathbf{b}^T\,\mathbf{G}^T$. Finally, using (3.44), transforms the term in the first parenthesis at the right-hand side of (3.47) into $\mathbf{h}^T\,\mathbf{M}\,\mathbf{G} = \mathbf{b}^T\,\mathbf{G}^T\,\mathbf{M}\,\mathbf{G} = \mathbf{b}^T$. In conclusion, the potential-energy term of (3.42) is recast in terms of $\mathbf{b}$ as $\mathbf{h}^T\,\mathbf{C}\,\mathbf{h} = \mathbf{b}^T\,\mathbf{L}\,\mathbf{b}$, which is the diagonal form sought. By a similar procedure one finds for the kinetic-energy term $\dot{\mathbf{h}}^T\,\mathbf{M}\,\dot{\mathbf{h}} = \dot{\mathbf{b}}^T\,\mathbf{G}^T\,\mathbf{M}\,\mathbf{G}\,\dot{\mathbf{b}} = \dot{\mathbf{b}}^T\,\dot{\mathbf{b}}$.

The terms $\dot{\mathbf{b}}^T\,\dot{\mathbf{b}}$ and $\mathbf{b}^T\,\mathbf{L}\,\mathbf{b}$ have the same units. As a consequence, the units of $\mathbf{L}$ are the inverse of a time squared. Remembering that the entries of $\mathbf{L}$ are positive, one introduces the new symbol $\omega_\sigma^2 = \lambda_\sigma$ for the eigenvalues, $\omega_\sigma > 0$. In conclusion, the diagonal form of (3.42) reads

$$H_a - V_{a0} = \sum_{\sigma=1}^{3N} H_\sigma\,, \qquad H_\sigma = \frac{1}{2}\,\dot{b}_\sigma^2 + \frac{1}{2}\,\omega_\sigma^2\,b_\sigma^2\,. \tag{3.48}$$

Apart from the constant V_{a0}, the Hamiltonian function H_a is given by a sum of terms, each associated with a single degree of freedom. A comparison with (3.1) shows that the individual summands H_σ are identical to the Hamiltonian function of a linear harmonic oscillator with $m = 1$. As a consequence, the two canonical

variables of the σth degree of freedom are $q_\sigma = b_\sigma, p_\sigma = \dot{b}_\sigma$, and the time evolution of b_σ is the same as that in (3.2),

$$b_\sigma(t) = \beta_\sigma \cos(\omega_\sigma t + \varphi_\sigma) = \frac{1}{2}\left[\tilde{\beta}_\sigma \exp(-\mathrm{i}\omega_\sigma t) + \tilde{\beta}_\sigma^* \exp(\mathrm{i}\omega_\sigma t)\right]. \tag{3.49}$$

The constants β_σ, φ_σ depend on the initial conditions $b_\sigma(0)$, $\dot{b}_\sigma(0)$. The complex coefficients are related to the above constants by $\tilde{\beta}_\sigma = \beta_\sigma \exp(-\mathrm{i}\varphi_\sigma)$. In turn, the initial conditions are derived from those of the displacements, $\mathbf{b}(0) = \mathbf{G}^{-1}\mathbf{h}(0)$, $\dot{\mathbf{b}}(0) = \mathbf{G}^{-1}\dot{\mathbf{h}}(0)$.

The $3N$ functions $b_\sigma(t)$ are called *normal coordinates* or *principal coordinates*. Once the normal coordinates have been found, the displacements of the particles are determined from $\mathbf{h} = \mathbf{G}\,\mathbf{b}$. It follows that such displacements are superpositions of oscillatory functions. Despite the complicacy of the system, the approximation of truncating the potential energy to the second order makes the Hamiltonian function completely separable in the normal coordinates. The problem then becomes a generalization of that of the linear harmonic oscillator (Sect. 3.3), and the frequencies of the oscillators are determined by combining the system parameters, specifically, the particle masses and elastic constants. The Hamiltonian function associated with each degree of freedom is a constant of motion, $H_\sigma = E_\sigma$, whose value is prescribed by the initial conditions. The total energy of the system is also a constant and is given by

$$E = V_{a0} + \sum_{\sigma=1}^{3N} E_\sigma\,, \qquad E_\sigma = \frac{1}{2}\dot{b}_\sigma^2(0) + \frac{1}{2}\omega_\sigma^2 b_\sigma^2(0)\,. \tag{3.50}$$

The oscillation of the normal coordinate of index σ is also called *mode* of the vibrating system.

3.11 Periodic Potential Energy

An interesting application of the action-angle variables introduced in Sect. 2.10 is found in the case of a conservative motion where the potential energy V is periodic. For simplicity a linear motion is considered (Sect. 2.9), whence $V(x + a) = V(x)$, with $a > 0$ the spatial period. Letting E be the total energy and m the mass of the particle, an unlimited motion is assumed, namely, $E > V$; it follows that the momentum $p = \sqrt{2\,m\,[E - V(x)]}$ is a spatially periodic function of period a whence, according to the definition of Sect. 2.10, the motion is a rotation. For any position g, the time τ necessary for the particle to move from g to $g + a$ is found from (2.47), where the positive sign is provisionally chosen:

$$\tau = \sqrt{\frac{m}{2}} \int_g^{g+a} \frac{\mathrm{d}x}{\sqrt{E - V(x)}} > 0\,. \tag{3.51}$$

The integral in (3.51) is independent of g due to the periodicity of V. As a consequence, for any g the position of the particle grows by a during the time τ. The action variable is found from (2.49):

$$J(E) = \int_g^{g+a} p\,\mathrm{d}x = \sqrt{2m} \int_g^{g+a} \sqrt{E - V(x)}\,\mathrm{d}x = \text{const}\,. \tag{3.52}$$

In turn, the derivative of the angle variable is found from (2.51). It reads $\dot{w} = \partial H/\partial J = 1/(\mathrm{d}J/\mathrm{d}E) = \text{const}$, with H the Hamiltonian function. The second form of $\dot{w}$ holds because H does not depend on w, and $H = E$. Using (3.52) and comparing with (3.51) one finds

$$\frac{1}{\dot{w}} = \frac{\mathrm{d}J}{\mathrm{d}E} = \int_g^{g+a} \frac{m}{[2\,m\,(E - V(x))]^{1/2}}\,\mathrm{d}x = \tau\,. \tag{3.53}$$

As expected, $1/\tau$ is the rotation frequency. In conclusion, the time evolution of the action-angle variables is given by $w = t/\tau + w_0$, $J = \text{const}$. Note that the relation (3.52) between E and J holds when the positive sign is chosen in (2.47); if the above calculations are repeated after choosing the negative sign, one finds that $-J$ is associated with the same E. As a consequence, E is an even function of J.

Another observation is that the action-angle variables can be scaled by letting, e.g., $w\,J = (a\,w)\,(J/a)$. In this way the property that the product of two canonically conjugate variables is dimensionally an action is still fulfilled. A comparison with (3.52) shows that, thanks to this choice of the scaling factor, $P = J/a$ is the average momentum over a period, while $X = a\,w$ is a length. The Hamilton equations and the time evolution of the new variables are then

$$\dot{X} = \frac{\partial H}{\partial P}\,, \quad \dot{P} = -\frac{\partial H}{\partial X} = 0\,, \qquad X = \frac{a}{\tau}\,t + X_0\,, \quad P = P_0 = \text{const}\,, \tag{3.54}$$

where a/τ is the average velocity of the particle over the spatial period, and $X_0 = X(0)$, $P_0 = P(0)$. In conclusion, in the new canonical variables no force is acting ($\dot{P} = 0$), and the motion of the new position X is uniform in time. However, the relation between E and $P = J/a$, given by (3.52), is not quadratic as it would be in free space.[3]

In many cases it is of interest to investigate the particle's dynamics when a perturbation δH is superimposed to the periodic potential energy V. It is assumed that δH depends on x only, and that E is the same as in the unperturbed case (the latter assumption is not essential). The Hamiltonian function of the perturbed case

[3] Compare with comments made in Sect. 19.6.1.

is then written as the sum of the unperturbed one and of the perturbation; also in this case an unlimited motion is assumed, specifically, $E > V$ and $E > V + \delta H$. Still using the positive sign for the momentum, one finds

$$H(x,p) = \frac{p^2}{2m} + V(x) + \delta H(x) = E\,, \qquad p(x,E) = \sqrt{2\,m\,(E - V - \delta H)}\,. \tag{3.55}$$

Like in the unperturbed case one defines the average momentum over a,

$$\tilde{P}(g,E) = \frac{\sqrt{2m}}{a} \int_g^{g+a} \sqrt{E - V - \delta H}\,\mathrm{d}x\,, \tag{3.56}$$

which depends also on g because δH is not periodic. Differentiating (3.56) with respect to E and comparing with (3.51) show that

$$\frac{\partial \tilde{P}}{\partial E} = \frac{\tilde{\tau}}{a}\,, \qquad \tilde{\tau}(g) = \int_g^{g+a} \frac{m}{[2\,m\,(E - V(x) - \delta H)]^{1/2}}\,\mathrm{d}x\,, \tag{3.57}$$

with $\tilde{\tau}$ the time necessary for the particle to move from g to $g + a$ in the perturbed case. Using $H = E$ in the above yields

$$\frac{\partial H}{\partial \tilde{P}} = \frac{a}{\tilde{\tau}} = \frac{(g+a) - g}{\tilde{\tau}}\,. \tag{3.58}$$

So far no hypothesis has been made about the perturbation. Now one assumes that δH is weak and varies little over the period a. The first hypothesis implies $|\delta H| \ll E - V$ so that, to first order, $[2\,m\,(E - V - \delta H)]^{1/2} \simeq [2\,m\,(E - V)]^{1/2} - m\,[2\,m\,(E - V)]^{-1/2}\,\delta H$. Using $P = J/a$ and (3.52), the average momentum (3.56) becomes

$$\tilde{P}(g,E) \simeq P(E) - \frac{1}{a} \int_g^{g+a} \frac{m\,\delta H}{[2\,m\,(E - V)]^{1/2}}\,\mathrm{d}x\,. \tag{3.59}$$

In turn, the hypothesis that the perturbation varies little over the period a implies that in the interval $[g, g + a]$ one can approximate $\delta H(x)$ with $\delta H(g)$, which transforms (3.59), due to (3.53), into

$$\tilde{P}(g,E) \simeq P(E) - \frac{\tau}{a}\,\delta H(g)\,. \tag{3.60}$$

If the procedure leading to (3.60) is repeated in the interval $[g + a, g + 2\,a]$ and the result is subtracted from (3.60), the following is found:

$$\frac{\tilde{P}(g+a,E) - \tilde{P}(g,E)}{\tau} = -\frac{\delta H(g+a) - \delta H(g)}{a}\,. \tag{3.61}$$

The above shows that the perturbed momentum $\tilde{P}$ varies between g and $g + a$ due to the corresponding variation in δH. Fixing the time origin at the position g and letting $\tau \simeq \tilde{\tau}$ in the denominator transform (3.61) into

$$\frac{\tilde{P}(\tilde{\tau}, E) - \tilde{P}(0, E)}{\tilde{\tau}} \simeq -\frac{\delta H(g + a) - \delta H(g)}{a} . \tag{3.62}$$

The relations (3.58, 3.62) are worth discussing. If one considers g as a position coordinate and $\tilde{P}$ as the momentum conjugate to it, (3.58, 3.62) become a pair of Hamilton equations where some derivatives are replaced with difference quotients. Specifically, (3.62) shows that the average momentum varies so that its "coarse-grained" variation with respect to time, $\Delta\tilde{P}/\Delta\tilde{\tau}$, is the negative coarse-grained variation of the Hamiltonian function with respect to space, $-\Delta H/\Delta g = -\Delta\delta H/\Delta g$. In turn, (3.58) shows that the coarse-grained variation of position with respect to time, $\Delta g/\Delta\tilde{\tau}$, is the derivative of the Hamiltonian function with respect to the average momentum. In conclusion, (3.58, 3.62) are useful when one is not interested in the details of the particle's motion within each spatial period, but wants to investigate on a larger scale how the perturbation influences the average properties of the motion.

3.12 Energy-Momentum Relation in a Periodic Potential Energy

It has been observed, with reference to the non-perturbed case, that the relation (3.52) between the total energy and the average momentum is not quadratic. In the perturbed case, as shown by (3.56), the momentum depends on both the total energy and the coarse-grained position. To investigate this case it is then necessary to fix g and consider the dependence of $\tilde{P}$ on E only. To proceed one takes a small interval of $\tilde{P}$ around a given value, say $\tilde{P}_s$, corresponding to a total energy E_s, and approximates the $E(\tilde{P})$ relation with a second-order Taylor expansion around $\tilde{P}$,

$$E \simeq E_s + \left(\frac{\mathrm{d}E}{\mathrm{d}\tilde{P}}\right)_s (\tilde{P} - \tilde{P}_s) + \frac{1}{2}\left(\frac{\mathrm{d}^2E}{\mathrm{d}\tilde{P}^2}\right)_s (\tilde{P} - \tilde{P}_s)^2 . \tag{3.63}$$

Although in general the $\tilde{P}(E)$ relation (3.56) cannot be inverted analytically, one can calculate the derivatives that appear in (3.63). The latter are worked out in the unperturbed case $\delta H = 0$ for simplicity. Using (3.53), the first derivative is found to be $(\mathrm{d}E/\mathrm{d}\tilde{P})_s \simeq (\mathrm{d}E/\mathrm{d}P)_s = a/\tau_s$, with $\tau_s = \tau(E_s)$. For the second derivative,

$$\frac{\mathrm{d}^2E}{\mathrm{d}\tilde{P}^2} \simeq \frac{\mathrm{d}^2E}{\mathrm{d}P^2} = \frac{\mathrm{d}(a/\tau)}{\mathrm{d}P} = -\frac{a}{\tau^2}\frac{\mathrm{d}\tau}{\mathrm{d}E}\frac{\mathrm{d}E}{\mathrm{d}P} = -\frac{a^3}{\tau^3}\frac{\mathrm{d}^2P}{\mathrm{d}E^2} . \tag{3.64}$$

On the other hand, using (3.53) again, it is

$$\frac{\mathrm{d}^2 P}{\mathrm{d}E^2} = \frac{\mathrm{d}(\tau/a)}{\mathrm{d}E} = -\frac{m^2}{a}\int_g^{g+a} K^3\,\mathrm{d}x\,, \qquad K = [2\,m\,(E - V(x))]^{-1/2}\,. \tag{3.65}$$

Combining (3.64) with (3.65) and defining the dimensionless parameter

$$r_s = \int_{g/a}^{g/a+1} K^3\,\mathrm{d}(x/a) \times \left[\int_{g/a}^{g/a+1} K\,\mathrm{d}(x/a)\right]^{-3}\,, \tag{3.66}$$

transforms (3.63) into

$$E \simeq E_s + \frac{a}{\tau_s}(\tilde{P} - \tilde{P}_s) + \frac{r_s}{2\,m}(\tilde{P} - \tilde{P}_s)^2\,, \tag{3.67}$$

where the coefficients, thanks to the neglect of the perturbation, do not depend on g. The linear term is readily eliminated by shifting the origin of the average momentum; in fact, letting $\tilde{P} - \tilde{P}_s = \tilde{p} - (m/r_s)\,(a/\tau_s)$ yields

$$E - E(0) = \frac{r_s}{2\,m}\tilde{p}^2\,, \qquad E(0) = E_s - \frac{1}{2}\frac{m}{r_s}\left(\frac{a}{\tau_s}\right)^2\,. \tag{3.68}$$

In conclusion, in a small interval of $\tilde{P}$ or $\tilde{p}$ the relation between energy and average momentum of a particle of mass m subjected to a periodic potential has the same form as that of a free particle of mass m/r_s. In other terms, the ratio m/r_s acts as an *effective mass* within the frame of the coarse-grained dynamics.

A bound for r_s is obtained from Hölder's inequality (C.119). Letting $|F| = K$, $G = 1$, $b = 3$, $x_1 = g/a$, $x_2 = g/a + 1$ in (C.119) and using the definition (3.66) yield $r_s \geq 1$, whence $m/r_s \leq m$: the effective mass can never exceed the true mass. The equality between the two masses is found in the limiting case $E - V_M \gg V_M - V - \delta H$, with V_M the maximum of V. In fact, (3.66) yields $r \simeq 1$ and, from (3.56), it is $\tilde{P} \simeq \sqrt{2mE}$. As expected, this limiting case yields the dynamics of a free particle.

3.13 Complements

3.13.1 Comments on the Linear Harmonic Oscillator

The paramount importance of the example of the linear harmonic oscillator, shown in Sect. 3.1, is due to the fact that in several physical systems the position of a particle at any instant happens to depart little from a point where the potential energy V has a minimum. As a consequence, the potential energy can be approximated with a second-order expansion around the minimum, that yields a positive-definite

quadratic form for the potential energy and a linear form for the force. The theory depicted in this section is then applicable to many physical systems, as shown by the examples of Sects. 3.9 and 5.6. The approximation of the potential energy with a second-order expansion, like that discussed in Sects. 3.9, 3.10, is called *harmonic approximation*. The terms beyond the second order in the expansion are called *anharmonic*.

3.13.2 Degrees of Freedom and Coordinate Separation

With reference to the analysis of the central motion carried out in Sect. 3.4, it is worth noting that the constancy of $\mathbf{M}$ reduces the number of degrees of freedom of the problem from three to two. Also, the form (3.6) of the Hamiltonian function is such as to provide a relation containing only r and the corresponding momentum p_r. Thus the coordinate r is separable according to the definition of Sect. 2.4. This allows one to independently find the time evolution (3.7) of r by solving an ordinary differential equation of the first order. Then one finds (3.8), that is, the trajectory $\varphi(r)$, through another equation of the same type. Finally, combining (3.7) with (3.8) yields the time evolution of the remaining coordinate φ.

It has been noted in Sect. 3.13.2 that, thanks to the constancy of the angular momentum, the adoption of spherical coordinates allows one to separate the radial coordinate r. This simplifies the problem, whose solution is in fact reduced to the successive solution of the evolution equations for r and φ. The same problem, instead, is not separable in the Cartesian coordinates. In other terms, separability may hold in some coordinate reference, but does not hold in general in an arbitrarily chosen reference.

Another example of separability is that illustrated in Sects. 3.9, 3.10. In general the Hamiltonian function is not separable in the Cartesian coordinates, whereas it is completely separable in the normal coordinates, no matter how large the number of the degrees of freedom is. Moreover, after the separation has been accomplished, one finds that all the equations related to the single degrees of freedom (the second relation in (3.48)) have the same form. In fact, they differ only in the numerical value of the angular frequency ω_σ. As a consequence, the expression of the solution is the same for all. Also, as the energy H_σ of each degree of freedom is independently conserved, no exchange of energy among the normal coordinates occurs: therefore, the distribution of energy among the normal coordinates that is present at $t = 0$ is maintained forever. This result is baffling because, for instance, it seems to prevent the condition of thermal equilibrium from being established; actually it is due to the fact that the system under investigation is isolated: if it were put in contact with a thermal reservoir, the exchanges of energy occurring with the reservoir would eventually bring the energy distribution of the system to the condition of thermal equilibrium.

Still with reference to the system discussed in Sects. 3.9, 3.10, it is important to underline the formal analogy between the modes of a mechanical, vibrating

system and those of the electromagnetic field *in vacuo* described in Sect. 5.6. In both cases the energy of each mode is that of a linear harmonic oscillator of unit mass (Eq. (3.48) and, respectively, (5.40)).

3.13.3 Comments on the Normal Coordinates

It has been shown in Sect. 3.9 that the elastic matrix $\mathbf{C}$ is positive definite. One may argue that in some cases the matrix is positive semi-definite. Consider, for instance, the case where the potential energy depends on the relative distance of the particles, $V_a = V_a(\mathbf{R}_1 - \mathbf{R}_2, \mathbf{R}_1 - \mathbf{R}_3, \ldots)$. For any set of positions $\mathbf{R}_1, \mathbf{R}_2, \ldots$, a uniform displacement $\mathbf{R}_\delta$ of all particles, that transforms each $\mathbf{R}_j$ into $\mathbf{R}_j + \mathbf{R}_\delta$, leaves V_a unchanged. As a consequence, if the positions prior to the displacement correspond to the equilibrium point $\mathbf{R}_{01}, \mathbf{R}_{02}, \ldots$, it is $V_a(\mathbf{R}_{01} + \mathbf{R}_\delta, \ldots) = V_a(\mathbf{R}_{01}, \ldots) = V_{a0}$. In such a case all terms beyond the zero-order term in the Taylor expansion of V_a around the equilibrium point vanish, which implies that the elastic matrix $\mathbf{C}$ is positive semi-definite. In the case examined in Sect. 3.9 the eigenvalues are real and positive; here, instead, they are real and nonnegative. Remembering (3.48), one finds that the Hamiltonian function of the degree of freedom corresponding to the null eigenvalue reads $H = \dot{b}_\sigma^2/2$, whence $\ddot{b}_\sigma = 0$, $b_\sigma = b_\sigma(0) + a\,t$, with a a constant.

The problem tackled in Sect. 3.10 is that of diagonalizing the right-hand side of (3.42). The diagonalization of a quadratic form entails a linear transformation over the original vector ($\mathbf{h}$ in this case) using a matrix formed by eigenvectors. One may observe that in (3.42), the kinetic energy $\dot{\mathbf{h}}^T \mathbf{M}\,\dot{\mathbf{h}}/2$ is already diagonal in the original vector, while the potential energy $\mathbf{h}^T \mathbf{C}\,\mathbf{h}/2$ is not. If the diagonalization was carried out using the matrix formed by the eigenvalues of $\mathbf{C}$ alone, the outcome of the process would be that of making the potential energy diagonal while making the kinetic energy non-diagonal (both in the transformed vector). The problem is solved by using the eigenvalue equation (3.43) that involves both matrices $\mathbf{M}$ and $\mathbf{C}$ in the diagonalization process. In fact, as shown in Sect. 3.10, in the transformed vector $\mathbf{b}$ the potential energy becomes diagonal, and the kinetic energy remains diagonal.

One may observe that, given the solutions of the eigenvalue equation (3.43), the process of diagonalizing (3.42) is straightforward. The real difficulty lies in solving (3.43). When the number of degrees of freedom is large, the solution of (3.43) must be tackled by numerical methods and may become quite cumbersome. In practical applications the elastic matrix $\mathbf{C}$ exhibits some structural properties, like symmetry or periodicity (e.g., Sect. 17.7.1), that are exploited to ease the problem of solving (3.43).

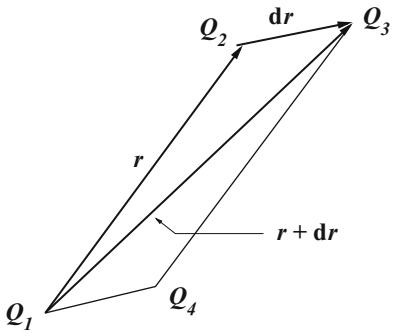


Fig. 3.6 Graphic representation of (3.69)

3.13.4 Areal Velocity in the Central-Motion Problem

Consider the central-motion problem discussed in Sect. 3.4. In the elementary time-interval dt the position vector changes from $\mathbf{r}$ to $\mathbf{r} + \mathrm{d}\mathbf{r}$. The area d$A$ of the triangle whose sides are $\mathbf{r}$, $\mathbf{r} + \mathrm{d}\mathbf{r}$, and d$\mathbf{r}$ is half the area of the parallelogram $Q_1Q_2Q_3Q_4$ whose consecutive sides are, e.g., $\mathbf{r}$ and d$\mathbf{r}$ (Fig. 3.6). Thus,

$$\mathrm{d}A = \frac{1}{2}\,|\mathbf{r} \wedge \mathrm{d}\mathbf{r}| = \frac{1}{2}\,|\mathbf{r} \wedge \dot{\mathbf{r}}\,\mathrm{d}t| = \frac{|\mathbf{M}|}{2\,m}\,\mathrm{d}t\,, \qquad \frac{\mathrm{d}A}{\mathrm{d}t} = \frac{|\mathbf{M}|}{2\,m}\,, \tag{3.69}$$

with $\mathbf{M}$ the angular momentum. The derivative d$A/$dt is called *areal velocity*. The derivation of (3.69) is based purely on definitions, hence it holds in general. For a central motion the angular momentum $\mathbf{M}$ is constant, whence the areal velocity is constant as well: the area swept out by the position vector $\mathbf{r}$ in a given time interval is proportional to the interval itself (*Kepler's second law*). If the particle's trajectory is closed, the time T taken by $\mathbf{r}$ to complete a revolution and the area A enclosed by the orbit are related by

$$A = \int_0^T \frac{\mathrm{d}A}{\mathrm{d}t}\,\mathrm{d}t = \frac{|\mathbf{M}|}{2\,m}\,T\,, \qquad T = \frac{2\,m\,A}{|\mathbf{M}|}\,. \tag{3.70}$$

3.13.5 Initial Conditions in the Central-Motion Problem

The theory of the central motion for a two-particle system has been worked out in Sects. 3.7, 3.8 without specifying the initial conditions. To complete the analysis it is convenient to use the same prescription as in Sect. 3.6, namely, to select an O reference where the particle of mass m_2 is initially at rest ($\mathbf{v}_{2a} = 0$). Moreover, here reference O is chosen in such a way as to make the initial position of the particle of mass m_2 to coincide with the origin ($\mathbf{r}_{2a} = 0$), and the initial velocity

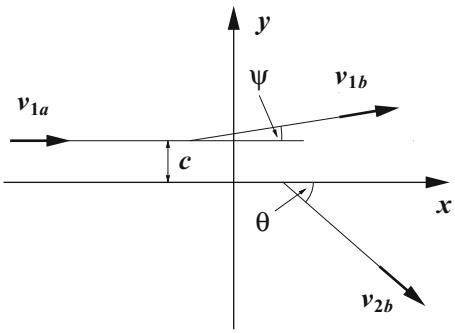


Fig. 3.7 Definition of the angles used in Sects. 3.6 and 3.13.5

$\mathbf{v}_a = \mathbf{v}_{1a}$ of the particle of mass m_1 to be parallel to the x axis (Fig. 3.7), so that $\mathbf{v}_{1a} = (\mathbf{v}_{1a} \cdot \mathbf{i})\,\mathbf{i}$. From (3.11) one finds $E = E_a = T_a = T_{1a} = m_1 v_{1a}^2/2$ and, from Sect. 3.5, $(m_1 + m_2)\,\dot{\mathbf{R}}_a = m_1\,\mathbf{v}_{1a}$. Using $\mathbf{r}_{1a} = x_{1a}\,\mathbf{i} + y_{1a}\,\mathbf{j}$ and (3.28) then yields

$$E_B = \frac{1}{2}\,m\,v_{1a}^2\,, \qquad \mathbf{M} = \mathbf{r}_{1a} \wedge m_1\,\mathbf{v}_{1a} = -m_1\,y_{1a}\,(\mathbf{v}_{1a} \cdot \mathbf{i})\,\mathbf{k}\,, \tag{3.71}$$

with $\mathbf{i}, \mathbf{j}, \mathbf{k}$ the unit vectors of the x, y, z axes and m the reduced mass. On the other hand, (3.29) shows that $\mathbf{M} = (m_1 + m_2)\,\mathbf{R}_a \wedge \dot{\mathbf{R}} + \mathbf{M}_B$ whence, writing $\mathbf{R}_a$, $\dot{\mathbf{R}}$ in terms of $\mathbf{r}_{1a}$, $\mathbf{v}_{1a}$ and equating the two expressions of $\mathbf{M}$ provides

$$M_B = -y_{1a}\,m\,\mathbf{v}_{1a} \cdot \mathbf{i}\,. \tag{3.72}$$

Replacing (3.72) and the first of (3.71) and in (3.32, 3.33) yields

$$\mu = \mp y_{1a}\,, \qquad s_0 = \lambda + \sqrt{\lambda^2 + c^2}\,, \qquad c = |y_{1a}|\,. \tag{3.73}$$

The distance c between the x axis and the direction of $\mathbf{v}_{1a}$ (Fig. 3.7) is called *impact parameter*. The outcome of the calculation demonstrates the usefulness of choosing reference O as described above. In fact, for a given form of the potential energy V, the angle χ defined in Sect. 3.6 becomes a function of two easily specified quantities: kinetic energy ($E = m_1\,v_{1a}^2/2$ or $E_B = m\,v_{1a}^2/2$) and impact parameter c (compare, e.g., with (3.36)). Once χ is determined, the final kinetic energies T_{1b}, T_{2b} and the angles ψ, θ are recovered from (3.22, 3.23) and (3.25, 3.26), respectively. Another property of reference O is that θ turns out to be the angle between the x axis and the final direction of the particle of mass m_2.

3.13.6 The Coulomb Field in the Attractive Case

To treat the attractive case one lets $\kappa = -1$ in (3.31). The trajectory lies in the x, y plane; in polar coordinates it is still given by (3.30), with

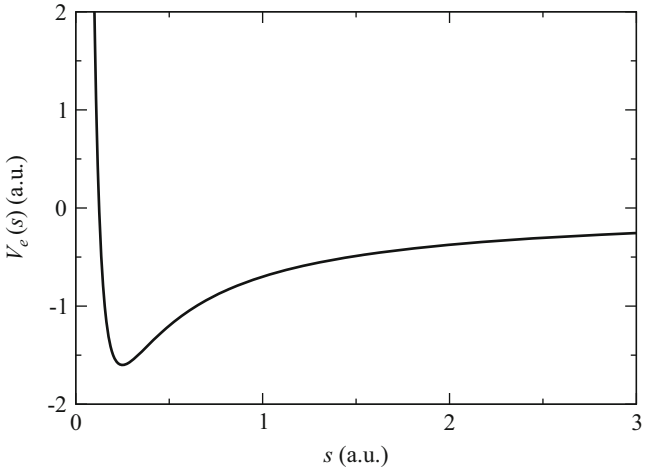


Fig. 3.8 Dependence of V_e on the distance s from the center of force, as given by (3.74) in arbitrary units

$$V_e(s) = \frac{M_B^2}{2ms^2} - \frac{Z_1\,Z_2\,q^2}{4\pi\,\varepsilon_0\,s}\,, \qquad s > 0\,. \tag{3.74}$$

In this case V_e becomes negative for some values of s. As a consequence, E_B may also be negative, provided the condition $E_B \geq V_e$ is fulfilled. Then, it is not possible to use the definitions (3.32) because E_B is positive there. The following will be used instead,

$$\alpha = \frac{Z_1\,Z_2\,q^2}{8\pi\,\varepsilon_0} > 0\,, \qquad \beta = \frac{M_B}{\sqrt{2m}}\,, \tag{3.75}$$

so that $V_e(s) = (\beta/s)^2 - 2\,\alpha/s$. Like in Sect. 3.8 it is assumed that M_B differs from zero and has either sign. It is found by inspection that V_e has only one zero at $s = s_c = \beta^2/(2\,\alpha)$ and only one minimum at $s = 2\,s_c$, with $\min(V_e) = V_e(2\,s_c) = -\alpha^2/\beta^2$. Also, it is $\lim_{s\to 0} V_e = \infty$, $\lim_{s\to\infty} V_e = 0$ (Fig. 3.8). The motion is unlimited when $E_B \geq 0$, while it is limited when $\min(V_e) \leq E_B < 0$. The case $E_B = \min(V_e)$ yields $s = 2\,s_c = \text{const}$, namely, the trajectory is a circumference. When $\min(V_e) = -\alpha^2/\beta^2 < E_B < 0$ it is $\alpha^2 > \beta^2\,|E_B|$. Then, the difference $E_B - V_e = -(|E_B|\,s^2 - 2\,\alpha\,s + \beta^2)/s^2$ has two real, positive zeros given by

$$s_0 = \frac{\alpha - \sqrt{\alpha^2 - \beta^2\,|E_B|}}{|E_B|}\,, \qquad s_1 = \frac{\alpha + \sqrt{\alpha^2 - \beta^2\,|E_B|}}{|E_B|}\,, \qquad s_0 < s_1\,. \tag{3.76}$$

Using the zeros one finds $s^2\sqrt{E_B - V_e} = \sqrt{|E_B|}\, s\sqrt{(s-s_0)(s_1-s)}$, that is replaced within (3.30) after letting $s \leftarrow \xi$. The upper limit of the integral belongs to the interval $s_0 \le s \le s_1$. To calculate the integral it is convenient to use a new variable w such that $2\,s_0\,s_1/\xi = (s_1 - s_0)\,w + s_1 + s_0$. The range of w corresponding to the condition $s_0 \le \xi \le s$ is

$$\frac{2\,s_0\,s_1 - (s_0 + s_1)\,s}{(s_1 - s_0)\,s} \le w \le 1\,, \qquad w(s_1) = -1\,. \tag{3.77}$$

From (3.30) the trajectory in the s, φ reference is thus found to be

$$\varphi(s) = \varphi_0 \pm \frac{M_B}{|M_B|}\,\arccos\left[\frac{2\,s_0\,s_1 - (s_0 + s_1)\,s}{(s_1 - s_0)\,s}\right]. \tag{3.78}$$

As noted in Sect. 3.4, the trajectory is symmetric with respect to φ_0. When (3.78) is inverted, the $\pm M_B/|M_B|$ factor is irrelevant because the cosine is an even function of the argument. Thus,

$$\frac{1}{s} = \frac{s_1 + s_0}{2\,s_0\,s_1}\left[1 + \frac{s_1 - s_0}{s_1 + s_0}\,\cos(\varphi - \varphi_0)\right]. \tag{3.79}$$

When $\varphi = \varphi_0$ it is $s = s_0$; when $\varphi = \varphi_0 + \pi$ it is $s = s_1$. The $s(\varphi)$ relation (3.79) is the equation of an ellipse of eccentricity $e = (s_1 - s_0)/(s_1 + s_0)$, where the center of force $s = 0$ is one of the foci. The distance between the foci is $s_1 - s_0$. With the aid of Fig. 3.9 one finds that the semimajor and semiminor axes are obtained, respectively, from $a = (s_1 + s_0)/2$, $b^2 = a^2 - (s_1 - s_0)^2/4$ whence, using (3.76),

$$a = \frac{\alpha}{|E_B|}\,, \qquad b = \frac{|\beta|}{|E_B|} = \frac{|M_B|}{\sqrt{2\,m\,|E_B|}}\,. \tag{3.80}$$

As the particle's trajectory is two-dimensional, the problem has four constants of motion (Sect. 2.11.2); the total energy E_B and the angular momentum M_B are two of such constants. As shown by (3.80), the semimajor axis of the elliptical trajectory depends only on E_B; the area of the ellipse in terms of the constants of motion is $A = \pi\,a\,b = (\pi\,\alpha/\sqrt{2\,m})\,|M_B|\,|E_B|^{-3/2}$. The position vector $\mathbf{s}$ completes a full orbit in a period T given by (3.70); combining the latter with the expression of A yields

$$T = \frac{\pi\,\alpha\,\sqrt{2\,m}}{|E_B|^{3/2}} = \pi\,\sqrt{\frac{2\,m}{\alpha}}\,a^{3/2}\,, \tag{3.81}$$

namely, the period depends on the total energy, but not on the angular momentum. Thus, the period is still given by (3.81) in the limiting case $M_B \to 0$, which makes the trajectory to shrink into a segment of length a crossing the origin of the s, φ reference, and the position vector to oscillate along this segment (compare with problem 3.2). The second form of (3.81) shows that $T^2 \propto a^3$ (*Kepler's third law*).

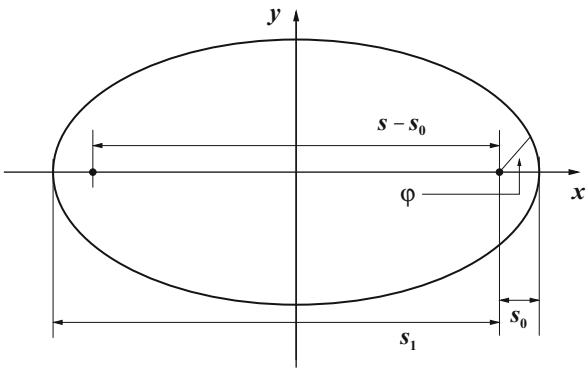


Fig. 3.9 The elliptical trajectory described by (3.79) with $\varphi_0 = 0$

3.13.7 Dynamic Relations of Special Relativity

The dynamic relations considered in this book refer almost invariably to situations where the particle velocity is small with respect to that of light. For this reason it is sufficient to use the nonrelativistic relations. The only exception where the velocities of the particles involved do not allow for such an approximation is considered in Sects. 3.13.8 and 7.4.3. For this reason a set of relations of the Special-Relativity Theory are given here, that apply to the case of a free particle. The first of them is the relation between velocity **u** and momentum **p**,

$$\mathbf{p} = \frac{m_0\,\mathbf{u}}{\sqrt{1 - u^2/c^2}}\,, \qquad u = |\mathbf{u}|\,, \tag{3.82}$$

with c the velocity of light and m_0 a constant mass. The second relation involves energy and velocity and reads

$$E = m\,c^2\,, \qquad m = \frac{m_0}{\sqrt{1 - u^2/c^2}}\,. \tag{3.83}$$

where E is a kinetic energy because a free particle is considered. In the above, $m = m(u)$ is called *relativistic mass* and $m_0 = m(0)$ is called *rest mass*. The latter is the mass measured in a reference where the particle is at rest, and is the value of the mass that is used in nonrelativistic mechanics. From (3.82, 3.83) it follows

$$\mathbf{p} = m\,\mathbf{u}\,, \qquad m^2\,c^2 - m^2\,u^2 = m_0^2\,c^2\,, \qquad m^2\,c^2 = E^2/c^2\,, \tag{3.84}$$

whence the elimination of u provides the relation between E and the modulus of **p**:

$$E^2/c^2 - p^2 = m_0^2\,c_2\,, \qquad p = \sqrt{E^2/c^2 - m_0^2\,c^2}\,. \tag{3.85}$$

In general the $p = p(E)$ relation is nonlinear. However, for particles with $m_0 = 0$ the expressions (3.85) simplify to the linear relation $p = E/c$. An example of this is found in the theory of the electromagnetic field, where the same momentum-energy relation is derived from the Maxwell equations (Sect. 5.7). It is worth observing that if a particle has $m_0 = 0$ and $p \neq 0$, its velocity is necessarily equal to c; in fact, the first of (3.84) yields $\lim_{m_0 \to 0} \mathbf{p} = 0$ when $u < c$. Another limiting case is found when $u/c \ll 1$. In fact, the second of (3.83) simplifies to

$$m \simeq \frac{m_0}{1 - u^2/(2\,c^2)} \simeq m_0 \left(1 + \frac{u^2}{2\,c^2}\right) . \tag{3.86}$$

Inserting the last form into the first of (3.83) yields

$$E \simeq m_0\,c^2 + \frac{1}{2}\,m_0\,u^2 . \tag{3.87}$$

The constant $m_0\,c^2$ is called *rest energy*. The limiting case $u/c \ll 1$ then renders for $E - m_0\,c^2$ the nonrelativistic expression of the kinetic energy.

3.13.8 Collision of Relativistic Particles

This section illustrates the collision between two relativistic particles that constitute an isolated system. The same approach of Sect. 3.5 is used here, namely, the asymptotic values only are considered. Also, the case where the particles' trajectories belong to the same plane, specifically, the x, y plane, is investigated. The initial conditions are the same as in Sect. 3.6: the asymptotic motion of the first particle before the collision is parallel to the x axis, while the second particle is initially at rest. Finally, it is assumed that the rest mass of the first particle is zero, so that the momentum-energy relation of this particle is $p = E/c$ as shown in Sect. 3.13.7, while the rest mass of the second particle is $m_0 \neq 0$.

The collision is treated by combining the conservation laws of energy and momentum of the two-particle system. Let E_a, E_b be the asymptotic energies of the first particle before and after the collision, respectively. As for the second particle, which is initially at rest, the energy before the collision is its rest energy $m_0\,c^2$, while that after the collision is $m\,c^2$ (Sect. 3.13.7). The conservation of energy then reads

$$E_a + m_0\,c^2 = E_b + m\,c^2 , \tag{3.88}$$

while the conservation of momentum reads, respectively for the x and y components,

$$\frac{E_a}{c} = \frac{E_b}{c}\,\cos\psi + m\,u\,\cos\theta , \qquad 0 = \frac{E_b}{c}\,\sin\psi - m\,u\,\sin\theta . \tag{3.89}$$

The angles ψ and θ in (3.89) are the same as in Fig. 3.7. Extracting $m\,c^2$ from (3.88) and squaring both sides yield

$$(E_a - E_b)^2 + 2\,m_0\,c^2\,(E_a - E_b) = m^2\,c^4 - m_0^2\,c^4 = m^2\,u^2\,c^2\,, \tag{3.90}$$

where the last equality derives from the second of (3.84). Then, the momentum-conservation relations (3.89) are used to eliminate θ by squaring and adding up the results, to find

$$m^2\,u^2\,c^2 = E_a^2 + E_b^2 - 2\,E_a\,E_b\,\cos\psi = (E_a - E_b)^2 + 4\,E_a\,E_b\,\sin^2(\psi/2)\,. \tag{3.91}$$

Eliminating $m^2\,u^2\,c^2 - (E_a - E_b)^2$ between (3.90) and (3.91) yields

$$\frac{1}{E_b} - \frac{1}{E_a} = \frac{2}{m_0\,c^2}\,\sin^2\left(\frac{\psi}{2}\right)\,, \tag{3.92}$$

that provides the asymptotic energy after the collision of the first particle, as a function of the asymptotic energy before the collision, the deflection angle of the same particle, and the rest energy of the second particle. Equation (3.92) is used in Sect. 7.4.3 for the explanation of the Compton effect.

The nonrelativistic analogue of the above procedure is illustrated in Sect. 3.6. It is interesting to note that the calculation carried out here seems rather less involved than the nonrelativistic one. This surprising fact is actually due to the special choice of the first particle, whose rest energy is zero. In this case, in fact, the relation between momentum and energy becomes linear. That of the second particle, which is nonlinear, is eliminated from the equations. On the contrary, in the nonrelativistic case treated in Sect. 3.6 the energy-momentum relations are nonlinear for both particles, thus making the calculation more laborious.

3.13.9 Energy Conservation in Charged-Particles' Interaction

The two-particle interaction considered in Sect. 3.8 involves charged particles. As the particles' velocity during the interaction is not constant, the particles radiate (Sect. 5.11.2) and, consequently, lose energy. This phenomenon is not considered in the analysis carried out in Sect. 3.8, where the total energy of the two-particle system is assumed constant. The assumption is justified on the basis that the radiated power is relatively small. This subject is further discussed in Sect. 5.11.3.

Problems

3.1 Given the Hamiltonian function of the one-dimensional harmonic oscillator of the general form $H = p^2/(2\,m) + (c/s)\,|x|^s$, $m, c, s > 0$, find the oscillator's period.

3.2 Given the Hamiltonian function of the one-dimensional harmonic oscillator of the general form $H = p^2/(2\,m) - (k/s)\,|x|^{-s} = E < 0$, $m, k, s > 0$, find the oscillator's period.

3.3 Consider the collision between two particles in the repulsive Coulomb case. Calculate the relation $T_{1b}(T_{1a}, c)$, with c the impact parameter (hint: follow the discussion of Sect. 3.13.5 and use (3.23, 3.36), (3.32, 3.33), and (3.73)).

Chapter 4
Electromagnetism

4.1 Introduction

This chapter outlines the basic principles of the electromagnetic theory *in vacuo*. First, the extension of the Lagrangian formalism to functions that depend on more than one variable is tackled: this yields useful tools for the analysis of continuous media. Next, the Maxwell equations are introduced along with the derivation of the electric and magnetic potentials, and the concept of gauge transformation is illustrated. The second part of the chapter is devoted to the Helmholtz and wave equations, both in a finite and infinite domain. The chapter finally introduces the Lorentz force that connects the electromagnetic field with the particles' dynamics. The complements discuss some invariance properties of the Euler equations, derive the wave equations for the electric and magnetic field, and clarify some issues related to the boundary conditions in the application of the Green method to the boundary-value problem.

4.2 Extension of the Lagrangian Formalism

In Sect. 1.2 the derivation of the extremum functions has been carried out with reference to a functional $G[w]$ of the form (1.1). Such a functional contains one unknown function w that, in turn, depends on one independent variable ξ. The result has been extended to the case where the functional depends on several unknown functions $w_1, w_2, \ldots$, each dependent on one variable only (compare with (1.6)). The extension to more than one independent variable is shown here.

To proceed it suffices to consider a single unknown function w that depends on two independent variables ξ, σ and is differentiable at least twice with respect to each. The first and second derivatives of w are indicated with w_ξ, w_σ, $w_{\xi\xi}$, $w_{\sigma\sigma}$, and $w_{\xi\sigma}$. Letting Ω be the domain over which w is defined, and g the generating function, the functional reads

M. Rudan, *Physics of Semiconductor Devices*, DOI 10.1007/978-3-319-63154-7_4

$$G[w] = \int_\Omega g(w, w_\xi, w_\sigma, \xi, \sigma)\, \mathrm{d}\Omega\,. \tag{4.1}$$

Then, let $\delta w = \alpha\,\eta$, with $\eta(\xi, \sigma)$ an arbitrary function defined in Ω and differentiable in its interior, and α a real parameter. Like in the case of one independent variable the choice is restricted to those functions η that vanish at the boundary of Ω, so that w and $w + \delta w$ coincide along the boundary for any value of α. If w is an extremum function of G, the extremum condition if found by replacing w with $w + \alpha\,\eta$ and letting $(\mathrm{d}G/\mathrm{d}\alpha)_0 = 0$, where suffix 0 indicates that the derivative is calculated at $\alpha = 0$ (compare with Sect. 1.2). Exchanging the integral with the derivative in (4.1) yields

$$\left(\frac{\mathrm{d}G}{\mathrm{d}\alpha}\right)_0 = \int_\Omega \left(\frac{\partial g}{\partial w}\,\eta + \frac{\partial g}{\partial w_\xi}\,\eta_\xi + \frac{\partial g}{\partial w_\sigma}\,\eta_\sigma\right) \mathrm{d}\Omega = 0\,. \tag{4.2}$$

The second and third terms of the integrand in (4.2) are recast in compact form by defining vector $\mathbf{u} = (\partial g/\partial w_\xi\,, \partial g/\partial w_\sigma)$ and using the second identity in (A.16), so that the sum of the two terms reads $\mathbf{u}\cdot\operatorname{grad}\eta = \operatorname{div}(\eta\mathbf{u}) - \eta\,\operatorname{div}\mathbf{u}$. Integrating over Ω and using the divergence theorem (A.23) yields

$$\int_\Omega \mathbf{u}\cdot\operatorname{grad}\eta\;\mathrm{d}\Omega = \int_\Sigma \eta\,\mathbf{u}\cdot\mathbf{n}\;\mathrm{d}\Sigma - \int_\Omega \eta\;\operatorname{div}\mathbf{u}\;\mathrm{d}\Omega\,, \tag{4.3}$$

where Σ is the boundary of Ω and $\mathbf{n}$ the unit vector normal to $\mathrm{d}\Sigma$, oriented in the outward direction with respect to Σ. The first term at the right-hand side of (4.3) is equal to zero because η vanishes over Σ.

It is important to clarify the symbols that will be used to denote the derivatives. In fact, to calculate $\operatorname{div}\mathbf{u}$ one needs, first, to differentiate $\partial g/\partial w_\xi$ with respect to ξ considering also the implicit ξ-dependence within w, w_ξ, and w_σ; then, one differentiates in a similar manner $\partial g/\partial w_\sigma$ with respect to σ. The two derivatives are summed up to form $\operatorname{div}\mathbf{u}$. For this type of differentiation the symbols $\mathrm{d}/\mathrm{d}\xi$ and $\mathrm{d}/\mathrm{d}\sigma$ are used, even if the functions in hand depend on two independent variables instead of one. The symbols $\partial/\partial\xi$ and $\partial/\partial\sigma$ are instead reserved to the derivatives with respect to the explicit dependence on ξ or σ only. With this provision, inserting (4.3) into (4.2) yields the extremum condition

$$\int_\Omega \left(\frac{\partial g}{\partial w} - \frac{\mathrm{d}}{\mathrm{d}\xi}\frac{\partial g}{\partial w_\xi} - \frac{\mathrm{d}}{\mathrm{d}\sigma}\frac{\partial g}{\partial w_\sigma}\right) \eta\;\mathrm{d}\Omega = 0\,. \tag{4.4}$$

As (4.4) holds for any η, the term in parentheses must vanish. In conclusion, the extremum condition is

$$\frac{\mathrm{d}}{\mathrm{d}\xi}\frac{\partial g}{\partial w_\xi} + \frac{\mathrm{d}}{\mathrm{d}\sigma}\frac{\partial g}{\partial w_\sigma} = \frac{\partial g}{\partial w}\,, \tag{4.5}$$

namely, a second-order partial-differential equation in the unknown function w, that must be supplemented with suitable boundary conditions. The equation is linear with respect to the second derivatives of w because g does not depend on such derivatives.

The result is readily extended to the case where g depends on several functions $w_1, w_2, \ldots, w_l$ and the corresponding derivatives. Defining the vectors $\mathbf{w}(\xi, \sigma) = (w_1, \ldots, w_l)$, $\mathbf{w}_\xi = (\partial w_1/\partial\xi, \ldots, \partial w_l/\partial\xi)$, $\mathbf{w}_\sigma = (\partial w_1/\partial\sigma, \ldots, \partial w_l/\partial\sigma)$, the set of the l extremum functions w_i of functional

$$G[\mathbf{w}] = \int_\Omega g(\mathbf{w}, \mathbf{w}_\xi, \mathbf{w}_\sigma, \xi, \sigma)\, \mathrm{d}\Omega \tag{4.6}$$

is found by solving the equations

$$\frac{\mathrm{d}}{\mathrm{d}\xi}\frac{\partial g}{\partial(\partial w_i/\partial\xi)} + \frac{\mathrm{d}}{\mathrm{d}\sigma}\frac{\partial g}{\partial(\partial w_i/\partial\sigma)} = \frac{\partial g}{\partial w_i}, \qquad i = 1, \ldots, l, \tag{4.7}$$

supplemented with the suitable boundary conditions. It follows that (4.7) are the Euler equations of G. Finally, the case where the independent variables are more than two is a direct extension of (4.7). For instance, for m variables $\xi_1, \ldots, \xi_m$ one finds

$$\sum_{j=1}^{m} \frac{\mathrm{d}}{\mathrm{d}\xi_j}\frac{\partial g}{\partial(\partial w_i/\partial\xi_j)} = \frac{\partial g}{\partial w_i}, \qquad i = 1, \ldots, l. \tag{4.8}$$

If g is replaced with $g' = g + \operatorname{div}\mathbf{h}$, where $\mathbf{h}$ is an arbitrary vector of length m whose entries depend on $\mathbf{w}$ and $\xi_1, \ldots, \xi_m$, but not on the derivatives of $\mathbf{w}$, then (4.8) is still fulfilled. The replacement, in fact, adds the same term to both sides. For instance, the term added to the left-hand side is

$$\sum_{j=1}^{m} \frac{\mathrm{d}}{\mathrm{d}\xi_j}\frac{\partial}{\partial(\partial w_i/\partial\xi_j)} \sum_{r=1}^{m}\left(\frac{\partial h_r}{\partial\xi_r} + \sum_{s=1}^{l}\frac{\partial h_r}{\partial w_s}\frac{\partial w_s}{\partial\xi_r}\right), \qquad i = 1, \ldots, l, \tag{4.9}$$

where the sum over r is the explicit expression of $\operatorname{div}\mathbf{h}$. Remembering that $\mathbf{h}$ does not depend on the derivatives of w_i one recasts (4.9) as

$$\sum_{j=1}^{m} \frac{\mathrm{d}}{\mathrm{d}\xi_j} \sum_{r=1}^{m}\sum_{s=1}^{l}\frac{\partial h_r}{\partial w_s}\frac{\partial(\partial w_s/\partial\xi_r)}{\partial(\partial w_i/\partial\xi_j)} = \sum_{j=1}^{m}\frac{\partial}{\partial\xi_j}\frac{\partial h_j}{\partial w_i}, \qquad i = 1, \ldots, l, \tag{4.10}$$

where the equality is due to the relation $\partial(\partial w_s/\partial\xi_r)/\partial(\partial w_i/\partial\xi_j) = \delta_{is}\delta_{jr}$, with $\delta_{is(jr)}$ the Kronecker symbol (A.18). Inverting the order of the derivatives at the right-hand side of (4.10) yields $\partial \operatorname{div}\mathbf{h}/\partial w_i$, that coincides with the term added to the right-hand

side of (4.8). Finally, (4.8) is recast in compact form by defining a vector $\mathbf{u}_i$ and a scalar s_i as

$$\mathbf{u}_i = \left[\frac{\partial g}{\partial(\partial w_i/\partial \xi_1)}, \ldots, \frac{\partial g}{\partial(\partial w_i/\partial \xi_m)} \right], \qquad s_i = \frac{\partial g}{\partial w_i} \tag{4.11}$$

to find

$$\mathrm{div}_\xi \mathbf{u}_i = s_i, \qquad i = 1, \ldots, l. \tag{4.12}$$

If w_i depends on one variable only, say ξ, (4.8, 4.12) reduce to (1.7). Using the language of the Lagrangian theory, the comparison between the one-dimensional and multi-dimensional case shows that in both cases the functions w_i play the role of generalized coordinates; in turn, the scalar parameter ξ of (1.7) becomes the vector $(\xi_1, \ldots, \xi_m)$ of (4.8) and, finally, each generalized velocity $\dot{w}_i$ becomes the set $\partial w_i/\partial \xi_1, \ldots, \partial w_i/\partial \xi_m$.

4.3 Lagrangian Function for the Wave Equation

It has been shown in Sect. 1.3 that the relations $\ddot{w}_i = \ddot{w}_i(\mathbf{w}, \dot{\mathbf{w}}, \xi)$, $i = 1, \ldots, n$, describing the motion of a system of particles with n degrees of freedom, are the Euler equations of a suitable functional. Then, the analysis of Sect. 4.2 has shown that when the unknown functions $w_1, \ldots, w_l$ depend on more than one variable, the Euler equations are the second-order partial-differential equations (4.8). The form (4.8) is typical of the problems involving continuous media (e.g., elasticity field, electromagnetic field). Following the same reasoning as in Sect. 1.3 it is possible to construct the Lagrangian function whence the partial-differential equation derives. This is done here with reference to the important case of the *wave equation*[1]

$$\nabla^2 w - \frac{1}{u^2} \frac{\partial^2 w}{\partial t^2} = s, \tag{4.13}$$

where $u = \text{const}$ is a velocity and, for the sake of simplicity, s is assumed to depend on x and t, but not on w or its derivatives. It is worth noting that when a differential equation other than Newton's law is considered, the corresponding Lagrangian function is not necessarily an energy. For this reason it will provisionally be indicated with L_e instead of L. To proceed one considers the one-dimensional form of (4.13), $\partial^2 w/\partial x^2 - (1/u^2)\, \partial^2 w/\partial t^2 = s$ and replaces ξ, σ, g with x, t, L_e, respectively. Then, one makes the one-dimensional form identical to (4.5) by letting

[1] Also called *D'Alembert equation* in the homogeneous case.

$$\frac{\partial^2 L_e}{\partial w_x^2} = 1\,, \qquad \frac{\partial^2 L_e}{\partial w_t^2} = -\frac{1}{u^2}\,, \qquad \frac{\partial L_e}{\partial w} = s\,, \tag{4.14}$$

with $w_x = \partial w/\partial x$ and $w_t = \partial w/\partial t$. The second derivatives of $L_e(w, w_x, w_t, x, t)$ with respect to the combinations of the arguments not appearing in the first two equations of (4.14) are set to zero. The third of (4.14) provides $L_e = s\,w + c$, with c independent of w. Replacing $L_e = s\,w + c$ into the first two equations in (4.14) and integrating the first one with respect to w_x yield $\partial c/\partial w_x = w_x + a_{01}$, with a_{01} independent of w_x. Similarly, from the second equation in (4.14), $\partial c/\partial w_t = -w_t/u^2 + a_{02}$, with a_{02} independent of w_t. Also, remembering that c is independent of w, one finds that a_{01} and a_{02} do not depend on w either. Considering that all the second derivatives of L_e not appearing in (4.14) are equal to zero shows that a_{01} depends on t at most, while a_{02} depends on x at most. Integrating $\partial c/\partial w_x = w_x + a_{01}$ and $\partial c/\partial w_t = -w_t/u^2 + a_{02}$ one finds

$$c = \frac{1}{2}\,w_x^2 + a_{01}(t)\,w_x + a_{11}\,, \qquad c = -\frac{1}{2u^2}\,w_t^2 + a_{02}(x)\,w_t + a_{12}\,, \tag{4.15}$$

where a_{11} does not depend on w or w_x, while a_{12} does not depend on w or w_t. Also, a_{11} cannot depend on both t and w_t due to $\partial^2 L_e/(\partial t\,\partial w_t) = 0$; similarly, a_{12} cannot depend on both x and w_x due to $\partial^2 L_e/(\partial x\,\partial w_x) = 0$. On the other hand, as both (4.15) hold, a_{11} must coincide (apart from an additive constant) with the first two terms at the right-hand side of the second equation in (4.15), and a_{12} must coincide with the first two terms at the right-hand side of the first equation. In conclusion,

$$c = \frac{1}{2}\,w_x^2 - \frac{1}{2u^2}\,w_t^2 + a_{01}(t)\,w_x + a_{02}(x)\,w_t\,, \tag{4.16}$$

with $a_{01}(t)$, $a_{02}(x)$ arbitrary functions. The last two terms in (4.16) are equal to $\mathrm{d}(a_{01}\,w)/\mathrm{d}x + \mathrm{d}(a_{02}\,w)/\mathrm{d}t$, namely, they form the divergence of a vector. As shown in Sect. 4.2 such a vector is arbitrary, so it can be eliminated by letting $a_{01} = 0$, $a_{02} = 0$. The relation $L_e = s\,w + c$ then yields

$$L_e = \frac{1}{2}\,w_x^2 - \frac{1}{2u^2}\,w_t^2 + sw\,. \tag{4.17}$$

The generalization to the three-dimensional case (4.13) is immediate,

$$L_e = \frac{1}{2}\,|\,\mathrm{grad}\,w|^2 - \frac{1}{2u^2}\left(\frac{\partial w}{\partial t}\right)^2 + s\,w\,. \tag{4.18}$$

with $|\,\mathrm{grad}\,w|^2 = w_x^2 + w_y^2 + w_z^2$.

4.4 Maxwell Equations

The *Maxwell equations*, that describe the electromagnetic field, lend themselves to an interesting application of the results of Sect. 4.3. The first group of Maxwell equations reads

$$\operatorname{div} \mathbf{D} = \varrho \,, \qquad \operatorname{rot} \mathbf{H} - \frac{\partial \mathbf{D}}{\partial t} = \mathbf{J} \,, \tag{4.19}$$

where $\mathbf{D}$ is the electric displacement and $\mathbf{H}$ the magnetic field.[2] The sources of the electromagnetic field are the charge density ϱ and the current density $\mathbf{J}$. When point-like charges are considered, they read

$$\varrho_c = \sum_j e_j \, \delta \left(\mathbf{r} - \mathbf{s}_j(t) \right) \,, \qquad \mathbf{J}_c = \sum_j e_j \, \delta \left(\mathbf{r} - \mathbf{s}_j(t) \right) \, \mathbf{u}_j(t) \,, \tag{4.20}$$

where index c is used to distinguish the case of point-like charges from that of a continuous charge distribution. In (4.20), e_j is the value of the jth charge, $\mathbf{s}_j$ and $\mathbf{u}_j$ its position and velocity at time t, respectively, and $\mathbf{r}$ the independent positional variable. If the spatial scale of the problem is such that one can replace the point-like charges with a continuous distribution, one applies the same procedure as in Sect. 23.2. The number of charges belonging to a cell of volume ΔV centered at $\mathbf{r}$ is $\int_{\Delta V} \varrho_c \, \mathrm{d}^3 s' = \sum_j' e_j$, where the prime indicates that the sum is limited to the charges that belong to Δ at time t. Then one defines $\varrho(\mathbf{r}, t) = \sum_j' e_j / \Delta V$. The continuous distribution of the current density is obtained in a similar manner,

$$\mathbf{J} = \frac{1}{\Delta V} \int_{\Delta V} \mathbf{J}_c \, \mathrm{d}^3 s' = \frac{1}{\Delta V} \sum_j{}' e_j \, \mathbf{u}_j = \varrho \, \mathbf{v} \,, \qquad \mathbf{v} = \frac{\sum_j' e_j \, \mathbf{u}_j}{\sum_j' e_j} \,, \tag{4.21}$$

with $\mathbf{v}(\mathbf{r}, t)$ the average velocity of the charges. If all charges are equal, $e_1 = e_2 = \ldots = e$, then $\varrho = e N$, with $N(\mathbf{r}, t)$ the concentration, and $\mathbf{J} = e N \, \mathbf{v} = e \, \mathbf{F}$, with $\mathbf{F}(\mathbf{r}, t)$ the flux density (compare with the definitions of Sect. 23.2). If the charges are different from each other it is convenient to distribute the sum $\sum_j$ over the groups made of equal charges. In this case the charge density and current density read

$$\varrho = \varrho_1 + \varrho_2 + \ldots \,, \qquad \mathbf{J} = \varrho_1 \, \mathbf{v}_1 + \varrho_2 \, \mathbf{v}_2 + \ldots \,, \tag{4.22}$$

[2] The units in (4.19, 4.23, 4.24) are: $[\mathbf{D}] = \mathrm{C\,m^{-2}}$, $[\varrho] = \mathrm{C\,m^{-3}}$, $[\mathbf{H}] = \mathrm{A\,m^{-1}}$, $[\mathbf{J}] = \mathrm{C\,s^{-1}\,m^{-2}} = \mathrm{A\,m^{-2}}$, $[\mathbf{B}] = \mathrm{V\,s\,m^{-2}} = \mathrm{Wb\,m^{-2}} = \mathrm{T}$, $[\mathbf{E}] = \mathrm{V\,m^{-1}}$, where "C," "A," "V," "Wb," and "T" stand for Coulomb, Ampère, Volt, Weber, and Tesla, respectively. The coefficients in (4.19, 4.23, 4.24) differ from those of [10] because of the different units adopted there. In turn, the units in (4.25) are $[\varepsilon_0] = \mathrm{C\,V^{-1}\,m^{-1}} = \mathrm{F\,m^{-1}}$, $[\mu_0] = \mathrm{s^2\,F^{-1}\,m^{-1}} = \mathrm{H\,m^{-1}}$, where "F" and "H" stand for Farad and Henry, respectively, and those in (4.26) are $[\varphi] = \mathrm{V}$, $[\mathbf{A}] = \mathrm{V\,s\,m^{-1}} = \mathrm{Wb\,m^{-1}}$.

where ϱ_1, $\mathbf{v}_1$ are the charge density and average velocity of the charges of the first group, and so on. Taking the divergence of the second equation in (4.19) and using the third identity in (A.35) yield the *continuity equation*

$$\frac{\partial \varrho}{\partial t} + \operatorname{div} \mathbf{J} = 0 \,. \tag{4.23}$$

Apart from the different units of the functions involved, the form of (4.23) is the same as that of (23.3). The meaning of (4.23) is that of conservation of the electric charge. The second group of Maxwell equations is

$$\operatorname{div} \mathbf{B} = 0 \,, \qquad \operatorname{rot} \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0 \,, \tag{4.24}$$

where $\mathbf{B}$ and $\mathbf{E}$ are the magnetic induction and the electric field, respectively. Here the Maxwell equations are considered *in vacuo*, so that the following hold

$$\mathbf{D} = \varepsilon_0 \, \mathbf{E} \,, \qquad \mathbf{B} = \mu_0 \, \mathbf{H} \,, \qquad \frac{1}{\sqrt{\varepsilon_0 \, \mu_0}} = c \,, \tag{4.25}$$

with $\varepsilon_0 \simeq 8.854 \times 10^{-12} \, \mathrm{F \, m^{-1}}$ and $\mu_0 \simeq 1.256 \times 10^{-6} \, \mathrm{H \, m^{-1}}$ the vacuum permittivity and permeability, respectively, and $c \simeq 2.998 \times 10^{8} \, \mathrm{m \, s^{-1}}$ the speed of light *in vacuo*.

4.5 Potentials and *Gauge* Transformations

Thanks to (4.25), the electromagnetic field *in vacuo* is determined by two suitably chosen vectors—typically, $\mathbf{E}$ and $\mathbf{B}$—out of the four ones appearing in (4.25). This amounts to using six scalar functions of position and time. However, the number of scalar functions is reduced by observing that while (4.19) provides relations between the electromagnetic field and its sources, (4.24) provides relations among the field vectors themselves; as a consequence, (4.24) reduces the number of independent vectors. In fact, using the properties illustrated in Sect. A.9, one finds that from $\operatorname{div} \mathbf{B} = 0$ one derives $\mathbf{B} = \operatorname{rot} \mathbf{A}$, where $\mathbf{A}$ is called *vector potential* or *magnetic potential*. In turn, the vector potential transforms the second of (4.24) into $\operatorname{rot}(\mathbf{E} + \partial \mathbf{A}/\partial t) = 0$; using again the results of Sect. A.9 shows that the term in parentheses is the gradient of a scalar function, that is customarily indicated with $-\varphi$. Such a function[3] is called *scalar potential* or *electric potential*. In summary,

[3]The minus sign in the definition of φ is used for consistency with the definition of the gravitational potential, where the force is opposite to the direction along which the potential grows.

$$\mathbf{B} = \operatorname{rot} \mathbf{A}\,, \qquad \mathbf{E} = -\operatorname{grad} \varphi - \frac{\partial \mathbf{A}}{\partial t}\,, \tag{4.26}$$

showing that for determining the electromagnetic field *in vacuo* it suffices to know four scalar functions, namely, φ and the three components of $\mathbf{A}$. To proceed, one replaces (4.26) into (4.19) and uses the third relation in (4.25), to find

$$\nabla^2 \varphi + \frac{\partial}{\partial t} \operatorname{div} \mathbf{A} = -\frac{\varrho}{\varepsilon_0}\,, \qquad -\operatorname{rot}\operatorname{rot} \mathbf{A} - \frac{1}{c^2}\frac{\partial^2 \mathbf{A}}{\partial t^2} = -\mu_0 \mathbf{J} + \frac{1}{c^2} \operatorname{grad} \frac{\partial \varphi}{\partial t}\,. \tag{4.27}$$

Thanks to the first identity in (A.36) the second equation in (4.27) becomes

$$\nabla^2 \mathbf{A} - \frac{1}{c^2}\frac{\partial^2 \mathbf{A}}{\partial t^2} = -\mu_0 \mathbf{J} + \operatorname{grad} \theta\,, \qquad \theta = \operatorname{div} \mathbf{A} + \frac{1}{c^2}\frac{\partial \varphi}{\partial t} \tag{4.28}$$

while, using the definition (4.28) of θ, one transforms the first equation in (4.27) into

$$\nabla^2 \varphi - \frac{1}{c^2}\frac{\partial^2 \varphi}{\partial t^2} = -\frac{\varrho}{\varepsilon_0} - \frac{\partial \theta}{\partial t}\,. \tag{4.29}$$

In conclusion, (4.29) and the first equation in (4.28) are a set of four scalar differential equations whose unknowns are φ and the components of $\mathbf{A}$. Such equations are coupled because θ contains all unknowns; however, they become decoupled after suitable transformations, shown below.

To proceed, one observes that only the derivatives of the potentials, not the potential themselves, appear in (4.26); as a consequence, while the fields $\mathbf{E}$, $\mathbf{B}$ are uniquely defined by the potentials, the opposite is not true. For instance, replacing $\mathbf{A}$ with $\mathbf{A}' = \mathbf{A} + \operatorname{grad} f$, where $f(\mathbf{r}, t)$ is any differentiable scalar function, and using the second identity in (A.35), yields $\mathbf{B}' = \operatorname{rot} \mathbf{A}' = \mathbf{B}$, namely, $\mathbf{B}$ is invariant with respect to such a replacement. If, at the same time, one replaces φ with a yet undetermined function φ', (4.26) yields $\mathbf{E}' = -\operatorname{grad}(\varphi' + \partial f/\partial t) - \partial \mathbf{A}/\partial t$. It follows that by choosing $\varphi' = \varphi - \partial f/\partial t$ one obtains $\mathbf{E}' = \mathbf{E}$. The transformation $(\varphi\,, \mathbf{A}) \to (\varphi'\,, \mathbf{A}')$ defined by

$$\varphi' = \varphi - \frac{\partial f}{\partial t}\,, \qquad \mathbf{A}' = \mathbf{A} + \operatorname{grad} f\,. \tag{4.30}$$

is called *gauge transformation*. As shown above, $\mathbf{E}$ and $\mathbf{B}$ are invariant with respect to such a transformation. One also finds that (4.29) and the first equation in (4.28) are invariant with respect to the transformation: all terms involving f cancel each other, so that the equations in the primed unknowns are identical to the original ones. However, the solutions φ', $\mathbf{A}'$ are different from φ, $\mathbf{A}$ because, due to (4.30), their initial and boundary conditions are not necessarily the same. The difference

between the primed and unprimed solutions is unimportant because the fields, as shown above, are invariant under the transformation. Using (4.30) in the second equation of (4.28) shows that θ transforms as

$$\theta' = \operatorname{div} \mathbf{A}' + \frac{1}{c^2}\frac{\partial \varphi'}{\partial t} = \theta + \nabla^2 f - \frac{1}{c^2}\frac{\partial^2 f}{\partial t^2}\,. \tag{4.31}$$

The arbitrariness of f may be exploited to give θ' a convenient form. For instance one may choose f such that $\theta' = (1/c^2)\,\partial\varphi/\partial t$, which is equivalent to letting

$$\operatorname{div} \mathbf{A}' = 0\,, \tag{4.32}$$

called *Coulomb gauge*. The latter yields

$$\nabla^2 \varphi' = -\frac{\varrho}{\varepsilon_0}\,, \qquad \nabla^2 \mathbf{A}' - \frac{1}{c^2}\frac{\partial^2 \mathbf{A}'}{\partial t^2} = -\mu_0\,\mathbf{J} + \frac{1}{c^2}\frac{\partial}{\partial t}\operatorname{grad}\varphi'\,, \tag{4.33}$$

the first of which (the *Poisson equation*) is decoupled from the second one. After solving the Poisson equation, the last term at the right-hand side of the second equation is not an unknown any more, thus showing that the equations resulting from the Coulomb gauge are indeed decoupled. Another possibility is choosing f such that $\theta' = 0$, which is equivalent to letting

$$\operatorname{div} \mathbf{A}' = -\frac{1}{c^2}\frac{\partial \varphi'}{\partial t}\,, \tag{4.34}$$

called *Lorentz gauge*. This transformation yields

$$\nabla^2 \varphi' - \frac{1}{c^2}\frac{\partial^2 \varphi'}{\partial t^2} = -\frac{\varrho}{\varepsilon_0}\,, \qquad \nabla^2 \mathbf{A}' - \frac{1}{c^2}\frac{\partial^2 \mathbf{A}'}{\partial t^2} = -\mu_0\,\mathbf{J}\,. \tag{4.35}$$

that are decoupled and have the form of the wave equation (4.13). Another interesting application of the gauge transformation is shown in Sect. 5.11.4.

4.6 Lagrangian Density for the Maxwell Equations

To apply the Lagrangian formalism to the Maxwell equations it is useful to use the expressions (4.26) of the fields in terms of the potentials. It follows that the functions playing the role of generalized coordinates and generalized velocities are φ, A_i, and, respectively, $\partial\varphi/\partial x_k$, $\partial A_i/\partial x_k$, $\partial A_i/\partial t$, with $i, k = 1, 2, 3$, $k \neq i$. The Lagrangian density, whose units are J m^{-3}, then reads

$$L_e = \frac{\varepsilon_0}{2}\,E^2 - \frac{1}{2\,\mu_0}\,B^2 - \varrho\,\varphi + \mathbf{J}\cdot\mathbf{A}\,, \tag{4.36}$$

with

$$E^2 = \left(\frac{\partial\varphi}{\partial x_1} + \frac{\partial A_1}{\partial t}\right)^2 + \left(\frac{\partial\varphi}{\partial x_2} + \frac{\partial A_2}{\partial t}\right)^2 + \left(\frac{\partial\varphi}{\partial x_3} + \frac{\partial A_3}{\partial t}\right)^2 \tag{4.37}$$

and

$$B^2 = \left(\frac{\partial A_3}{\partial x_2} - \frac{\partial A_2}{\partial x_3}\right)^2 + \left(\frac{\partial A_1}{\partial x_3} - \frac{\partial A_3}{\partial x_1}\right)^2 + \left(\frac{\partial A_2}{\partial x_1} - \frac{\partial A_1}{\partial x_2}\right)^2 , \tag{4.38}$$

To show that (4.36) is in fact the Lagrangian function of the Maxwell equations one starts with the generalized coordinate φ, to find $\partial L_e/\partial\varphi = -\varrho$. Then, considering the kth component,

$$\frac{\partial L_e/\varepsilon_0}{\partial(\partial\varphi/\partial x_k)} = \frac{\partial E^2/2}{\partial(\partial\varphi/\partial x_k)} = \frac{\partial E^2/2}{\partial(\partial A_k/\partial t)} = \frac{\partial\varphi}{\partial x_k} + \frac{\partial A_k}{\partial t} = -E_k = -\frac{D_k}{\varepsilon_0} . \tag{4.39}$$

Using (4.8) after replacing g with L_e, ξ_j with x_j, and w_i with φ yields $\operatorname{div}\mathbf{D} = \varrho$, namely, the first equation in (4.19). Turning now to another generalized coordinate, say, A_1, one finds $\partial L_e/\partial A_1 = J_1$. As L_e depends on the spatial derivatives of A_1 only through B^2, (4.38) and the first of (4.26) yield

$$\frac{\partial B^2/2}{\partial(\partial A_1/\partial x_3)} = \frac{\partial A_1}{\partial x_3} - \frac{\partial A_3}{\partial x_1} = B_2 , \quad \frac{\partial B^2/2}{\partial(\partial A_1/\partial x_2)} = \frac{\partial A_1}{\partial x_2} - \frac{\partial A_2}{\partial x_1} = -B_3 . \tag{4.40}$$

In contrast, L_e depends on the time derivative of A_1 only through E^2, as shown by (4.39). To use (4.8) one replaces g with L_e and w_i with A_1, then takes the derivative with respect to x_3 in the first relation in (4.40), the derivative with respect to x_2 in the second relation, and the derivative with respect to t of the last term in (4.39). In summary this yields

$$\frac{1}{\mu_0}\left(\frac{\partial B_3}{\partial x_2} - \frac{\partial B_2}{\partial x_3}\right) - \frac{\partial D_1}{\partial t} = J_1 , \tag{4.41}$$

namely, the first component of the second equation in (4.19).

4.7 Helmholtz Equation

Consider the wave equations (4.35) and assume that the charge density ϱ and current density $\mathbf{J}$ are given as functions of position and time. In the following, the apex in φ and $\mathbf{A}$ will be dropped for the sake of conciseness. The four scalar equations (4.35) are linear with respect to the unknowns and have the same structure; also, their coefficients and unknowns are all real. The solution of (4.35) will be tackled in this

section and in the following ones, basing upon the Fourier transform whose general properties are depicted in Sect. C.2. This solution procedure involves the use of complex functions. The starting assumption is that the condition for the existence of the Fourier transform with respect to time holds (such a condition is found by replacing x with t in (C.19)). Then one obtains

$$\nabla^2 \mathscr{F}_t\varphi + \frac{\omega^2}{c^2}\,\mathscr{F}_t\varphi = -\frac{1}{\varepsilon_0}\,\mathscr{F}_t\varrho\,, \qquad \nabla^2 \mathscr{F}_t\mathbf{A} + \frac{\omega^2}{c^2}\,\mathscr{F}_t\mathbf{A} = -\mu_0\,\mathscr{F}_t\mathbf{J}\,. \tag{4.42}$$

Indicating with f the transform of φ or A_i, and with b the transform of $-\varrho/\varepsilon_0$ or $-\mu_0 J_i$, $i = 1, 2, 3$, and letting $k^2 = \omega^2/c^2$, each scalar equation in (4.42) has the form of the *Helmholtz equation*

$$\nabla^2 f + k^2 f = b\,. \tag{4.43}$$

The solution of (4.43) is sought within a finite domain V (Fig. 4.1), for a given set of boundary conditions defined over the boundary S of V, and for a given right-hand side b defined within V and over S. Let $\mathbf{r} = (x, y, z)$ be a point external to V, $\mathbf{q} = (\xi, \eta, \zeta)$ a point internal to V, and

$$\mathbf{g} = \mathbf{r} - \mathbf{q}\,, \qquad g = \left[(x-\xi)^2 + (y-\eta)^2 + (z-\zeta)^2\right]^{1/2} \tag{4.44}$$

where, by construction, it is $g > 0$. In the following calculation, $\mathbf{r}$ is kept fixed while $\mathbf{q}$ varies. As a consequence, the derivatives of g are calculated with respect to ξ, η, and ζ. It is easily shown that in a three-dimensional space the auxiliary function[4]

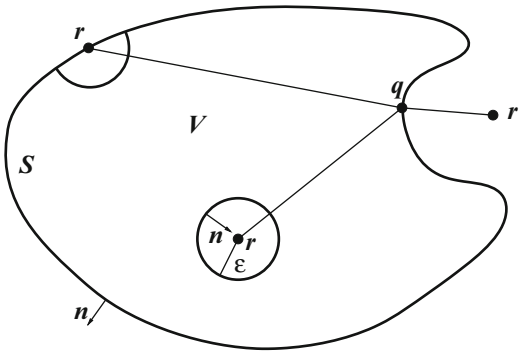


Fig. 4.1 The domain V used for the solution of the Helmholtz equation (4.43). The three possible positions of point $\mathbf{r}$ are shown: external to V, internal to V, or on the boundary S

[4]Function G is also called *Green function*. If $k = 0$, this function fulfills the Laplace equation $\nabla^2 G = 0$. As shown in Sect. 4.12.4, the form of the Green function that solves, e.g., the Laplace equation depends on the number of spatial dimensions considered.

$$G(g) = \frac{1}{g} \exp(-\mathrm{i}\, k\, g)\,, \qquad k \text{ real}, \tag{4.45}$$

fulfills the homogeneous Helmholtz equation $\nabla^2 G + k^2 G = 0$. Using the procedure that leads to the second Green theorem (Sect. A.5, Eq. (A.25)) yields the integral relation

$$\int_S \left(G \frac{\partial f}{\partial n} - f \frac{\partial G}{\partial n} \right) \mathrm{d}S = \int_V G\, b\, \mathrm{d}V\,, \tag{4.46}$$

where the unit vector $\mathbf{n}$ over S is oriented in the outward direction and, by construction, point $\mathbf{r}$ is external to V.

4.8 Helmholtz Equation in a Finite Domain

The relation (4.46) would not be applicable if $\mathbf{r}$ were internal to V, because G diverges for $g \to 0$ and, as a consequence, is not differentiable in $\mathbf{q} = \mathbf{r}$. On the other hand, in many applications $\mathbf{r}$ happens to be internal to V. In such cases one must exclude from the integral a suitable portion of volume V; this is achieved by considering a spherical domain of radius ϵ centered on $\mathbf{r}$ and internal to V (Fig. 4.1). Letting V_ϵ, S_ϵ be, respectively, the volume and surface of such a sphere, and considering the new volume $V' = V - V_\epsilon$, having $S' = S \cup S_\epsilon$ as boundary, makes (4.46) applicable to V', to yield

$$\int_S \left(G \frac{\partial f}{\partial n} - f \frac{\partial G}{\partial n} \right) \mathrm{d}S + \int_{S_\epsilon} (\ldots)\, \mathrm{d}S_\epsilon = \int_V G\, b\, \mathrm{d}V - \int_{V_\epsilon} G\, b\, \mathrm{d}V_\epsilon\,, \tag{4.47}$$

where the dots indicate that the integrand is the same as in the first integral at the left-hand side. Over S_ϵ it is $G = (1/\epsilon) \exp(-\mathrm{i}\, k\, \epsilon)$, with the unit vector $\mathbf{n}$ pointing from the surface towards the center of the sphere, namely, opposite to the direction along which ϵ increases. It follows $\partial G/\partial n = -\partial G/\partial \epsilon = (\mathrm{i}\, k + 1/\epsilon)\, G$. Letting $[f]$ and $[\partial f/\partial n]$ be the average values of f and, respectively, $\partial f/\partial n$ over S_ϵ, and observing that G and $\partial G/\partial \epsilon$ are constant there, yields

$$\int_{S_\epsilon} \left(G \frac{\partial f}{\partial n} - f \frac{\partial G}{\partial n} \right) \mathrm{d}S_\epsilon = 4\,\pi\, \exp(-\mathrm{i}\, k\, \epsilon) \left(\epsilon \left[\frac{\partial f}{\partial n} \right] - (1 + \mathrm{i}\, k\, \epsilon)\, [f] \right). \tag{4.48}$$

As for the integral $I = \int_{V_\epsilon} G\, b\, \mathrm{d}V_\epsilon$ it is useful to adopt the spherical coordinates (B.1) after shifting the origin to the center of the sphere. In the new reference it is $\mathbf{r} = 0$, so that the radial coordinate coincides with g. It follows

$$I = \int_0^{\epsilon} \int_0^{\pi} \int_0^{2\pi} g \, \sin\vartheta \, \exp(-\mathrm{i}\, k\, g)\, b(g, \vartheta, \varphi)\, \mathrm{d}g \, \mathrm{d}\vartheta \, \mathrm{d}\varphi \,. \tag{4.49}$$

Taking the absolute value of I and observing that g and $\sin\vartheta$ are positive yields $|I| \leq 2\,\pi\,\epsilon^2 \, \sup_{V_\epsilon} |b|$. To proceed, one assumes that f and b are sufficiently smooth as to fulfill the conditions

$$\lim_{\epsilon \to 0} \epsilon\, [f] = 0 \,, \qquad \lim_{\epsilon \to 0} \epsilon \left[\frac{\partial f}{\partial n} \right] = 0 \,, \qquad \lim_{\epsilon \to 0} \epsilon^2 \sup_{V_\epsilon} |b| = 0 \,. \tag{4.50}$$

Thanks to (4.50) one restores the original volume V by taking the limit $\epsilon \to 0$. Observing that $\lim_{\epsilon \to 0} [f] = f(\mathbf{r})$, one finds

$$4\,\pi\, f(\mathbf{r}) = \int_S \left(G \, \frac{\partial f}{\partial n} - f \, \frac{\partial G}{\partial n} \right) \mathrm{d}S - \int_V G\, b \, \mathrm{d}V \,, \tag{4.51}$$

that renders $f(\mathbf{r})$ as a function of b, of the boundary values of f and $\partial f / \partial n$, and of the auxiliary function G. It is easily found that if $\mathbf{r}$ were on the boundary S instead of being internal to V, the left-hand side of (4.51) would be replaced by $2\,\pi\, f(\mathbf{p})$. Similarly, if $\mathbf{r}$ were external to V, the left-hand side would be zero. In conclusion one generalizes (4.51) to

$$\omega_r f(\mathbf{r}) = \int_S \left(G \, \frac{\partial f}{\partial n} - f \, \frac{\partial G}{\partial n} \right) \mathrm{d}S - \int_V G\, b \, \mathrm{d}V \,, \tag{4.52}$$

where ω_r is the solid angle under which the surface S is seen from $\mathbf{r}$ considering the orientation of the unit vector $\mathbf{n}$. Namely, $\omega_r = 0$, $\omega_r = 2\,\pi$, or $\omega_r = 4\,\pi$ when $\mathbf{r}$ is external to V, on the boundary of V, or internal to V, respectively.

Letting $k = 0$ in (4.45), namely, taking $G = 1/g$, makes the results of this section applicable to the Poisson equation $\nabla^2 f = b$. It must be noted that (4.52) should be considered as an integral relation for f, not as the solution of the differential equation whence it derives. In fact, for actually calculating (4.52) it is necessary to prescribe both f and $\partial f / \partial n$ over the boundary. This is an overspecification of the problem: in fact, the theory of boundary-value problems shows that the solution of an equation of the form (4.43) is found by specifying over the boundary either the unknown function only (*Dirichlet boundary condition*), or its normal derivative only (*Neumann boundary condition*). To find a solution starting from (4.52) it is necessary to carry out more steps, by which either f or $\partial f / \partial n$ is eliminated from the integral at the right-hand side [69, Sect. 1.8–1.10]. In contrast, when the solution is sought in a domain whose boundary extends to infinity, and the contribution of the boundary conditions vanish as shown in Sect. 4.9, the limiting case of (4.52) provides a solution proper. More comments about this issue are made in Sect. 4.12.3.

4.9 Solution of the Helmholtz Equation in an Infinite Domain

The procedure shown in Sect. 4.8 is readily extended to the case $V \to \infty$. Here one may replace S with a spherical surface of radius $R \to \infty$, centered on $\mathbf{r}$; this makes the calculation of the integral over S similar to that over the sphere of radius ε outlined in Sect. 4.8, the only difference being that the unit vector $\mathbf{n}$ now points in the direction where R increases. Shifting the origin to $\mathbf{r}$ and observing that $\omega_r = 4\,\pi$ yield

$$\int_S \left(G\,\frac{\partial f}{\partial n} - f\,\frac{\partial G}{\partial n} \right) \mathrm{d}S = 4\,\pi\,\exp(-\mathrm{i}\,k\,R)\,\left(R\left[\frac{\partial f}{\partial n}\right] + (1 + \mathrm{i}\,k\,R)\,[f] \right)\,, \tag{4.53}$$

where the averages are calculated over S. To proceed one assumes that the following relations hold,

$$\lim_{R\to\infty} [f] = 0\,, \qquad \lim_{R\to\infty} R\left(\left[\frac{\partial f}{\partial n}\right] + \mathrm{i}\,k\,[f] \right) = 0\,, \tag{4.54}$$

that are called *Sommerfeld asymptotic conditions*. Due to (4.54) the surface integral (4.53) vanishes. Shifting the origin back from $\mathbf{r}$ to the initial position, the solution of the Helmholtz equation (4.43) over an infinite domain finally reads

$$f(\mathbf{r}) = -\frac{1}{4\,\pi} \int_\infty b(\mathbf{q})\,\frac{\exp(-\mathrm{i}\,k\,|\mathbf{r}-\mathbf{q}|)}{|\mathbf{r}-\mathbf{q}|}\,\mathrm{d}^3 q\,, \tag{4.55}$$

where $\int_\infty$ indicates the integral over the whole three-dimensional $\mathbf{q}$ space. The $k = 0$ case yields the solution of the Poisson equation $\nabla^2 f = b$ in an infinite domain,

$$f(\mathbf{r}) = -\frac{1}{4\,\pi} \int_\infty b(\mathbf{q})\,\frac{1}{|\mathbf{r}-\mathbf{q}|}\,\mathrm{d}^3 q\,. \tag{4.56}$$

4.10 Solution of the Wave Equation in an Infinite Domain

The solutions of the Helmholtz equation found in Sects. 4.8, 4.9 allow one to calculate that of the wave equation. In fact, it is worth reminding that the Helmholtz equation (4.43) was deduced in Sect. 4.7 by Fourier transforming the wave equations (4.35) and *i)* letting f indicates the transform of the scalar potential φ or of any component A_i of the vector potential, *ii)* letting b indicates the transform of $-\varrho/\varepsilon_0$ or $-\mu_0 J_i$, $i = 1, 2, 3$. As a consequence, f and b depend on the angular frequency ω besides the spatial coordinates. From the definition $k^2 = \omega^2/c^2$ one may also assume that both k and ω have the same sign, so that $k = \omega/c$. Considering for simplicity the case $V \to \infty$, applying (C.17) to antitransform (4.56), and interchanging the order of integrals yields

$$\mathscr{F}^{-1}f = -\frac{1}{4\pi}\int_{\infty}\frac{1}{g}\left[\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} b(\mathbf{q},\omega)\,\exp[\mathrm{i}\,\omega\,(t-g/c)]\,\mathrm{d}\omega\right]\mathrm{d}^3q\,, \tag{4.57}$$

with $g = |\mathbf{r}-\mathbf{q}|$. Now, denote with a the antitransform of b, $a(\mathbf{q},t) = \mathscr{F}^{-1}b$. It follows that the function between brackets in (4.57) coincides with $a(\mathbf{q}, t-g/c)$. As remarked above, when f represents φ, then a stands for $-\varrho/\varepsilon_0$; similarly, when f represents a component of $\mathbf{A}$, then a stands for the corresponding component of $-\mu_0\,\mathbf{J}$. In conclusion,

$$\varphi(\mathbf{r},t) = \frac{1}{4\,\pi\,\varepsilon_0}\int_{\infty}\frac{\varrho(\mathbf{q},t-|\mathbf{r}-\mathbf{q}|/c)}{|\mathbf{r}-\mathbf{q}|}\,\mathrm{d}^3q\,, \tag{4.58}$$

$$\mathbf{A}(\mathbf{r},t) = \frac{\mu_0}{4\,\pi}\int_{\infty}\frac{\mathbf{J}(\mathbf{q},t-|\mathbf{r}-\mathbf{q}|/c)}{|\mathbf{r}-\mathbf{q}|}\,\mathrm{d}^3q\,, \tag{4.59}$$

that express the potentials in terms of the field sources ϱ and $\mathbf{J}$, when the asymptotic behavior of the potentials fulfills the Sommerfeld conditions (4.54). The functions rendered by the antitransforms are real, as should be. Note that $|\mathbf{r}-\mathbf{q}|/c > 0$ is the time necessary for a signal propagating with velocity c to cross the distance $|\mathbf{r}-\mathbf{q}|$. As $t-|\mathbf{r}-\mathbf{q}|/c < t$, the above expressions of φ and $\mathbf{A}$ are called *retarded potentials*.[5]

4.11 Lorentz Force

It has been assumed so far that the sources of the electromagnetic field, namely charge density and current density, are prescribed functions of position and time. This is not necessarily so, because the charges are in turn acted upon by the electromagnetic field, so that their dynamics is influenced by it. Consider a test charge of value e immersed in an electromagnetic field described by the vectors $\mathbf{E}$, $\mathbf{B}$ generated by other charges. The force acting upon the test charge is the *Lorentz force* [10, Vol. I, Sect. 44]

$$\mathbf{F} = e\,(\mathbf{E} + \mathbf{u}\wedge\mathbf{B})\,, \tag{4.60}$$

where $\mathbf{u}$ is the velocity of the test charge and $\mathbf{E}$, $\mathbf{B}$ are independent of $\mathbf{u}$. The expression of the Lorentz force does not derive from assumptions separate from Maxwell's equations; in fact, it follows from Maxwell's equations and Special

[5] Expressions of φ and $\mathbf{A}$ obtained from (4.58,4.59) after replacing $t-|\mathbf{r}-\mathbf{q}|/c$ with $t+|\mathbf{r}-\mathbf{q}|/c$ are also solutions of the wave equations (4.35). This is due to the fact that the Helmholtz equation (4.43) can also be solved by using G^* instead of G, which in turn reflects the time reversibility of the wave equation. However, the form with $t-|\mathbf{r}-\mathbf{q}|/c$ better represents the idea that an electromagnetic perturbation, that is present in $\mathbf{r}$ at the time t, is produced by a source acting in $\mathbf{q}$ at a time prior to t.

Relativity [52, 135]. The extension of (4.60) to the case of a number of point-like charges follows the same line as in Sect. 4.4: considering the charges belonging to a cell of volume ΔV centered at $\mathbf{r}$, one writes (4.60) for the jth charge and takes the sum over j, to find

$$\mathbf{f} = \frac{\sum_j' \mathbf{F}_j}{\Delta V} = \varrho \, (\mathbf{E} + \mathbf{v} \wedge \mathbf{B}) \,, \tag{4.61}$$

where ϱ, $\mathbf{v}$ are defined in (4.21) and $\mathbf{f}$ is the force density ($[\mathbf{f}] = \mathrm{N\ m^{-3}}$). The fields in (4.61) are calculated in $\mathbf{r}$ and t.

Consider a small time interval δt during which the charge contained within ΔV is displaced by $\delta \mathbf{r} = \mathbf{v}\, \delta t$. The work per unit volume exchanged between the charge and the electromagnetic field due to such a displacement is

$$\delta w = \mathbf{f} \cdot \delta \mathbf{r} = \varrho \ (\mathbf{E} + \mathbf{v} \wedge \mathbf{B}) \cdot \mathbf{v}\, \delta t = \mathbf{E} \cdot \mathbf{J}\, \delta t \,, \qquad [w] = \mathrm{J\ m^{-3}} \,, \tag{4.62}$$

where (A.32) and (4.21) have been used. When the scalar product is positive, the charge acquires kinetic energy from the field, and vice versa. Letting $\delta t \to 0$ yields

$$\frac{\partial w}{\partial t} = \mathbf{E} \cdot \mathbf{J} \,, \tag{4.63}$$

where the symbol of partial derivative is used because (4.63) is calculated with $\mathbf{r}$ fixed.

4.12 Complements

4.12.1 Invariance of the Euler Equations

It has been shown in Sect. 4.2 that the Euler equations (4.8) are still fulfilled if the generating function g is replaced with $g' = g + \operatorname{div} \mathbf{h}$, where $\mathbf{h}$ is an arbitrary vector of length m whose entries depend on $\mathbf{w}$ and $\xi_1, \ldots, \xi_m$, but not on the derivatives of $\mathbf{w}$. This property is a generalization of that illustrated in Sect. 1.2 with reference to a system of particles, where it was shown that the solutions $w_i(\xi)$ are invariant under addition to g of the total derivative of an arbitrary function that depends on $\mathbf{w}$ and ξ only.

4.12.2 Wave Equations for the $\mathbf{E}$ *and* $\mathbf{B}$ *Fields*

The Maxwell equations can be rearranged in the form of wave equations for the electric and magnetic fields. To this purpose, one takes the rotational of both sides

of the second equation in (4.24). Using the first identity in (A.36) and the relation $\mathbf{D} = \varepsilon_0 \mathbf{E}$ provides $-\partial \operatorname{rot} \mathbf{B}/\partial t = \operatorname{rot}\operatorname{rot} \mathbf{E} = \operatorname{grad}\operatorname{div}(\mathbf{D}/\varepsilon_0) - \nabla^2 \mathbf{E}$. Replacing $\operatorname{div} \mathbf{D}$ and $\operatorname{rot} \mathbf{H} = \operatorname{rot} \mathbf{B}/\mu_0$ from (4.19) and using $\varepsilon_0 \mu_0 = 1/c^2$ then yield

$$\nabla^2 \mathbf{E} - \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} = \frac{1}{\varepsilon_0} \operatorname{grad} \varrho + \mu_0 \frac{\partial \mathbf{J}}{\partial t} . \tag{4.64}$$

Similarly, one takes the rotational of both sides of the second equation in (4.19). Using the relation $\mathbf{B} = \mu_0 \mathbf{H}$ provides $\varepsilon_0 \, \partial \operatorname{rot} \mathbf{E}/\partial t + \operatorname{rot} \mathbf{J} = \operatorname{rot}\operatorname{rot} \mathbf{H} = \operatorname{grad}\operatorname{div}(\mathbf{B}/\mu_0) - \nabla^2 \mathbf{H}$. Replacing $\operatorname{div} \mathbf{B}$ and $\operatorname{rot} \mathbf{E}$ from (4.24) yields

$$\nabla^2 \mathbf{H} - \frac{1}{c^2} \frac{\partial^2 \mathbf{H}}{\partial t^2} = -\operatorname{rot} \mathbf{J} . \tag{4.65}$$

4.12.3 Comments on the Boundary-Value Problem

Considering relation (4.52) derived in Sect. 4.8, one notes that the right-hand side is made of the difference between two terms; the first one depends on the boundary values of f, $\partial f/\partial n$, but not on b, while the second one depends only on the values of b within V and over the boundary. In these considerations it does not matter whether point $\mathbf{r}$ is external to V, on the boundary of V, or internal to it. In latter case the two terms at the right-hand side of (4.52) balance each other.

If b is replaced with a different function $\tilde{b}$, and thereby the value of the second integral changes, it is possible to modify the boundary values in such a way as to balance the variation of the second integral with that of the first one; as a consequence, $f(\mathbf{r})$ is left unchanged. A possible choice for the modified b is $\tilde{b} = 0$; by this choice one eliminates the data of the differential equation and suitably modifies the boundary values, leaving the solution unaffected. An observer placed at $\mathbf{r}$ would be unable to detect that the data have disappeared. The same process can also be carried out in reverse, namely, by eliminating the boundary values and suitably changing the data.

An example is given in Prob. 4.4 with reference to a one-dimensional Poisson equation where the original charge density differs from zero in a finite interval $[a, b]$. The charge density is removed and the boundary values at a are modified so that the electric potential φ is unaffected for $x \geq b$. Obviously φ changes for $a < x < b$ because both the charge density and boundary conditions are different, and also for $x \leq a$ because the boundary conditions are different.

4.12.4 Green Function for the Upper Half Plane

The solution of the Laplace equation in two dimensions is considered in this section, using the method of the Green function introduced in Sect. 4.7. Using the Cartesian coordinates, let $P = (x, y)$ be a point of the upper half plane $y > 0$, $Q = (\xi, 0)$ a point of the x axis, and $r = |P - Q| = [(\xi - x)^2 + y^2]^{1/2} > 0$ the distance between them. Treating ξ as a parameter, the function

$$G(x, y; \xi) = \log(r^2) \tag{4.66}$$

is harmonic in the upper half plane. Indeed, using a suffix to indicate a partial derivative, it is $G_x = 2\,(x-\xi)/r^2$ and $G_y = 2\,y/r^2$, whence $G_{xx} = 2\,[y^2-(x-\xi)^2]/r^4$ and $G_{yy} = 2\,[(x-\xi)^2-y^2]/r^4$. Adding the last two relations yields $\nabla^2 G = 0$. From the above it follows that also $\log r = G/2$ is harmonic in the upper half plane. On the contrary, G is not everywhere harmonic along the x axis because $G(x, 0)$ diverges as $|\xi - x| \to 0$.

Let $E(x)$ be a piecewise continuous function defined along the x axis. It is also assumed that E is bounded, namely, that some number $M > 0$ exists such that $|E| \le M$ for all x. The integral

$$\varphi(x, y) = \varphi_0 - \frac{1}{2\pi} \int_{-\infty}^{+\infty} E(\xi)\, G(x, y; \xi)\, \mathrm{d}\xi\,, \tag{4.67}$$

with φ_0 an arbitrary constant, is harmonic in the upper half plane due to the properties of G. In addition, its derivative with respect to y along the x axis is $-E$. More precisely, if E is continuous at x, then $\lim_{y\to 0+} \partial\varphi/\partial y = -E(x)$, whereas if E has a first-order discontinuity at x, $E(x)$ in the above must be replaced with $[E(x^+) + E(x^-)]/2$. In fact, from the definition of G it follows

$$\lim_{y\to 0+} \frac{\partial\varphi}{\partial y} = -\lim_{y\to 0+} \int_{-\infty}^{+\infty} \frac{E(\xi)\, y/\pi}{(\xi - x)^2 + y^2}\, \mathrm{d}\xi\,, \tag{4.68}$$

where x is kept fixed. Consider an interval of length s centered at x. The contribution to the integral of this interval is

$$\int_{x-s/2}^{x+s/2} \frac{E(\xi)\, y/\pi}{(\xi - x)^2 + y^2}\, \mathrm{d}\xi = \langle E\rangle_s \int_{x-s/2}^{x+s/2} \frac{y/\pi}{(\xi - x)^2 + y^2}\, \mathrm{d}\xi\,, \tag{4.69}$$

with $\langle E\rangle_s$ the average value of E over the interval. Letting $\lambda = (\xi - x)/y$, the integral at the right-hand side of (4.69) becomes

$$\int_{-s/(2\,y)}^{+s/(2\,y)} \frac{1/\pi}{1 + \lambda^2}\, \mathrm{d}\lambda = \frac{2}{\pi} \arctan\left(\frac{s}{2\,y}\right), \tag{4.70}$$

whose limit for $y \to 0^+$ is 1. The contribution to the integral of the interval from $-\infty$ to $x - s/2$ is zero because

$$\left| \int_{-\infty}^{x-s/2} \frac{E(\xi)\, y/\pi}{(\xi - x)^2 + y^2} \, \mathrm{d}\xi \right| \leq M \int_{-\infty}^{x-s/2} \frac{y/\pi}{(\xi - x)^2 + y^2} \, \mathrm{d}\xi \,. \tag{4.71}$$

The integral at the right-hand side of (4.71) is equal to $(M/\pi)\,[\arctan(-s/(2\,y)) - \arctan(-\infty)]$, whose limit for $y \to 0^+$ is 0. Similarly, the integral over the interval from $x + s/2$ to $+\infty$ does not contribute. In conclusion,

$$\lim_{y \to 0^+} \frac{\partial \varphi}{\partial y} = -\langle E \rangle_s \,. \tag{4.72}$$

The proof is completed by observing that s can be taken as small as we please. In conclusion, the integral (4.67) defining φ provides the solution of the Laplace equation $\nabla^2 \varphi = 0$ in the upper half plane, supplemented with the prescription of the normal derivative along the x axis (*Neumann boundary conditions*).

The results found in this section have a useful application to a method for measuring the conductivity of a simply connected, flat sample of material of arbitrary shape. The method is described in Sect. 25.7.

Problems

4.1 Solve the one-dimensional Poisson equation $\mathrm{d}^2\varphi/\mathrm{d}x^2 = -\varrho(x)/\varepsilon_0$, with ϱ given, using the integration by parts to avoid a double integral. The solution is prescribed at $x = a$ while the first derivative is prescribed at $x = c$.

4.2 Let $c = a$ in the solution of Prob. 4.1 and assume that the charge density ϱ differs from zero only in a finite interval $a \leq x \leq b$. Find the expression of φ for $x > b$ when both the solution and the first derivative are prescribed at $x = a$.

4.3 In Prob. 4.2 replace the charge density ϱ with a different one, say, $\tilde{\varrho}$. Discuss the conditions that leave the solution unchanged.

4.4 In Prob. 4.2 remove the charge density ϱ and modify the boundary conditions at a so that the solution for $x > b$ is left unchanged.

4.5 Using the results of Probs. 4.2 and 4.3, and assuming that both M_0 and M_1 are different from zero, replace the ratio ϱ/ε_0 with $\mu\,\delta(x - h)$ and find the parameters μ, h that leave M_0, M_1 unchanged. Noting that h does not necessarily belong to the interval $[a, b]$, discuss the outcome for different positions of h with respect to a.

4.6 A particle of mass m and charge q enters at $t = 0$ a region where a constant magnetic induction $\mathbf{B} = B\,\mathbf{k}$ is present, with $B > 0$. The velocity $\mathbf{u}_0$ at $t = 0$ is normal to $\mathbf{B}$. Find the particle's trajectory for $t > 0$.

Chapter 5
Applications of the Concepts of Electromagnetism

5.1 Introduction

This chapter provides a number of important applications of the concepts of Electromagnetism. The solution of the wave equation found in Chap. 4 is used to calculate the potentials generated by a point-like charge; this result is exploited later to analyze the decay of atoms in the frame of the classical model, due to the radiated power. Next, the continuity equations for the energy and momentum of the electromagnetic field are found. As an application, the energy and momentum of the electromagnetic field are calculated in terms of modes in a finite domain, showing that the energy of each mode has the same expression as that of a linear harmonic oscillator. The analysis is extended also to an infinite domain. The chapter is concluded by the derivation of the eikonal equation, leading to the approximation of Geometrical Optics, followed by the demonstration that the eikonal equation is generated by a variational principle, namely, the Fermat principle. The complements show the derivation of the fields generated by a point-like charge and the power radiated by it. It is found that the planetary model of the atom is inconsistent with electromagnetism because it contradicts the atom's stability. Finally, a number of analogies are outlined and commented between Mechanics and Geometrical Optics, based on the comparison between the Maupertuis and Fermat principles. The reasoning deriving from the comparison hints at the possibility that mechanical laws more general than Newton's law exist.

5.2 Potentials Generated by a Point-Like Charge

The calculation of φ and $\mathbf{A}$ based upon (4.58, 4.59) has the inconvenience that as $\mathbf{q}$ varies over the space, it is necessary to consider the sources ϱ and $\mathbf{J}$ at different time instants. This may be avoided by recasting a in the form

M. Rudan, *Physics of Semiconductor Devices*, DOI 10.1007/978-3-319-63154-7_5

$$a(\mathbf{q}, t - |\mathbf{r} - \mathbf{q}|/c) = \int_{-\infty}^{+\infty} a(\mathbf{q}, t')\, \delta(t' - t + |\mathbf{r} - \mathbf{q}|/c)\, \mathrm{d}t' \,, \tag{5.1}$$

with δ the Dirac delta (Sect. C.4), and interchanging the integration over $\mathbf{q}$ in (4.58, 4.59) with that over t'. This procedure is particularly useful when the source of the field is a single point-like charge. Remembering (4.20), one replaces ϱ and $\mathbf{J}$ with

$$\varrho_c(\mathbf{q}, t') = e\, \delta\left(\mathbf{q} - \mathbf{s}(t')\right) \,, \qquad \mathbf{J}_c(\mathbf{q}, t') = e\, \delta\left(\mathbf{q} - \mathbf{s}(t')\right)\, \mathbf{u}(t') \,, \tag{5.2}$$

where e is the value of the point-like charge, $\mathbf{s} = \mathbf{s}(t')$ its trajectory, and $\mathbf{u}(t') = \mathrm{d}\mathbf{s}/\mathrm{d}t'$ its velocity. First, the integration over space fixes $\mathbf{q}$ at $\mathbf{s}' = \mathbf{s}(t')$, thus yielding

$$\varphi(\mathbf{r}, t) = \frac{e}{4\,\pi\,\varepsilon_0} \int_{-\infty}^{+\infty} \frac{\delta[\beta(t')]}{|\mathbf{r} - \mathbf{s}'|}\, \mathrm{d}t' \,, \quad \mathbf{A}(\mathbf{r}, t) = \frac{e\,\mu_0}{4\pi} \int_{-\infty}^{+\infty} \frac{\delta[\beta(t')]\, \mathbf{u}(t')}{|\mathbf{r} - \mathbf{s}'|}\, \mathrm{d}t' \,, \tag{5.3}$$

with $\beta(t') = t' - t + |\mathbf{r} - \mathbf{s}'|/c$. Next, the integration over t' fixes the latter to the value that makes the argument of δ to vanish. Such a value is the solution of

$$|\mathbf{r} - \mathbf{s}(t')| = c\,(t - t') \,, \tag{5.4}$$

where t, $\mathbf{r}$, and the function $\mathbf{s}(t')$ are prescribed. As $|\mathbf{u}| < c$ it can be shown that the solution of (5.4) exists and is unique [85, Sect. 63]. Observing that the argument of δ in (5.3) is a function of t', to complete the calculation one must follow the procedure depicted in Sect. C.5, which involves the derivative

$$\frac{\mathrm{d}\beta}{\mathrm{d}t'} = 1 + \frac{1}{c}\, \frac{\mathrm{d}|\mathbf{r} - \mathbf{s}'|}{\mathrm{d}t'} = 1 + \frac{\mathrm{d}[(\mathbf{r} - \mathbf{s}') \cdot (\mathbf{r} - \mathbf{s}')]}{2\,c\,|\mathbf{r} - \mathbf{s}'|\,\mathrm{d}t'} = 1 - \frac{\mathbf{r} - \mathbf{s}'}{|\mathbf{r} - \mathbf{s}'|}\, \frac{\mathbf{u}}{c} \,. \tag{5.5}$$

Then, letting $t' = \tau$ be the solution of (5.4) and $\dot{\beta} = (\mathrm{d}\beta/\mathrm{d}t')_{t'=\tau}$, one applies (C.57) to (5.3). The use of the absolute value is not necessary here, in fact one has $[(\mathbf{r} - \mathbf{s}')/|\mathbf{r} - \mathbf{s}'|] \cdot (\mathbf{u}/c) \le u/c < 1$, whence $|\dot{\beta}| = \dot{\beta}$. In conclusion one finds

$$\varphi(\mathbf{r}, t) = \frac{e/(4\,\pi\,\varepsilon_0)}{|\mathbf{r} - \mathbf{s}(\tau)| - (\mathbf{r} - \mathbf{s}(\tau)) \cdot \mathbf{u}(\tau)/c} \,, \tag{5.6}$$

$$\mathbf{A}(\mathbf{r}, t) = \frac{e\,\mu_0/(4\,\pi)\, \mathbf{u}(\tau)}{|\mathbf{r} - \mathbf{s}(\tau)| - (\mathbf{r} - \mathbf{s}(\tau)) \cdot \mathbf{u}(\tau)/c} \,, \tag{5.7}$$

that provide the potentials generated in $\mathbf{r}$ and at time t by a point-like charge that follows the trajectory $\mathbf{s} = \mathbf{s}(\tau)$. The relation between t and τ is given by $t = \tau + |\mathbf{r} - \mathbf{s}(\tau)|/c$, showing that $t - \tau$ is the time necessary for the electromagnetic perturbation produced by the point-like charge at $\mathbf{s}$ to reach the position $\mathbf{r}$. The expressions (5.6, 5.7) are called *Liénard and Wiechert potentials*. In the case $\mathbf{u} = 0$ they become

$$\varphi(\mathbf{r}) = \frac{e}{4\,\pi\,\varepsilon_0\,|\mathbf{r}-\mathbf{s}|}\,, \qquad \mathbf{A}(\mathbf{r}) = 0\,, \tag{5.8}$$

$\mathbf{s} = \text{const}$, the first of which is the Coulomb potential. The fields $\mathbf{E}$, $\mathbf{B}$ generated by a point-like charge are obtained from (5.6, 5.7) using (4.26). The calculation is outlined in Sect. 5.11.1.

5.3 Energy Continuity—Poynting Vector

The right-hand side of (4.63) is recast in terms of the fields by replacing $\mathbf{J}$ with the left-hand side of the second equation in (4.19); using $\mathbf{D} = \varepsilon_0\,\mathbf{E}$,

$$\frac{\partial w}{\partial t} = \mathbf{E}\cdot\left(\operatorname{rot}\mathbf{H} - \varepsilon_0\,\frac{\partial \mathbf{E}}{\partial t}\right) = \mathbf{E}\cdot\operatorname{rot}\mathbf{H} - \frac{\varepsilon_0}{2}\,\frac{\partial E^2}{\partial t}\,. \tag{5.9}$$

The above expression is given a more symmetric form by exploiting the first equation in (4.19). In fact, a scalar multiplication of the latter by $\mathbf{H}$ along with the relation $\mathbf{B} = \mu_0\,\mathbf{H}$ provides $0 = \mathbf{H}\cdot\operatorname{rot}\mathbf{E} + \mu_0\,\partial(H^2/2)/\partial t$ which, subtracted from (5.9), finally yields

$$\frac{\partial w}{\partial t} = \mathbf{E}\cdot\operatorname{rot}\mathbf{H} - \mathbf{H}\cdot\operatorname{rot}\mathbf{E} - \frac{\partial w_{\text{em}}}{\partial t}\,, \qquad w_{\text{em}} = \frac{1}{2}\left(\varepsilon_0\,E^2 + \mu_0\,H^2\right)\,. \tag{5.10}$$

Then, using the second identity in (A.36) transforms (5.10) into

$$\frac{\partial}{\partial t}\,(w + w_{\text{em}}) + \operatorname{div}\mathbf{S} = 0\,, \qquad \mathbf{S} = \mathbf{E}\wedge\mathbf{H}\,. \tag{5.11}$$

As $w + w_{\text{em}}$ is an energy density, $\mathbf{S}$ (called *Poynting vector*) is an energy-flux density ($[\mathbf{S}] = \text{J m}^{-2}\,\text{s}^{-1}$). To give w_{em} and $\mathbf{S}$ a physical meaning one notes that (5.11) has the form of a continuity equation (compare, e.g., with (23.3) and (4.23)) where two interacting systems are involved, namely, the charges and the electromagnetic field. Integrating (5.11) over a volume V yields

$$\frac{\mathrm{d}}{\mathrm{d}t}(W + W_{\text{em}}) = -\int_\Sigma \mathbf{S}\cdot\mathbf{n}\,\mathrm{d}\Sigma\,, \qquad W = \int_V w\,\mathrm{d}V\,, \quad W_{\text{em}} = \int_V w_{\text{em}}\,\mathrm{d}V\,, \tag{5.12}$$

where Σ is the boundary of V, $\mathbf{n}$ is the unit vector normal to $\mathrm{d}\Sigma$ oriented in the outward direction with respect to V, and W, W_{em} are energies. If V is let expand to occupy all space, the surface integral in (5.12) vanishes because the fields $\mathbf{E}$, $\mathbf{H}$ vanish at infinity; it follows that for an infinite domain the sum $W + W_{\text{em}}$ is conserved in time, so that $\mathrm{d}W_{\text{em}}/\mathrm{d}t = -\mathrm{d}W/\mathrm{d}t$. Observing that W is the kinetic energy of the charge, and that the latter exchanges energy with the electromagnetic field, gives

W_{em} the meaning of energy of the electromagnetic field; as a consequence, w_{em} is the energy density of the electromagnetic field, and the sum $W + W_{\text{em}}$ is the constant energy of the two interacting systems.

When V is finite, the surface integral in (5.12) may be different from zero, hence the sum $W + W_{\text{em}}$ is not necessarily conserved. This allows one to give the surface integral the meaning of energy per unit time that crosses the boundary Σ, carried by the electromagnetic field. In this reasoning it is implied that when V is finite, it is chosen in such a way that no charge is on the boundary at time t. Otherwise the kinetic energy of the charges crossing Σ during dt should also be accounted for.

5.4 Momentum Continuity

The procedure used in Sect. 5.3 to derive the continuity equation for the charge energy can be replicated to obtain the continuity equation for the charge momentum per unit volume, $\mathbf{m}$. Remembering (4.61) one finds the relation $\mathbf{f} = \dot{\mathbf{m}} = \sum_j' \dot{\mathbf{p}}_j / \Delta V$, with $\mathbf{p}_j$ the momentum of the jth charge contained within ΔV. Using (4.19) along with $\mathbf{J} = \varrho\, \mathbf{v}$ yields

$$\dot{\mathbf{m}} = \varrho\, \mathbf{E} + \mathbf{J} \wedge \mathbf{B} = \mathbf{E}\ \mathrm{div}\, \mathbf{D} + \left(\mathrm{rot}\, \mathbf{H} - \frac{\partial \mathbf{D}}{\partial t} \right) \wedge \mathbf{B}\,. \tag{5.13}$$

Adding $\mathbf{D} \wedge \partial \mathbf{B}/\partial t$ to both sides of (5.13), using $\partial \mathbf{B}/\partial t = -\,\mathrm{rot}\, \mathbf{E}$, and rearranging:

$$\dot{\mathbf{m}} + \frac{\partial \mathbf{D}}{\partial t} \wedge \mathbf{B} + \mathbf{D} \wedge \frac{\partial \mathbf{B}}{\partial t} = \mathbf{E}\ \mathrm{div}\, \mathbf{D} + (\mathrm{rot}\, \mathbf{H}) \wedge \mathbf{B} + (\mathrm{rot}\, \mathbf{E}) \wedge \mathbf{D}\,. \tag{5.14}$$

Poynting vector's definition (5.11) transforms the left-hand side of (5.14) into $\dot{\mathbf{m}} + \varepsilon_0\, \mu_0\, \partial(\mathbf{E} \wedge \mathbf{H})/\partial t = \partial(\mathbf{m} + \mathbf{S}/c^2)/\partial t$. In turn, the kth component of $\mathbf{E}\ \mathrm{div}\, \mathbf{D} + (\mathrm{rot}\, \mathbf{E}) \wedge \mathbf{D}$ can be recast as

$$\varepsilon_0\ (E_k\ \mathrm{div}\, \mathbf{E} + \mathbf{E} \cdot \mathrm{grad}\, E_k) - \frac{\varepsilon_0}{2} \frac{\partial E^2}{\partial x_k} = \varepsilon_0\ \mathrm{div}(E_k\, \mathbf{E}) - \frac{\varepsilon_0}{2} \frac{\partial E^2}{\partial x_k}\,. \tag{5.15}$$

Remembering that $\mathrm{div}\, \mathbf{B} = 0$, the kth component of the term $(\mathrm{rot}\, \mathbf{H}) \wedge \mathbf{B} = \mathbf{H}\ \mathrm{div}\, \mathbf{B} + (\mathrm{rot}\, \mathbf{H}) \wedge \mathbf{B}$ is treated in the same manner. Adding up the contributions of the electric and magnetic parts and using the definition (5.10) of w_{em} yield

$$\frac{\partial}{\partial t} \left(m_k + \frac{1}{c^2}\, S_k \right) + \mathrm{div}\, \mathbf{T}_k = 0\,, \qquad \mathbf{T}_k = w_{\text{em}}\, \mathbf{i}_k - \varepsilon_0\, E_k\, \mathbf{E} - \mu_0\, H_k \mathbf{H}\,, \tag{5.16}$$

with $\mathbf{i}_k$ the unit vector of the kth axis. As $m_k + S_k/c^2$ is a momentum density, $\mathbf{T}_k$ is a momentum-flux density ($[\mathbf{T}_k] = \mathrm{J\ m^{-3}}$). Following the same reasoning as in Sect. 5.3 one integrates (5.16) over a volume V, to find

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_V \left(m_k + \frac{1}{c^2}\,S_k\right)\mathrm{d}V = -\int_\Sigma \mathbf{T}_k\cdot\mathbf{n}\,\mathrm{d}\Sigma\,, \tag{5.17}$$

where Σ and $\mathbf{n}$ are defined as in (5.12). If V is let expand to occupy all space, the surface integral in (5.17) vanishes because the fields $\mathbf{E}$, $\mathbf{H}$ vanish at infinity; it follows that for an infinite domain the sum

$$\int_V \left(\mathbf{m} + \frac{1}{c^2}\,\mathbf{S}\right)\mathrm{d}V = \mathbf{p} + \int_V \frac{1}{c^2}\,\mathbf{S}\,\mathrm{d}V\,, \qquad \mathbf{p} = \int_V \mathbf{m}\,\mathrm{d}V \tag{5.18}$$

is conserved in time. As $\mathbf{p}$ is the momentum of the charge, $\int_V \mathbf{S}/c^2\,\mathrm{d}^3r$ takes the meaning of momentum of the electromagnetic field within V. As a consequence, $\mathbf{S}/c^2$ takes the meaning of momentum per unit volume of the electromagnetic field.

When V is finite, the surface integral in (5.17) may be different from zero, hence the sum (5.18) is not necessarily conserved. This allows one to give the surface integral in (5.17) the meaning of momentum per unit time that crosses the boundary Σ, carried by the electromagnetic field. In this reasoning it is implied that when V is finite, it is chosen in such a way that no charge is on the boundary at time t. Otherwise the momentum of the charges crossing Σ during $\mathrm{d}t$ should also be accounted for.

5.5 Modes of the Electromagnetic Field

The expressions of the energy and momentum of the electromagnetic field, worked out in Sect. 5.3 and 5.4, take a particularly interesting form when a spatial region free of charges is considered. In fact, if one lets $\varrho = 0$, $\mathbf{J} = 0$, Eqs. (4.33) or (4.35) that provide the potentials become homogeneous. To proceed one takes a finite region of volume V; the calculation will be extended in Sect. 5.8 to the case of an infinite domain. As the shape of V is not essential for the considerations illustrated here, it is chosen as that of a box whose sides d_1, d_2, d_3 are aligned with the coordinate axes and start from the origin (Fig. 5.1). The volume of the box is $V = d_1\,d_2\,d_3$.

The calculation is based on (4.33) that are the equations for the potentials deriving from the Coulomb gauge (4.32). Letting $\varrho = 0$, $\mathbf{J} = 0$ and dropping the primes yield

$$\nabla^2\varphi = 0\,, \qquad \nabla^2\mathbf{A} - \frac{1}{c^2}\,\frac{\partial^2\mathbf{A}}{\partial t^2} = \frac{1}{c^2}\,\frac{\partial}{\partial t}\,\mathrm{grad}\,\varphi\,, \tag{5.19}$$

the first of which is a Laplace equation. It is shown in Sect. 5.11.4 that a gauge transformation exists such that $\varphi = 0$ here. The system (5.19) then reduces to the linear, homogeneous wave equation for the vector potential $\mathbf{A} = \mathbf{A}(\mathbf{r}, t)$,

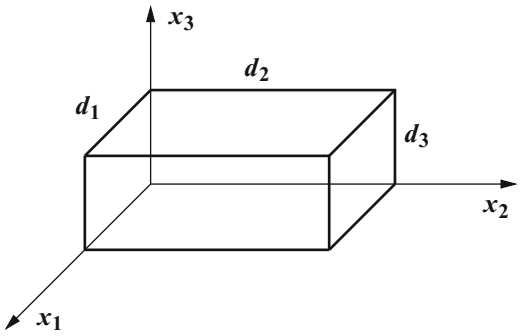


Fig. 5.1 The domain used for the expansion of the vector potential into a Fourier series (Sect. 5.5)

$$\nabla^2 \mathbf{A} - \frac{1}{c^2} \frac{\partial^2 \mathbf{A}}{\partial t^2} = 0\,. \tag{5.20}$$

As the vector potential is defined within a finite volume and has a finite module as well, one can expand it into the Fourier series

$$\mathbf{A} = \sum_{\mathbf{k}} \mathbf{a}_{\mathbf{k}} \exp(\mathrm{i}\,\mathbf{k}\cdot\mathbf{r})\,, \qquad \mathbf{a}_{\mathbf{k}} = \frac{1}{V}\int_V \mathbf{A} \exp(-\mathrm{i}\,\mathbf{k}\cdot\mathbf{r})\,\mathrm{d}V\,, \tag{5.21}$$

where $\mathbf{a}_{\mathbf{k}} = \mathbf{a}(\mathbf{k}, t)$ is complex and the *wave vector* $\mathbf{k}$ is given by

$$\mathbf{k} = n_1 \frac{2\pi}{d_1}\,\mathbf{i}_1 + n_2 \frac{2\pi}{d_2}\,\mathbf{i}_2 + n_3 \frac{2\pi}{d_3}\,\mathbf{i}_3\,, \qquad n_i = 0, \pm 1, \pm 2, \ldots \tag{5.22}$$

The symbol $\sum_{\mathbf{k}}$ indicates a triple sum over all integers n_1, n_2, n_3. The definition of $\mathbf{a}_{\mathbf{k}}$ yields

$$\mathbf{a}_{-\mathbf{k}} = \mathbf{a}_{\mathbf{k}}^*\,, \qquad \mathbf{a}_0 = \frac{1}{V}\int_V \mathbf{A}\,\mathrm{d}V\,, \tag{5.23}$$

with $\mathbf{a}_0$ real. Applying Coulomb's gauge $\operatorname{div}\mathbf{A} = 0$ to the expansion (5.21) provides

$$\operatorname{div}\mathbf{A} = \sum_{\mathbf{k}} \sum_{m=1}^{3} a_{\mathbf{k}m}\,\mathrm{i}\,k_m \exp(\mathrm{i}\,\mathbf{k}\cdot\mathbf{r}) = \mathrm{i} \sum_{\mathbf{k}} \mathbf{a}_{\mathbf{k}}\cdot\mathbf{k} \exp(\mathrm{i}\,\mathbf{k}\cdot\mathbf{r}) = 0\,, \tag{5.24}$$

that is, a linear combination of functions of $\mathbf{r}$. As such functions are linearly independent from each other, (5.24) vanishes only if the coefficients vanish, so it is $\mathbf{a}_{\mathbf{k}}\cdot\mathbf{k} = 0$. Replacing $\mathbf{k}$ with $-\mathbf{k}$ and using (5.23) show that $\mathbf{a}_{-\mathbf{k}}\cdot\mathbf{k} = \mathbf{a}_{\mathbf{k}}^*\cdot\mathbf{k} = 0$. In conclusion, $\mathbf{a}_{\mathbf{k}}$ has no components in the direction of $\mathbf{k}$, namely, it has only two independent (complex) components that lie on the plane normal to $\mathbf{k}$: letting $\mathbf{e}_1$, $\mathbf{e}_2$ be unit vectors belonging to such a plane and normal to each other, one has

$$\mathbf{a}_\mathbf{k} = \mathbf{a}_\mathbf{k} \cdot \mathbf{e}_1 \, \mathbf{e}_1 + \mathbf{a}_\mathbf{k} \cdot \mathbf{e}_2 \, \mathbf{e}_2 \ . \tag{5.25}$$

Clearly the reasoning above does not apply to $\mathbf{a}_0$; however, it is shown below that eventually this term does not contribute to the fields. The Fourier series (5.21) is now inserted into the wave equation (5.20), whose two summands become

$$\nabla^2 \mathbf{A} = \sum_\mathbf{k} \mathbf{a}_\mathbf{k} \sum_{m=1}^{3} (\mathrm{i}\, k_m)^2 \, \exp(\mathrm{i}\, \mathbf{k} \cdot \mathbf{r}) = - \sum_\mathbf{k} \mathbf{a}_\mathbf{k} \, k^2 \, \exp(\mathrm{i}\mathbf{k} \cdot \mathbf{r}) \,, \tag{5.26}$$

$$- \frac{1}{c^2} \frac{\partial^2 \mathbf{A}}{\partial t^2} = - \frac{1}{c^2} \sum_\mathbf{k} \ddot{\mathbf{a}}_\mathbf{k} \, \exp(\mathrm{i}\, \mathbf{k} \cdot \mathbf{r}) \,. \tag{5.27}$$

Adding up yields $\sum_\mathbf{k} \left(\ddot{\mathbf{a}}_\mathbf{k} + c^2 \, k^2 \, \mathbf{a}_\mathbf{k} \right) \exp(\mathrm{i}\, \mathbf{k} \cdot \mathbf{r}) = 0$ whence, using the same reasoning as that used for discussing (5.24),

$$\ddot{\mathbf{a}}_\mathbf{k} + \omega^2 \, \mathbf{a}_\mathbf{k} = 0 \,, \qquad \omega(\mathbf{k}) = c \, k \geq 0 \,, \qquad \omega(-\mathbf{k}) = \omega(\mathbf{k}) \,. \tag{5.28}$$

The case $\mathbf{k} = 0$ yields $\ddot{\mathbf{a}}_0 = 0$ whence $\mathbf{a}_0(t) = \mathbf{a}_0(t = 0) + \dot{\mathbf{a}}_0(t = 0)\, t$. The constant $\dot{\mathbf{a}}_0(t = 0)$ must be set to zero to prevent $\mathbf{a}_0$ from diverging. When $\mathbf{k} \neq 0$ the solution of (5.28) is readily found to be $\mathbf{a}_\mathbf{k}(t) = \mathbf{c}_\mathbf{k} \, \exp(-\mathrm{i}\, \omega t) + \mathbf{c}'_\mathbf{k} \, \exp(\mathrm{i}\, \omega t)$, where the complex vectors $\mathbf{c}_\mathbf{k}$, $\mathbf{c}'_\mathbf{k}$ depend on $\mathbf{k}$ only and lie on the plane normal to it. Using the first relation in (5.23) yields $\mathbf{c}'_\mathbf{k} = \mathbf{c}^*_{-\mathbf{k}}$ and, finally,

$$\mathbf{a}_\mathbf{k} = \mathbf{s}_\mathbf{k} + \mathbf{s}^*_{-\mathbf{k}} \,, \qquad \mathbf{s}_\mathbf{k}(t) = \mathbf{c}_\mathbf{k} \, \exp(-\mathrm{i}\, \omega \, t) \,, \qquad \mathbf{k} \neq 0 \,. \tag{5.29}$$

Thanks to (5.29) one reconstructs the vector potential $\mathbf{A}$ in a form that shows its dependence on space and time explicitly. To this purpose one notes that the sum (5.21) contains all possible combinations of indices n_1, n_2, n_3, so that a summand corresponding to $\mathbf{k}$ is paired with another summand corresponding to $-\mathbf{k}$. One can then rearrange (5.21) as $\mathbf{A} = (1/2) \sum_\mathbf{k} [\mathbf{a}_\mathbf{k} \, \exp(\mathrm{i}\, \mathbf{k} \cdot \mathbf{r}) + \mathbf{a}_{-\mathbf{k}} \, \exp(-\mathrm{i}\, \mathbf{k} \cdot \mathbf{r})]$, where the factor $1/2$ is introduced to eliminate a double counting. Using (5.29), and remembering from (5.28) that $\omega(\mathbf{k})$ is even render $\mathbf{A}$ as a sum of real terms,

$$\mathbf{A} = \sum_\mathbf{k} \Re \left\{ \mathbf{c}_\mathbf{k} \, \exp[\mathrm{i}\, (\mathbf{k} \cdot \mathbf{r} - \omega \, t)] + \mathbf{c}^*_{-\mathbf{k}} \, \exp[\mathrm{i}\, (\mathbf{k} \cdot \mathbf{r} + \omega \, t)] \right\} \,, \tag{5.30}$$

The summands of (5.30) corresponding to $\mathbf{k}$ and $-\mathbf{k}$ describe two plane and monochromatic waves that propagate in the $\mathbf{k}$ and $-\mathbf{k}$ direction, respectively. The two waves together form a *mode* of the electromagnetic field, whose angular frequency is $\omega = c \, k$. The summands corresponding to $\mathbf{k} = 0$ yield the real constant $\mathbf{c}_0 + \mathbf{c}^*_0 = \mathbf{a}_0$. Finally, the $\mathbf{E}$ and $\mathbf{B}$ fields are found by introducing the expansion of $\mathbf{A}$ into (4.26) after letting $\varphi = 0$. For this calculation it is convenient to use the form of the expansion bearing the factor $1/2$ introduced above: from the definition (5.29) of $\mathbf{s}_\mathbf{k}$ and the first identity in (A.35) one finds

$$\mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t} = \frac{1}{2} \sum_{\mathbf{k}} \mathrm{i}\,\omega \left[\left(\mathbf{s}_{\mathbf{k}} - \mathbf{s}^*_{-\mathbf{k}} \right) \exp(\mathrm{i}\,\mathbf{k} \cdot \mathbf{r}) + \left(\mathbf{s}_{-\mathbf{k}} - \mathbf{s}^*_{\mathbf{k}} \right) \exp(-\mathrm{i}\,\mathbf{k} \cdot \mathbf{r}) \right] , \tag{5.31}$$

$$\mathbf{B} = \mathrm{rot}\,\mathbf{A} = \frac{1}{2} \sum_{\mathbf{k}} \mathrm{i}\,\mathbf{k} \wedge \left[\left(\mathbf{s}_{\mathbf{k}} + \mathbf{s}^*_{-\mathbf{k}} \right) \exp(\mathrm{i}\,\mathbf{k} \cdot \mathbf{r}) - \left(\mathbf{s}_{-\mathbf{k}} + \mathbf{s}^*_{\mathbf{k}} \right) \exp(-\mathrm{i}\,\mathbf{k} \cdot \mathbf{r}) \right] . \tag{5.32}$$

As anticipated, the constant term $\mathbf{a}_0$ does not contribute to the fields. Also, due to the second relation in (5.29), the vectors $\mathbf{s}_{\mathbf{k}}$, $\mathbf{s}^*_{-\mathbf{k}}$, $\mathbf{s}_{-\mathbf{k}}$, and $\mathbf{s}^*_{\mathbf{k}}$ lie over the plane normal to $\mathbf{k}$. Due to (5.31, 5.32) the $\mathbf{E}$ and $\mathbf{B}$ fields lie on the same plane as well, namely, they have no component in the propagation direction. For this reason they are called *transversal*.

5.6 Energy of the Electromagnetic Field in Terms of Modes

The expressions of the $\mathbf{E}$, $\mathbf{B}$ fields within a finite volume V free of charges have been calculated in Sect. 5.5 as superpositions of modes, each of them associated with a wave vector $\mathbf{k}$ and an angular frequency $\omega = c\,k$. Basing upon such expressions one is able to determine the electromagnetic energy within V in terms of modes. To this purpose one calculates from (5.31, 5.32) the squares $E^2 = \mathbf{E} \cdot \mathbf{E}$ and $B^2 = \mathbf{B} \cdot \mathbf{B}$, inserts the resulting expression into the second relation of (5.10) to obtain the energy per unit volume, and, finally, integrates the latter over V (last relation in (5.12)). Letting $\mathbf{I}_{\mathbf{k}}$ be the quantity enclosed within brackets in (5.31), it is $E^2 = -(1/4) \sum_{\mathbf{k}} \sum_{\mathbf{k}'} \omega\,\omega'\,\mathbf{I}_{\mathbf{k}} \cdot \mathbf{I}_{\mathbf{k}'}$, where $\omega' = c\,k'$. The integration over V avails itself of the integrals (C.130) to yield

$$-\frac{1}{4} \sum_{\mathbf{k}'} \omega\,\omega' \int_V \mathbf{I}_{\mathbf{k}} \cdot \mathbf{I}_{\mathbf{k}'}\,\mathrm{d}V = V\,\omega^2 \left(\mathbf{s}_{\mathbf{k}} - \mathbf{s}^*_{-\mathbf{k}} \right) \cdot \left(\mathbf{s}^*_{\mathbf{k}} - \mathbf{s}_{-\mathbf{k}} \right) , \tag{5.33}$$

so that the part of the electromagnetic energy deriving from $\mathbf{E}$ reads

$$\int_V \frac{\varepsilon_0}{2} E^2\,\mathrm{d}V = \frac{\varepsilon_0}{2} V \sum_{\mathbf{k}} \omega^2 \left(\mathbf{s}_{\mathbf{k}} - \mathbf{s}^*_{-\mathbf{k}} \right) \cdot \left(\mathbf{s}^*_{\mathbf{k}} - \mathbf{s}_{-\mathbf{k}} \right) . \tag{5.34}$$

By the same token one lets $\mathbf{Y}_{\mathbf{k}}$ be the quantity enclosed within brackets in (5.32), whence $B^2 = -(1/4) \sum_{\mathbf{k}} \sum_{\mathbf{k}'} (\mathbf{k} \wedge \mathbf{Y}_{\mathbf{k}}) \cdot (\mathbf{k}' \wedge \mathbf{Y}_{\mathbf{k}'})$ and

$$-\frac{1}{4} \sum_{\mathbf{k}'} \int_V (\mathbf{k} \wedge \mathbf{Y}_{\mathbf{k}}) \cdot (\mathbf{k}' \wedge \mathbf{Y}_{\mathbf{k}'})\,\mathrm{d}V = V \left[\mathbf{k} \wedge \left(\mathbf{s}_{\mathbf{k}} + \mathbf{s}^*_{-\mathbf{k}} \right) \right] \cdot \left[\mathbf{k} \wedge \left(\mathbf{s}_{-\mathbf{k}} + \mathbf{s}^*_{\mathbf{k}} \right) \right] . \tag{5.35}$$

The expression at the right-hand side of (5.35) simplifies because, due to (5.29), $\mathbf{k}$ is normal to the plane where $\mathbf{s}_{\mathbf{k}}$, $\mathbf{s}^*_{-\mathbf{k}}$, $\mathbf{s}_{-\mathbf{k}}$, and $\mathbf{s}^*_{\mathbf{k}}$ lie, so that $[\mathbf{k} \wedge (\mathbf{s}_{\mathbf{k}} + \mathbf{s}^*_{-\mathbf{k}})] \cdot [\mathbf{k} \wedge (\mathbf{s}_{-\mathbf{k}} + \mathbf{s}^*_{\mathbf{k}})] = k^2 \left(\mathbf{s}_{\mathbf{k}} + \mathbf{s}^*_{-\mathbf{k}} \right) \cdot \left(\mathbf{s}_{-\mathbf{k}} + \mathbf{s}^*_{\mathbf{k}} \right)$. Using the relation $k^2 = \omega^2/c^2 = \varepsilon_0\,\mu_0\,\omega^2$

yields the part of the electromagnetic energy deriving from $\mathbf{H} = \mathbf{B}/\mu_0$,

$$\int_V \frac{1}{2\,\mu_0} B^2 \,\mathrm{d}V = \frac{\varepsilon_0}{2} V \sum_{\mathbf{k}} \omega^2 \,(\mathbf{s}_{\mathbf{k}} + \mathbf{s}^*_{-\mathbf{k}}) \cdot (\mathbf{s}_{-\mathbf{k}} + \mathbf{s}^*_{\mathbf{k}}) \,. \tag{5.36}$$

Adding up (5.34) and (5.36) one finally obtains

$$W_{\mathrm{em}} = \varepsilon_0 \, V \sum_{\mathbf{k}} \omega^2 \,\left(\mathbf{s}_{\mathbf{k}} \cdot \mathbf{s}^*_{\mathbf{k}} + \mathbf{s}_{-\mathbf{k}} \cdot \mathbf{s}^*_{-\mathbf{k}}\right) = 2\,\varepsilon_0 \, V \sum_{\mathbf{k}} \omega^2 \,\mathbf{s}_{\mathbf{k}} \cdot \mathbf{s}^*_{\mathbf{k}} \,. \tag{5.37}$$

This result shows that the energy of the electromagnetic field within V is the sum of individual contributions, each associated with a wave vector $\mathbf{k}$ through the complex vector $\mathbf{s}_{\mathbf{k}}$. As the latter lies on the plane normal to $\mathbf{k}$, it is expressed in terms of two scalar components as $\mathbf{s}_{\mathbf{k}} = s_{\mathbf{k}1}\,\mathbf{e}_1 + s_{\mathbf{k}2}\,\mathbf{e}_2$. Such components are related to the polarization of the electromagnetic field over the plane [15, Sect. 1.4.2]. These considerations allow one to count the number of indices that are involved in the representation (5.37) of W_{em}: in fact, the set of $\mathbf{k}$ vectors is described by the triple infinity of indices $n_1, n_2, n_3 \in \mathbf{Z}$ that appear in (5.22), while the two scalar components require another index $\sigma = 1, 2$. The $\mathbf{s}_{\mathbf{k}}$ vectors describe the electromagnetic field through (5.31) and (5.32), hence one may think of each scalar component $s_{\mathbf{k}\sigma}$ as a degree of freedom of the field; the counting outlined above shows that the number of degrees of freedom is $2 \times \mathbf{Z}^3$. In turn, each degree of freedom is made of a real and an imaginary part, $s_{\mathbf{k}\sigma} = R_{\mathbf{k}\sigma} + \mathrm{i}\, I_{\mathbf{k}\sigma}$, thus yielding

$$W_{\mathrm{em}} = \sum_{\mathbf{k}\sigma} W_{\mathbf{k}\sigma} \,, \qquad W_{\mathbf{k}\sigma} = 2\,\varepsilon_0 \, V \,\omega^2 \,\left(R^2_{\mathbf{k}\sigma} + I^2_{\mathbf{k}\sigma}\right) \,. \tag{5.38}$$

As $\omega = c\,k$, the mode with $\mathbf{k} = 0$ does not contribute to the energy. In (5.38) it is $R_{\mathbf{k}\sigma} = |c_{\mathbf{k}\sigma}| \,\cos[\mathrm{i}\,\omega\,(t_0 - t)]$, $I_{\mathbf{k}\sigma} = |c_{\mathbf{k}\sigma}| \,\sin[\mathrm{i}\,\omega\,(t_0 - t)]$, where the polar form $|c_{\mathbf{k}\sigma}| \,\exp(\mathrm{i}\,\omega\,t_0)$ has been used for $c_{\mathbf{k}\sigma}$. One notes that each summand in (5.38) is related to a single degree of freedom and has a form similar to the Hamiltonian function of the linear harmonic oscillator discussed in Sect. 3.3. To further pursue the analogy one defines the new pair

$$q_{\mathbf{k}\sigma}(t) = 2\,\sqrt{\varepsilon_0 \, V}\, R_{\mathbf{k}\sigma} \,, \qquad p_{\mathbf{k}\sigma}(t) = 2\,\omega\,\sqrt{\varepsilon_0 \, V}\, I_{\mathbf{k}\sigma} \,, \tag{5.39}$$

whence

$$W_{\mathbf{k}\sigma} = \frac{1}{2}\,\left(p^2_{\mathbf{k}\sigma} + \omega^2 \, q^2_{\mathbf{k}\sigma}\right) \,, \qquad \frac{\partial W_{\mathbf{k}\sigma}}{\partial p_{\mathbf{k}\sigma}} = p_{\mathbf{k}\sigma} \,, \qquad \frac{\partial W_{\mathbf{k}\sigma}}{\partial q_{\mathbf{k}\sigma}} = \omega^2 \, q_{\mathbf{k}\sigma} \,. \tag{5.40}$$

On the other hand, the time dependence of $R_{\mathbf{k}\sigma}, I_{\mathbf{k}\sigma}$ is such that

$$\dot{q}_{\mathbf{k}\sigma} = p_{\mathbf{k}\sigma} = \frac{\partial W_{\mathbf{k}\sigma}}{\partial p_{\mathbf{k}\sigma}} \,, \qquad \dot{p}_{\mathbf{k}\sigma} = -\omega^2 \, q_{\mathbf{k}\sigma} = -\frac{\partial W_{\mathbf{k}\sigma}}{\partial q_{\mathbf{k}\sigma}} \,. \tag{5.41}$$

Comparing (5.41) with (1.42) shows that $q_{\mathbf{k}\sigma}$, $p_{\mathbf{k}\sigma}$ are canonically conjugate variables and $W_{\mathbf{k}\sigma}$ is the Hamiltonian function of the degree of freedom associated with $\mathbf{k}\sigma$. Then, comparing (5.40) with (3.1, 3.2) shows that $W_{\mathbf{k}\sigma}$ is indeed the Hamiltonian function of a linear harmonic oscillator of unit mass.

The energy associated with each degree of freedom is constant in time. In fact, from the second relation in (5.29) one derives $W_{\mathbf{k}\sigma} = 2\,\varepsilon_0\,V\,\omega^2\,\mathbf{c}_{\mathbf{k}\sigma}\cdot\mathbf{c}^*_{\mathbf{k}\sigma}$. The same result can be obtained from the properties of the linear harmonic oscillator (Sect. 3.3). It follows that the total energy W_{em} is conserved. As shown in Sect. 5.11.4 this is due to the periodicity of the Poynting vector (5.11): in fact, the electromagnetic energies the cross per unit time two opposite faces of the boundary of V are the negative of each other.

5.7 Momentum of the Electromagnetic Field in Terms of Modes

It has been shown in Sect. 5.4 that the momentum per unit volume of the electromagnetic field is $\mathbf{S}/c^2 = \mathbf{E}\wedge\mathbf{B}/(\mu_0\,c^2) = \varepsilon_0\mathbf{E}\wedge\mathbf{B}$. Using the symbols defined in Sect. 5.6 one finds $\varepsilon_0\mathbf{E}\wedge\mathbf{B} = -(\varepsilon_0/4)\,\sum_{\mathbf{k}}\sum_{\mathbf{k}'}\omega\,\mathbf{I}_{\mathbf{k}}\wedge(\mathbf{k}'\wedge\mathbf{Y}_{\mathbf{k}'})$ and

$$-\frac{\varepsilon_0}{4}\sum_{\mathbf{k}'}\int_V \omega\,\mathbf{I}_{\mathbf{k}}\wedge(\mathbf{k}'\wedge\mathbf{Y}_{\mathbf{k}'})\,\mathrm{d}V = \frac{\varepsilon_0}{2}\,\omega\,V\,(\mathbf{Z}_{\mathbf{k}}+\mathbf{Z}_{-\mathbf{k}}), \tag{5.42}$$

with $\mathbf{Z}_{\mathbf{k}} = (\mathbf{s}_{\mathbf{k}}-\mathbf{s}^*_{-\mathbf{k}})\wedge[\mathbf{k}\wedge(\mathbf{s}_{-\mathbf{k}}+\mathbf{s}^*_{\mathbf{k}})]$. The expression of $\mathbf{Z}_{\mathbf{k}}$ simplifies because, due to (5.29), $\mathbf{k}$ is normal to the plane where $\mathbf{s}_{\mathbf{k}}$, $\mathbf{s}^*_{-\mathbf{k}}$, $\mathbf{s}_{-\mathbf{k}}$, and $\mathbf{s}^*_{\mathbf{k}}$ lie, so that $\mathbf{Z}_{\mathbf{k}} = \mathbf{k}\,(\mathbf{s}_{\mathbf{k}}-\mathbf{s}^*_{-\mathbf{k}})\cdot(\mathbf{s}_{-\mathbf{k}}+\mathbf{s}^*_{\mathbf{k}})$ and $\mathbf{Z}_{\mathbf{k}}+\mathbf{Z}_{-\mathbf{k}} = 2\,\mathbf{k}\,\mathbf{s}_{\mathbf{k}}\cdot\mathbf{s}^*_{\mathbf{k}}+2\,(-\mathbf{k})\,\mathbf{s}_{-\mathbf{k}}\cdot\mathbf{s}^*_{-\mathbf{k}}$. In conclusion, observing that $\omega\,\mathbf{k} = (\omega^2/c)\,\mathbf{k}/k$,

$$\int_V \frac{\mathbf{S}}{c^2}\,\mathrm{d}V = 2\,\varepsilon_0\,V\sum_{\mathbf{k}}\omega\,\mathbf{s}_{\mathbf{k}}\cdot\mathbf{s}^*_{\mathbf{k}}\,\mathbf{k} = \sum_{\mathbf{k}\sigma}\frac{1}{c}\,W_{\mathbf{k}\sigma}\,\frac{\mathbf{k}}{k} = \frac{1}{2c}\sum_{\mathbf{k}\sigma}\left(p^2_{\mathbf{k}\sigma}+\omega^2 q^2_{\mathbf{k}\sigma}\right)\,\frac{\mathbf{k}}{k}\,, \tag{5.43}$$

where the last two equalities derive from (5.37, 5.38, 5.40). One notes from (5.43) that the momentum of the electromagnetic field is the sum of individual momenta, each related to a single degree of freedom. The modulus of the individual momentum is equal to the energy $W_{\mathbf{k}\sigma}$ pertaining to the same degree of freedom divided by c. The same relation between momentum and energy has been derived in Sect. 3.13.7 with reference to the dynamic relations of Special Relativity. Each summand in (5.43) is constant in time, so the electromagnetic momentum is conserved; as noted in Sects. 5.6,5.11.4 this is due to the periodicity of the Poynting vector.

5.8 Modes of the Electromagnetic Field in an Infinite Domain

The treatment of Sects. 5.5, 5.6, 5.7 is extended to the case of an infinite domain by means of the Fourier transform (Sect. C.2)[1]

$$\mathbf{A} = \iiint_{-\infty}^{+\infty} \mathbf{b}_\mathbf{k} \frac{\exp(\mathrm{i}\,\mathbf{k}\cdot\mathbf{r})}{(2\,\pi)^{3/2}}\,\mathrm{d}^3k\,, \qquad \mathbf{b}_\mathbf{k} = \iiint_{-\infty}^{+\infty} \mathbf{A}\,\frac{\exp(-\mathrm{i}\,\mathbf{k}\cdot\mathbf{r})}{(2\,\pi)^{3/2}}\,\mathrm{d}^3r\,. \tag{5.44}$$

where $\mathbf{b}_\mathbf{k} = \mathbf{b}(\mathbf{k}, t)$ is complex, with $\mathbf{b}_{-\mathbf{k}} = \mathbf{b}_\mathbf{k}^*$, and the components of the wave vector $\mathbf{k}$ are continuous. Relations of the same form as (5.25) and (5.28) hold for $\mathbf{b}_\mathbf{k}$, yielding

$$\mathbf{b}_\mathbf{k} = \tilde{\mathbf{s}}_\mathbf{k} + \tilde{\mathbf{s}}_{-\mathbf{k}}^*\,, \qquad \tilde{\mathbf{s}}_\mathbf{k}(t) = \mathbf{d}_\mathbf{k}\,\exp(-\mathrm{i}\,\omega\,t)\,, \qquad \mathbf{k} \neq 0\,. \tag{5.45}$$

where the complex vector $\mathbf{d}_\mathbf{k}$ depend on $\mathbf{k}$ only and lies on the plane normal to it. Relations similar to (5.30, 5.31, 5.32) hold as well, where $\mathbf{c}_\mathbf{k}$ and $\mathbf{s}_\mathbf{k}$ are replaced with $\mathbf{d}_\mathbf{k}$ and $\tilde{\mathbf{s}}_\mathbf{k}$, respectively, and the sum is suitably replaced with an integral over $\mathbf{k}$. To determine the energy of the electromagnetic field one must integrate over the whole space the energy density w_em. Using (C.56) and the second relation in (5.45) one finds

$$W_\mathrm{em} = \iiint_{-\infty}^{+\infty} w_\mathrm{em}\,\mathrm{d}^3r = 2\,\varepsilon_0 \iiint_{-\infty}^{+\infty} \omega^2\,\mathbf{d}_\mathbf{k}\cdot\mathbf{d}_\mathbf{k}^*\,\mathrm{d}^3k\,. \tag{5.46}$$

It is sometimes useful to consider the frequency distribution of the integrand at the right-hand side of (5.46). For this one converts the integral into spherical coordinates k, ϑ, γ and uses the relation $k = \omega/c = 2\,\pi\,\nu/c$ to obtain $\mathbf{d}_\mathbf{k} = \mathbf{d}(\nu, \vartheta, \gamma)$; then, from (B.3),

$$W_\mathrm{em} = \int_{-\infty}^{+\infty} U_\mathrm{em}(\nu)\,\mathrm{d}\nu\,, \quad U_\mathrm{em} = \frac{2\,\varepsilon_0}{c^2}\,(2\,\pi\,\nu)^4 \int_0^\pi \int_0^{2\,\pi} |\mathbf{d}|^2\,\sin\vartheta\,\mathrm{d}\vartheta\,\mathrm{d}\gamma\,, \tag{5.47}$$

where U_em (whose units are J s) is called *spectral energy* of the electromagnetic field. By a similar procedure one finds the total momentum that reads

$$\iiint_{-\infty}^{+\infty} \frac{\mathbf{S}}{c^2}\,\mathrm{d}^3r = 2\,\varepsilon_0 \iiint_{-\infty}^{+\infty} \omega\,|\mathbf{d}|^2\,\mathbf{k}\,\mathrm{d}^3k\,. \tag{5.48}$$

[1] For the existence of (5.44) it is implied that the three-dimensional equivalent of condition (C.19) holds.

5.9 Eikonal Equation

Consider the case of a monochromatic electromagnetic field with angular frequency ω. For the calculation in hand it is convenient to consider the Maxwell equations in complex form; specifically, $\operatorname{rot}\mathbf{H} = \partial\mathbf{D}/\partial t$ yields, *in vacuo*,

$$\Re\left[(\operatorname{rot}\mathbf{H}_c + \mathrm{i}\,\omega\,\varepsilon_0\,\mathbf{E}_c)\,\exp(-\mathrm{i}\,\omega\,t)\right] = 0\,, \tag{5.49}$$

while $\operatorname{rot}\mathbf{E} = -\mathbf{B}/\partial t$ yields

$$\Re\left[(\operatorname{rot}\mathbf{E}_c - \mathrm{i}\,\omega\,\mu_0\,\mathbf{H}_c)\,\exp(-\mathrm{i}\,\omega\,t)\right] = 0\,. \tag{5.50}$$

The solution of (5.49, 5.50) has the form $\mathbf{E}_c = \mathbf{E}_{c0}\,\exp(\mathrm{i}\,\mathbf{k}\cdot\mathbf{r})$, $\mathbf{H}_c = \mathbf{H}_{c0}\,\exp(\mathrm{i}\,\mathbf{k}\cdot\mathbf{r})$, with $\mathbf{E}_{c0}\,, \mathbf{H}_{c0} = \text{const}$, i.e., a planar wave propagating along the direction of $\mathbf{k}$. In a nonuniform medium it is $\varepsilon = \varepsilon(\mathbf{r})$, $\mu = \mu(\mathbf{r})$, and the form of the solution differs from the planar wave. The latter can tentatively be generalized as

$$\mathbf{E}_c = \mathbf{E}_{c0}\,(\mathbf{r})\,\exp[\mathrm{i}\,k\,S(\mathbf{r})]\,, \qquad \mathbf{H}_c = \mathbf{H}_{c0}(\mathbf{r})\,\exp[\mathrm{i}\,k\,S(\mathbf{r})]\,, \tag{5.51}$$

with $k = \omega\sqrt{\varepsilon_0\,\mu_0} = \omega/c$. Function $k\,S$ is called *eikonal* ($[S] = \text{cm}$). Replacing (5.51) into (5.49, 5.50) and using the first identity in (A.35) yield

$$\operatorname{grad} S \wedge \mathbf{H}_{c0} + c\,\varepsilon\,\mathbf{E}_{c0} = -\frac{c}{\mathrm{i}\,\omega}\,\operatorname{rot}\mathbf{H}_{c0}\,, \tag{5.52}$$

$$\operatorname{grad} S \wedge \mathbf{E}_{c0} - c\,\mu\,\mathbf{H}_{c0} = -\frac{c}{\mathrm{i}\,\omega}\,\operatorname{rot}\mathbf{E}_{c0}\,. \tag{5.53}$$

Now it is assumed that ω is large enough to make the right-hand side of (5.52, 5.53) negligible; in this case $\operatorname{grad} S$, $\mathbf{E}_{c0}$, $\mathbf{H}_{c0}$ become normal to each other. Vector multiplying (5.52) by $\operatorname{grad} S$ and using (5.53) then yield

$$\operatorname{grad} S \wedge (\operatorname{grad} S \wedge \mathbf{H}_{c0}) + \frac{\varepsilon\,\mu}{\varepsilon_0\,\mu_0}\,\mathbf{H}_{c0} = 0\,. \tag{5.54}$$

Remembering that $c = 1/\sqrt{\varepsilon_0\,\mu_0}$ one defines the *phase velocity*, *refraction index*, and *wavelength* of the medium as

$$u_f(\mathbf{r}) = \frac{1}{\sqrt{\varepsilon\,\mu}}\,, \qquad n(\mathbf{r}) = \frac{c}{u_f}\,, \qquad \lambda(\mathbf{r}) = \frac{u_f}{\nu}\,, \tag{5.55}$$

respectively, so that $\varepsilon\,\mu/(\varepsilon_0\,\mu_0) = n^2$. Using the first identity in (A.33) and remembering that $\operatorname{grad} S \cdot \mathbf{H}_{c0} = 0$ transforms (5.54) into $(|\operatorname{grad} S|^2 - n^2)\,\mathbf{H}_{c0} = 0$. As $\mathbf{H}_{c0} \neq 0$ it follows

$$|\operatorname{grad} S|^2 = n^2\,, \qquad n = n(\mathbf{r})\,, \tag{5.56}$$

that is, a partial-differential equation, called *eikonal equation*, in the unknown S. The equation has been derived in the hypothesis that $\omega = 2\pi\nu$ is large, hence $\lambda = u_f/\nu$ is small; it is shown below that this condition is also described by stating that $\mathbf{E}_{c0}(\mathbf{r})$, $\mathbf{H}_{c0}(\mathbf{r})$, and $S(\mathbf{r})$ vary little over a distance of the order of λ.

The form of (5.51) is such that $S(\mathbf{r}) = \text{const}$ defines the constant-phase surface (the same concept has been encountered in Sect. 2.5 for the case of a system of particles). It follows that the normal direction at each point $\mathbf{r}$ of the surface is that of $\operatorname{grad} S$. Let $\mathbf{t} = \mathrm{d}\mathbf{r}/\mathrm{d}s$ be the unit vector parallel to $\operatorname{grad} S$ in the direction of increasing S. A *ray* is defined as the envelope of the $\mathbf{t}$ vectors, taken starting from a point A in a given direction. The description of rays obtained through the approximation of the eikonal equation is called *Geometrical Optics*.

The eikonal equation (5.56) can be given a different form by observing that from the definition of $\mathbf{t}$ it follows $\operatorname{grad} S = n\,\mathbf{t}$ and $\mathbf{t}\cdot\operatorname{grad} S = \mathrm{d}S/\mathrm{d}s = n$, whence

$$\operatorname{grad} n = \operatorname{grad}\frac{\mathrm{d}S}{\mathrm{d}s} = \frac{\mathrm{d}\operatorname{grad} S}{\mathrm{d}s} = \frac{\mathrm{d}\,(n\mathbf{t})}{\mathrm{d}s} = \frac{\mathrm{d}}{\mathrm{d}s}\left(n\,\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}s}\right). \tag{5.57}$$

This form of the eikonal equation is more often used. It shows that the equation is of the second order in the unknown function $\mathbf{r}(s)$, where $\mathbf{r}$ is the point of the ray corresponding to the curvilinear abscissa s along the ray itself. The equation's coefficient and data are given by the refraction index n. As the equation is of the second order, two boundary conditions are necessary to completely define the solution; for instance, the value of $\mathbf{r}(s = 0)$ corresponding to the initial point A, and the direction $\mathbf{t} = \mathrm{d}\mathbf{r}/\mathrm{d}s$ of the ray at the same point. Remembering that $\mathrm{d}\mathbf{t}/\mathrm{d}s = \mathbf{n}/\varrho_c$, where ϱ_c is the *curvature radius* of the ray at $\mathbf{r}$, and $\mathbf{n}$ the *principal normal unit vector*, the eikonal equation may also be recast as $\operatorname{grad} n = (\mathrm{d}n/\mathrm{d}s)\,\mathbf{t} + (n/\varrho_c)\,\mathbf{n}$. Using the curvature radius one can specify in a quantitative manner the approximation upon which the eikonal equation is based; in fact, for the approximation to hold it is necessary that the electromagnetic wave can be considered planar, namely, that its amplitude and direction do not significantly change over a distance of the order of λ. This happens if at each point $\mathbf{r}$ along the ray it is $\varrho_c \gg \lambda$.

5.10 Fermat Principle

It is worth investigating whether the eikonal equation (5.57) worked out in Sect. 5.9 is derivable from a variational principle. In fact it is shown below that the *Fermat* (or *least time*) principle holds, stating that if A and B are two different points belonging to a ray, the natural ray (that is, the actual path followed by the radiation between the given points) is the one that minimizes the time $\int_{AB} \mathrm{d}t$. The principle thus reads

$$\delta \int_{AB} \mathrm{d}t = 0\,, \tag{5.58}$$

where the integral is carried out along the trajectory. The analysis is similar to that carried out in Sect. 2.7 with reference to the Maupertuis principle.

Using the relations (5.55) and observing that $\mathrm{d}t = \mathrm{d}s/u_f = n\,\mathrm{d}s/c$ transform (5.58) into $\delta \int_{AB} n\,\mathrm{d}s = 0$. Introducing a parametric description $\mathbf{r} = \mathbf{r}(\xi)$ of the ray, with $\xi = a$ when $\mathbf{r} = A$ and $\xi = b$ when $\mathbf{r} = B$, yields

$$\int_{AB} n\,\mathrm{d}s = \int_a^b g\,\mathrm{d}\xi\,, \qquad g = n\,\frac{\mathrm{d}s}{\mathrm{d}\xi} = n(x_1, x_2, x_3)\,\sqrt{\dot{x}_1^2 + \dot{x}_2^2 + \dot{x}_3^2}\,, \tag{5.59}$$

$$\frac{\partial g}{\partial \dot{x}_i} = n\,\frac{2\dot{x}_i}{2\,\mathrm{d}s/\mathrm{d}\xi} = n\,\frac{\mathrm{d}x_i}{\mathrm{d}s}\,, \qquad \frac{\partial g}{\partial x_i} = \frac{\partial n}{\partial x_i}\,\frac{\mathrm{d}s}{\mathrm{d}\xi}\,. \tag{5.60}$$

Remembering (1.7), the Euler equation for the ith coordinate reads

$$\frac{\mathrm{d}}{\mathrm{d}\xi}\left(n\,\frac{\mathrm{d}x_i}{\mathrm{d}s}\right) = \frac{\partial n}{\partial x_i}\,\frac{\mathrm{d}s}{\mathrm{d}\xi}\,, \qquad i = 1, 2, 3\,, \tag{5.61}$$

whence

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(n\,\frac{\mathrm{d}x_i}{\mathrm{d}s}\right) = \frac{\partial n}{\partial x_i}\,. \tag{5.62}$$

As (5.62) is the ith component of the eikonal equation (5.57), such an equation is indeed derivable from the variational principle (5.58). Some comments about the formal analogy between the Maupertuis and Fermat principles are made in Sect. 5.11.6.

5.11 Complements

5.11.1 Field Generated by a Point-Like Charge

The Liénard and Wiechert expressions (5.6, 5.7) provide the potentials generated in $\mathbf{r}$ at time t by a point-like charge that follows a trajectory $\mathbf{s}$. More specifically, if $\mathbf{s} = \mathbf{s}(\tau)$ is the position occupied by the charge at the instant τ, and $\mathbf{r}$ is the position where the potentials produced by the charge are detected at time $t > \tau$, the relation (5.4) holds, namely, $|\mathbf{r} - \mathbf{s}(\tau)| = c\,(t - \tau)$, that links the spatial coordinates with the time instants. Letting

$$\mathbf{g} = \mathbf{r} - \mathbf{s}(\tau)\,, \qquad g = |\mathbf{g}|\,, \qquad \mathbf{u} = \frac{\mathrm{d}\mathbf{s}}{\mathrm{d}\tau}\,, \qquad \dot{\mathbf{u}} = \frac{\mathrm{d}\mathbf{u}}{\mathrm{d}\tau}\,, \tag{5.63}$$

the fields **E**, **B** are determined by applying (4.26) to (5.6, 5.7), which amounts to calculating the derivatives with respect to t and the components of **r**. This is somewhat complicate because (4.26) introduces a relation of the form $\tau = \tau(\mathbf{r}, t)$, so that $\varphi = \varphi(\mathbf{r}, \tau(\mathbf{r}, t))$ and $\mathbf{A} = \mathbf{A}(\mathbf{r}, \tau(\mathbf{r}, t))$. It is therefore convenient to calculate some intermediate steps first. To this purpose, (5.4) is recast in implicit form as

$$\sigma(x_1, x_2, x_3, t, \tau) = \left[\sum_{i=1}^{3} (x_i - s_i(\tau))^2\right]^{1/2} + c\,(\tau - t) = 0\,, \tag{5.64}$$

whence $\operatorname{grad}\sigma = \mathbf{g}/g$, $\partial\sigma/\partial t = -c$, $\partial\sigma/\partial\tau = c - \mathbf{u}\cdot\mathbf{g}/g$. The gradient indicates derivation with respect to the x_i coordinates. The differentiation rule of the implicit functions then yields

$$\frac{\partial\tau}{\partial t} = -\frac{\partial\sigma/\partial t}{\partial\sigma/\partial\tau} = \frac{c}{c - \mathbf{u}\cdot\mathbf{g}/g}\,. \tag{5.65}$$

Basing on (5.64, 5.65) and following the calculation scheme reported, e.g., in [120, Chap. 6] and in Probs. 5.3–5.9, one obtains

$$\mathbf{E} = \frac{e\,(\partial\tau/\partial t)^3}{4\,\pi\,\varepsilon_0\,g^3}\left\{\left(1 - \frac{u^2}{c^2}\right)\left(\mathbf{g} - g\,\frac{\mathbf{u}}{c}\right) + \mathbf{g}\wedge\left[\left(\mathbf{g} - g\,\frac{\mathbf{u}}{c}\right)\wedge\frac{\dot{\mathbf{u}}}{c^2}\right]\right\}\,, \tag{5.66}$$

$$\mathbf{B} = \frac{\mathbf{g}}{g}\wedge\frac{\mathbf{E}}{c}\,. \tag{5.67}$$

This result shows that **E** and **B** are the sum of two terms, the first of which decays at infinity like g^{-2}, while the second decays like g^{-1}. The latter term differs from zero only if the charge is accelerated ($\dot{\mathbf{u}} \neq 0$); its contribution is called *radiation field*. Also, **E** and **B** are orthogonal to each other, while **g** is orthogonal to **B** but not to **E**; however, if g is large enough to make the second term in (5.66) dominant, **g** becomes orthogonal also to **E** and (5.67) yields $B = E/c$. In the case $\mathbf{u} = 0$ the relations (5.66, 5.67) simplify to

$$\mathbf{E} = \frac{e}{4\,\pi\,\varepsilon_0\,g^2}\,\frac{\mathbf{g}}{g}\,, \qquad \mathbf{B} = 0\,, \tag{5.68}$$

that hold approximately also for $\mathbf{u} = \text{const}$, $u/c \ll 1$.

5.11.2 Power Radiated by a Point-Like Charge

The expressions of the **E** and **B** fields worked out in Sect. 5.11.1 are readily exploited to determine the power radiated by a point-like charge. Remembering the results of

Sect. 5.3, it suffices to integrate the Poynting vector over a surface Σ surrounding the charge. Introducing (5.67) into the definition (5.11) of Poynting's vector and using the first identity in (A.33) yield

$$\mathbf{S} = \frac{1}{\mu_0\, c}\, \mathbf{E} \wedge \left(\frac{\mathbf{g}}{g} \wedge \mathbf{E} \right) = \frac{\varepsilon_0\, c}{g}\, \left(E^2\, \mathbf{g} - \mathbf{E} \cdot \mathbf{g}\, \mathbf{E} \right) . \tag{5.69}$$

The case $\dot{\mathbf{u}} \neq 0$, $u/c \ll 1$ is considered, which is typical of a bound particle. As the surface Σ can be chosen arbitrarily, it is convenient to select it at a large distance from the charge in order to make the second term in (5.66) dominant and $\mathbf{E}$ practically normal to $\mathbf{g}$. This simplifies (5.66) and (5.69) to

$$\mathbf{S} \simeq \varepsilon_0\, c\, E^2\, \frac{\mathbf{g}}{g} , \qquad \mathbf{E} \simeq \frac{e}{4\,\pi\,\varepsilon_0\, g} \left(\frac{\mathbf{g}}{g} \cdot \frac{\dot{\mathbf{u}}}{c^2}\, \frac{\mathbf{g}}{g} - \frac{\dot{\mathbf{u}}}{c^2} \right) , \tag{5.70}$$

where the first identity in (A.33) has been used. Letting ϑ be the angle between $\mathbf{g}$ and $\dot{\mathbf{u}}$, one combines the two expressions in (5.70) to find

$$\mathbf{S} \simeq \varepsilon_0\, c\, \mathbf{E} \cdot \mathbf{E}\, \frac{\mathbf{g}}{g} = \frac{1}{4\,\pi\,\varepsilon_0}\, \frac{e^2\, \dot{u}^2}{4\,\pi\, c^3}\, \frac{\sin^2 \vartheta}{g^2}\, \frac{\mathbf{g}}{g} . \tag{5.71}$$

To proceed one chooses for Σ a spherical surface centered at $\mathbf{s}(\tau)$ and shifts the origin to its center. This yields $\mathbf{g} = \mathbf{r}$ at time τ, whence the unit vector normal to Σ becomes $\mathbf{n} = \mathbf{g}/g$. The radiation emitted by the charge reaches Σ at a later time $t = \tau + g/c$; however, thanks to the hypothesis $u/c \ll 1$, during the interval $t - \tau$ the charge moves little with respect to center of the sphere. For this reason, the surface integral can be calculated by keeping the charge fixed in the center of the spherical surface, so that the integral $\int_\Sigma (\sin^2 \vartheta / g^2)\, \mathrm{d}\Sigma$ must be evaluated with $g =$ const. Such an integral is easily found to equal $8\,\pi/3$: first, one turns to spherical coordinates and expresses the volume element as $J\, \mathrm{d}\vartheta\, \mathrm{d}\varphi\, \mathrm{d}g = \mathrm{d}\Sigma\, \mathrm{d}g$; then, one finds from (B.3) the ratio $\mathrm{d}\Sigma/g^2 = \sin\vartheta\, \mathrm{d}\theta\, \mathrm{d}\varphi$ and replaces it in the integral. In conclusion, combining the above result with (5.12, 5.71),

$$-\, \frac{\mathrm{d}(W + W_{\mathrm{em}})}{\mathrm{d}t} = \frac{1}{4\,\pi\,\varepsilon_0}\, \frac{e^2\, \dot{u}^2}{4\,\pi\, c^3} \int_\Sigma \frac{\sin^2 \vartheta}{g^2}\, \mathbf{n} \cdot \mathbf{n}\, \mathrm{d}\Sigma = \frac{2\, e^2/3}{4\,\pi\,\varepsilon_0\, c^3}\, \dot{u}^2 . \tag{5.72}$$

The expression at the right-hand side of (5.72), called *Larmor formula*, gives an approximate expression of the power emitted by a point-like charge, that is applicable when $u/c \ll 1$. As shown by the left-hand side, part of the emitted power ($-\mathrm{d}W/\mathrm{d}t$) is due to the variation in the charge's mechanical energy, while the other part ($-\mathrm{d}W_{\mathrm{em}}/\mathrm{d}t$) is due to the variation in the electromagnetic energy within the volume enclosed by Σ.

5.11.3 Decay of Atoms According to the Classical Model

The power radiated by a point-like charge has been determined in Sect. 5.11.2 under the approximations $\dot{\mathbf{u}} \neq 0$, $u/c \ll 1$, typical of a bound particle. The radiated power (5.72) is proportional to the square of the particle's acceleration: this result is of a paramount importance, because it shows that the so-called *planetary model* of the atom is not stable. Considering for instance the simple case of hydrogen, the model describes the atom as a planetary system whose nucleus is fixed in the reference's origin while the electron orbits around it. If no power was emitted the motion's description would be that given, e.g., in Sect. 3.13.6, where the Hamiltonian function is a constant of motion. In other terms, the total energy would be conserved. In fact, in the planetary motion the electron's acceleration, hence the emitted power, differ from zero; the emission produces an energy loss which was not considered in the analysis of Sect. 3.13.6. Some comments about this problem were anticipated in Sect. 3.13.9.

To proceed it is useful to carry out a quantitative estimate of the emitted power. The outcome of it is that in the case of a bound electron the emission is relatively weak, so that one can consider it as a perturbation with respect to the conservative case analyzed in Sect. 3.13.6. The estimate starts from the experimental observation of the emission of electromagnetic radiation by excited atoms; here the datum that matters is the minimum angular frequency ω_0 of the emitted radiation, which is found to be in the range $[10^{15}, 10^{16}]$ rad s^{-1}. The simplest model for describing the unperturbed electron's motion is that of the linear harmonic oscillator [10, Vol. II, Sect. 4]

$$\mathbf{s}(\tau) = \mathbf{s}_0 \cos(\omega_0 \tau)\,, \tag{5.73}$$

with $s_0 = |\mathbf{s}_0|$ the maximum elongation with respect to the origin, where the nucleus is placed. Equation (5.73) may be thought of as describing the projection over the direction of $\mathbf{s}_0$ of the instantaneous position of an electron that follows a circular orbit. The product $e\,\mathbf{s}$ is called *electric dipole moment* of the oscillator. Other experimental results, relative to the measure of the atom's size, show that s_0 is of the order of 10^{-10} m so that, calculating $\mathbf{u} = \mathrm{d}\mathbf{s}/\mathrm{d}\tau = -\mathbf{s}_0\,\omega_0\,\sin(\omega_0\tau)$ from (5.73) and letting $\omega_0 = 5 \times 10^{15}$, one finds $u/c \leq s_0\,\omega_0/c \simeq 2 \times 10^{-3}$. This shows that the approximations of Sect. 5.11.2 are applicable.

It is worth noting that the type of motion (5.73) is energy conserving, hence it must be understood as describing the unperturbed dynamics of the electron. Remembering the discussion of Sect. 3.3 one finds for the total, unperturbed energy the expression $E_u = m\,\omega_0^2\,\mathbf{s}_0^2/2$, with $m = 9.11 \times 10^{-31}$ kg the electron mass. To tackle the perturbative calculation it is now necessary to estimate the energy E_r lost by the electron during an oscillation period $2\,\pi/\omega_0$ and compare it with E_u. From $\dot{\mathbf{u}} = \mathrm{d}\mathbf{u}/\mathrm{d}\tau = \omega_0^2\,\mathbf{s}$ one obtains the maximum square modulus of the electron's acceleration, $\dot{u}_M^2 = \omega_0^4\,s_0^2$; inserting the latter into (5.72) and using $e = -q = -1.602 \times 10^{-19}$ C for the electron charge provide the upper bounds

$$E_r \leq \frac{2\,\pi}{\omega_0}\,\frac{2\,e^2/3}{4\,\pi\,\varepsilon_0\,c^3}\,\omega_0^4\,s_0^2 = \frac{e^2\,s_0^2\,\omega_0^3}{3\,\varepsilon_0\,c^3}\,, \qquad \frac{E_r}{E_u} \leq \frac{2\,e^2\omega_0}{\varepsilon_0\,m\,c^3} \simeq 4\times 10^{-7}\,. \tag{5.74}$$

This result shows that the energy lost during an oscillation period is indeed small, so that the electron's motion is only slightly perturbed with respect to the periodic case. The equation of motion of the perturbed case can now tentatively be written as

$$m\,\ddot{\mathbf{s}} + m\,\omega_0^2\,\mathbf{s} = \mathbf{F}_r\,, \tag{5.75}$$

where $\mathbf{F}_r$ is a yet unknown force that accounts for the emitted power. A scalar multiplication of (5.75) by $\mathbf{u}$ yields $m\,\mathbf{u}\cdot\dot{\mathbf{u}} + m\,\omega_0^2\,\mathbf{s}\cdot\mathbf{u} = \mathrm{d}W/\mathrm{d}\tau = \mathbf{u}\cdot\mathbf{F}_r$, with $W = (m/2)\,(u^2 + \omega_0^2\,s^2)$. One notes that W has the same expression as the total energy E_u of the unperturbed case; however, W is not conserved due to the presence of $\mathbf{F}_r \neq 0$ at the right-hand side. In fact, $-\mathrm{d}W/\mathrm{d}\tau = -\mathbf{u}\cdot\mathbf{F}_r > 0$ is the power emitted by the electron, and its time average over $2\,\pi/\omega_0$,

$$-\,\langle\mathbf{u}\cdot\mathbf{F}_r\rangle = -\frac{\omega_0}{2\,\pi}\int_0^{2\,\pi/\omega_0}\mathbf{u}\cdot\mathbf{F}_r\,\mathrm{d}\tau > 0\,, \tag{5.76}$$

is the variation in the oscillator's energy during a period; a part of it crosses the surface Σ, while the other part is the variation in the electromagnetic energy within Σ (Sects. 5.3 and 5.11.2). The part that crosses Σ is the time average of (5.72)[2]; for the sake of simplicity it is assumed that it is dominant with respect to the other one. The factor $\langle\dot{u}^2\rangle$ that appears in (5.72) is worked out by taking the time average of the identity $\mathrm{d}(\mathbf{u}\cdot\dot{\mathbf{u}})/\mathrm{d}\tau = \dot{u}^2 + \mathbf{u}\cdot\ddot{\mathbf{u}}$ and observing that $\langle\mathrm{d}(\mathbf{u}\cdot\dot{\mathbf{u}})/\mathrm{d}\tau\rangle$ is negligibly small, whence $\langle\dot{u}^2\rangle = -\langle\mathbf{u}\cdot\ddot{\mathbf{u}}\rangle > 0$. In conclusion, defining a time τ_0 such that $e^2/(6\,\pi\,\varepsilon_0\,c^3) = m\,\tau_0$ and equating (5.76) to the time average of (5.72) yield $\langle\mathbf{u}\cdot m\,\tau_0\,\ddot{\mathbf{u}}\rangle = \langle\mathbf{u}\cdot\mathbf{F}_r\rangle$. It is found $\tau_0 \simeq 6\times 10^{-24}$ s.

As a crude approximation one finally converts the equality of the averages just found into an equality of the arguments, whence $\mathbf{F}_r \simeq m\,\tau_0\,\ddot{\mathbf{u}}$. Replacing the latter into (5.75) yields $\ddot{\mathbf{s}} + \omega_0^2\,\mathbf{s} = \tau_0\,\ddot{\mathbf{u}}$, that is, a linear, homogeneous equation of the third order in $\mathbf{s}$ with constant coefficients. The equation is solved by letting $\mathbf{s} = \mathbf{s}(\tau = 0)\,\exp(\alpha\,\tau)\,\cos(\omega\,\tau)$, with α, ω undetermined. Using the tentative solution provides the system of characteristic algebraic equations

$$\tau_0\,\omega^2 = 3\,\tau_0\,\alpha^2 - 2\,\alpha\,, \qquad \alpha^2 + \omega_0^2 = \tau_0\,\alpha^3 + (1 - 3\,\tau_0\,\alpha)\,\omega^2\,, \tag{5.77}$$

whence the elimination of ω^2 yields $8\,\alpha^2 - 2\,\alpha/\tau_0 - \omega_0^2 = 8\,\tau_0\alpha^3$. Thanks to the smallness of τ_0 the latter equation may be solved by successive approximations starting from the zeroth-order solution $\alpha^{(0)} \simeq -\tau_0\,\omega_0^2/2$ (this solution is found by

[2]Remembering the discussion of Sect. 5.11.2, the use of (5.72) implies that the particle's position departs little from the center of the spherical surface. Thus the radius of Σ must be much larger than the size of the atom.

solving $8\,\alpha^2 - 2\,\alpha/\tau_0 - \omega_0^2 = 0$ and using the binomial approximation; the other possible value of $\alpha^{(0)}$ is positive and must be discarded to prevent $\mathbf{s}$ from diverging). Replacing $\alpha^{(0)}$ into the first equation in (5.77) yields $\omega^2 = \omega_0^2\,(1+3\,\tau_0^2\,\omega_0^2/4) \simeq \omega_0^2$. In conclusion, the zeroth-order solution of the differential equation for $\mathbf{s}$ reads

$$\mathbf{s}(\tau) \simeq \mathbf{s}(\tau = 0)\,\cos(\omega_0\,\tau)\,\exp(-\tau_0\,\omega_0^2\,\tau/2)\,. \tag{5.78}$$

Basing upon (5.78) one can identify the decay time of the atom with the time necessary for the modulus of $\mathbf{s}$ to reduce by a factor $1/\mathrm{e}$ with respect to the initial value. The decay time is thus found to be $2/(\tau_0\,\omega_0^2) \simeq 13 \times 10^{-9}$ s. As the ratio of the decay time to the period $2\,\pi/\omega_0$ is about 10^7, the perturbative approach is indeed justified.

As anticipated at the beginning of this section, the planetary model of the atom is not stable. The approximate solution (5.78) of the electron's dynamics shows that according to this model the electron would collapse into the nucleus in a very short time due to the radiation emitted by the electron. This behavior is not observed experimentally: in fact, the experiments show a different pattern in the energy-emission or absorption behavior of the atoms. The latter are able to absorb energy from an external radiation and subsequently release it: an absorption event brings the atom to a higher-energy state called *excited state*; the absorbed energy is then released by radiation in one or more steps (*emissions*) until, eventually, the atom reaches the lowest energy state (*ground state*). However, when the atom is in the ground state and no external perturbations is present, the atom is stable and no emission occurs. In conclusion, the experimental evidence shows that the planetary model is not applicable to the description of atoms.[3]

5.11.4 Comments About the Field's Expansion into Modes

The homogeneous wave equation (5.20) used in Sects. 5.5, 5.8 as a starting point for the derivation of the field's expansion into modes is based on the hypothesis that a gauge transformation exists such that $\varphi = 0$. In turn, (5.20) derives from (5.19), that implies the Coulomb gauge $\mathrm{div}\,\mathbf{A} = 0$. To show that these conditions are mutually compatible one chooses f in (4.30) such that $\varphi' = 0$, whence $\mathbf{E}' = -\partial\mathbf{A}'/\partial t$ due to the second relation in (4.26). In a charge-free space it is $\mathrm{div}\,\mathbf{D}' = \varepsilon_0\,\mathrm{div}\,\mathbf{E}' = 0$; it follows $\partial\,\mathrm{div}\,\mathbf{A}'/\partial t = 0$, namely, $\mathrm{div}\,\mathbf{A}'$ does not depend on time. The second equation in (4.19) with $\mathbf{J} = 0$ yields $(1/c^2)\,\partial\mathbf{E}'/\partial t = \mathrm{rot}\,\mathbf{B}'$, so that $-(1/c^2)\,\partial^2\mathbf{A}'/\partial t^2 = \mathrm{rot}\,\mathrm{rot}\,\mathbf{A}'$. Now let $\mathbf{A}' = \mathbf{A}'' + \mathrm{grad}\,g$, where g is an arbitrary function of the coordinates only; the second identity in (A.35) and the

[3] In this respect one might argue that the inconsistency between calculation and experiment is due to some flaw in the electromagnetic equations. However, other sets of experiments show that it is not the case.

first identity in (A.36) then yield $-(1/c^2)\,\partial^2\mathbf{A}''/\partial t^2 = \operatorname{grad}\operatorname{div}\mathbf{A}'' - \nabla^2\mathbf{A}''$, with $\operatorname{div}\mathbf{A}'' = \operatorname{div}\mathbf{A}' - \nabla^2 g$. Choosing g such that $\operatorname{div}\mathbf{A}'' = 0$ and dropping the double apex finally yield (5.20) [85, Sect. 46].

The vector potential $\mathbf{A}$ has been expressed in (5.21) as a Fourier series and in (5.44) as a Fourier integral. Such expressions produce a separation of the spatial coordinates from the time coordinate: the former appear only in the terms $\exp(\mathrm{i}\,\mathbf{k}\cdot\mathbf{r})$, while the latter appears only in the terms $\mathbf{a}_\mathbf{k}$ and, respectively, $\mathbf{b}_\mathbf{k}$.

The Fourier series (5.21) applies to the case of a finite domain of the form shown in Fig. 5.1 and prescribes the spatial periodicity of $\mathbf{A}$ at all times. By way of example, let $0 \le x_2 \le d_2$, $0 \le x_3 \le d_3$ and consider the point $\mathbf{r}_A = x_2\,\mathbf{i}_2 + x_3\,\mathbf{i}_3$; then, consider a second point $\mathbf{r}_B = d_1\,\mathbf{i}_1 + x_2\,\mathbf{i}_2 + x_3\,\mathbf{i}_3$. By construction, $\mathbf{r}_A$ and $\mathbf{r}_B$ belong to two opposite faces of the domain of Fig. 5.1 and are aligned with each other in the x_1 direction. From (5.22) one obtains for any $\mathbf{k}$

$$\exp(\mathrm{i}\,\mathbf{k}\cdot\mathbf{r}_B) = \exp(\mathrm{i}\,2\,\pi\,n_1)\,\exp(\mathrm{i}\,\mathbf{k}\cdot\mathbf{r}_A) = \exp(\mathrm{i}\,\mathbf{k}\cdot\mathbf{r}_A) \tag{5.79}$$

which, combined with (5.21), yields $\mathbf{A}(\mathbf{r}_B, t) = \mathbf{A}(\mathbf{r}_A, t)$. Clearly an equality of this form is found for any pair of opposite boundary points that are aligned along the coordinate direction normal to the faces where the points lie. On the other hand, such an equality is a homogeneous relation among the boundary values of the solution of the differential equation (5.20), namely, it is a homogeneous boundary condition of the Dirichlet type.

The reasoning based on (5.79) is applicable also to the expressions (5.31, 5.32) to yield $\mathbf{E}(\mathbf{r}_B, t) = \mathbf{E}(\mathbf{r}_A, t)$ and $\mathbf{B}(\mathbf{r}_B, t) = \mathbf{B}(\mathbf{r}_A, t)$, namely, the fields have the same periodicity as $\mathbf{A}$. The Poynting vector $\mathbf{S} = \mathbf{E} \wedge \mathbf{H}$ has this property as well, whence $\mathbf{S}(\mathbf{r}_B, t)\cdot\mathbf{n}_B = -\mathbf{S}(\mathbf{r}_A, t)\cdot\mathbf{n}_A$; in fact, the unit vector $\mathbf{n}$ is oriented in the outward direction with respect to the domain (Sect. 5.3), so that when two opposite faces of V are considered it is $\mathbf{n}_B = -\mathbf{n}_A$. Using (5.12) with $W = 0$ shows that $\mathrm{d}W_{\mathrm{em}}/\mathrm{d}t = 0$ namely, as noted in Sect. 5.6, the electromagnetic energy within V is conserved. The same reasoning applies to the conservation of the electromagnetic momentum found in Sect. 5.7. As for the initial condition on $\mathbf{A}$, from (5.21) and (5.29) one derives $\mathbf{A}(\mathbf{r}, t = 0) = \sum_\mathbf{k}(\mathbf{c}_\mathbf{k} + \mathbf{c}^*_{-\mathbf{k}})\,\exp(\mathrm{i}\,\mathbf{k}\cdot\mathbf{r})$. It follows that the initial condition is provided by the vectors $\mathbf{c}_\mathbf{k}$.

5.11.5 Finiteness of the Total Energy

The differential equation (5.20) is linear and homogeneous with respect to the unknown $\mathbf{A}$; when the Fourier series (5.21) is replaced in it, the resulting equation (5.28) is linear and homogeneous with respect to $\mathbf{a}_\mathbf{k}$, hence (due to (5.29)) with respect to $\mathbf{s}_\mathbf{k}$ and $\mathbf{c}_\mathbf{k}$ as well. It follows that the fields (5.31, 5.32) are linear and homogeneous functions of these quantities. The same applies in the case of an infinite domain (Sect. 5.8), in which the fields $\mathbf{E}$, $\mathbf{B}$ are linear and homogeneous functions of $\tilde{\mathbf{s}}_\mathbf{k}$ and $\mathbf{d}_\mathbf{k}$.

In turn, the energy density w_{em} of the electromagnetic field, given by the second relation in (5.11), is a quadratic and homogeneous function of the fields; this explains why the expressions (5.37) and (5.46) are quadratic and homogeneous functions of $\mathbf{s_k}$, $\mathbf{c_k}$ or, respectively, $\tilde{\mathbf{s}}_\mathbf{k}$, $\mathbf{d_k}$.

When (5.37) is considered, the energy associated with the individual degree of freedom is $W_{\mathbf{k}\sigma} = 2\,\varepsilon_0\,V\,\omega^2\,|\mathbf{c}_{\mathbf{k}\sigma}|^2$; as the sum $\sum_{\mathbf{k}\sigma} W_{\mathbf{k}\sigma}$ spans all wave vectors, the factor $\omega^2 = c^2\,k^2$ diverges. On the other hand, the energy of the electromagnetic field within a finite region of space is finite; this means that the term $|\mathbf{c}_{\mathbf{k}\sigma}|^2$ becomes vanishingly small as $|\mathbf{k}| \to \infty$, in such a way as to keep the sum $\sum_{\mathbf{k}\sigma} W_{\mathbf{k}\sigma}$ finite. The same reasoning applies to the term $|\mathbf{d}_{\mathbf{k}\sigma}|^2$ in (5.46); in this case the finiteness of the total energy W_{em} is due to the fact that the vanishing of the fields at infinity makes the Fourier transform in (5.44) to converge.

5.11.6 Analogies Between Mechanics and Geometrical Optics

A number of analogies exist between the Maupertuis principle, discussed in Sect. 2.7, and the Fermat principle discussed in Sect. 5.10. The principles read, respectively,

$$\delta \int_{AB} \sqrt{E - V}\,\mathrm{d}s = 0\,, \qquad \delta \int_{AB} n\,\mathrm{d}s = 0\,, \tag{5.80}$$

and the analogies are:

1. A constant parameter is present, namely, the total energy E on one side, the frequency ν on the other side (in fact, the Fermat principle generates the eikonal equation which, in turn, applies to a monochromatic electromagnetic field, Sect. 5.9).
2. Given the constant parameter, the integrand is uniquely defined by a property of the medium where the physical phenomenon occurs: the potential energy $V(\mathbf{r})$ and the refraction index $n(\mathbf{r})$, respectively.
3. The outcome of the calculation is a curve of the three-dimensional space: the particle's trajectory and the optical ray, respectively. In both cases the initial conditions are the starting position and direction (in the mechanical case the initial velocity is obtained by combining the initial direction with the momentum extracted from $E - V$).

In summary, by a suitable choice of the units, the same concept is applicable to both mechanical and optical problems. In particular it is used for realizing devices able to obtain a trajectory or a ray of the desired form: the control of the ray's shape is achieved by prescribing the refraction index n by means of, e.g., a lens or a system of lenses; similarly, the trajectory of a particle of charge e is controlled by a set of electrodes (*electrostatic lenses*) that prescribe the electric potential $\varphi = V/e$. The design of *electron guns* and of the equipments for *electron-beam lithography* and *ion-beam lithography* is based on this analogy.

It must be emphasized that the Maupertuis principle is derived without approximations: as shown in Sect. 2.7, the principle is equivalent to Newton's law applied to a particle of constant energy. The Fermat principle, instead, is equivalent to the eikonal equation; the latter, in turn, is derived from the Maxwell equations in the hypothesis that at each position along the ray the curvature radius of the ray is much larger than the wavelength. In other terms, the mechanical principle is exact whereas the optical principle entails an approximation. If the exact formulation of electromagnetism given by the Maxwell equation was used, the analogy discussed here would be lost.

The rather surprising asymmetry outlined above could be fixed by speculating that Newton's law is in fact an approximation deriving from more general laws, possibly similar to the Maxwell equations. In this case one could identify in such laws a parameter analogue of the wavelength, and deduce Newton's law as the limiting case in which the parameter is small. It will be shown later that mechanical laws more general than Newton's laws indeed exist: they form the object of Quantum Mechanics.[4]

The analogy between Mechanics and Geometrical Optics discussed here is one of the possible courses of reasoning useful for introducing the quantum-mechanical concepts; however, in this reasoning the analogy should not be pushed too far. In fact, one must observe that the Maupertuis principle given by the first expression in (5.80) provides the nonrelativistic form of Newton's law, whereas the Maxwell equations, of which the Fermat principle is an approximation, are intrinsically relativistic. As a consequence, the analogy discussed in this section is useful for generalizing the geometrical properties of the motion, but not the dynamic properties.

Problems

5.1 Solve the eikonal equation (5.57) in a medium whose refraction index depends on one coordinate only, say, $n = n(x_1)$.

5.2 Use the solution of problem 5.1 to deduce the Descartes law of refraction.

5.3 With reference to the fields generated by a point-like charge (Sect. 5.11.1), express $\partial^2\tau/\partial t^2$ in terms of $\partial\tau/\partial t$, $\mathbf{g}$, $\mathbf{u}$, and $\dot{\mathbf{u}} = \partial\mathbf{u}/\partial\tau$ using the same procedure as that leading to (5.65).

5.4 Show that (5.6) and (5.7) can be recast in the form

$$\varphi = \frac{\partial\tau}{\partial t}\,\frac{e}{4\,\pi\,\varepsilon_0\,g}\,, \qquad \mathbf{A} = \frac{\partial\tau}{\partial t}\,\frac{e\,\mathbf{u}}{4\,\pi\,\varepsilon_0\,c^2\,g}\,,$$

and find the corresponding expressions of $\mathbf{E}$ and $\mathbf{B}$.

[4] Once Quantum Mechanics is introduced, Newtonian Mechanics is distinguished from it by the designation *Classical Mechanics*.

5.5 Starting from the result of Prob. 5.4, determine the expression of the electric field $\mathbf{E}$ in terms of $\partial\tau/\partial t$, $\mathbf{g}$, $\mathbf{u}$, and $\dot{\mathbf{u}} = \partial\mathbf{u}/\partial\tau$.

5.6 Using the result of Prob. 5.5, show that

$$\frac{4\,\pi\,\varepsilon_0\,g}{e\,(\partial\tau/\partial t)^3}\,\mathbf{E}\wedge\left(\mathbf{g}-g\,\frac{\mathbf{u}}{c}\right)=\left(1-\frac{\mathbf{u}}{c}\cdot\frac{\mathbf{g}}{g}\right)\left(\mathbf{g}-g\,\frac{\mathbf{u}}{c}\right)\wedge\frac{\dot{\mathbf{u}}}{c^2}\,.$$

5.7 With reference to the fields generated by a point-like charge (Sect. 5.11.1), and to the expression of the magnetic induction $\mathbf{B}$ worked out in Prob. 5.4, an application of the first identity in (A.35) yields

$$\frac{4\,\pi\,\varepsilon_0\,c^2}{e}\,\mathbf{B}=\frac{1}{g}\,\frac{\partial\tau}{\partial t}\,\mathrm{rot}\,\mathbf{u}-\mathbf{u}\wedge\mathrm{grad}\left(\frac{1}{g}\,\frac{\partial\tau}{\partial t}\right).$$

Rework the above expression in terms of $\partial\tau/\partial t$, $\mathbf{g}$, $\mathbf{u}$, and $\dot{\mathbf{u}} = \partial\mathbf{u}/\partial\tau$.

5.8 Starting from the results of Probs. 5.4 and 5.7, determine the expression of the magnetic induction $\mathbf{B}$ in terms of $\partial\tau/\partial t$, $\mathbf{g}$, $\mathbf{u}$, and $\dot{\mathbf{u}} = \partial\mathbf{u}/\partial\tau$.

5.9 Starting from the results of Probs. 5.6 and 5.8, derive the compact expression (5.67) of the magnetic induction $\mathbf{B}$.

Part II
Introductory Concepts to Statistical and Quantum Mechanics

Chapter 6
Classical Distribution Function and Transport Equation

6.1 Introduction

When a system made of a large number of molecules is considered, the description of the dynamics of each individual member of the system is practically impossible, and it is necessary to resort to the methods of Statistical Mechanics. The chapter introduces the concept of distribution function in the phase space and provides the definition of statistical average (over the phase space and momentum space) of a dynamic variable. The derivation of the equilibrium distribution in the classical case follows, leading to the Maxwell-Boltzmann distribution. The analysis proceeds with the derivation of the continuity equation in the phase space: the collisionless case is treated first, followed by the more general case where the collisions are present, thus leading to the Boltzmann Transport Equation. In the Complements, after a discussion about the condition of a vanishing total momentum and angular momentum in the equilibrium case, and the derivation of statistical averages based on the Maxwell-Boltzmann distribution, the Boltzmann H-theorem is demonstrated. This is followed by an illustration of the apparent paradoxes brought about by Boltzmann's Transport Equation and H-theorem: the violation of the symmetry of the laws of mechanics with respect to time reversal, and the violation of Poincaré's time recurrence. The illustration is carried out basing on Kac's ring model. The chapter is completed by the derivation of the equilibrium limit of the Boltzmann Transport Equation.

6.2 Distribution Function

Consider a system made of N identical particles with no constraints. For the sake of simplicity, point-like particles are assumed, so that the total number of degrees of freedom is $3N$. The dynamics of the jth particle is described by the canonical

M. Rudan, *Physics of Semiconductor Devices*, DOI 10.1007/978-3-319-63154-7_6

coordinates q_{1j}, q_{2j}, q_{3j}, p_{1j}, p_{2j}, p_{3j} that belong to the 6-dimensional μ-space introduced in Sect. 1.9.

If the number of particles is large, the description of the dynamics of each individual belonging to the system is in fact impossible. For instance, the number density of air at 20 °C and 1 atm is about 2.5×10^{19} cm^{-3}. Even if the measurement of the initial conditions was possible, it would not be feasible in practice [83, Sect. 1]. This aspect is present also when the number of particles is much lower than in the example above.

The problem is faced by adopting the viewpoint of *statistical mechanics*, whose object is not the description of the dynamic behavior of the individual particles but, rather, that of the distribution of the dynamic properties over the phase space. To this purpose one identifies each point of the μ-space with the pair of vectors $\mathbf{q} = (q_1, q_2, q_3)$, $\mathbf{p} = (p_1, p_2, p_3)$ pertaining to it, and considers the elementary volume $\mathrm{d}\omega = \mathrm{d}^3q\, \mathrm{d}^3p$ of the μ-space centered at $(\mathbf{q}, \mathbf{p})$, with $\mathrm{d}^3q = \mathrm{d}q_1\, \mathrm{d}q_2\, \mathrm{d}q_3$ and the like for d^3p. Then, the number of particles $\mathrm{d}N$ that at time t belong to $\mathrm{d}\omega$ is given by

$$\mathrm{d}N = f_\mu(\mathbf{q}, \mathbf{p}, t)\, \mathrm{d}\omega\,, \tag{6.1}$$

with f_μ the concentration in the μ-space. The procedure here is similar to that carried out in Sect. 23.2, the difference being that the space considered here is the phase space instead of the configuration space of Sect. 23.2.[1] In both cases, the motion of the particles is described as that of a continuous fluid: in fact, index j is dropped from the canonical coordinates, which do not indicate any more a specific particle, but the center of the elementary cell of the phase space where the concentration f_μ is considered. As in Sect. 23.2, this procedure is legitimate if the cells of volume $\mathrm{d}\omega$ into which the phase space is partitioned can be treated as infinitesimal quantities in the scale of the problem that is being investigated, and the number of particles within each cell is large enough to make their average properties significant. The concentration f_μ is also called *distribution function*. By definition it fulfills the normalization condition

$$\int f_\mu(\mathbf{q}, \mathbf{p}, t)\, \mathrm{d}\omega = N\,, \tag{6.2}$$

where the integral is 6-dimensional and extends over the whole μ-space. As the order of integration is immaterial, the calculation can be split into two steps, namely,

$$n(\mathbf{q}, t) = \iiint_{-\infty}^{+\infty} f_\mu(\mathbf{q}, \mathbf{p}, t)\, \mathrm{d}^3p\,, \qquad N = \iiint_{-\infty}^{+\infty} n(\mathbf{q}, t)\, \mathrm{d}^3q\,. \tag{6.3}$$

[1] Note that here the symbol N indicates the number of particles; instead, in Sect. 23.2 the number of particles is indicated with $\mathcal{N}$, whereas N indicates the concentration.

The function $n(\mathbf{q}, t)$ defined by the first of (6.3) is the concentration of the particles in the configuration space.

Basing on the distribution function it is possible to define the average of a dynamic function. For the sake of generality the dynamic function is considered to be a vector that depends on all canonical coordinates and time, say, $\mathbf{a} = \mathbf{a}(\mathbf{q}, \mathbf{p}, t)$. Due to the smallness of the cell size one assumes that the dynamic function takes the same value for all particles within the same cell. As a consequence, the product $\mathbf{a} f_\mu \, \mathrm{d}\omega$ is the cell value of the function weighed by the number of particles belonging to the cell. The *statistical average* of $\mathbf{a}$ over the phase space is then

$$\mathrm{Av}[\mathbf{a}](t) = \frac{1}{N} \int \mathbf{a}(\mathbf{q}, \mathbf{p}, t) f_\mu(\mathbf{q}, \mathbf{p}, t) \, \mathrm{d}\omega \,, \tag{6.4}$$

where the integral is 6-dimensional, while the average over the momentum space is

$$\overline{\mathbf{a}}(\mathbf{q}, t) = \frac{1}{n(\mathbf{q}, t)} \iiint_{-\infty}^{+\infty} \mathbf{a}(\mathbf{q}, \mathbf{p}, t) f_\mu(\mathbf{q}, \mathbf{p}, t) \, \mathrm{d}^3 p \,. \tag{6.5}$$

Using the expression of N given by (6.2), and that of n given by the first of (6.3), shows that the definitions (6.4, 6.5) indeed provide the weighed averages of interest. By way of example, the dynamic function may be identified with the particle velocity $\mathbf{u}$: using the Cartesian coordinates one finds for the average velocity $\mathbf{v}$ in the configuration space the expression

$$\mathbf{v}(\mathbf{r}, t) = \frac{1}{n(\mathbf{r}, t)} \iiint_{-\infty}^{+\infty} \mathbf{u}(\mathbf{r}, \mathbf{p}, t) f_\mu(\mathbf{r}, \mathbf{p}, t) \, \mathrm{d}^3 p \,. \tag{6.6}$$

Similarly, the average Hamiltonian function in the configuration space reads

$$\overline{H}(\mathbf{r}, t) = \frac{1}{n(\mathbf{r}, t)} \iiint_{-\infty}^{+\infty} H(\mathbf{r}, \mathbf{p}, t) f_\mu(\mathbf{r}, \mathbf{p}, t) \, \mathrm{d}^3 p \,. \tag{6.7}$$

6.3 Statistical Equilibrium

This section deals with the properties of a system of particles in a condition of macroscopic equilibrium. Considering that in general the systems that are considered are composed of a large number of particles or molecules, the statistical concepts introduced in Sect. 6.2 will be used. Generally speaking, the condition of statistical equilibrium is fulfilled if the distribution function is independent of time. This condition may be achieved in different ways: for instance, $f_\mu = \text{const}$ fulfills the required condition. A more general definition of a distribution function fulfilling the condition of statistical equilibrium is $f_\mu = f_\mu(c)$, where c is any constant of motion of the system. In case of a conservative system, energy is the most natural constant of motion to be used.

To proceed, consider a conservative system having a total energy E_S, enclosed in a stationary container of volume Ω. Let the walls of the container be such that no energy flows across them. Also, the container is assumed to be sufficiently massive so that it can be regarded as stationary despite the transfer of kinetic energy due to the molecules' collisions with the walls. If any external force acts on the molecules, it is assumed to be independent of time and conservative [136, Sect. 26]. Finally, the total momentum and angular momentum of the system are assumed to vanish; this condition is by no means obvious and requires some reasoning, as detailed in Sect. 6.6.1.

So far it has not been explicitly indicated whether the molecules that form the system are identical to each other or not; in the practical cases it is to be expected that the system under consideration be made of a mixture of different atoms or molecules. As the extension to a mixture is straightforward [136, Sect. 30], the analysis is limited here to a system made of identical molecules. It should be noted that the molecules are identical, but distinguishable from each other: from the point of view of Classical Mechanics a continuous observation of the trajectory of each molecule is in fact possible, without disturbing its motion. As a consequence, two systems that differ by the exchange of two molecules are to be considered as different from each other.

To proceed one assumes that the number of molecules forming the system is N, and that each molecule has R degrees of freedom. The canonical coordinates that describe the motion of a single molecule are then $q_1, \ldots, q_R, p_1, \ldots, p_R$, so that the number of dimensions of the μ-space is $2R$. As anticipated in Sect. 6.2, the description of the precise state of each molecule is impossible in practice; the state of the system will then be specified in a somewhat less precise manner, as detailed below. First, each q axis of the μ-space is divided into equal intervals of size $\Delta q_1, \ldots, \Delta q_R$ and, similarly, each p axis is divided into equal intervals of size $\Delta p_1, \ldots, \Delta p_R$. As a consequence, the μ-space is partitioned into elements, called *cells*, whose volume and units are, respectively,

$$\Delta M = (\Delta q_1 \, \Delta p_1) \ldots (\Delta q_R \, \Delta p_R) , \quad [\Delta M] = (\mathrm{J\,s})^R . \tag{6.8}$$

The partitioning of the μ-space into cells has the advantage, first, that the set of cells is countable. Besides that, the partitioning into cells of finite size has a deeper meaning, that becomes apparent when the theory outlined in this section is extended to the quantum-mechanical case. In fact, due to the Heisenberg uncertainty relation (Sect. 10.6), the precision by which two conjugate variables can simultaneously be known is limited, so that each product $\Delta q_i \, \Delta p_i$ in (6.8) is bounded below, the bound being of the order of the Planck constant.

After the partitioning is accomplished, the state of the system is assigned by specifying the numbers $N_1, N_2, \ldots \geq 0$ of molecules that belong, respectively, to the cell labeled $1, 2, \ldots$; such numbers are subjected to the constraint $N_1 + N_2 + \ldots = N$ which, in view of the calculations that follow, is more conveniently expressed as

$$F_N(N_1, N_2, \ldots) = 0\,, \qquad F_N = N - \sum_i N_i\,. \tag{6.9}$$

The sum in (6.9) may be thought of as extending to all cells, due to the fact that only in a finite number of cases the cell population N_i differs from zero. Clearly, the description using the numbers $N_1, N_2, \ldots$ is less precise than that given by the molecules' trajectories, and provides a partial specification of the state of each molecule within the limits of the size of ΔM. This means, among other things, that identical molecules belonging to different cells are distinguishable from each other, whereas identical molecules belonging to the same cell are not distinguishable.

As mentioned above, the total energy of the system is E_S, which provides a second constraint to be vested with mathematical form. It is provisionally assumed that the system is *dilute*, namely, that the energy of the interaction among the molecules is negligible. It follows that one can approximately assign to each molecule a total energy that corresponds to its position, momentum, and internal configuration, in other terms, to the cell where the molecule belongs. Letting the energy corresponding to the ith cell be E_i, the constraint on energy reads $N_1 E_1 + N_2 E_2 + \ldots = E_S$, namely,

$$F_E(N_1, N_2, \ldots) = 0\,, \qquad F_E = E_S - \sum_i N_i E_i\,. \tag{6.10}$$

The above reasoning does not apply to concentrated systems, where the interaction energy is strong. However, it can be shown that (6.10) still holds, albeit with a different interpretation of E_i [136, Sect. 29]. Another observation is that, given the constraints to which the system is subjected, the set of numbers $N_1, N_2, \ldots$ may not be unique. It is therefore necessary to extract the set, that actually describes the system, from a larger number of sets made of all possible sets that are compatible with the constraints.

To proceed, let $N_1, N_2, \ldots$ be a set that provides a molecules' distribution compatible with the constraints; such a set is called *accessible state*. It is postulated that no accessible state is privileged with respect to any other; this is in fact the fundamental hypothesis of equal *a priori* probability of the accessible states, upon which Statistical Mechanics is based [136, Sect. 23]. Remembering that the particles are identical to each other, any system obtained from the original distribution by exchanging two molecules is also compatible with the constraints. However, the system resulting from the exchange of molecules belonging to different cells is different from the original one because such molecules are distinguishable; in contrast, an exchange within the same cell does not produce a different system. As the total number of possible exchanges of N molecules of the system as a whole is $N!$, and the number of possible exchanges within the ith cell is $N_i!$, the total number of different systems corresponding to a set $N_1, N_2, \ldots$ is

$$W(N_1, N_2, \ldots) = \frac{N!}{N_1!\, N_2! \ldots} \tag{6.11}$$

As time evolves, the interactions among the molecules make the numbers $N_1, N_2, \ldots$ to change, so that, if the system is inspected at successive times, it may be found to belong to different accessible states (as the principles of Classical Mechanics apply here, such an inspection does not perturb the state of the system); in principle, given enough time, the system will go through all accessible states. Now, the dependence of W on $N_1, N_2, \ldots$ is quite strong, so that some accessible states correspond to a large number W of systems, others to a much smaller number.[2] As no accessible state is privileged with respect to any other, the majority of inspections carried out onto the system will provide the accessible state that maximizes W. Such an accessible state corresponds, by definition, to the condition of *statistical equilibrium* of the system under consideration.

6.4 Maxwell-Boltzmann Distribution

The analysis carried out in Sect. 6.3 led to the conclusion that the condition of statistical equilibrium of the system is found by maximizing the expression of W given by (6.11), under the constraints (6.9) and (6.10). The calculation is based upon the Lagrange method that determines the free maximum of an auxiliary function F embedding the constraints (Sect. B.6). It is convenient to maximize $\log W$, which is a monotonic function of W, instead of W itself, so that the auxiliary function reads

$$F(N_1, N_2, \ldots, \alpha, \beta) = \log W + \alpha\, F_N + \beta\, F_E\,, \tag{6.12}$$

where α, β are the *Lagrange multipliers*, respectively related to the total number of molecules and total energy of the system. In a typical system the total number of molecules and the populations of the majority of nonempty cells are very large,[3] so that the Stirling approximation (C.104) is applicable; it follows, after neglecting terms of the form $\log \sqrt{2\,\pi\,N}$ and $\log \sqrt{2\,\pi\,N_i}$,

$$\log W = \log(N!) - \sum_i \log(N_i!) \simeq N\, \log(N) - N - \sum_i N_i\, \log(N_i) + \sum_i N_i\,, \tag{6.13}$$

where $-N$ and $\sum_i N_i$ cancel each other due to (6.9). The function to maximize then becomes $F = N\, \log(N) - \sum_i N_i\, \log(N_i) + \alpha\, F_N + \beta\, F_E$. Here the property of N_i of being very large is again of help, because, on account of the fact that a change

[2] This is apparent even if the numbers $N_1, N_2, \ldots$ are much smaller than in realistic systems. Let for instance $N = 8$: the combination $N_1 = 8$, $N_2 = N_3 = \ldots = 0$ yields $W = 1$, whereas the combination $N_2 = N_3 = N_4 = N_5 = 2$, $N_1 = N_6 = N_7 = \ldots = 0$ yields $W = 2.520$. It is implied that $8\,E_1 = 2\,(E_2 + E_3 + E_4 + E_5) = E_S$.

[3] The hypothesis that the populations are large is not essential. A more complicate calculation, in which such a hypothesis is not present, leads to the same result [32].

of N_i by one unit is negligibly small with respect to N_i itself, in calculating the maximum one treats the integers N_i as continuous variables. Taking the derivative of F with respect to, say, N_r and equating it to zero yields $\log N_r + 1 = -\alpha - \beta E_r$. Neglecting the unity at the left-hand side eventually yields the *Maxwell-Boltzmann distribution law*

$$N_r = \exp(-\alpha - \beta E_r)\,. \tag{6.14}$$

The Lagrange multipliers are then determined from (6.9, 6.10); the first one yields

$$\sum_r \exp(-\alpha - \beta E_r) = N\,, \qquad N \exp(\alpha) = \sum_r \exp(-\beta E_r) = Z\,, \tag{6.15}$$

where Z denotes the *partition function*.[4] Extracting $\exp(\alpha)$ from (6.15) and replacing it into (6.10) provide the relation

$$\frac{E_S}{N} = \frac{1}{Z} \sum_r E_r \exp(-\beta E_r) = -\frac{\partial}{\partial \beta} \log Z\,, \tag{6.16}$$

with β the only unknown. This procedure is able to express β and, consequently, α, in terms of the dynamic properties of the molecules. A different method to determine the parameters that relates them with macroscopic quantities typical of Thermodynamics, like pressure and temperature, is shown in the following. First, one introduces a constant C such that

$$C N \Delta M = \exp(-\alpha)\,, \tag{6.17}$$

where ΔM is given by (6.8). After replacing (6.17) in (6.14), the cell's volume ΔM is made smaller and smaller so that, after dropping index r, (6.14) is recast in differential form as

$$\mathrm{d}N = C N \exp(-\beta E)\,\mathrm{d}M\,, \qquad \frac{1}{C} = \int \exp(-\beta E)\,\mathrm{d}M\,, \tag{6.18}$$

where the integral is extended over the μ-space and energy E is expressed in terms of the $2R$ coordinates $q_1, \ldots, q_R, p_1, \ldots, p_R$. The nature of the system is now specified as that of a monatomic gas of mass m, described by the Cartesian coordinates x_i and conjugate momenta p_i. Integrating (6.18) over the container's volume Ω yields the number $\mathrm{d}N_p$ of atoms belonging to the elementary volume of the momentum space. As the gas is assumed to be dilute, the energy in (6.18) is substantially of the kinetic type,

[4] Symbol Z comes from the German term *Zustandssumme* ("sum over states") [119, Chap. II].

$$E \simeq \frac{1}{2\,m}\left(p_1^2 + p_2^2 + p_3^2\right)\,, \tag{6.19}$$

so that the integrand is independent of the spatial coordinates. It follows

$$\mathrm{d}N_p = C\,N\,\Omega\,\exp(-\beta\,E)\,\mathrm{d}p_1\,\mathrm{d}p_2\,\mathrm{d}p_3\,. \tag{6.20}$$

The integral of (6.20) over the momentum space yields N, whence

$$\frac{1}{C\,\Omega} = \iiint_{-\infty}^{+\infty} \exp(-\beta\,E)\,\mathrm{d}p_1\,\mathrm{d}p_2\,\mathrm{d}p_3\,. \tag{6.21}$$

The pressure P exerted by the gas is uniform over all the container's walls, hence it can be calculated with reference to any surface element $\mathrm{d}\Sigma$ belonging to them. One can then choose a surface element normal to the x_1 axis, placed in such a way that the atoms impinging on it travel in the positive direction. Let p_1 be the component of momentum of one of such atoms before hitting the wall; after the atom is reflected by the wall, the component transforms into $-p_1$, so that the variation in the momentum component along the x_1 direction is $2\,p_1$. The p_2, p_3 components, instead, are left unchanged. The product of $2\,p_1$ by the number of atoms hitting $\mathrm{d}\Sigma$ in the unit time provides the force $\mathrm{d}F$ exerted by the gas on the surface element, whence the pressure is obtained as $P = \mathrm{d}F/\Delta\Sigma$. Now, consider an elementary cylinder whose base and height are $\mathrm{d}\Sigma$ and $\mathrm{d}x_1$ respectively; the number $\mathrm{d}\tilde{N}_1$ of atoms that belong to such a cylinder and whose momentum p_1 belongs to $\mathrm{d}p_1$ is obtained as the product of the atoms' concentration $\mathrm{d}N_p/\Omega$ times the cylinder's volume $\mathrm{d}\Sigma\,\mathrm{d}x_1 = \mathrm{d}\Sigma\,\mathrm{d}t\,\mathrm{d}x_1/\mathrm{d}t = \mathrm{d}\Sigma\,(p_1/m)\,\mathrm{d}t$, integrated over the other two momenta p_2, p_3:

$$\mathrm{d}\tilde{N}_1 = \mathrm{d}\Sigma\,\frac{p_1}{m}\,\mathrm{d}t\,\mathrm{d}p_1\,C\,N\iint_{-\infty}^{+\infty} \exp(-\beta\,E)\,\mathrm{d}p_2\,\mathrm{d}p_3\,. \tag{6.22}$$

As each atom in (6.22) undergoes a change $2\,p_1$ in the momentum component due to the reflexion at the wall, the force $\mathrm{d}F = P\,\mathrm{d}\Sigma$ is obtained by multiplying (6.22) by $2\,p_1$, dividing it by $\mathrm{d}t$, and integrating over p_1 between 0 and $+\infty$; in fact, only the atoms that move towards the wall must be accounted for. Eliminating $\mathrm{d}\Sigma$ yields

$$\frac{P}{C\,N} = \frac{2}{m}\int_0^{+\infty}\iint_{-\infty}^{+\infty} p_1^2\,\exp(-\beta\,E)\,\mathrm{d}p_1\,\mathrm{d}p_2\,\mathrm{d}p_3\,. \tag{6.23}$$

Due to the form of (6.19), the integrals in (6.21) and (6.23) are products of one-dimensional integrals over p_1, p_2, and p_3. As a consequence, in the ratio between (6.23) and (6.21) the integrals over p_2 and p_3 cancel each other to yield

$$\frac{P\,\Omega}{N} = \frac{2}{m}\,\frac{\int_0^{+\infty} p_1^2\,\exp[-\beta\,p_1^2/(2\,m)]\,\mathrm{d}p_1}{\int_{-\infty}^{+\infty}\exp[-\beta\,p_1^2/(2\,m)]\,\mathrm{d}p_1}\,. \tag{6.24}$$

Letting Y indicates the integral at the denominator of (6.24), the integral at the numerator is found to be $-m\,\mathrm{d}Y/\mathrm{d}\beta$. Using (C.27) one finds

$$P\,\Omega = -2\,\frac{\mathrm{d}\sqrt{2\,\pi\,m/\beta}/\mathrm{d}\beta}{\sqrt{2\,\pi\,m/\beta}} = \frac{N}{\beta}\,. \tag{6.25}$$

The assumption that the gas used in the derivation of (6.25) is dilute makes it possible to consider it as a perfect gas, for which the phenomenological relation $P\,\Omega = N\,k_B\,T$ holds, with $k_B = 1.38 \times 10^{-23}$ J/K the *Boltzmann constant* and T the gas temperature. Comparing with (6.25) yields

$$\beta = \frac{1}{k_B\,T}\,. \tag{6.26}$$

It can be shown that the validity of (6.26) is not limited to the case where the simple derivation shown here applies. Actually, (6.26) is found to hold for any system that follows the Maxwell-Boltzmann distribution law [136, Sect. 32] and also, as shown in Sects. 15.9.1, 15.9.5, for quantum systems in equilibrium.

6.5 Boltzmann Transport Equation

The expressions worked out in this section show the important role of the distribution function. It is then necessary to determine the equation fulfilled by it when the system is not in an equilibrium condition. The derivation is made in two steps; in the first one the interactions between molecules are neglected, in the second one they are accounted for. To start with the first step one observes that, due to the neglect of collisions, the only force acting on the molecules is that of an external field. To denote the position in the phase space it is convenient to use the symbol $\mathbf{s}$ introduced in Sect. 1.8. Here the symbol has a slightly different meaning, because the space is 6-dimensional instead of being n-dimensional. However, the relations (1.57, 1.58, 1.59) still hold. Applying to the μ-space the same reasoning used in Sect. 23.2 to find the continuity equation (23.3), and considering the case where no particles are generated or destroyed, yield

$$\frac{\partial f_\mu}{\partial t} + \mathrm{div}_{\mathbf{s}}\left(\dot{\mathbf{s}}\,f_\mu\right) = 0\,. \tag{6.27}$$

From (1.58, 1.59) it follows $\mathrm{div}_{\mathbf{s}}(\dot{\mathbf{s}}\,f_\mu) = \dot{\mathbf{s}} \cdot \mathrm{grad}_{\mathbf{s}} f_\mu$. Replacing the latter into (6.27) and using the Cartesian coordinates yields the *Boltzmann collisionless equation*[5]

[5] In plasma physics, (6.28) is also called *Vlasov equation* [107, Sect. 13.2].

$$\frac{\partial f_\mu}{\partial t} + \dot{\mathbf{r}} \cdot \mathrm{grad}_{\mathbf{r}} f_\mu + \dot{\mathbf{p}} \cdot \mathrm{grad}_{\mathbf{p}} f_\mu = 0\,. \tag{6.28}$$

From the meaning of a continuity equation it follows that $-\mathrm{div}_{\mathbf{s}}\left(\dot{\mathbf{s}} f_\mu\right)$ is the time variation of f_μ per unit volume $\mathrm{d}\omega$ of the μ-space. As f_μ depends on $\mathbf{r}$, $\mathbf{p}$, and t, (6.28) is recast in compact form as $\mathrm{d}f/\mathrm{d}t = 0$.

To accomplish the second step, namely, adding the effects of collisions, one observes that the latter produce a further time change in f_μ. In principle, one might incorporate such effects into $-\mathrm{div}_{\mathbf{s}}\left(\dot{\mathbf{s}} f_\mu\right)$; however, it is more convenient to keep the effects of collisions separate from those of the external field. In fact, assuming as before that the system under consideration is dilute, each molecule spends a relatively large fraction of time far enough from the other molecules not to suffer any interaction; in other terms, the time during which a molecule is subjected to the external field is much longer than that involved in a collision. For this reason it is preferable to write the continuity equation, when the collisions are accounted for, as

$$\frac{\partial f_\mu}{\partial t} + \dot{\mathbf{r}} \cdot \mathrm{grad}_{\mathbf{r}} f_\mu + \dot{\mathbf{p}} \cdot \mathrm{grad}_{\mathbf{p}} f_\mu = C\,, \tag{6.29}$$

called *Boltzmann Transport Equation*. In (6.29), term C indicates the time variation of f_μ per unit volume $\mathrm{d}\omega$ due to collisions, whereas $-\mathrm{div}_{\mathbf{s}}\left(\dot{\mathbf{s}} f_\mu\right)$ keeps the meaning of variation due to the external field. The compact form of (6.29) reads in this case

$$\frac{\mathrm{d}f_\mu}{\mathrm{d}t} = C\,, \qquad \int C\,\mathrm{d}\omega = 0\,. \tag{6.30}$$

where the second relation is due to the normalization condition (6.2). In the equilibrium condition the distribution function has no explicit dependence on time ($\partial f_\mu/\partial t = 0$) and depends on constants of motion only, so that $C = 0$. The condition $C = 0$ does not prevent collisions from happening; in fact, in the equilibrium condition the change in the state of two colliding particles is balanced by simultaneous state changes of other particles that occur in the same elementary volume $\mathrm{d}\omega$, in such a way that the distribution function is left unchanged (*principle of detailed balance*).

In the calculation of C it is assumed that collisions are of the *binary* type, namely, that they involve only two particles at the time because the probability of a simultaneous interaction of more than two particles is negligibly small. This hypothesis, along with the assumption of a short duration of the interactions, greatly simplifies the calculation of C. This issue will not be pursued further here, because it will be developed directly in the quantum case (Sect. 19.3.1). It is worth observing that in a general nonequilibrium condition it is $C \neq 0$; the second relation in (6.30) then indicates that the form of C must be such, that in the integration over the μ-space every elementary contribution to the integral is balanced by an opposite contribution.

When the system under consideration is made of charged particles, the external field that matters is the electromagnetic one; if the particles are identical to each other, (6.29) takes the form

$$\frac{\partial f_\mu}{\partial t} + \mathbf{u} \cdot \mathrm{grad}_{\mathbf{r}} f_\mu + e\,(\mathbf{E} + \mathbf{u} \wedge \mathbf{B}) \cdot \mathrm{grad}_{\mathbf{p}} f_\mu = C\,, \tag{6.31}$$

with $\mathbf{E}(\mathbf{r}, t)$ the electric field, $\mathbf{B}(\mathbf{r}, t)$ the magnetic induction, e the common value of the particles' charge, and $\mathbf{u}$ the velocity (Sect. 1.3.2).

6.6 Complements

6.6.1 Momentum and Angular Momentum at Equilibrium

In the introductory discussion about statistical equilibrium, carried out in Sect. 6.3, it has been assumed that the total momentum and angular momentum of the system vanish. To discuss this issue, consider the box-shaped container whose cross-section is shown in Fig. 6.1, filled with molecules identical to each other, having initial positions near the right wall of the container. If the initial velocities are normal to the wall and equal to each other, as schematically indicated by the arrows, the total momentum $\mathbf{P}$ of the particles at $t = 0$ is different from zero. If the left and right walls are perfectly reflecting and parallel to each other, the particles keep bouncing back and forth between the two walls, and the total momentum alternates between $\mathbf{P}$ and $-\mathbf{P}$. As the container's mass is large, absorbing the momentum $2\,\mathbf{P}$ leaves the stationary condition of the container unaltered. On the other hand, as remarked in Sect. 1.3, this picture should not be regarded as describing a "system of particles," because the latter have no mutual interaction. To establish the interaction one must assume that the initial velocities are slightly misaligned, or the walls are not exactly parallel, or both; in this way the molecules will start colliding with each other and, after some time, their velocities will not be parallel anymore. If each collision, and reflection at the walls, is energy conserving, the total energy of the system does not change; in contrast, opposite velocity components of different molecules compensate each other in the calculation of the total momentum, so that the latter will vanish after a sufficiently large number of collisions.[6] A similar argument is applicable to the case of the cylindrical container whose cross-section is shown in Fig. 6.2, where the initial positions and velocities of the molecules are such that all molecules would move along the square described by the arrows. In this case the total angular momentum with respect to the cylinder's axis is different from zero.

[6] If the walls are perfectly reflecting, and the collisions are elastic (Sect. 3.5), the molecular motions are reversible so that, in both examples of this section, the initial condition is recovered by reversing the direction of time. More comments about this are made in Sect. 6.6.4.

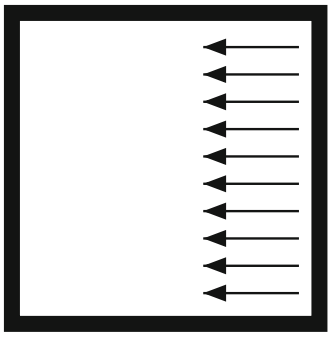


Fig. 6.1 Schematic picture used for discussing the issue of the total momentum of identical molecules within a container

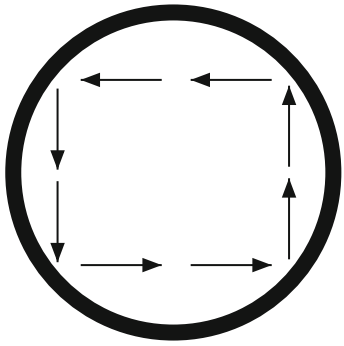


Fig. 6.2 Schematic picture used for discussing the issue of the total angular momentum of identical molecules within a container

A slight misalignment of the initial conditions, or a deviation of the container's wall from the perfectly cylindrical form, or both, will eventually make the molecules to collide with each other.

6.6.2 Averages Based on the Maxwell-Boltzmann Distribution

In a system made of classical particles in equilibrium at temperature T, each having R degrees of freedom, the average occupation number at energy E_r is given by (6.14). In general, the number of energy levels is large and their separation small, so that one disposes of the index and considers the number of particles belonging to the infinitesimal interval $\mathrm{d}q_1\,\mathrm{d}p_1 \ldots \mathrm{d}q_R\,\mathrm{d}p_R$ centered at $(q_1, p_1, \ldots, q_R, p_R)$. After dropping the index, the energy becomes a function of the position $(q_1, p_1, \ldots, q_R, p_R)$ of the interval's center, $E = E(q_1, p_1, \ldots, q_R, p_R)$; in turn, the Maxwell-Boltzmann distribution (6.14) takes the form $\exp(-\alpha - \beta E)$. Given these premises, and extending to the case of R degrees of freedom the definitions of Sect. 6.2, the statistical average over the Maxwell-Boltzmann distribution of a function $\zeta(q_1, p_1, \ldots, q_R, p_R)$ is

$$\mathrm{Av}[\zeta] = \frac{\int \ldots \int \zeta \, \exp(-\beta\, E)\, \mathrm{d}q_1\, \mathrm{d}p_1 \ldots \mathrm{d}q_R\, \mathrm{d}p_R}{\int \ldots \int \exp(-\beta\, E)\, \mathrm{d}q_1\, \mathrm{d}p_1 \ldots \mathrm{d}q_R\, \mathrm{d}p_R}\,, \tag{6.32}$$

where the factor $\exp(-\alpha)$ has been canceled out. An interesting case occurs when ζ depends on the generalized coordinates through the total energy only ($\zeta = \zeta(E)$) and, in turn, the energy is a positive-definite quadratic form of the coordinates,

$$E = a_1\, q_1^2 + b_1\, p_1^2 + \ldots + a_R\, q_R^2 + b_R\, p_R^2\,, \qquad a_i\,, b_i > 0\,. \tag{6.33}$$

Letting $n = 2\,R$ and using the Herring-Vogt transformation (17.66) yield

$$E = \eta/\beta\,, \qquad \eta = u_1^2 + \ldots + u_n^2\,, \tag{6.34}$$

where

$$u_1 = \sqrt{\beta\, a_1}\, q_1\,,\ u_2 = \sqrt{\beta\, b_1}\, p_1\,, \ldots\,, u_{n-1} = \sqrt{\beta\, a_R}\, q_R\,,\ u_n = \sqrt{\beta\, b_R}\, p_R\,. \tag{6.35}$$

In turn it is $\mathrm{d}u_1 \ldots \mathrm{d}u_n = c\, \mathrm{d}q_1\, \mathrm{d}p_1 \ldots \mathrm{d}q_R\, \mathrm{d}p_R$, with $c = \beta^R \sqrt{a_1\, b_1 \ldots a_R\, b_R}$. Using the procedure involving the density of states illustrated in Sect. B.5 yields

$$\mathrm{Av}[\zeta] = \frac{\int_0^{+\infty} \zeta(\eta)\, \exp(-\eta)\, b(\eta)\, \mathrm{d}\eta}{\int_0^{+\infty} \exp(-\eta)\, b(\eta)\, \mathrm{d}\eta} = \frac{\int_0^{+\infty} \zeta(\eta)\, \exp(-\eta)\, \eta^{n/2-1}\, \mathrm{d}\eta}{\int_0^{+\infty} \exp(-\eta)\, \eta^{n/2-1}\, \mathrm{d}\eta}\,, \tag{6.36}$$

where the last expression is obtained after canceling out the numerical factors appearing in (B.40). An important case of (6.36) is $\zeta = E = \eta/\beta$, which yields the average energy of the particles. Remembering (C.95, C.96), and using (6.26), one finds

$$\mathrm{Av}[E] = k_B\, T\, \frac{\Gamma(n/2+1)}{\Gamma(n/2)} = \frac{n}{2}\, k_B\, T = R\, k_B\, T\,. \tag{6.37}$$

The physical systems where the energy is a quadratic form of the type (6.33) are made of linear-harmonic oscillators, like those deriving from the diagonalization of the Hamiltonian function of a system of particles near the mechanical-equilibrium point (Sect. 3.10), or from the expression of the energy of the electromagnetic field *in vacuo* in terms of modes (Sect. 5.6). For systems of this type the average energy of the particles equals $k_B\, T$ times the number R of degrees of freedom of each particle.

Another important system is the dilute one, where the energy is essentially kinetic. In this case the form of the latter is found by letting $a_i \to 0$ in (6.33), so that the energy is made of a sum of R terms instead of $n = 2\,R$. Thus, the average energy of the particles is given by an expression of the form (6.37) where $n/2$ is replaced with $R/2$:

$$\mathrm{Av}[E] = k_B\, T\, \frac{\Gamma(R/2+1)}{\Gamma(R/2)} = R\, \frac{k_B\, T}{2}\,. \tag{6.38}$$

The above shows that in a dilute system the average energy of the particles equals $k_B T/2$ times the number R of degrees of freedom of each particle. The result expressed by (6.37) or (6.38) is called *principle of equipartition of energy*.

6.6.3 *Boltzmann's* H-*Theorem*

A measure of the extent by which the condition of a system departs from the equilibrium case is given by Boltzmann's H_B quantity, whose definition is that of statistical average of $\log f_\mu$. The case considered here refers to a spatially homogeneous system in the absence of external forces, so that $\mathrm{d}f_\mu/\mathrm{d}t = \partial f_\mu/\partial t$. Remembering (6.4) one finds

$$H_B(t) = \mathrm{Av}[\log f_\mu] = \frac{1}{N} \int f_\mu \, \log f_\mu \, \mathrm{d}\omega \,, \tag{6.39}$$

whose time derivative, using (6.30), reads

$$\frac{\mathrm{d}H_B}{\mathrm{d}t} = \frac{1}{N} \int \left(\frac{\partial f_\mu}{\partial t} \, \log f_\mu + \frac{\partial f_\mu}{\partial t} \right) \mathrm{d}\omega = \int C \, \log f_\mu \, \mathrm{d}\omega \,. \tag{6.40}$$

As indicated in Sect. 6.5, the collision term will be worked out directly in the quantum case. However, it is worth anticipating that the analysis of the collision term C leads, both in the classical and quantum cases, to an important conclusion: the time derivative $\mathrm{d}H_B/\mathrm{d}t$ is negative for any distribution function f_μ different from the equilibrium one, while it vanishes in the equilibrium case. This result is the *Boltzmann H-theorem*. This implies that if a uniform system is initially set in a condition described by a nonequilibrium distribution function, and the external forces are removed, then the initial distribution cannot be stationary: an equilibration process occurs, that brings the distribution to the equilibrium one, and whose duration is dictated by the time constants typical of C. The decrease of H_B with respect to time while the system reaches the equilibrium condition reminds one of the behavior of entropy. In fact, it can be shown that H_B is the entropy apart from a negative multiplicative factor and an additive constant [141, Sect. 18.3].[7]

The demonstration of the Boltzmann H-theorem is given below considering for simplicity the partitioning of the phase space into a discrete set of cells, like that used in Sect. 6.3. As before, the number of particles of the system is N; the cells are labeled by a single index. Letting $N_r = N_r(t)$ be the number of particles belonging to the rth cell at time t, the fraction $P_r(t)$ of particles belonging to the same cell is

[7] Compare with the definition of entropy given in Sect. 15.9.1 which, at first sight, looks different. The equivalence between the two definitions is justified in Sects. 47 and 102 of [136].

$$0 \leq P_r = \frac{N_r}{N} \leq 1\,. \tag{6.41}$$

From $\sum_r N_r = N$, the normalization condition $\sum_r P_r = 1$ follows. Using the above definitions, the statistical average of any dynamic quantity G in the discrete case is (compare with (6.4))

$$\mathrm{Av}[G](t) = \frac{\sum_r N_r\, G_r}{\sum_r N_r} = \sum_r P_r\, G_r\,, \tag{6.42}$$

with G_r the value of G in the rth cell. Here it is $G = \log P$, and the corresponding average is (compare with (6.39))

$$H_B(t) = \mathrm{Av}[\log P] = \sum_r P_r\, \log P_r \leq 0\,. \tag{6.43}$$

Note that in some cells it is $P_r = 0$; however, in (6.43) the logarithm of P_r appears only in the product $P_r \log P_r$, which vanishes when $P_r \to 0$. Similarly, the product vanishes for $P_r = 1$, whereas it is negative for $0 < P_r < 1$ (Fig. 6.5). It follows that H_B is by construction a non-positive quantity. Observing that $N > 0$, the only condition for H_B to vanish occurs when all particles belong to the same cell. To examine the time evolution of H_B one calculates the derivative

$$\frac{\mathrm{d}H_B}{\mathrm{d}t} = \sum_r (1 + \log P_r)\, \frac{\mathrm{d}P_r}{\mathrm{d}t}\,. \tag{6.44}$$

In turn, the time derivative of P_r is obtained from a balance equation: let $W_{sr} \geq 0$ be the unconditional probability per unit time that a particle makes a transition from the sth to the rth cell, $s \neq r$. Thus, the number of particles that in the unit time make a transition from the sth to the rth cell is $W_{sr} N_s$; adding up the above over all cells different from r provides the increase per unit time in the population of the rth cell. In parallel to this, there are particles that make a transition from the rth cell to any different cell, thus contributing to the decrease of N_r; letting W_{rs} be the unconditional probability per unit time of a transition from the rth to the sth cell, the balance equation at time t reads

$$\frac{\mathrm{d}N_r}{\mathrm{d}t} = \sum_s W_{sr}\, N_s - \sum_s W_{rs}\, N_r\,. \tag{6.45}$$

In the sums at the right-hand side of (6.45) the prescription $s \neq r$ is not necessary because the two summands corresponding to $s = r$ cancel each other. The quantity W_{sr} depends on the interactions of the particles among each other and with external perturbing agents. If the system is assumed to be isolated, W_{sr} is determined only by the mutual interactions of the particles. The calculation shows that in this case W_{sr} is independent of time; also, it can be shown that W_{sr} is invariant upon the exchange

of the indices,[8] namely, $W_{rs} = W_{sr}$. Using this result in (6.45) and dividing both sides by N yield, due to (6.41),

$$\frac{\mathrm{d}P_r}{\mathrm{d}t} = \sum_s W_{sr} \left(P_s - P_r\right) . \tag{6.46}$$

Inserting the above into (6.44),

$$\frac{\mathrm{d}H_B}{\mathrm{d}t} = \sum_{rs} \left(1 + \log P_r\right) \, W_{sr} \left(P_s - P_r\right) . \tag{6.47}$$

Now, considering that both indices in (6.47) span over all cells of the phase space, the derivative $\mathrm{d}H_B/\mathrm{d}t$ is equally well represented by an expression derived from (6.47) after exchanging r and s. Observing that the order of the indices in the sum and in W_{sr} is immaterial, one finds

$$\frac{\mathrm{d}H_B}{\mathrm{d}t} = \sum_{rs} \left(1 + \log P_s\right) \, W_{sr} \left(P_r - P_s\right) . \tag{6.48}$$

Finally, $\mathrm{d}H_B/\mathrm{d}t$ is given another expression, obtained as half the sum of (6.47) and (6.48); such a procedure has the advantage of eliminating the unity at the right-hand side, and yields

$$\frac{\mathrm{d}H_B}{\mathrm{d}t} = \frac{1}{2} \sum_{rs} \left(\log P_r - \log P_s\right) \, W_{sr} \left(P_s - P_r\right) . \tag{6.49}$$

Remembering that W_{sr} is nonnegative, and that the logarithm is a monotonic function of the argument, one draws the expected conclusion $\mathrm{d}H_B/\mathrm{d}t \leq 0$.

6.6.4 Paradoxes—Kac-Ring Model

It is known that the Boltzmann Transport Equation (6.29), and the H-theorem derived from it, bring about two apparent paradoxes: the first one is that the equation contains irreversibility, because any initial distribution function, different from the equilibrium one, evolves towards the equilibrium distribution when the external forces are removed, whereas an opposite evolution never occurs. This outcome is in contrast with the symmetry of the laws of mechanics with respect to time reversal (Sect. 2.6.1). The second paradox is the violation of Poincaré's time recurrence, which states that every finite mechanical system returns to a state arbitrarily close to

[8] The demonstration is carried out in the quantum case (Chap. 14) using the first-order perturbation method.

the initial one after a sufficiently long time (called *Poincaré cycle*); this is forbidden by the H-theorem that prevents entropy from decreasing back to the initial value.

A discussion about the mathematical origin of the paradoxes can be found, e.g., in [141, Sect. 18.4]; a qualitative insight into the question is given by a simple model, called *Kac's ring model*, also reported in [141] and taken from [76]. In the model, N objects are uniformly distributed over a circle, so that at time $t = 0$ each object is ascribed to a specific arc. The objects have two possible states, say, either "0" or "1." The time variable is discrete so that, when time evolves from $k\,\Delta t$ to $(k+1)\,\Delta t$, $k = 0, 1, 2, \ldots$, each object moves clockwise from the arc it occupied at time $k\,\Delta t$ to the next arc. A number $n < N$ of markers is present along the circle: specifically, the markers' positions are at the junctions between two neighboring arcs. The objects that cross the position of a marker change the state from "0" to "1" or vice versa; those that do not cross the position of a marker keep their state.

Given the number of objects and markers, the initial state of each object, and the markers' positions along the circle, one wants to investigate the time evolution of the states. Such an evolution is obviously time reversible and fulfills Poincaré's time recurrence; in fact, the set of objects goes back into the initial condition after N time steps if n is even, and after $2N$ time steps if n is odd.

Providing the time evolution of the individual object's state is in fact a microscopic description of the system; as remarked in Sect. 6.2, such a description becomes impossible when the number of objects in the system is large. A less detailed, macroscopic description of the Kac ring consists, for instance, in providing the time evolution of the number of "0" states. However, the outcome of the latter analysis seems to indicate that an irreversible process takes place; for instance, Fig. 6.3 shows a computer calculation of the time evolution of the number of "0" states in two samples made of $N = 4{,}000$ objects, which at time $t = 0$ were all set to "0." The markers of the two samples are $n = 4$ and $n = 8$, respectively, and the number of time steps is much smaller than N. Both curves tend to decrease and, after some fluctuations (that depend on the markers' positions), stabilize to a constant value; the same behavior occurs at a larger number of markers, although the number of time steps necessary to reach a constant value increases (curve $n = 16$ in Fig. 6.4). A further increase in the number of markers makes the fluctuations more pronounced (curve $n = 32$ in the same figure) (Fig. 6.5).

On the other hand, a similar calculation using a number of time steps larger than the number of objects shows that the stabilization at or around a constant value is eventually lost: the system fulfills Poincaré's time recurrence and recovers the initial condition (Fig. 6.6). Such an outcome is not detectable in real many-body systems, because the Poincaré cycle is enormously long with respect to the typical time scales of experiments.[9]

A detailed analysis of the dilemma is given in [58], where the time evolution of the Kac-ring model is calculated as follows: let F_k be the number of objects that at

[9] A crude estimate of the Poincaré cycle yields $\sim \exp(N)$, with N the total number of molecules in the system [65, Sect. 4.5]. In typical situations such a time is longer than the age of the universe.

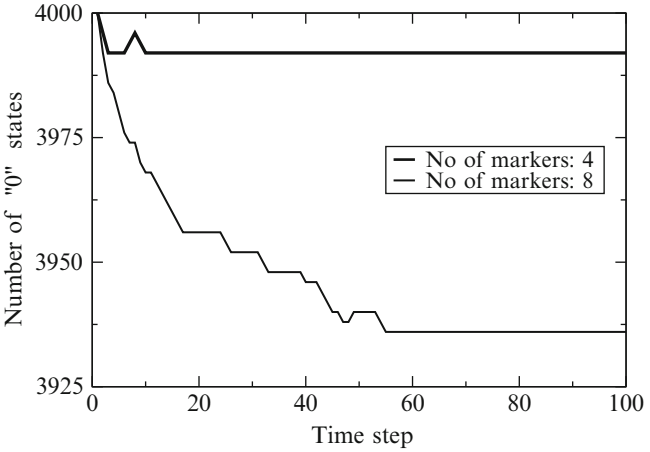


Fig. 6.3 Kac-ring model: computer calculation of the time evolution of the number of "0" states in two samples made of $N = 4{,}000$ objects, which at time $t = 0$ were all set to "0." The markers of the two samples are $n = 4$ and $n = 8$, respectively, and the number of time steps is much smaller than N

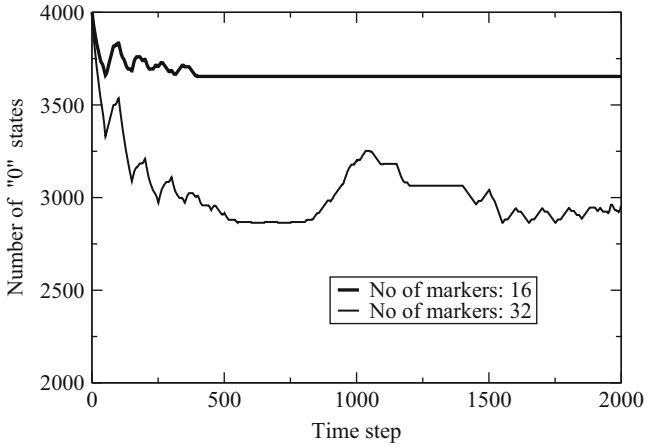


Fig. 6.4 Kac-ring model: computer calculation of the time evolution of the number of "0" states in two samples made of $N = 4{,}000$ objects, which at time $t = 0$ were all set to "0." The markers of the two samples are $n = 16$ and $n = 32$, respectively, and the number of time steps is much smaller than N

time $k\,\Delta t$ are in the "0" state, and let $f_k \leq F_k$ be the number of objects that at time $k\,\Delta t$ are in the "0" state and sit in front of a marker; similarly, let $V_k = N - F_k$ be the number of objects that at time $k\,\Delta t$ are in the "1" state, and let $v_k \leq V_k$ be the number of objects that at time $k\,\Delta t$ are in the "1" state and sit in front of a marker. From the definitions it follows

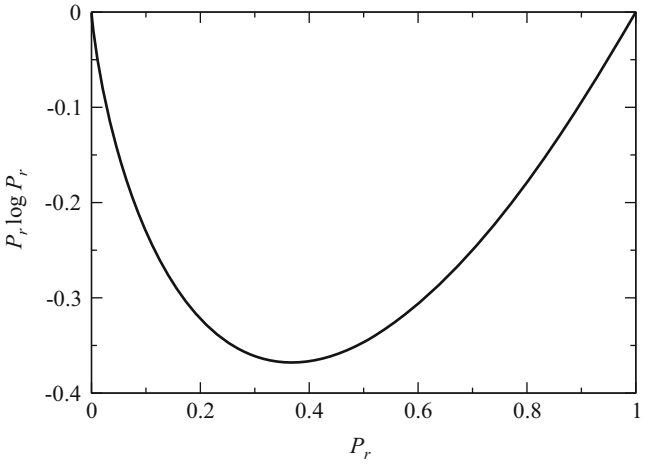


Fig. 6.5 Graph of $P_r \log P_r$

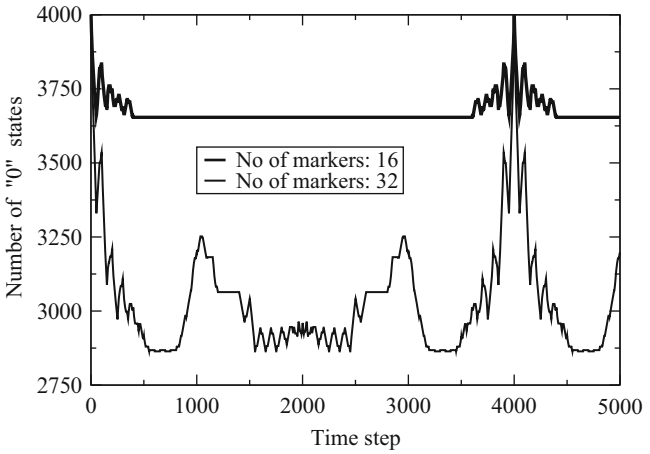


Fig. 6.6 Kac-ring model: computer calculation of the time evolution of the number of "0" states in two samples made of $N = 4{,}000$ objects, which at time $t = 0$ were all set to "0." The markers of the two samples are $n = 16$ and $n = 32$, respectively, and the number of time steps is larger than N

$$F_{k+1} = F_k + v_k - f_k\,, \qquad V_{k+1} = V_k - v_k + f_k\,, \qquad D_{k+1} = D_k + 2\,v_k - 2f_k\,, \tag{6.50}$$

where $D_k = F_k - V_k$. As remarked above, F_k, V_k are *macroscopic* quantities, describing global features of the system under examination. In contrast, f_k, v_k are *microscopic* quantities; in fact, they can be computed only if the positions of the markers and the state of the objects at every arc are known. The point is that there is no way to eliminate f_k, v_k from (6.50) as they stand; as the set of available equations

is not closed, the time evolution of the macroscopic quantities cannot be computed without introducing further relations.

To proceed, one introduces the hypothesis of a random distribution of markers, so that the probability that a marker is present at the edge of an arc, in the clockwise direction, is $P = n/N$. It follows $f_k = P\,F_k$ and $v_k = P\,V_k$, whence the third relation in (6.50) becomes $D_{k+1} = C\,D_k$, with $C = 1 - 2\,P$. As the result holds for any index, the time evolution follows in the form

$$D_k = C^k\,D_0\,, \qquad |C| < 1\,. \tag{6.51}$$

where the inequality derives from $0 < P < 1$. In (6.51) it is $\lim_{k\to\infty} D_k = 0$, namely, the additional hypothesis makes the system to reach an equilibrium configuration in which the populations of "0" and "1" states are equal: in fact, the dynamics described by (6.51) is irreversible. In conclusion, the hypothesis of a random distribution of markers has provided the closure condition of the problem, at the cost of disregarding the past history of the system's evolution (besides, D_k in (6.51) is not an integer any more). The basic idea is that such a random distribution of markers represents the typical behavior of large rings. More discussion on this issue is found in [58].

6.6.5 *Equilibrium Limit of the Boltzmann Transport Equation*

As remarked in Sect. 6.5, in the equilibrium condition the distribution function has no explicit dependence on time ($\partial f_\mu/\partial t = 0$) and depends on constants of motion only, so that $C = 0$. From (6.29) it then follows that $\dot{\mathbf{r}}\cdot\mathrm{grad}_{\mathbf{r}} f_\mu + \dot{\mathbf{p}}\cdot\mathrm{grad}_{\mathbf{p}} f_\mu = 0$. In case of a conservative system, energy is the most natural constant of motion to be used; in fact it is $H(\mathbf{r},\mathbf{p}) = E = \text{const}$, with H the Hamiltonian function (Sect. 1.6). From $f_\mu = f_\mu(H)$ one derives $\mathrm{grad}_{\mathbf{r}} f_\mu = (\mathrm{d}f_\mu/\mathrm{d}H)\,\mathrm{grad}_{\mathbf{r}} H$ and $\mathrm{grad}_{\mathbf{p}} f_\mu = (\mathrm{d}f_\mu/\mathrm{d}H)\,\mathrm{grad}_{\mathbf{p}} H$, so that the equilibrium limit of the Boltzmann Transport Equation reads

$$\frac{\mathrm{d}f_\mu}{\mathrm{d}H}\left(\dot{\mathbf{r}}\cdot\mathrm{grad}_{\mathbf{r}} H + \dot{\mathbf{p}}\cdot\mathrm{grad}_{\mathbf{p}} H\right) = 0\,, \tag{6.52}$$

Apart from the trivial case $f_\mu = \text{const}$, it is $\mathrm{d}f_\mu/\mathrm{d}H \neq 0$. On the other hand, recasting the Hamilton equations (1.42) as $\dot{\mathbf{r}} = \mathrm{grad}_{\mathbf{p}} H$, $\dot{\mathbf{p}} = -\mathrm{grad}_{\mathbf{r}} H$, and replacing them in (6.52), reduces the term in parentheses to $-\dot{\mathbf{r}}\cdot\dot{\mathbf{p}} + \dot{\mathbf{p}}\cdot\dot{\mathbf{r}}$, thus showing that the equation is fulfilled identically regardless of the explicit form of $f_\mu(H)$. This result implies that the equilibrium distribution (6.14) cannot be extracted solely from the equilibrium limit of (6.29); its derivation requires also the maximization procedure described in Sects. 6.3 and 6.4.

The equilibrium limit of the Boltzmann Transport Equation described in this section applies in general; as a consequence, it includes also cases like that of (6.31), where a magnetic force is involved. This seems to contradict one of the hypotheses used in Sect. 6.3 to derive the equilibrium distribution, namely, that the system is acted upon by conservative forces. To clarify the issue one uses the concepts illustrated in Sects. 1.5,1.6; first, one notes that the equilibrium limit is achieved by making the external force independent of time, so that the scalar and vector potentials whence **E** and **B** derive are independent of time as well: $\varphi = \varphi(\mathbf{r})$, $\mathbf{A} = \mathbf{A}(\mathbf{r})$ (the dependence on space is kept to account for the possibility that the system under consideration is not uniform). It follows that the Hamiltonian function, which in this case is given by (1.35), is still a constant of motion; as a consequence, the procedure leading to (6.52) is applicable. One notes in passing that each summand in the resulting identity $-\dot{\mathbf{r}} \cdot \dot{\mathbf{p}} + \dot{\mathbf{p}} \cdot \dot{\mathbf{r}}$ becomes in this case

$$\dot{\mathbf{p}} \cdot \mathbf{u} = e\,(\mathbf{E} + \mathbf{u} \wedge \mathbf{B}) \cdot \mathbf{u} = e\,\mathbf{E} \cdot \mathbf{u}\,, \tag{6.53}$$

where the mixed product vanishes due to (A.32).

Problem

6.1 Calculate the average energy like in (6.37) assuming that energy, instead of being a continuous variable, has the discrete form $E_n = n h \nu$, $n = 0, 1, 2, \ldots$, $h\nu = \text{const}$. This is the hypothesis from which Planck deduced the black-body's spectral energy density (Sect. 7.4.1).

Chapter 7
From Classical Mechanics to Quantum Mechanics

7.1 Introduction

This chapter tackles the difficult problem of bridging the concepts of Classical Mechanics and Electromagnetism with those of Quantum Mechanics. The subject, which is fascinating *per se*, is illustrated within a historical perspective, covering the years from 1900, when Planck's solution of the black-body radiation was given, to 1926, when Schrödinger's paper was published.

At the end of the 1800s, the main branches of physics (mechanics, thermodynamics, kinetic theory, optics, electromagnetic theory) had been established firmly. The ability of the physical theories to interpret the experiments was such, that many believed that all the important laws of physics had been discovered: the task of physicists in the future years would be that of clarifying the details and improving the experimental methods. Fortunately, it was not so: the elaboration of the theoretical aspects and the refinement in the experimental techniques showed that the existing physical laws were unable to explain the outcome of some experiments, and could not be adjusted to incorporate the new experimental findings. In some cases, the theoretical formulations themselves led to paradoxes: a famous example is the *Gibbs entropy paradox* [87]. It was then necessary to elaborate new ideas that eventually produced a consistent body generally referred to as *modern physics*. The elaboration stemming from the investigation of the microscopic particles led to the development of Quantum Mechanics, that stemming from investigations on high-velocity dynamics led to Special Relativity.

This chapter starts with the illustration of the planetary model of the atom, showing that the model is able to justify a number of experimental findings; this is followed by the description of experiments that cannot be justified in full by the physical theories existing in the late 1800s: stability of the atoms, spectral lines of excited atoms, photoelectric effect, spectrum of the black-body radiation, Compton effect. The solutions that were proposed to explain such phenomena are then illustrated; they proved to be correct, although at the time they were suggested a comprehensive theory was still lacking. This part is concluded by a heuristic

M. Rudan, *Physics of Semiconductor Devices*, DOI 10.1007/978-3-319-63154-7_7

derivation of the time-independent Schrödinger equation, based upon the analogy between the variational principles of Mechanics and Geometrical Optics.

In the final part of this chapter the meaning of the wave function is given: for this, an analysis of the measuring process is carried out first, showing the necessity of describing the statistical distribution of the measured values of dynamic quantities when microscopic particles are dealt with; the connection with the similar situations involving massive bodies is also analyzed in detail. This chapter is concluded with the illustration of the probabilistic interpretation of the wave function.

7.2 Planetary Model of the Atom

Several experiments were carried out in the late 1800s and early 1900s, whose outcome was the determination of a number of fundamental constants of atomic physics; among them, the *electron charge-to-mass ratio* was measured by J. J. Thomson in 1897, and *electron charge* was measured by R. Millikan in 1909 (Table D.1). A theory of the atom was proposed by E. Rutherford after a series of experiments in 1909–1914, that led to the measurement of the atomic radius r_a. The experiments consisted in measuring the broadening of beams of finely collimated α particles[1] passing through thin metal foils. The latter were typically made of gold sheets with a thickness of a few thousand atomic layers; the dynamics of the interaction between an α particle and the metal foil was treated by Classical Mechanics, using an interaction of the Coulomb type (Sect. 3.8). The outcome of the experiments led Rutherford to conceive the *planetary model* of the atom; this model depicts the atom as made of a small nucleus of atomic charge $Z\,q$ surrounded by Z electrons, where q is the absolute value of the electron charge and Z indicates the position of the element in the periodic table. The model assumes that as the α-particles are rather heavy, they are deflected mainly by the nuclei[2] of the foil; the type of deflection implies that the majority of the α particles is deflected only once when crossing the foil, thus indicating that the foil's atoms are placed far apart from each other. In fact, the deflection experiments made it possible to estimate[3] the atom's and nucleus' diameters, respectively, as

$$r_a \approx 0.1\ \text{nm}\,, \qquad r_e \approx 2 \times 10^{-6}\,\sqrt{Z}\ \text{nm}\,, \qquad Z = 1, 2, \ldots\,. \tag{7.1}$$

[1] These particles are obtained by ionizing helium atoms.

[2] The meaning of the term "nucleus" in this context needs to be clarified. Throughout this book the term is used to indicate the system made of protons, neutrons, and *core electrons*, namely, those electrons that do not belong to the outer shell of the atom and therefore do not participate in the chemical bonds. In solid-state materials, core electrons are negligibly perturbed by the environment, in contrast to the electrons that belong to the outer shell (*valence electrons*).

[3] The estimate means that for scale lengths equal or larger than those indicated in (7.1), an atom or a nucleus can be considered as geometrical points having no internal structure. The electron's radius can be determined in a similar way using X-ray diffusion.

As Z ranges from 1 to about 100, the second relation in (7.1) shows that $r_e \ll r_a$ (compare with Table D.1).

The simplest atom is that of hydrogen. The planetary model depicts it as an electron moving nearby the proton under the effect of a Coulomb potential V, where the energy reference is chosen in such a way as to make $V(\infty) = 0$. Then, the theory of the Coulomb interaction in the attractive case applies (Sect. 3.13.6). As the proton is much more massive than the electron, its position can be approximated with that of the atom's center of mass, and placed in the origin; in summary, it is

$$V(\mathbf{r}) = V(r) = -\frac{q^2}{4\,\pi\,\varepsilon_0\,r}\,, \qquad T+V = \frac{1}{2}\,m\,u^2 - \frac{q^2}{4\,\pi\,\varepsilon_0\,r} = E = \text{const}\,, \tag{7.2}$$

where $\mathbf{r}$ is the electron position, u its velocity's module, T and E its kinetic and total energies, respectively, and ε_0 the permittivity of vacuum. The planetary model is extended to more complicate atoms by considering an outer electron moving nearby a core of net charge q embedding Z protons and $Z-1$ electrons, or to even more complicate cases (*hydrogenic-like systems*). Observing that $T = E - V = E + |V| \geq 0$, and remembering the analysis of Sect. 3.8, one finds that two cases are possible: the first one, shown in Fig. 7.1, is $E \geq 0$, corresponding to $V \leq 0 \leq E$ and $r_{\max} = \infty$ (*free electron*). The second case is $E < 0$, corresponding to $V \leq E < 0$ and $r_{\max} < \infty$ (*bound electron*). For the qualitative reasoning to be carried out here, it is sufficient to consider the simpler case of a bound electron whose trajectory is circular (Fig. 7.2). In such a case, using the results of Sects. 3.7 and 3.13.6 after letting $Z_1 = Z_2 = 1$ and replacing s with r, yields $r = 4\,\pi\,\varepsilon_0\,M_B^2/(m\,q^2) = \text{const}$, where $M_B^2 = m^2\,r^2\,u^2$. Combining these relations with the first one in (7.2) shows that $F = |\mathbf{F}| = m\,|\mathbf{a}| = m\,u^2/r = 2\,T/r$ whence, using $F = q^2/(4\,\pi\,\varepsilon_0\,r^2) = -V/r$, one finds

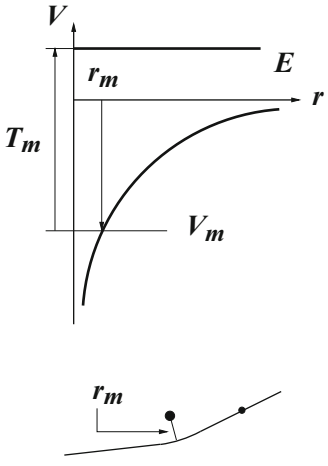


Fig. 7.1 Classical description of the electron's orbit for $E \geq 0$

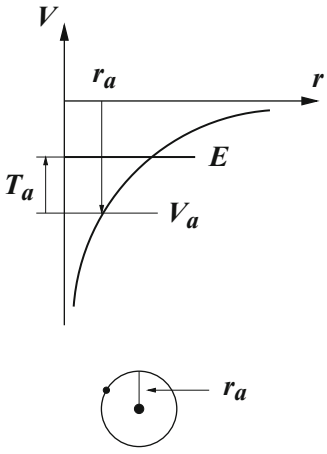


Fig. 7.2 Classical description of the electron's orbit for $E < 0$. For simplicity, a circular orbit is considered

$$T = -\frac{V}{2}\,, \qquad E = T + V = \frac{V}{2} = -\frac{q^2}{8\,\pi\,\varepsilon_0\,r} = \text{const} < 0\,. \tag{7.3}$$

It follows $\mathrm{d}E/\mathrm{d}r = |E|/r > 0$, that is, the total energy is larger at larger orbits (this is a particular case of the general theory worked out in Prob. 3.2).

Despite its simplicity, the planetary model is able to explain phenomena like the excitation and ionization of atoms; *excitation* corresponds to the absorption of energy from an external electromagnetic field, such that an initially bound electron increases its energy from E_1 to E_2, where $E_1 < E_2 < 0$: the electron in the final state is still bound. The inverse process is the emission of energy in the form of an electromagnetic radiation, so that the electron's total energy decreases from E_2 to E_1. In turn, *ionization* corresponds to the absorption of energy such that an initially bound electron becomes free: $E_1 < 0$ and $E_2 \geq 0$. The inverse process is the capture of a free electron by an ionized atom, with an energy emission equal to $E_2 - E_1$.

The above reasoning can also be used to explain the behavior of systems more complicate than single atoms. For instance, consider a finite *linear monatomic chain*, namely, a system made of a finite number of identical atoms placed along a line, at equal mutual distances (Fig. 7.3), which can be thought of as a rudimental version of a crystal. The positions of the atoms are marked by the dots visible in the upper part of the figure. Let the chain be aligned with the x axis; if each nucleus is made to coincide with a local origin of the reference, the distance $r = \sqrt{x^2 + y^2 + z^2}$ in the expression of the potential energy becomes $|x|$; the potential energy pertaining to each nucleus is proportional to $1/|x|$ and is indicated with a dashed line in the figure. An electron placed at some position in the chain is subjected to the sum of the potential energies; being the latter negative, the sum is lower than the individual contributions: in the figure, it is indicated by the continuous line, which for simplicity is drawn by adding up the contributions of the nearest-neighboring nuclei only. The chain, however, has a finite length; when the leftmost nucleus is considered, the potential energy on its left does not suffer any

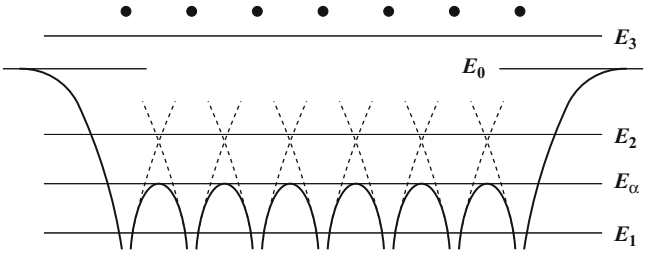


Fig. 7.3 Schematic description of the potential energy in a linear monatomic chain

lowering: this creates the energy step visible in the figure. The same happens on the right side of the rightmost nucleus. The shape of the potential energy thus obtained is able to qualitatively explain several features of crystals. For instance, consider by way of example the case where the only force acting on an electron inside the crystal derives from the potential energy of Fig. 7.3, namely, it is a conservative force. If the total energy of the electron under consideration is E_1, the electron's position is confined within the potential well dictated by the initial position of its motion. Thus, the electron oscillates within the well like in the example of Prob. 3.2, and its motion cannot extend out of it; if all electrons of the crystal are bound, the material is an insulator. This reasoning implies that the situation where all electrons are bound is maintained also under the application of an external voltage; due to this, no electric current ensues.

If the total energy of the electron under consideration is E_2, the electron can move within the whole crystal; finally, if the total energy is E_3, the electron overcomes one or the other of the energy steps and moves into vacuum: for this reason, the minimum energy E_0 necessary for the electron to leave the crystal is called *vacuum level*. If the two ends of the crystal are connected to a voltage generator by suitable contacts, and an external voltage is applied, the material can carry an electric current, whose amplitude depends also on the number of electrons whose energy is sufficiently high.[4] It is worth noting that although this is prohibited in the frame of Classical Mechanics, the electrons whose energy is of the type E_2 may also contribute to the current; in fact, Quantum Mechanics shows that they have a finite probability to penetrate the energy step and reach the contact. This phenomenon is called *tunnel effect* (Sect. 11.3.1).

To proceed it is convenient to give Fig. 7.3 a simpler appearance: in fact, considering that the interatomic distance is a fraction of a nanometer, the spatial extent in the x direction where each peak is placed is hardly visible in a macroscopic representation; for this reason, it is sufficient to graphically indicate the envelope E_α of the peaks. By the same token, the steps on the sides are described as discontinuities (Fig. 7.4). The electrons with $E < E_\alpha$ or $E \geq E_\alpha$ are called,

[4]The combination of the number of such electrons with other factors also determines whether the material is a conductor or a semiconductor (Chap. 18).

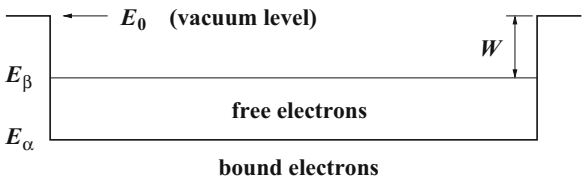


Fig. 7.4 The same structure of Fig. 7.3, where the peaks are replaced with the envelope

respectively, *bound electrons* and *free electrons*.[5] In the equilibrium condition the total energy of the electrons is prescribed; it follows that the majority of the electrons has an energy E lower than a given value E_β (which is not necessarily larger than E_α). Thus, the difference $W = E_0 - E_\beta$ is a measure of the energy that is necessary to extract an electron from the material: among other things, the model explains the existence of a minimum extraction energy of the electrons.[6]

7.3 Experiments Contradicting the Classical Laws

About 1900, experimental evidence was found for a number of phenomena that contradict the calculations based on the known physical laws, that is, the laws of Analytical Mechanics, Electromagnetism, and Statistical Mechanics. A number of such phenomena are listed in this section.

7.3.1 Stability of the Atom

The solution of the electromagnetic equations shows that an accelerated electron radiates a power given by (5.72), namely, $q^2\dot{v}^2/(6\,\pi\,\varepsilon_0\,c^3)$, with $\dot{v}$ the electron's acceleration. As discussed in Sect. 5.11.2, this is in contradiction with the planetary model of the atom (7.3), in which $T + V = E = \text{const}$: due to the radiated power, the electron should lose energy and, as a consequence, the atom should shrink. The possibility of an extremely slow, non-detectable shrinking must be ruled out: in fact, a perturbative calculation (Sect. 5.11.3) shows that due to the energy loss the atomic radius should decrease from the initial value, say, r, to the value r/e in about 10^{-8} s. This, however, is not observed (Sect. 9.7.2).

[5]The description is qualitative; for instance, it does not consider the band structure of the solid (Sect. 17.6).

[6]When the material is a conductor, E_β coincides with the Fermi level (Sect. 15.8.1), and W is called *work function*; in a semiconductor, E_β coincides with the lower edge E_C of the conduction band (Sect. 17.6.5) and the minimum extraction energy (typically indicated with a symbol different from W) is called *electron affinity* (Sect. 22.2).

7.3.2 *Spectral Lines of Excited Atoms*

The planetary model explains the emission of electromagnetic radiation by excited atoms. The explanation, however, is qualitative only, because the model does not impose any constraint on the frequency of the emitted waves. In contrast, the experiments show that the waves emitted by, e.g., hydrogen atoms have frequencies ν of the form (*Balmer law*, 1885),[7]

$$\nu_{nm} = \nu_R \left(\frac{1}{n^2} - \frac{1}{m^2} \right) , \qquad \nu_R \simeq 3.3 \times 10^{15} \ \mathrm{s}^{-1} , \tag{7.4}$$

where n, m are integers, $m > n \geq 1$. The emissions described by (7.4) are also called *spectral lines*. The lower series of spectral lines are shown in Fig. 7.5 along with their designations; the numbers in the figure correspond to n, m in (7.4). Another experimental finding shows that instead of occurring with a single emission of frequency ν_{nm}, the release of electromagnetic energy by the atom may be accomplished in steps; if that happens, the frequencies associated with the individual steps fulfill a relation called *Ritz emission rule*: considering, e.g., two steps, it reads

$$\nu_{nm} = \nu_{nk} + \nu_{km} , \tag{7.5}$$

with ν_{nk}, ν_{km} the frequencies of the individual steps.

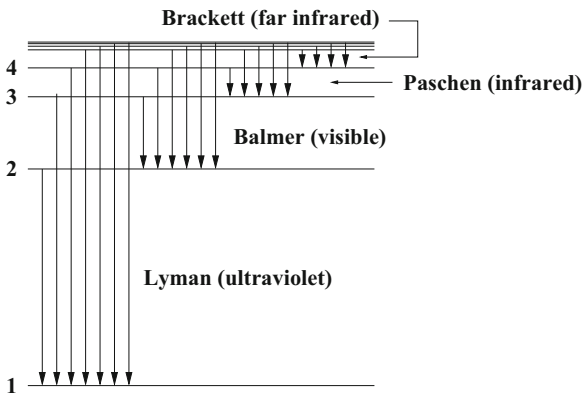


Fig. 7.5 Designation of the lower series of spectral lines (7.4)

[7]The ratio $R = \nu_R / c \simeq 1.1 \times 10^5 \ \mathrm{cm}^{-1}$ is called *Rydberg constant*. The formula was generalized in the 1880s to the hydrogenic-like atoms by Rydberg: the expression (7.4) of the frequencies must be multiplied by a constant that depends on the atom under consideration.

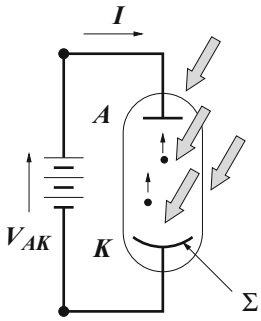


Fig. 7.6 Schematic cross-section of the apparatus used for measuring the photoelectric effect

7.3.3 Photoelectric Effect

It is found that an impinging electromagnetic radiation extracts charges from a metal (H. Hertz, 1887) and that these charges are electrons (J. J. Thomson, 1899). The phenomenon is ascribed to the absorption of energy from the radiation: the electron absorbs an energy sufficiently large to be extracted from the metal. An electron thus extracted is also called *photoelectron*. A sketch of the measuring apparatus is given in Fig. 7.6, where two electrodes, anode (A) and cathode (K), are placed inside a vacuum tube in order to prevent interactions between the photoelectrons and the atmosphere. A voltage V_{AK} is applied to the electrodes, such that $V_{AK} > 0$ when the electric potential at the anode is larger than that at the cathode. A monochromatic radiation of a given intensity is made to impinge on the cathode, whose surface is marked with Σ, and the current I flowing in the tube is recorded. Important parameters are the radiation's frequency ν, the spectral intensity of the radiation, $\eta = \mathrm{d}E/(\mathrm{d}\Sigma\,\mathrm{d}t\,\mathrm{d}\nu)$, where $\mathrm{d}\Sigma$ is the surface element of the cathode, and the spectral power[8]

$$\Pi = \int_{\Sigma} \eta \, \mathrm{d}\Sigma = \frac{\mathrm{d}E}{\mathrm{d}t\,\mathrm{d}\nu} \,. \tag{7.6}$$

The outcome of the experiment is shown in arbitrary units in Figs. 7.7 and 7.8. The first one shows a set of the $I = I(V_{AK})$ curves at constant ν, with Π a parameter. When V_{AK} is positive and sufficiently high, it is expected that practically all electrons extracted from the cathode be driven to the anode; as a consequence, the slope of the curves should be negligible. Also, when the intensity of the radiation increases, the number of extracted electrons, and the current with it, should also increase. This is in fact confirmed by the curves of Fig. 7.7. When, instead, V_{AK} is negative, only the electrons leaving the cathode with a sufficiently high kinetic energy are able to reach the anode, whereas those whose initial kinetic energy is low are repelled

[8]The units of η and Π are $[\eta] = \mathrm{J\,cm^{-2}}$ and $[\Pi] = \mathrm{J}$, respectively.

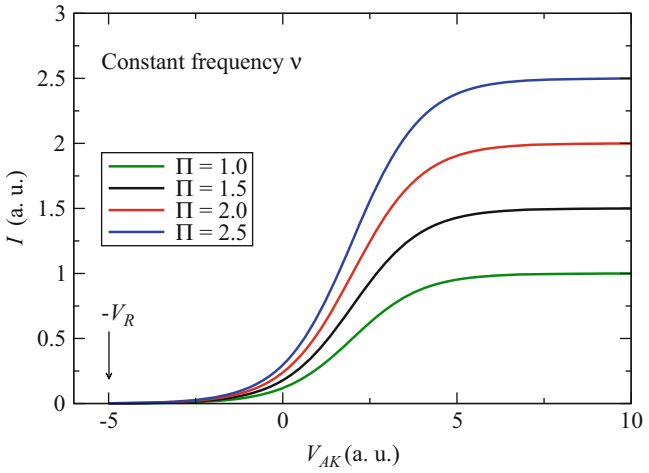


Fig. 7.7 The $I = I(V_{AK})$ curves, in arbitrary units, obtained from the photoelectric effect at constant frequency of the radiation, with the spectral power used as a parameter

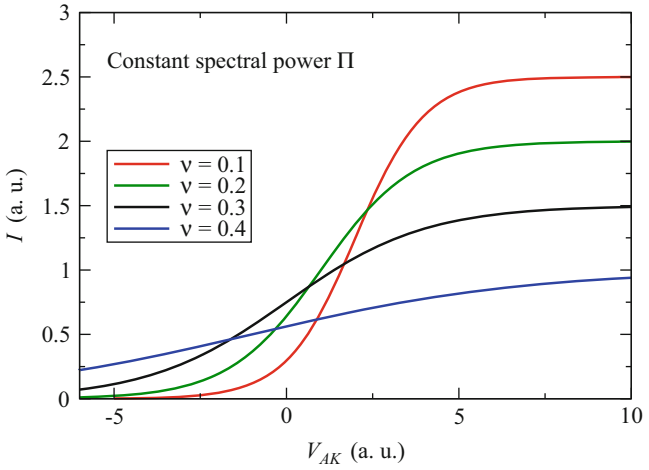


Fig. 7.8 The $I = I(V_{AK})$ curves, in arbitrary units, obtained from the photoelectric effect at constant spectral power of the radiation, with frequency used as a parameter

towards the cathode by the electric field imposed by the reverse bias.[9] Considering for simplicity a one-dimensional case, energy conservation yields for an electron traveling from cathode to anode,

[9]The concentration of electrons in the vacuum tube is small enough not to influence the electric field; thus, the latter is due only to the value of V_{AK} and to the form of the electrodes.

$$\frac{1}{2}\,m\,u_A^2 - \frac{1}{2}\,m\,u_K^2 = q\,V_{AK}\,, \tag{7.7}$$

where u_K is the electron's velocity at the cathode and u_A that at the anode. The *blocking voltage* $V_R > 0$ is the value $V_{AK} = -V_R$ such that $u_A = 0$; from (7.7) one obtains the relation

$$\frac{1}{2}\,m\,u_K^2 = q\,V_R\,, \tag{7.8}$$

which allows one to measure the kinetic energy of the most energetic electrons that are extracted from the cathode at given spectral power and frequency of the radiation.[10] Such electrons are those that inside the cathode have an energy in the vicinity of E_β (Fig. 7.4) and do not suffer energy losses while being extracted. If E_L is the energy that the most energetic electron absorbs from the radiation, its kinetic energy at the cathode is $(1/2)\,m\,u_K^2 = E_L - W$, with W the metal's work function, whence

$$q\,V_R = E_L - W\,, \tag{7.9}$$

so that the photoelectric effect provides in fact a method for measuring E_L. The classical model predicts that the blocking voltage should increase with Π; this, however, does not happen: as shown in Fig. 7.7, at a given frequency the blocking voltage is the same for all values of Π. In addition, it is unexpectedly found that both I and V_R depend on the frequency ν (Fig. 7.8). In fact, the comparison between the experimental blocking voltages and (7.9) shows that the energy E_L that the electron absorbs from the electromagnetic field is proportional to the frequency,

$$E_L = h\,\nu\,, \tag{7.10}$$

with $h \simeq 6.626 \times 10^{-34}$ J s the *Planck constant*. If $h\,\nu < W$, no current is measured; this provides a threshold value for the frequency to be used in the experiment.

7.3.4 Spectrum of the Black-Body Radiation

Consider a body at temperature T in equilibrium with an electromagnetic field. Due to the detailed-balance principle, the spectral intensity η_B emitted by the body, that is, the electromagnetic power emitted by it per unit surface element $\mathrm{d}\Sigma$ and unit frequency $\mathrm{d}\nu$, in the direction normal to $\mathrm{d}\Sigma$, fulfills the relation

[10] The most energetic electrons succeed in overcoming the effect of the reverse bias and reach the vicinity of the anode; they constantly slow down along the trajectory to the point that their velocity at the anode vanishes. Then, their motion reverses and they are driven back to the cathode.

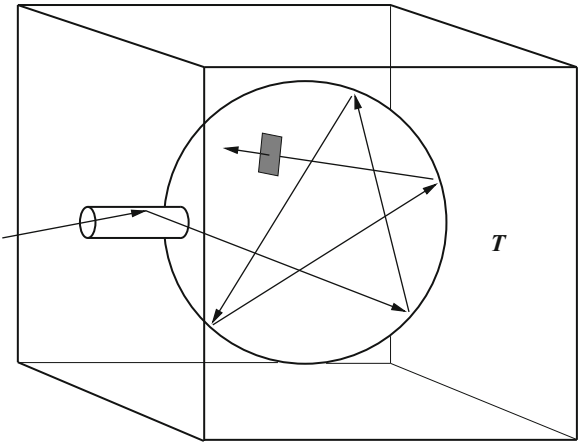


Fig. 7.9 The approximation to a black body consisting in a small hole in the wall of an enclosure kept at constant temperature. If a thermometer (represented by the *shaded area*) was suspended within the enclosure, it would indicate the same temperature T as the walls, irrespective of its position or orientation

$$\eta_B = \alpha\, \eta\,, \tag{7.11}$$

where η is the spectral intensity of the radiation (compare with (7.6)), and $0 \leq \alpha \leq 1$ the fraction of η absorbed by the body at frequency ν. By Kirchhoff's law (1859), for any body in thermal equilibrium with radiation it is

$$\frac{\eta_B}{\alpha} = K(\nu, T)\,, \tag{7.12}$$

where K is a universal function of ν and T [107, Sect. 9–15]. A *black body* is a body such that $\alpha = 1$ at all frequencies; thus, for a black body at equilibrium with radiation it is $\eta_B = K$. A good approximation to a black body is a small hole in the wall of an enclosure kept at constant temperature, like that illustrated in Fig. 7.9: any radiation entering the hole has a negligible probability of escaping, due to multiple reflections at the walls; as a consequence, the hole acts like a perfect absorber. Thanks to $\eta_B = K$, the spectral intensity emitted by any black body has the same characteristics: in particular, it is not influenced by the form of the enclosure, the material of which the walls are made, or other bodies present in the enclosure. As a consequence, η_B, or any other function related to it, can be calculated by considering a convenient geometry of the problem and assuming that the radiation propagates *in vacuo*. It is found experimentally that

$$\int_0^\infty \eta_B(\nu, T)\, \mathrm{d}\nu = \sigma\, T^4\,, \tag{7.13}$$

where $\sigma = 5.67 \times 10^{-12}\ \mathrm{W\ cm^{-2}\ K^{-4}}$ is the Stefan-Boltzmann constant.

One of the functions related to η_B is the *spectral energy density* u of the black body, that is, the energy per unit volume and frequency. The integral of u over the frequencies yields the energy density; remembering that equilibrium is assumed,

one finds[11]

$$w_{\text{em}}^{\text{eq}}(T) = \int_0^\infty u(\nu, T)\, \mathrm{d}\nu\,. \tag{7.14}$$

In turn, the integral of u over the coordinates gives the equilibrium value of the spectral energy, whose general definition is given by (5.47).

The spectral energy density u can be calculated as the product of the number of monochromatic components of the electromagnetic field per unit volume and frequency, times the energy of each monochromatic component. The first factor is readily found by taking an enclosure of prismatic form like that of Sect. 15.9.4; the calculation yields $8\,\pi\,\nu^2/c^3$, which is obtained by dividing both sides of (15.74) by the enclosure's volume V. As for the energy of each monochromatic component, the only assumption possible in the frame of Classical Mechanics is that the energy of the electromagnetic field at equilibrium is distributed over the frequencies according to the Maxwell-Boltzmann distribution (6.14). Assuming that each monochromatic component is equivalent to a one-dimensional linear-harmonic oscillator, the energy to be associated with it is the average energy of a system with one degree of freedom; thus, letting $R = 1$ in (6.37) yields for the average energy the value $k_B\,T$. The product of the two factors thus found yields for the spectral energy density of the black body the expression

$$u(\nu, T) = 8\,\pi\,\frac{k_B T}{c^3}\,\nu^2\,, \tag{7.15}$$

called *Rayleigh-Jeans law*. Experimental results for u as a function of frequency are shown in Fig. 7.10, with temperature a parameter. The comparison with experiments shows that the parabolic behavior of (7.15) approximates the correct form of the curves only at low frequencies; clearly the result expressed by (7.15) cannot be correct, because it makes the equilibrium energy density (7.14) to diverge.[12]

7.3.5 Compton Effect

When X-rays of a given frequency interact with atoms and are scattered with an angle ψ with respect to the direction of incidence, the frequency of the emerging rays is found to depend on ψ. This outcome is in contrast with the prediction of the electromagnetic theory, according to which the frequency of the scattered rays should be equal to that of the impinging ones. The dependence of frequency on the scattering angle is also called *Compton effect*.

[11] Compare with the general definition (5.10) of w_{em}, where the assumption of equilibrium is not made.

[12] This unphysical outcome is also called *ultraviolet catastrophe*.

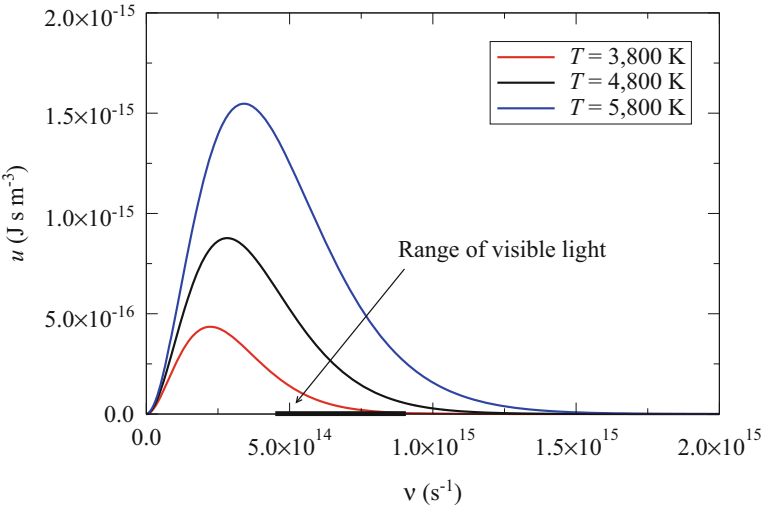


Fig. 7.10 Spectral energy density of the black body at different temperatures. The value $T = 5{,}800$ K corresponds to the surface temperature of the sun

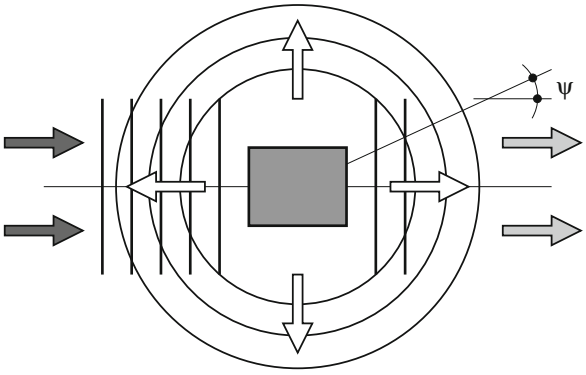


Fig. 7.11 Scheme of the experimental setup for measuring the Compton effect

The experimental setup for measuring the Compton effect is schematically shown in Fig. 7.11. The gray box in the middle of the figure is a piece of solid material, onto which the radiation impinges from the left (dark arrows); the vertical lines are the intersections of the constant-phase planes with the plane of the figure. The gray arrows on the right represent the part of the radiation that does not interact with the material and exits from it unaltered, while the white arrows indicate some of the directions of the rays scattered by the material. The circumferences are the intersections with the figure's plane of the spherical waves produced by the scattering. The origin of the measuring angle is aligned with the direction of the incoming radiation, so that $\psi = 0$ corresponds to the absence of scattering, $\psi = \pi$ to reflection.

7.4 Quantum Hypotheses

In the early 1900s, a number of hypotheses were made to solve the contradictions between the experimental evidence and the calculations based on the physical laws known at that time. The solutions thus found and the new concepts originating from them were eventually combined into a unified and coherent theory, Quantum Mechanics.

In essence, the contradictions with the physical laws known in the early 1900s were found thanks to the refinement of the experimental techniques. Such refinements were in turn made available by the general advancement of science that had taken place in the preceding decades. Thanks to them, it was possible to start investigating the microscopic world, namely, the dynamics of elementary particles. A parallel improvement took place in the same years in the investigation of the dynamics at high velocities, and led to the concepts of Special Relativity (1905).[13]

7.4.1 Planck's Solution of the Black-Body Problem

To explain the features of the black-body radiation, Planck made in 1900 the hypothesis that a monochromatic electromagnetic energy is absorbed or emitted only in quantities that are integer multiples of a fixed quantity $h\nu$, where h is a suitable constant [104]. The occupation number then becomes

$$P_n = P_0 \, \exp(-n\beta \, h \, \nu) \, , \qquad \beta = 1/(k_B \, T) \, . \tag{7.16}$$

As a consequence, using the same procedure as in Sect. 6.6.2 after replacing the integrals with sums, yield for the average energy $\mathrm{Av}[n \, h \, \nu]$ the expression[14]

$$\mathrm{Av}[n \, h \, \nu] = \frac{\sum_{n=0}^{\infty} n \, h \, \nu \, P_n}{\sum_{n=0}^{\infty} P_n} = \frac{h \, \nu}{\exp(\beta \, h \, \nu) - 1} \, . \tag{7.17}$$

In contrast with the constant value $k_B \, T$ used in the determination of the Rayleigh-Jeans law, here the average energy of each monochromatic component depends on the component's frequency. Multiplying (7.17) by the number $8 \, \pi \, \nu^2 / c^3$ of monochromatic components of the electromagnetic field per unit volume and frequency, found in Sect. 7.3, yields for the spectral energy density of the black body the expression

[13] As the particles' velocities that occur in solid-state physics are low, Special Relativity is not used in this book; the only exception is in the explanation of the Compton effect, illustrated in Sect. 7.4.3.

[14] The detailed calculation leading to (7.17) is shown in Prob. 6.1.

$$u(\nu, T) = 8\,\pi \frac{h\,\nu^3/c^3}{\exp[h\,\nu/(k_B\,T)] - 1}\,, \tag{7.18}$$

called *Planck law* (1900). The derivation of (7.18) involves one undetermined parameter, h. If the latter is made equal to the Planck constant introduced in the description of the photoelectric effect (Sect. 7.3), the resulting expression fits perfectly the experimental data like those of Fig. 7.10. Remembering that the spectral energy density of a black body in equilibrium is a universal function, it follows that h does not depend on the specific experiment, namely, it is a universal constant.

The low-frequency limit of (7.18), $h\,\nu \ll k_B\,T$, is independent of h and renders the Rayleigh-Jeans law (7.15).

7.4.2 Einstein's Solution of the Photoelectric Effect

In 1905, Einstein proposed the following explanation of the photoelectric effect: the transport of electromagnetic energy is *quantized*; specifically, a monochromatic electromagnetic wave of frequency ν is made of the flux of identical objects, called *photons*, each carrying the energy $h\,\nu$. In the interaction with a photon, an electron may absorb an energy up to $h\,\nu$. If the absorbed energy is exactly $h\nu$, the photon is annihilated [45].[15] This theory provides a correct explanation of the photoelectric effect: with reference to Fig. 7.7, the photoelectric current increases as the spectral power Π increases at constant ν, because the number of photons is larger: as a consequence, the number of photoelectrons is larger as well. In turn, with reference to Fig. 7.8, the blocking voltage V_R increases as ν increases at constant Π, because the photons are more energetic; however, they are fewer, which explains why the curves intersect each other: the spectral power, in fact, can be written as $\Pi = \mathrm{d}E/(\mathrm{d}t\,\mathrm{d}\nu) = h\,\nu\,[\mathrm{d}N/(\mathrm{d}t\,\mathrm{d}\nu)]$, where the quantity in brackets is the number of photons per unit time and frequency; as a consequence, the constraint $\Pi = \text{const}$ of the experiment of Fig. 7.8 makes the quantity in brackets to decrease when the photon energy $h\,\nu$ increases.

7.4.3 Explanation of the Compton Effect

The concept of photon, introduced in Sect. 7.4.2, explains the Compton effect by describing the interaction of the electron with the electromagnetic field as the collision between the electron and a photon [29]. As the photon's velocity

[15] Einstein's hypothesis is more general than Planck's: the latter, in fact, assumes that energy is quantized only in the absorption or emission events.

is c, its rest mass is zero (Sect. 3.13.7); in turn, the modulus of the photon's momentum is $p = E/c$, which is consistent with classical electromagnetism (compare with (5.43)).

The analysis of the electron-phonon collision is worked out assuming that the system made of the two particles under consideration is isolated; thus, the calculation is based upon the energy- and momentum-conservation equations, and the results of Sect. 3.13.8 hold. The dynamic quantities for the photon are given by

$$E = h\,\nu\,, \qquad p = \frac{E}{c} = \frac{h\,\nu}{c} = \frac{h}{\lambda}\,, \tag{7.19}$$

the second of which derives from (5.55) expressed *in vacuo*. Defining the *reduced Planck constant* $\hbar = h/(2\,\pi) \simeq 1.055 \times 10^{-34}$ J s, and using the modulus k of the wave vector, (7.19) becomes

$$E = \hbar\, 2\pi\, \nu = \hbar\,\omega\,, \qquad p = \frac{\hbar}{\lambda/(2\,\pi)} = \hbar\,k\,. \tag{7.20}$$

The second relation of (7.20) in vector form reads

$$\mathbf{p} = \hbar\,\mathbf{k}\,. \tag{7.21}$$

Here the useful outcome of the analysis of Sect. 3.13.8 is (3.92), that relates the photon's energies prior and after the collision (E_a and E_b, respectively) with the deflection angle ψ (Fig. 3.7). Using $E = c\,h/\lambda$ in (3.92) yields

$$\lambda_b - \lambda_a = 2\,\lambda_0\,\sin^2\left(\frac{\psi}{2}\right)\,, \qquad \lambda_0 = \frac{h}{m_0\,c}\,, \tag{7.22}$$

with $\lambda_0 \simeq 2.43 \times 10^{-12}$ m the *Compton wavelength* (1923). The frequency corresponding to it is $\nu_0 = c/\lambda_0 \simeq 1.2 \times 10^{20}$ Hz. The maximum difference in wavelength corresponds to the case of reflection, $\max(\lambda_b - \lambda_a) = 2\,\lambda_0$. Even in this case, the smallness of λ_0 makes the effect difficult to measure; in practice, the shift in wavelength is detectable only for sufficiently small values of λ_a, typically in the range of 10^{-10} m corresponding to the X-ray frequencies ($\nu \sim 10^{18}$ s^{-1}). Due to the large energy of the photon, the energy transferred to the electron brings the latter into a high-velocity regime; this, in turn, imposes the use of the relativistic expressions for describing the electron's dynamics.

7.4.4 Bohr's Hypothesis

The description of the monochromatic components of the electromagnetic field as a flow of identical photons with energy $h\,\nu$ lends itself to the explanation of the Balmer law (7.4). Such an explanation (*Bohr's hypothesis*, 1913) is based on the

idea that if ν_{nm} is the frequency of the emitted radiation, the corresponding energy of the emitted photon is $h\,\nu_{nm}$; multiplying (7.4) by h and remembering that $m > n$ then yields

$$h\,\nu_{nm} = h\,\nu_R \left(\frac{1}{n^2} - \frac{1}{m^2}\right) = \left(-\frac{h\,\nu_R}{m^2}\right) - \left(-\frac{h\nu_R}{n^2}\right) . \tag{7.23}$$

As the left-hand side is the energy of the emitted photon, the terms on the right-hand side can be recast as

$$E_m = -\frac{h\nu_R}{m^2} , \qquad E_n = -\frac{h\nu_R}{n^2} , \qquad E_n < E_m < 0 ; \tag{7.24}$$

then, if E_m (E_n) is interpreted as the atom's energy before (after) emitting the photon, Balmer's law becomes the expression of energy conservation. From this, the emission rule of Ritz is easily explained; in fact, (7.5) is equivalent to

$$E_m - E_n = (E_m - E_k) + (E_k - E_n) . \tag{7.25}$$

Bohr's hypothesis is expressed more precisely by the following statements:

1. The energy variations of the atom are due to the electrons of the outer shell, that exchange energy with the electromagnetic field.
2. The total energy of a non-radiative state is quantized, namely, it is associated with an integer index: $E_n = -h\,\nu_R/n^2$, $n = 1, 2, \ldots$; the values of energy thus identified are called *energy levels*. The lowest level corresponds to $n = 1$ and is called *ground level* or *ground state*.
3. The total energy can vary only between the quantized levels by exchanging with the electromagnetic field a photon of energy $\nu_{nm} = (E_m - E_n)/h$.

It is interesting to note that by combining Bohr's hypothesis with the planetary model of the atom, the quantization of the other dynamic quantities follows from that of energy; again, the case of a circular orbit is considered. By way of example, using $E_n = -h\,\nu_R/n^2$ in the second relation of (7.3) provides the quantization of the orbit's radius:

$$r = r_n = -\frac{q^2}{8\,\pi\,\varepsilon_0\,E_n} = \frac{q^2}{8\,\pi\,\varepsilon_0}\,\frac{n^2}{h\,\nu_R} . \tag{7.26}$$

The smallest radius r_1 corresponds to the ground state $n = 1$; taking ν_R from (7.4) and the other constants from Table D.1 one finds $r_1 \simeq 0.05$ nm; despite the simplicity of the model, r_1 is fairly close to the experimental value r_a given in (7.1).

In turn, the velocity is quantized by combining (7.3) to obtain $T = -V/2 = -E$; replacing the expressions of T and E then yields

$$\frac{1}{2}\,m\,u^2 = \frac{h\,\nu_R}{n^2} , \qquad u = u_n = \sqrt{\frac{2\,h\,\nu_R}{m\,n^2}} . \tag{7.27}$$

The largest velocity is found from (7.27) by letting $n = 1$ and using the minimum value for the mass, that is, the rest mass $m = m_0$. It turns out $u_1 \simeq 7 \times 10^{-3}\, c$; as a consequence, the velocity of a bound electron belonging to the outer shell of the atom can be considered nonrelativistic. Thanks to this result, from now on the electron's mass will be identified with the rest mass. Finally, for the angular momentum $M = r\,p = r\,m\,u$ one finds

$$M = M_n = \frac{q^2\, n^2}{8\,\pi\,\varepsilon_0\, h\,\nu_R}\; m\; \sqrt{\frac{2\,h\,\nu_R}{m\,n^2}} = \frac{1}{2\,\pi}\; \left[\frac{q^2}{\varepsilon_0}\; \sqrt{\frac{m}{8\,h\,\nu_R}}\right] n\,. \tag{7.28}$$

The quantity in brackets in (7.28) has the same units as M, namely, an action (Sect. 1.5) and, replacing the constants, it turns out[16] to be equal to h. Using the reduced Planck constant it follows

$$M_n = n\,\hbar\,. \tag{7.29}$$

The Bohr hypothesis provides a coherent description of some atomic properties; yet it does not explain, for instance, the fact that the electron belonging to an orbit of energy $E_n = -h\,\nu_R/n^2$ does not radiate, in contrast to what is predicted by the electromagnetic theory (compare with the discussion in Sect. 7.3). Another phenomenon not explained by the hypothesis is the fact that only the ground state of the atom is stable, whereas the excited states are unstable and tend to decay to the ground state.

7.4.5 De Broglie's Hypothesis

The explanation of the Compton effect (Sect. 7.4.3) involves a description of the photon's dynamics in which the latter is treated like a particle having energy and momentum. Such mechanical properties are obtained from the wave properties of a monochromatic component of the electromagnetic field: the relations involved are (7.19) (or (7.20)), by which the photon energy is related to the frequency, and its momentum to the wave vector. It is worth specifying that such relations are applied to the asymptotic part of the motion, namely, when the photon behaves like a free particle. In 1924, de Broglie postulated that analogous relations should hold for the free motion of a real particle: in this case, the fundamental dynamic properties are energy and momentum, to which a frequency and a wavelength (or a wave vector) are associated by relations identical to (7.19), (7.20),[17]

[16] This result shows that the physical constants appearing in (7.28) are not independent from each other. Among them, ν_R is considered the dependent one, while q, $m = m_0$, ε_0, and h are considered fundamental.

[17] The wavelength associated with the particle's momentum is called *de Broglie's wavelength*.

$$\omega = 2\,\pi\,\nu = 2\,\pi\,\frac{E}{h} = \frac{E}{\hbar}\,, \qquad k = \frac{2\,\pi}{\lambda} = \frac{2\,\pi}{h/p} = \frac{p}{\hbar}\,, \qquad \mathbf{k} = \frac{\mathbf{p}}{\hbar}\,. \tag{7.30}$$

The usefulness of associating, e.g., a wavelength to a particle's motion lies in the possibility of qualitatively justifying the quantization of the mechanical properties illustrated in Sect. 7.4.4. For this, consider the case of the circular orbit of the planetary motion, and associate a wavelength with the particle's momentum, $\lambda = h/p$. Such an association violates the prescription that (7.30) apply only to a free motion; however, if the orbit's radius is very large, such that $\lambda \ll r$, the orbit may be considered as locally linear and the concept of wavelength is applicable. Replacing $\lambda = h/p$ in (7.29) yields

$$2\,\pi\,r = n\,\lambda\,, \tag{7.31}$$

namely, the quantization of the mechanical properties implies that the orbit's length is an integer multiple of the wavelength associated with the particle. This outcome suggests that the formal description of quantization should be sought in the field of eigenvalue equations.

De Broglie also postulated that a function $\psi = \psi(\mathbf{r}, t)$, containing the parameters ω, $\mathbf{k}$ defined in (7.30), and called *wave function*, is associated with the particle's motion. Its meaning is provisionally left indefinite; as for its form, it is sensible to associate with the free motion, which is the simplest one, the simplest wave function, that is, the planar monochromatic wave. The latter is conveniently expressed in complex form as

$$\psi = A\,\exp[\mathrm{i}\,(\mathbf{k}\cdot\mathbf{r} - \omega\,t)]\,, \tag{7.32}$$

where $A \neq 0$ is a complex constant, not specified. Due to (7.30), the constant wave vector $\mathbf{k}$ identifies the momentum of the particle, and the angular frequency ω identifies its total energy, which in a free motion coincides with the kinetic energy. It is worth pointing out that despite its form, the wave function is not of electromagnetic nature; in fact, remembering that in a free motion it is $H = p^2/(2\,m) = E$, with H the Hamiltonian function, it follows

$$\hbar\,\omega = \frac{1}{2\,m}\,\hbar^2\,k^2\,, \qquad \omega(\mathbf{k}) = \frac{\hbar}{2\,m}\,k^2\,, \tag{7.33}$$

which is different from the electromagnetic relation $\omega = c\,k$. By the same token it would not be correct to identify the particle's velocity with the phase velocity u_f derived from the electromagnetic definition; in fact, one has

$$u_f = \frac{\omega}{k} = \frac{E/\hbar}{p/\hbar} = \frac{p^2/(2\,m)}{p} = \frac{p}{2\,m}\,. \tag{7.34}$$

The proper definition of velocity is that deriving from Hamilton's equations (1.42); its ith component reads in this case

$$u_i = \dot{x}_i = \frac{\partial H}{\partial p_i} = \frac{1}{\hbar}\frac{\partial H}{\partial k_i} = \frac{\partial \omega}{\partial k_i} = \frac{\hbar k_i}{m} = \frac{p_i}{m}\,. \tag{7.35}$$

The concepts introduced so far must now be extended to motions of a more general type. A sensible generalization is that of the conservative motion of a particle subjected to the force deriving from a potential energy $V(\mathbf{r})$. In this case the association described by (7.30) works only partially, because in a conservative motion the total energy is a constant, whereas momentum is generally not so. As a consequence, letting $\omega = E/\hbar$ yields for the wave function the form

$$\psi = w(\mathbf{r})\, \exp(-\mathrm{i}\,\omega\, t)\,, \tag{7.36}$$

which is still monochromatic but, in general, not planar. Its *spatial part* $w(\mathbf{r})$ reduces to $A\, \exp(\mathrm{i}\,\mathbf{k}\cdot\mathbf{r})$ for the free motion. The function of the form (7.36) is postulated to be the wave function associated with the motion of a particle at constant energy $E = \hbar\omega$. While the time dependence of ψ is prescribed, its space dependence must be worked out, likely by solving a suitable equation involving the potential energy V, the particle's mass and, possibly, other parameters.

7.5 Heuristic Derivation of the Schrödinger Equation

The concept of wave function introduced in Sect. 7.4.5 has been extended from the case of a free motion, where the wave function is fully prescribed apart from the multiplicative constant A, to the case of a conservative motion (7.36), where only the time dependence of the wave function is known. It is then necessary to work out a general method for determining the spatial part $w(\mathbf{r})$. The starting point is the observation that w is able at most to provide information about the particle's trajectory, not about the particle's dynamics along the trajectory. One of the methods used in Classical Mechanics to determine the trajectories is based on the Maupertuis principle (Sect. 2.7); moreover, from the discussion carried out in Sect. 5.11.6 it turns out that the analogy between the Maupertuis principle and the Fermat principle of Geometrical Optics (compare with (5.80)) provides the basis for a generalization of the mechanical laws. The first of the two principles applies to a particle (or system of particles) subjected to a conservative force field prescribed by a potential energy $V(\mathbf{r})$, with E a given constant; the second one applies to a monochromatic ray (ν = const) propagating in a medium whose properties are prescribed by the refraction index $n(\mathbf{r})$. The latter is related to frequency and wavelength by $n = c/(\lambda\,\nu)$ (compare with (5.55)); as a consequence, (5.80) can be rewritten as

$$\delta \int_{AB} \sqrt{E-V}\, \mathrm{d}s = 0\,, \qquad \delta \int_{AB} \frac{1}{\lambda}\, \mathrm{d}s = 0\,. \tag{7.37}$$

Considering that the variational principles hold apart from a multiplicative constant, the two expressions in (7.37) transform into each other by letting

$$\sqrt{E-V} = \frac{\alpha}{\lambda}\,, \tag{7.38}$$

where α is a constant that must not depend on the form of V or λ, nor on other parameters of the problem. For this reason, α is left unchanged also after removing the Geometrical-Optics approximation; when this happens, the Fermat principle is replaced with the Maxwell equations or, equivalently, with the wave equations for the electric field (4.64) and magnetic field (4.65). For simplicity, the latter equations are solved in a uniform medium with no charges in it, on account of the fact that α is not influenced by the medium's properties. Also, considering that in the uniform case (4.64) and (4.65) have the same structure, and that the function w under investigation is scalar, the analysis is limited to any scalar component C of $\mathbf{E}$ or $\mathbf{H}$; such a component fulfills the equation

$$\nabla^2 C - \frac{1}{u_f^2} \frac{\partial^2 C}{\partial t^2} = 0\,, \tag{7.39}$$

with $u_f = \text{const}$ the medium's phase velocity. Solving (7.39) by separation with $C(\mathbf{r}, t) = \eta\,(\mathbf{r})\,\theta(t)$ yields

$$\theta\, \nabla^2 \eta = \frac{1}{u_f^2}\, \ddot{\theta}\, \eta\,, \qquad \frac{\nabla^2 \eta}{\eta} = \frac{1}{u_f^2}\, \frac{\ddot{\theta}}{\theta} = -k^2\,, \tag{7.40}$$

where the separation constant $-k^2$ must be negative to prevent θ from diverging. As a consequence, k is real and can be assumed to be positive. The solution for the time factor is $\theta = \cos(\omega t + \varphi)$, where the phase φ depends on the initial conditions, and $\omega = 2\pi\, \nu = u_f\, k > 0$. It follows $k = 2\,\pi\, \nu / u_f = 2\,\pi/\lambda$, whence the spatial part of (7.39) reads

$$\nabla^2 \eta + \frac{(2\,\pi)^2}{\lambda^2}\, \eta = 0\,, \tag{7.41}$$

namely, a Helmholtz equation (Sect. 4.7). By analogy, the equation for the spatial part w of the wave function is assumed to be

$$\nabla^2 w + \frac{(2\,\pi)^2}{\alpha^2}\, E\, w = 0\,, \tag{7.42}$$

which is obtained from (7.41) by replacing η with w and using (7.38) with $V = 0$. The value of α is determined by expressing E in (7.42) in terms of the de Broglie wavelength; using the symbol λ_{dB} for the latter to avoid confusion with the electromagnetic counterpart, one finds $E = p^2/(2\,m) = h^2/(2\,m\,\lambda_{\mathrm{dB}}^2)$, namely,

$$\nabla^2 w + \frac{(2\,\pi)^2}{\alpha^2}\,\frac{h^2}{2\,m}\,\frac{1}{\lambda_{\mathrm{dB}}^2}\,w = 0\,. \tag{7.43}$$

Equation (7.43) becomes identical to (7.41) by letting $\alpha^2 = h^2/(2m)$ whence, using the reduced Planck constant $\hbar$, (7.42) becomes $\nabla^2 w + (2\,m\,E/\hbar^2)\,w = 0$. Such a differential equation holds in a uniform medium; hence, the dynamic property involved is the kinetic energy of the particle. The extension to the case of a nonuniform medium is then obtained by using the general form $E - V$ of the kinetic energy in terms of the coordinates; in conclusion, the equation for the spatial part of the wave function in a conservative case is

$$\nabla^2 w + \frac{2\,m}{\hbar^2}\,(E - V)\,w = 0\,, \qquad -\frac{\hbar^2}{2\,m}\,\nabla^2 w + V\,w = E\,w\,. \tag{7.44}$$

The above is the *time-independent Schrödinger equation*. It is a homogenous equation, with E the eigenvalue and w the eigenfunction.[18] Although the derivation based on the analogy between mechanical and optical principles is easy to follow, it must be remarked that the step leading from (7.43) to (7.44) is not rigorous; in the electromagnetic case, in fact, Eq. (7.41) for the spatial part holds only in a uniform medium; when the latter is nonuniform, instead, the right-hand side of (7.41) is different from zero, even in a charge-free case, because it contains the gradient of the refraction index. As shown in Sect. 1.10.4, the actual method used in 1926 by Schrödinger for deriving (7.44) consists in seeking the constrained extremum of a functional generated by the Hamilton-Jacobi equation; in such a procedure, the hypothesis of a uniform medium is not necessary.

It is also worth noting that in the analogy between mechanical and optical principles the spatial part of the wave function, and also the wave function as a whole, is the analogue of a component of the electromagnetic field. From this standpoint, the analogue of the field's intensity is the wave function's square modulus. In the monochromatic case, the latter reads $|\psi|^2 = |w|^2$. This reasoning is useful in the discussion about the physical meaning of ψ.

[18] The structure of (7.44) is illustrated in detail in Chap. 8.

7.6 Measurement

To make the wave function a useful tool for the description of the particles' dynamics it is necessary to connect the value taken by ψ, at a specific position and time, with some physical property of the particle (or system of particles) under investigation. To make such a connection it is in turn necessary to measure the property of interest; otherwise, the hypotheses illustrated in the previous sections would be relegated to a purely abstract level. In other terms, the meaning of the wave function can be given only by discussing the measuring process in some detail. The analysis is carried out below, following the line of [86]; in particular, a general formalism is sought which applies to both the macroscopic and microscopic bodies; the specific features brought about by the different size of the objects that are being measured are made evident by suitable examples.

The measurement of a dynamic variable A pertaining to a physical body is performed by making the body to interact with a measuring apparatus and recording the reading shown by the latter. For simplicity it is assumed that there is a finite number of possible outcomes of the measurement, say, $A_1, \ldots, A_M$. The extension to the case of a continuous, infinitely extended set of outcomes can be incorporated into the theory at the cost of a more awkward notation. Letting A_i be the outcome of the measurement of A, consider the case where the body is later subjected to the measurement of another dynamic variable B. Assume that the outcome of such a measurement is B_j, out of the possible outcomes $B_1, \ldots, B_N$. Next, the body is subjected to the measurement of a third variable C, thus yielding the value C_k, and so on.

As in general the dynamic variables depend on time, it is necessary to specify the time of each measurement. The most convenient choice is to assume that the time interval between a measurement and the next one is negligibly small, namely, that the measurement of B takes place immediately after that of A, similarly for that of C, and so on. The duration of each measurement is considered negligible as well. A special consequence of this choice is the following: if the measurement of A yielded A_i, and the measurement is repeated (namely, $B = A$), the outcome of the second measurement is again A_i.

Consider now the case where, after finding the numbers A_i, B_j, and C_k from the measurements of A, B, and C, respectively, the three variables are measured again, in any order. The experiments show that the results depend on the size of the body being measured. For a massive body the three numbers A_i, B_j, and C_k are always found. One concludes that the dynamic state of a massive body is not influenced by the interaction with the measuring apparatus or, more precisely, that if such an influence exists, it is so small that it cannot be detected. As a consequence one may also say that the values of the dynamic variables are properties of the body that exist prior, during, and after each measurement.

The situation is different for a microscopic body. By way of example, consider the case of a measurement of B followed by a measurement of A, the first one yielding B_n, the second one yielding A_i. If the measurement of B is carried out

again after that of A, the result is still one of the possible outcomes $B_1, \ldots, B_N$, but it is not necessarily equal to B_n. In other terms, the individual outcome turns out to be unpredictable. For a microscopic body one concludes that the interaction with the measuring apparatus is not negligible. It is worth observing that the apparatus able to measure the dynamic variable A may also be conceived in such a way as to block all outcomes that are different from a specific one, say, A_i. In such a case the apparatus is termed *filter*. Using the concept of filter one may build up the statistical distribution of the outcomes, for instance by repeating a large number of times the experiment in which the measurement of B is carried out after filtering A_i. The statistics is built up by recording the fraction of cases in which the measurement of B carried out on an A_i-filtered body yields the result $B_j, j = 1, \ldots, N$.

7.6.1 Probabilities

The fraction of measurements of the type described above, namely, of those that yield B_j after a measurement of A that has yielded A_i, will be indicated with the symbol $P(A_i \to B_j)$. Obviously the following hold:

$$0 \leq P(A_i \to B_j) \leq 1\,, \qquad \sum_{j=1}^{N} P(A_i \to B_j) = 1\,. \tag{7.45}$$

The first relation in (7.45) is due to the definition of $P(A_i \to B_j)$, the second one to the fact that the set of values $B_1, \ldots, B_N$ encompasses all the possible outcomes of the measurement of B. It follows that $P(A_i \to B_j)$ is the probability that a measurement of the dynamic variable B, made on a particle that prior to the measurement is in the state A_i of the dynamic variable A, yields the value B_j. The possible combinations of $P(A_i \to B_j)$ are conveniently arranged in the form of an M-row $\times$ N-column matrix:

$$\mathbf{P}_{AB} = \begin{bmatrix} P(A_1 \to B_1) & \ldots & P(A_1 \to B_N) \\ \vdots & & \vdots \\ P(A_i \to B_1) & \ldots & P(A_i \to B_N) \\ \vdots & & \vdots \\ P(A_M \to B_1) & \ldots & P(A_M \to B_N) \end{bmatrix}. \tag{7.46}$$

Due to (7.45), each row of $\mathbf{P}_{AB}$ adds up to unity. As the number of rows is M, the sum of all entries of matrix (7.46) is M. The same reasoning can also be made when the measurement of B is carried out prior to that of A. In this case the following N-row $\times$ M-column matrix is obtained:

$$\mathbf{P}_{BA} = \begin{bmatrix} P(B_1 \to A_1) & \dots & P(B_1 \to A_M) \\ \vdots & & \vdots \\ P(B_j \to A_1) & \dots & P(B_j \to A_M) \\ \vdots & & \vdots \\ P(B_N \to A_1) & \dots & P(B_N \to A_M) \end{bmatrix}, \tag{7.47}$$

with

$$0 \le P(B_j \to A_i) \le 1\,, \qquad \sum_{i=1}^{M} P(B_j \to A_i) = 1\,. \tag{7.48}$$

As the number of rows in (7.48) is N, the sum of all entries of matrix $\mathbf{P}_{BA}$ is N. It can be proven that it must be

$$P(B_j \to A_i) = P(A_i \to B_j) \tag{7.49}$$

for any pair of indices ij. In fact, if (7.49) did not hold, thermodynamic equilibrium would not be possible [86, Chap. V-21]. Equality (7.49) makes $\mathbf{P}_{BA}$ the transpose of $\mathbf{P}_{AB}$. As a consequence, the sum of all entries of the two matrices must be the same, namely, $N = M$. In other terms the outcomes of the measurements have the same multiplicity, and the matrices (7.46), (7.47) are square matrices of order $N = M$. Combining (7.49) with the second of (7.48) yields

$$\sum_{i=1}^{M} P(A_i \to B_j) = \sum_{i=1}^{M} P(B_j \to A_i) = 1\,, \tag{7.50}$$

showing that in the matrices (7.46), (7.47) not only each row, but also each column adds up to unity. A square matrix where all entries are nonnegative and all rows and columns add up to unity is called *doubly stochastic matrix*. Some properties of this type of matrices are illustrated in [94, Chap. II-1.4] and in Sect. A.11.4.

Note that (7.49) does not imply any symmetry of $\mathbf{P}_{AB}$. In fact, symmetry would hold if $P(A_j \to B_i) = P(A_i \to B_j)$. If the filtered state is A_i and the measurement of the dynamic variable A is repeated, the result is A_i again. In other terms,

$$P(A_i \to A_i) = 1\,, \qquad P(A_i \to A_k) = 0\,, \quad k \neq i\,. \tag{7.51}$$

This result can be recast in a more compact form as $\mathbf{P}_{AA} = \mathbf{I}$, with $\mathbf{I}$ the identity matrix.

7.6.2 Massive Bodies

It is useful to consider the special case where the measurement of B does not change the outcome of a previous measurement of A, and vice versa. In other terms, assume that the measurement of A has yielded A_i and the subsequent measurement of B has yielded B_j; then, another measure of A yields A_i again, a later measure of B yields B_j again, and so on. It follows that in $\mathbf{P}'_{AB}$ it is $P'(A_i \to B_j) = 1$, while all remaining entries in the ith row and jth column are equal to zero. This situation is typical of the bodies that are sufficiently massive, such that the interference suffered during the measurement of a dynamic variable is not detectable. For the sake of clarity an apex is used here to distinguish the probabilities from those of the general case where the body's mass can take any value. Considering a 4×4 matrix by way of example, a possible form of the matrix would be

$$\mathbf{P}'_{AB} = \begin{bmatrix} 0\,1\,0\,0 \\ 0\,0\,1\,0 \\ 1\,0\,0\,0 \\ 0\,0\,0\,1 \end{bmatrix}, \tag{7.52}$$

that is, one of the 4! possible permutation matrices of order 4. Clearly all the other permutation matrices of order 4 different from (7.52) are equally possible. The meaning of a matrix like (7.52) is that the successive measurements of A and B yield either the pair A_1, B_2, or the pair A_2, B_3, or A_3, B_1, or A_4, B_4. Matrix (7.52) may be thought of as a limiting case: starting from a microscopic body described by a 4×4, doubly stochastic matrix whose entries are in general different from zero, the size of the body is increased by adding one atom at a time, and the set of measurements of A and B is repeated at each time. As the reasoning that prescribes the doubly stochastic nature of the matrix holds at each step, the successive matrices must tend to the limit of a permutation matrix. Which of the 4! permutation matrices will be reached by this process depends on the initial preparation of the experiments. One may wonder why a matrix like

$$\mathbf{P}'_{AB} = \begin{bmatrix} 0\,1\,0\,0 \\ 0\,1\,0\,0 \\ 1\,0\,0\,0 \\ 0\,0\,0\,1 \end{bmatrix}, \tag{7.53}$$

should not be reached. In fact, such a matrix is not acceptable because it is not doubly stochastic: its transpose implies that the outcomes B_1 and B_2 are simultaneously associated with A_2 with certainty, which is obviously impossible not only for a massive body, but for any type of body. This reasoning is associated with another argument, based on the theorem mentioned in Sect. A.11.4, stating that a doubly stochastic matrix is a convex combination of permutation matrices. Letting $\theta_1, \ldots, \theta_M$ be the combination's coefficients as those used in (A.70), in the process

of transforming the microscopic body into a macroscopic one all the coefficients but one vanish, and the nonvanishing one tends to unity. As a consequence, out of the original combination, only one permutation matrix is left.

7.6.3 Need of a Description of Probabilities

The non-negligible influence of the measuring apparatus on the dynamic state of a microscopic body makes it impossible to simultaneously measure the dynamic variables that constitute the initial conditions of the motion. As a consequence, the possibility of using the Hamiltonian theory for describing the dynamics is lost. As outlined in the above sections, the distinctive mark of experiments carried out on microscopic objects is the statistical distribution of the outcomes; thus, a theory that adopts the wave function as the basic tool must identify the connection between the wave function and such a statistical distribution. The theory must also contain the description of the massive bodies as a limiting case.

7.7 Born's Interpretation of the Wave Function

Basing on the optical analogy and the examination of experiments, the *probabilistic interpretation* of the wave function introduced by Born states that the integral

$$\int_\tau |\psi(\mathbf{r}, t)|^2 \, \mathrm{d}^3 r \tag{7.54}$$

is proportional to the probability that a measuring process finds the particle within the volume τ at the time t.[19] Note that the function used in (7.54) is the square modulus of ψ, namely, as noted in Sect. 7.5, the counterpart of the field's intensity in the optical analogy. Also, considering that by definition the integral of (7.54) is dimensionless, the units[20] of ψ are $\mathrm{m}^{-3/2}$.

When $\tau \to \infty$ the integral in (7.54) may, or may not, converge. In the first case, ψ is said to be *normalizable*, and a suitable constant σ can be found such that the integral of $|\sigma\,\psi|^2$ over the entire space equals unity. The new wave function provides a probability proper,

$$\int_\tau |\sigma\,\psi|^2 \, \mathrm{d}^3 r \le 1\,, \qquad \sigma^{-2} = \int_\infty |\psi|^2 \, \mathrm{d}^3 r\,. \tag{7.55}$$

[19] From this interpretation it follows that $|\psi|^2 \, \mathrm{d}^3 r$ is proportional to an infinitesimal probability, and $|\psi|^2$ to a probability density.

[20] A more detailed discussion about the units of the wave function is carried out in Sect. 9.7.1.

In the second case ψ is not normalizable[21]: a typical example is the wave function of a free particle, $\psi = A \exp[\mathrm{i}\,(\mathbf{k}\cdot\mathbf{r} - \omega t)]$; however, it is still possible to define a probability ratio

$$\int_{\tau_1} |\psi|^2 \,\mathrm{d}^3 r \left(\int_{\tau_2} |\psi|^2 \,\mathrm{d}^3 r \right)^{-1} , \tag{7.56}$$

where both volumes τ_1 and τ_2 are finite. Relation (7.56) gives the ratio between the probability of finding the particle within τ_1 and that of finding it within τ_2.

Consider a particle whose wave function at time t differs from zero within some volume τ, and assume that a process of measuring the particle's position is initiated at t and completed at some later time t'; let the outcome of the experiment be an improved information about the particle's location, namely, at t' the wave function differs from zero in a smaller volume $\tau' \subset \tau$. This event is also called *contraction* of the wave function.

Problem

7.1 Considering the sun as a black body whose surface temperature is 5,780 K (compare with Fig. 7.10), find the *solar constant*, that is, the power per unit area that impinges normally onto the outer edge of the earth's atmosphere. Assume that the earth's orbit is circular with a radius equal to R, and use $r = R/216$ for the radius of the sun.

[21] This issue is further discussed in Sect. 8.2.

Chapter 8
Time-Independent Schrödinger Equation

8.1 Introduction

The properties of the time-independent Schrödinger equation are introduced step by step, starting from a short discussion about its boundary conditions. Considering that the equation is seldom amenable to analytical solutions, two simple cases are examined first: that of a free particle and that of a particle in a box. The determination of the lower energy bound follows, introducing more general issues that build up the mathematical frame of the theory: norm of a function, scalar product of functions, Hermitean operators, eigenfunctions and eigenvalues of operators, orthogonal functions, and completeness of a set of functions. This chapter is concluded with the important examples of the Hamiltonian operator and momentum operator. The complements provide examples of Hermitean operators, a collection of operators' definitions and properties, examples of commuting operators, and a further discussion about the free-particle case.

8.2 Properties of the Time-Independent Schrödinger Equation

A number of properties of the time-independent Schrödinger equation are discussed in this section. The equation holds only when the force is conservative; however, as will be shown later, many of its properties still hold in more complicate cases. Equation (7.44) is a linear, homogeneous partial-differential equation of the second order, with the zero-order coefficient depending on **r**. As shown in Prob. 8.1, it is a very general form of linear, second-order equation. The boundary conditions are specified on a case-by-case basis depending on the problem under consideration. More details about the boundary conditions are discussed below. One notes that:

M. Rudan, *Physics of Semiconductor Devices*, DOI 10.1007/978-3-319-63154-7_8

1. The coefficients of (7.44) are real. As a consequence, the solutions are real. In same cases, however, it is convenient to express them in complex form. An example is given in Sect. 8.2.1.
2. The equation is linear and homogeneous and, as shown below, its boundary conditions are homogeneous as well. It follows that its solution is defined apart from a multiplicative constant. The function $w = 0$ is a solution of (7.44); however, it has no physical meaning and is not considered.
3. As the equation is of the second order, its solution w and first derivatives $\partial w/\partial x_i$ are continuous. These requirements are discussed from the physical standpoint in Sect. 9.4. The second derivatives may or may not be continuous, depending on the form of the potential energy V.
4. The solution of (7.44) may contain terms that diverge as $|\mathbf{r}| \to \infty$. In this case such terms must be discarded because they are not compatible with the physical meaning of w (examples are given in Sect. 8.2.1).

Given the above premises, to discuss the boundary conditions of (7.44) it is convenient to distinguish a few cases:

A. The domain Ω of w is finite; in other terms, some information about the problem in hand is available, from which it follows that w vanishes identically outside a finite domain Ω. The continuity of w (see point 3 above) then implies that w vanishes over the boundary of Ω, hence the boundary conditions are homogeneous. After discarding possible diverging terms from the solution, the integral $\int_\Omega |w|^2 \, \mathrm{d}\Omega$ is finite (the use of the absolute value is due to the possibility that w is expressed in complex form, see point 1 above).
B. The domain of w is infinite in all directions, but the form of w is such that $\int_\Omega |w|^2 \, \mathrm{d}\Omega$ is finite. When this happens, w necessarily vanishes as $|\mathbf{r}| \to \infty$. Thus, the boundary conditions are homogeneous also in this case.[1]
C. The domain of w is infinite, and the form of w is such that $\int_\Omega |w|^2 \, \mathrm{d}\Omega$ diverges. This is not due to the fact that $|w|^2$ diverges (in fact, divergent terms in w must be discarded beforehand), but to the fact that w, e.g., asymptotically tends to a constant different from zero, or oscillates (an example of asymptotically oscillating behavior is given in Sect. 8.2.1). These situations must be tackled separately; one finds that w is still defined apart from a multiplicative constant.

As remarked above, the time-independent Schrödinger equation is a second-order differential equation of a very general form. For this reason, an analytical solution can seldom be obtained, and in the majority of cases it is necessary to resort to numerical-solution methods. The typical situations where the problem can be tackled analytically are those where the equation is separable (compare with

[1] It may happen that the domain is infinite in some direction and finite in the others. For instance, one may consider the case where w vanishes identically for $x \geq 0$ and differs from zero for $x < 0$. Such situations are easily found to be a combination of cases A and B illustrated here.

Sect. 10.3), so that it can be split into one-dimensional equations. Even when this occurs, the analytical solution can be found only for some forms of the potential energy. The rest of this chapter provides examples that are solvable analytically.

8.2.1 Schrödinger Equation for a Free Particle

The equation for a free particle is obtained by letting $V = \text{const}$ in (7.44). Without loss of generality one may let $V = 0$, thus yielding $\nabla^2 w = -(2\,m\,E/\hbar^2)\,w$. As the above can be solved by separating the variables, it is sufficient to consider here only the one-dimensional form

$$\frac{\mathrm{d}^2 w}{\mathrm{d}x^2} = -\frac{2\,m\,E}{\hbar^2}\,w\,. \tag{8.1}$$

The case $E < 0$ must be discarded as it gives rise to divergent solutions, which are not acceptable from the physical standpoint. The case $E = 0$ yields $w = a_1 x + a_2$, where a_1 must be set to zero to prevent w from diverging. As a consequence, the value $E = 0$ yields $w = a_2 = \text{const}$, that is one of the possibilities anticipated at point C of Sect. 8.2. The integral of $|w|^2$ diverges. Finally, the case $E > 0$ yields

$$w = c_1\,\exp(\mathrm{i}\,k\,x) + c_2\,\exp(-\mathrm{i}\,k\,x)\,, \quad k = \sqrt{2\,m\,E/\hbar^2} = p/\hbar > 0\,, \tag{8.2}$$

where c_1, c_2 are constants to be determined. Thus, the value $E > 0$ yields the asymptotically oscillating behavior that has also been anticipated at point C of Sect. 8.2. The integral of $|w|^2$ diverges. One notes that w is written in terms of two complex functions; it could equally well be expressed in terms of the real functions $\cos(k\,x)$ and $\sin(k\,x)$. The time-dependent, monochromatic wave function $\psi = w\,\exp(-\mathrm{i}\,\omega\,t)$ corresponding to (8.2) reads

$$\psi = c_1\,\exp[\mathrm{i}\,(k\,x - \omega\,t)] + c_2\,\exp[-\mathrm{i}\,(k\,x + \omega\,t)]\,, \qquad \omega = E/\hbar\,. \tag{8.3}$$

The relations $k = p/\hbar$, $\omega = E/\hbar$ stem from the analogy described in Sect. 7.4.5. The total energy E and momentum's modulus p are fully determined; this outcome is the same as that found for the motion of a free particle in Classical Mechanics: for a free particle the kinetic energy equals the total energy; if the latter is prescribed, the momentum's modulus is prescribed as well due to $E = p^2/(2\,m)$. The direction of the motion, instead, is not determined because both the forward and backward motions, corresponding to the positive and negative square root of p^2 respectively, are possible solutions. To ascertain the motion's direction it is necessary to acquire more information; specifically, one should prescribe the initial conditions which, in turn, would provide the momentum's sign.

The quantum situation is similar, because the time-dependent wave function (8.2) is a superposition of a planar wave $c_1\,\exp[\mathrm{i}\,(k\,x - \omega\,t)]$ whose front moves in the

positive direction, and of a planar wave $c_2\ \exp[-\mathrm{i}\,(k\,x+\omega\,t)]$ whose front moves in the negative direction. Here to ascertain the motion's direction one must acquire the information about the coefficients in (8.2): the forward motion corresponds to $c_1 \neq 0$, $c_2 = 0$, the backward motion to $c_1 = 0$, $c_2 \neq 0$. Obviously (8.2) in itself does not provide any information about the coefficients, because such an expression is the general solution of (8.1) obtained as a combination of the two linearly independent, particular solutions $\exp(\mathrm{i}\,k\,x)$ and $\exp(-\mathrm{i}\,k\,x)$; so, without further information about c_1 and c_2, both the forward and backward motions are possible.

Another similarity between the classical and quantum cases is that no constraint is imposed on the total energy, apart from the prescription $E \geq 0$. From this viewpoint one concludes that (8.1) is an eigenvalue equation with a continuous distribution of eigenvalues in the interval $E \geq 0$.

8.2.2 *Schrödinger Equation for a Particle in a Box*

Considering again the one-dimensional case of (7.44),

$$\frac{\mathrm{d}^2 w}{\mathrm{d}x^2} = -\frac{2\,m}{\hbar^2}\,(E-V)\,w\,, \qquad V = V(x)\,, \tag{8.4}$$

let $V = \text{const} = 0$ for $x \in [0,a]$ and $V = V_0 > 0$ elsewhere. The form of the potential energy is that of a square well whose counterpart of Classical Mechanics is illustrated in Sect. 3.2. Here, however, the limit $V_0 \to \infty$ is considered for the sake of simplicity. This limiting case is what is referred to with the term *box*. As shown in Sect. 11.5, here w vanishes identically outside the interval $[0,a]$: this is one of the possibilities that were anticipated in Sect. 8.2 (point A). The continuity of w then yields $w(0) = w(a) = 0$. It is easily found that if $E \leq 0$ the only solution of (8.4) is $w = 0$, which is not considered because it has no physical meaning. When $E > 0$, the solution reads

$$w = c_1\,\exp(\mathrm{i}\,k\,x) + c_2\,\exp(-\mathrm{i}\,k\,x)\,, \qquad k = \sqrt{2\,m\,E/\hbar^2} > 0\,. \tag{8.5}$$

Letting $w(0) = 0$ yields $c_1 + c_2 = 0$ and $w = 2\,\mathrm{i}\,c_1\,\sin(k\,x)$. Then, $w(a) = 0$ yields $k\,a = n\,\pi$ with n an integer whence, using the relation $k = k_n = n\,\pi/a$ within those of E and w,

$$E = E_n = \frac{\hbar^2\,\pi^2}{2\,m\,a^2}\,n^2\,, \qquad w = w_n = 2\,\mathrm{i}\,c_1\,\sin\left(\frac{n\,\pi}{a}\,x\right)\,. \tag{8.6}$$

This result shows that (8.4) is an eigenvalue equation with a discrete distribution of eigenvalues, given by the first relation in (8.6). For this reason, the energy is said to be *quantized*. To each index n it corresponds one and only one eigenvalue E_n, and one and only one eigenfunction w_n; as a consequence, this case provides

a one-to-one-correspondence between eigenvalues and eigenfunctions.[2] Not every integer should be used in (8.6) though; in fact, $n = 0$ must be discarded because the corresponding eigenfunction identically vanishes. Also, the negative indices are to be excluded because $E_{-n} = E_n$ and $|w_{-n}|^2 = |w_n|^2$, so they do not add information with respect to the positive ones. In conclusion, the indices to be used are $n = 1, 2, \ldots$.

As expected, each eigenfunction contains a multiplicative constant; here the integral of $|w|^2$ converges, so the constant can be exploited to normalize the eigenfunction by letting $\int_0^a |w_n|^2 \, \mathrm{d}x = 1$. One finds

$$\int_0^a |w_n|^2 \, \mathrm{d}x = 4\,|c_1|^2 \int_0^a \sin^2\left(\frac{n\,\pi}{a}\,x\right) \mathrm{d}x = \frac{4\,|c_1|^2 a}{n\,\pi} \int_0^{n\,\pi} \sin^2(y)\,\mathrm{d}y\,. \tag{8.7}$$

Integrating by parts shows that the last integral equals $n\,\pi/2$, whence the normalization condition yields $4\,|c_1|^2 = 2/a$. Choosing $2\,c_1 = -j\,\sqrt{2/a}$ provides the eigenfunctions

$$w_n = \sqrt{\frac{2}{a}}\,\sin\left(\frac{n\,\pi}{a}\,x\right)\,. \tag{8.8}$$

The first eigenfunctions are shown in Fig. 8.1. Remembering that $w = 0$ outside the interval $[0, a]$, one notes that $\mathrm{d}w/\mathrm{d}x$ is discontinuous at $x = 0$ and $x = a$. This apparently contradicts the continuity property of the first derivative mentioned in

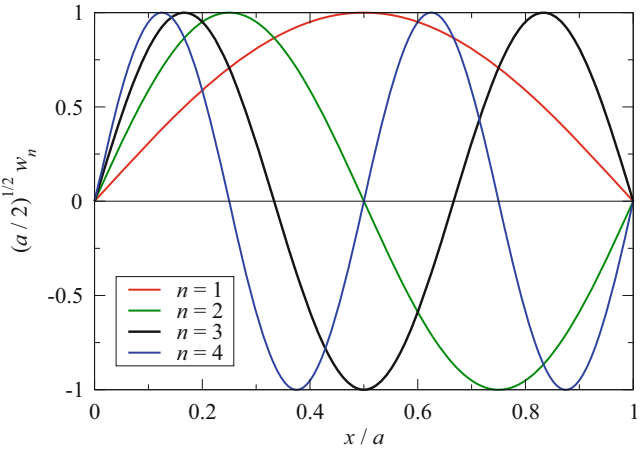


Fig. 8.1 The first eigenfunctions of the Schrödinger equation in the case of a particle in a box

[2] The one-to-one correspondence does not occur in general. Examples of the Schrödinger equation are easily given (Sect. 9.6) where to each eigenvalue there corresponds more than one—even infinite—eigenfunctions.

Sect. 8.2, point 3. However, in the case considered here the limit $V_0 \to \infty$ has introduced a discontinuity of the second kind into the potential energy; for this reason, the property mentioned above does not apply.

8.2.3 Lower Energy Bound in the Schrödinger Equation

In the example of Sect. 8.2.1, where the free particle is considered, the lower bound for the particle's total energy is $E \geq V_{\min}$, with $V_{\min}$ the minimum[3] of the potential energy; in contrast, in the example of the particle in a box illustrated in Sect. 8.2.2, the lower bound is $E > V_{\min}$. A more general analysis of the lower bound for the total energy in the Schrödinger equation is carried out here.

Consider the time-independent Schrödinger equation in a conservative case, (7.44), and let Ω be the domain of w (which may extend to infinity), with Σ the boundary of Ω. Recasting (7.44) as $-\nabla^2 w = 2\,m\,(E - V)\,w/\hbar^2$ and integrating it over Ω after multiplying both sides by w^* yields

$$-\int_\Omega w^*\,\nabla^2 w\,\mathrm{d}\Omega = \frac{2m}{\hbar^2}\int_\Omega (E - V)\,|w|^2\,\mathrm{d}\Omega\,. \tag{8.9}$$

It is implied that w is a physically meaningful solution of (7.44), whence w does not vanish identically within Ω. Thanks to the identity (A.17) and the divergence theorem (A.23) the above becomes

$$\frac{2\,m}{\hbar^2}\int_\Omega (E - V)\,|w|^2\,\mathrm{d}\Omega = \int_\Omega |\,\mathrm{grad}\,w|^2\,\mathrm{d}\Omega - \int_\Sigma w^*\,\frac{\partial w}{\partial n}\,\mathrm{d}\Sigma\,, \tag{8.10}$$

with $\partial w/\partial n$ the derivative of w in the direction normal to Σ. Consider now the case where w vanishes over Σ; as w^* vanishes as well, the boundary integral in (8.10) is equal to zero. In contrast, the other integral at the right-hand side of (8.10) is strictly positive: in fact, as w vanishes at the boundary while it is different from zero inside the domain, its gradient does not vanish identically in Ω. It follows

$$\int_\Omega (E - V)\,|w|^2\,\mathrm{d}\Omega > 0\,, \qquad E > \frac{\int_\Omega V\,|w|^2\,\mathrm{d}\Omega}{\int_\Omega |w|^2\,\mathrm{d}\Omega} \geq V_{\min}\,, \tag{8.11}$$

where the last inequality stems from the fact that $|w|^2$ is strictly positive. In conclusion, when V is such that w vanishes at the boundary, then the strict inequality $E > V_{\min}$ holds. When V does not vanish at the boundary, the reasoning leading

[3] Such a minimum is set to zero in the example of Sect. 8.2.1.

to (8.11) does not apply and the lower bound for E must be sought by a direct examination of the solutions. An example of this examination is that of the free-particle case shown in Sect. 8.2.1.

8.3 Norm of a Function—Scalar Product

The functions $f, g, \ldots$ that are considered in this section are *square-integrable* complex functions,[4] namely, they have the property that the integrals

$$||f||^2 = \int_{\Omega'} |f|^2 \, \mathrm{d}\Omega' \,, \qquad ||g||^2 = \int_{\Omega''} |g|^2 \, \mathrm{d}\Omega'' \,, \ldots \tag{8.12}$$

converge. In (8.12), Ω' is the domain of f, Ω'' that of g, and so on. The variables in the domains $\Omega', \Omega'', \ldots$ are real. The nonnegative numbers $||f||$ and $||g||$ are the *norm* of f and g, respectively. If f, g are square integrable over the same domain Ω, a linear combination $\lambda f + \mu g$, with λ, μ arbitrary complex constants, is also square integrable over Ω [96, Chap. V.2].

If a square-integrable function f is defined apart from a multiplicative constant, for instance because it solves a linear, homogeneous differential equation with homogeneous boundary conditions, it is often convenient to choose the constant such that the norm equals unity. This is accomplished by letting $\varphi = cf$ and $||\varphi|| = 1$, whence $|c|^2 = 1/||f||^2$.

Consider two square-integrable functions f and g defined over the same domain Ω; their *scalar product* is defined as

$$\langle g|f\rangle = \int_{\Omega} g^* f \, \mathrm{d}\Omega \,. \tag{8.13}$$

From (8.13) it follows

$$\langle f|g\rangle = \int_{\Omega} f^* g \, \mathrm{d}\Omega = \left(\int_{\Omega} f g^* \, \mathrm{d}\Omega \right)^* = \langle g|f\rangle^* \,. \tag{8.14}$$

It is implied that f, g are regular enough to make the integral in (8.13) to exist; in fact, this is proved by observing that for square-integrable functions the *Schwarz inequality* holds, analogous to that found in the case of vectors (Sect. A.2): if f and g are square integrable, then

$$|\langle g|f\rangle| \leq ||f|| \times ||g|| \,, \tag{8.15}$$

[4] Some definitions and properties illustrated in this section have an analogue in the theory of matrices outlined in Sect. A.11.

where the equality holds if and only if f is proportional to g (compare with (A.5)). In turn, to prove (8.15) one observes that $\sigma = f + \mu\, g$, where μ is an arbitrary constant, is also square integrable. Then [62],

$$||\sigma||^2 = ||f||^2 + |\mu|^2\, ||g||^2 + \mu\, \langle f|g\rangle + \mu^*\, \langle g|f\rangle \geq 0\,. \tag{8.16}$$

The relation (8.15) is obvious if $f = 0$ or $g = 0$. Let $g \neq 0$ and choose $\mu = -\langle g|f\rangle/||g||^2$. Replacing in (8.16) yields (8.15). For the equality to hold it must be $\sigma = 0$, which implies that f and g are proportional to each other; conversely, from $f = c\, g$ the equality follows.

The symbol $\langle g|f\rangle$ for the scalar product is called *Dirac's notation*.[5] If $\langle g|f\rangle = 0$, the functions f, g are called *orthogonal*. For any complex constants b, b_1, b_2 the following hold:

$$\langle g|b f\rangle = b\, \langle g|f\rangle\,, \qquad \langle g|b_1 f_1 + b_2 f_2\rangle = b_1\, \langle g|f_1\rangle + b_2\, \langle g|f_2\rangle\,, \tag{8.17}$$

$$\langle b\, g|f\rangle = b^*\, \langle g|f\rangle\,, \qquad \langle b_1\, g_1 + b_2\, g_2|f\rangle = b_1^*\, \langle g_1|f\rangle + b_2^*\, \langle g_2|f\rangle\,, \tag{8.18}$$

namely, the scalar product is distributive and bilinear. The properties defined here are the counterpart of those defined in Sect. A.1 for vectors.

8.3.1 Adjoint Operators and Hermitean Operators

A function appearing within a scalar product may result from the application of a linear operator, say, $\mathscr{A}$, onto another function.[6] For instance, if $s = \mathscr{A} f$, then from (8.13, 8.14) it follows

$$\langle g|s\rangle = \int_\Omega g^*\, \mathscr{A} f\, \mathrm{d}\Omega\,, \qquad \langle s|g\rangle = \int_\Omega (\mathscr{A} f)^*\, g\, \mathrm{d}\Omega = \langle g|s\rangle^*\,. \tag{8.19}$$

Given an operator $\mathscr{A}$ it is possible to find another operator, typically indicated with $\mathscr{A}^\dagger$, having the property that, for any pair f, g of square-integrable functions,

$$\int_\Omega \left(\mathscr{A}^\dagger g\right)^*\, f\, \mathrm{d}\Omega = \int_\Omega g^*\, \mathscr{A} f\, \mathrm{d}\Omega \tag{8.20}$$

[5] The two terms $\langle g|$ and $|f\rangle$ of the scalar product $\langle g|f\rangle$ are called *bra vector* and *ket vector*, respectively.

[6] In this context the term *operator* has the following meaning: if an operation brings each function f of a given function space into correspondence with one and only one function s of the same space, one says that this is obtained through the action of a given operator $\mathscr{A}$ onto f and writes $s = \mathscr{A} f$. A *linear* operator is such that $\mathscr{A}\,(c_1 f_1 + c_2 f_2) = c_1\, \mathscr{A} f_1 + c_2\, \mathscr{A} f_2$ for any pair of functions f_1, f_2 and of complex constants c_1, c_2 [96, Chap. II.11].

or, in Dirac's notation, $\langle \mathscr{A}^\dagger g|f\rangle = \langle g|\mathscr{A} f\rangle$. Operator $\mathscr{A}^\dagger$ is called the *adjoint* [7] of $\mathscr{A}$. In general it is $\mathscr{A}^\dagger \neq \mathscr{A}$; however, for some operators it happens that $\mathscr{A}^\dagger = \mathscr{A}$. In this case, $\mathscr{A}$ is called *Hermitean*. Thus, for Hermitean operators the following holds:

$$\langle g|\mathscr{A} f\rangle = \langle \mathscr{A} g|f\rangle = \langle g|\mathscr{A}|f\rangle \,. \tag{8.21}$$

The notation on the right of (8.21) indicates that one can consider the operator as applied onto f or g. Examples of Hermitean operators are given in Sect. 8.6.1. It is found by inspection that for any operator $\mathscr{C}$, the operators $\mathscr{S} = \mathscr{C} + \mathscr{C}^\dagger$ and $\mathscr{D} = -\mathrm{i}\,(\mathscr{C} - \mathscr{C}^\dagger)$ are Hermitean.

The following property is of use: a linear combination of Hermitean operator with real coefficients is Hermitean; considering, e.g., two Hermitean operators $\mathscr{A}$, $\mathscr{B}$ and two real numbers λ, μ, one finds

$$\int_\Omega g^* \,(\lambda\,\mathscr{A} + \mu\,\mathscr{B}) f \,\mathrm{d}\Omega = \int_\Omega \left[(\lambda\,\mathscr{A} + \mu\,\mathscr{B})\, g\right]^* f \,\mathrm{d}\Omega \,. \tag{8.22}$$

8.4 Eigenvalues and Eigenfunctions of an Operator

A linear operator $\mathscr{A}$ may be used to generate a homogeneous equation (*eigenvalue equation*) in the unknown v, having the form[8]

$$\mathscr{A} v = A\, v \,, \tag{8.23}$$

with A a parameter. Clearly (8.23) admits the solution $v = 0$ which, however, is of no interest; it is more important to find whether specific values of A exist (*eigenvalues*), such that (8.23) admits nonvanishing solutions (*eigenfunctions*). In general (8.23) must be supplemented with suitable boundary or regularity conditions on v.

The set of the eigenvalues of an operator found from (8.23) is the operator's *spectrum*. It may happen that the eigenvalues are distinguished by an index, or a set of indices, that take only discrete values; in this case the spectrum is called *discrete*. If, instead, the eigenvalues are distinguished by an index, or a set of indices, that vary continuously, the spectrum is *continuous*. Finally, it is *mixed* if a combination of discrete and continuous indices occurs.

An eigenvalue is *simple* if there is one and only one eigenfunction corresponding to it, while it is *degenerate of order s* if there are s linearly independent eigenfunc-

[7] The adjoint operator is the counterpart of the conjugate-transpose matrix in vector algebra.

[8] Some definitions and properties illustrated in this section have an analogue in the theory of matrices outlined in Sect. A.11.

tions corresponding to it. The order of degeneracy may also be infinite. By way of example, the Schrödinger equation for a free particle in one dimension discussed in Sect. 8.2.1 has a continuous spectrum of eigenvalues $E = \hbar^2 k^2/(2m)$ of index k, namely, $E = E_k$. Each eigenvalue is degenerate of order 2 because to each E there correspond two linearly independent eigenfunctions $\exp(\mathrm{i}\,k\,x)$, $\exp(-\mathrm{i}\,k\,x)$, with $k = \sqrt{2mE}/\hbar$. Instead, the Schrödinger equation for a particle in a box discussed in Sect. 8.2.2 has a discrete spectrum of eigenvalues E_n given by the first relation in (8.6). Each eigenvalue is simple as already indicated in Sect. 8.2.2.

Let $v^{(1)}, \ldots, v^{(s)}$ be the linearly independent eigenfunctions belonging to an eigenvalue A degenerate of order s; then a linear combination of such eigenfunctions is also an eigenfunction belonging to A. In fact, letting $\alpha_1, \ldots, \alpha_s$ be the coefficients of the linear combination, from $\mathscr{A} v^{(k)} = A\, v^{(k)}$ it follows

$$\mathscr{A} \sum_{k=1}^{s} \alpha_k\, v^{(k)} = \sum_{k=1}^{s} \alpha_k \mathscr{A} v^{(k)} = \sum_{k=1}^{s} \alpha_k\, A\, v^{(k)} = A \sum_{k=1}^{s} \alpha_k\, v^{(k)} \,. \tag{8.24}$$

8.4.1 Eigenvalues of Hermitean Operators

A fundamental property of the Hermitean operators is that their eigenvalues are real. Consider, first, the case where the eigenfunctions are square integrable, so that $\langle v|v\rangle$ is different from zero and finite. To proceed one considers the discrete spectrum, where the eigenvalues are A_n. Here n indicates a single index or also a set of indices. If the eigenvalue is simple, let v_n be the eigenfunction belonging to A_n; if it is degenerate, the same symbol v_n is used here to indicate any eigenfunction belonging to A_n. Then, two operations are performed: in the first one, the eigenvalue equation $\mathscr{A} v_n = A_n\, v_n$ is scalarly multiplied by v_n on the left, while in the second one the conjugate equation $(\mathscr{A} v_n)^* = A_n^*\, v_n^*$ is scalarly multiplied by v_n on the right. The operations yield, respectively,

$$\langle v_n|\mathscr{A} v_n\rangle = A_n\, \langle v_n|v_n\rangle \,, \qquad \langle \mathscr{A} v_n|v_n\rangle = A_n^*\, \langle v_n|v_n\rangle \,. \tag{8.25}$$

The left-hand sides in (8.25) are equal to each other due to the hermiticity of $\mathscr{A}$; as a consequence, $A_n^* = A_n$, that is, A_n is real.

Another fundamental property of the Hermitean operators is that two eigenfunctions belonging to different eigenvalues are orthogonal to each other. Still considering the discrete spectrum, let A_m, A_n be two different eigenvalues and let v_m (v_n) be an eigenfunction belonging to A_m (A_n). The two eigenvalues are real as demonstrated earlier. Then, the eigenvalue equation $\mathscr{A} v_n = A_n\, v_n$ is scalarly multiplied by v_m on the left, while the conjugate equation for the other eigenvalue, $(\mathscr{A} v_m)^* = A_m\, v_m^*$, is scalarly multiplied by v_n on the right. The operations yield, respectively,

$$\langle v_m|\mathscr{A}v_n\rangle = A_n\,\langle v_m|v_n\rangle\,, \qquad \langle \mathscr{A}v_m|v_n\rangle = A_m\,\langle v_m|v_n\rangle\,. \tag{8.26}$$

The left-hand sides in (8.26) are equal to each other due to the hermiticity of $\mathscr{A}$; as a consequence, $(A_m - A_n)\,\langle v_m|v_n\rangle = 0$. But $A_n \neq A_m$, so it is $\langle v_m|v_n\rangle = 0$.

8.4.2 Gram-Schmidt Orthogonalization

When two eigenfunctions belonging to a degenerate eigenvalue are considered, the reasoning that proves their orthogonality through (8.26) is not applicable because $A_n = A_m$. In fact, linearly independent eigenfunctions of an operator $\mathscr{A}$ belonging to the same eigenvalue are not mutually orthogonal in general. However, it is possible to form mutually orthogonal linear combinations of the eigenfunctions. As shown by (8.24), such linear combinations are also eigenfunctions, so their norm is different from zero. The procedure (*Gram-Schmidt orthogonalization*) is described here with reference to the case of the nth eigenfunction of a discrete spectrum, with a degeneracy of order s. Let the non-orthogonal eigenfunctions be $v_n^{(1)}, \ldots, v_n^{(s)}$, and let $u_n^{(1)}, \ldots, u_n^{(s)}$ be the linear combinations to be found. Then one prescribes $u_n^{(1)} = v_n^{(1)}$, $u_n^{(2)} = v_n^{(2)} + a_{21}\,u_n^{(1)}$ where a_{21} is such that $\langle u_n^{(1)}|u_n^{(2)}\rangle = 0$; thus

$$\langle u_n^{(1)}|v_n^{(2)}\rangle + a_{21}\,\langle u_n^{(1)}|u_n^{(1)}\rangle = 0\,, \qquad a_{21} = -\frac{\langle u_n^{(1)}|v_n^{(2)}\rangle}{\langle u_n^{(1)}|u_n^{(1)}\rangle}\,. \tag{8.27}$$

The next function is found by letting $u_n^{(3)} = v_n^{(3)} + a_{31}\,u_n^{(1)} + a_{32}\,u_n^{(2)}$, with $\langle u_n^{(1)}|u_n^{(3)}\rangle = 0$, $\langle u_n^{(2)}|u_n^{(3)}\rangle = 0$, whence

$$\langle u_n^{(1)}|v_n^{(3)}\rangle + a_{31}\,\langle u_n^{(1)}|u_n^{(1)}\rangle = 0\,, \qquad a_{31} = -\frac{\langle u_n^{(1)}|v_n^{(3)}\rangle}{\langle u_n^{(1)}|u_n^{(1)}\rangle}\,, \tag{8.28}$$

$$\langle u_n^{(2)}|v_n^{(3)}\rangle + a_{32}\,\langle u_n^{(2)}|u_n^{(2)}\rangle = 0\,, \qquad a_{32} = -\frac{\langle u_n^{(2)}|v_n^{(3)}\rangle}{\langle u_n^{(2)}|u_n^{(2)}\rangle}\,. \tag{8.29}$$

Similarly, the kth linear combination is built up recursively from the combinations of indices $1, \ldots, k-1$:

$$u_n^{(k)} = v_n^{(k)} + \sum_{i=1}^{k-1} a_{ki}\,u_n^{(i)}\,, \qquad a_{ki} = -\frac{\langle u_n^{(i)}|v_n^{(k)}\rangle}{\langle u_n^{(i)}|u_n^{(i)}\rangle}\,. \tag{8.30}$$

The denominators in (8.30) are different from zero because they are the squared norms of the previously defined combinations.

8.4.3 Completeness

As discussed in Sect. 8.2.1, the eigenfunctions of the Schrödinger equation for a free particle, for a given $k = \sqrt{2\,m\,E}/\hbar$ and apart from a multiplicative constant, are $w_{+k} = \exp(\mathrm{i}\,k\,x)$ and $w_{-k} = \exp(-\mathrm{i}\,k\,x)$. They may be written equivalently as $w(x,k) = \exp(\mathrm{i}\,k\,x)$, with $k = \pm\sqrt{2\,m\,E}/\hbar$. Taking the multiplicative constant equal to $1/\sqrt{2\,\pi}$, and considering a function f that fulfills the condition (C.19) for the Fourier representation, one applies (C.16) and (C.17) to find

$$f(x) = \int_{-\infty}^{+\infty} \frac{\exp(\mathrm{i}\,k\,x)}{\sqrt{2\,\pi}}\, c(k)\,\mathrm{d}k\,, \qquad c(k) = \int_{-\infty}^{+\infty} \frac{\exp(-\mathrm{i}\,k\,x)}{\sqrt{2\,\pi}} f(x)\,\mathrm{d}x\,. \tag{8.31}$$

Using the definition (8.13) of scalar product one recasts (8.31) as

$$f(x) = \int_{-\infty}^{+\infty} c(k)\, w(x,k)\,\mathrm{d}k\,, \qquad c(k) = \langle w|f\rangle\,. \tag{8.32}$$

In general the shorter notation $w_k(x)$, c_k is used instead of $w(x,k)$, $c(k)$. A set of functions like $w_k(x)$ that allows for the representation of f given by the first relation in (8.32) is said to be *complete*. Each member of the set is identified by the value of the continuous parameter k ranging from $-\infty$ to $+\infty$. To each k it corresponds a *coefficient* of the expansion, whose value is given by the second relation in (8.32).

Expressions (8.31) and (8.32) hold true because they provide the Fourier transform or antitransform of a function that fulfills (C.19). On the other hand, $w_k(x)$ is also the set of eigenfunctions of the free particle. In conclusion, the eigenfunctions of the Schrödinger equation for a free particle form a complete set.

The same conclusion is readily found for the eigenfunctions of the Schrödinger equation for a particle in a box. To show this, one considers a function $f(x)$ defined in an interval $[-\alpha/2, +\alpha/2]$ and fulfilling $\int_{-\alpha/2}^{+\alpha/2} |f(x)|\,\mathrm{d}x < \infty$. In this case the expansion into a Fourier series holds:

$$f(x) = \frac{1}{2}\,a_0 + \sum_{n=1}^{\infty} [a_n\, \cos(2\,\pi\,n\,x/\alpha) + b_n\, \sin(2\,\pi\,n\,x/\alpha)]\,, \tag{8.33}$$

with $a_0/2 = \bar{f} = (1/\alpha)\,\int_{-\alpha/2}^{+\alpha/2} f(x)\,\mathrm{d}x$ the average of f over the interval, and

$$\left\{\begin{matrix} a_n \\ b_n \end{matrix}\right\} = \frac{2}{\alpha} \int_{-\alpha/2}^{+\alpha/2} \left\{\begin{matrix} \cos \\ \sin \end{matrix}\right\} \left(\frac{2\,\pi\,n\,x}{\alpha}\right) f(x)\,\mathrm{d}x\,, \qquad n = 1, 2, \ldots \tag{8.34}$$

Equality (8.33) indicates convergence in the mean, namely, using $g = f - \bar{f}$ for the sake of simplicity, (8.33) is equivalent to

$$\lim_{N\to\infty} \int_{-\alpha/2}^{+\alpha/2} \left\{ g - \sum_{n=1}^{N} [a_n \cos(2\pi n x/\alpha) + b_n \sin(2\pi n x/\alpha)] \right\}^2 \mathrm{d}x = 0 \,. \tag{8.35}$$

Defining the auxiliary functions

$$\chi_n = \sqrt{2/\alpha}\, \cos(2\pi n x/\alpha)\,, \qquad \sigma_n = \sqrt{2/\alpha}\, \sin(2\pi n x/\alpha) \tag{8.36}$$

a more compact notation is obtained, namely, $f = \bar{f} + \sum_{n=1}^{\infty} (\langle \chi_n | f \rangle\, \chi_n + \langle \sigma_n | f \rangle\, \sigma_n)$ or, observing that $\langle \sigma_n | \mathrm{const} \rangle = \langle \chi_n | \mathrm{const} \rangle = 0$,

$$g = \sum_{n=1}^{\infty} (\langle \chi_n | g \rangle\, \chi_n + \langle \sigma_n | g \rangle\, \sigma_n) \,. \tag{8.37}$$

The norm of the auxiliary functions (8.36) is unity, $\langle \chi_n | \chi_n \rangle = \langle \sigma_n | \sigma_n \rangle = 1$ for $n = 1, 2, \ldots$, and all auxiliary functions are mutually orthogonal: $\langle \chi_m | \chi_n \rangle = \langle \sigma_m | \sigma_n \rangle = 0$ for $n, m = 0, 1, 2, \ldots$, $m \neq n$, and $\langle \sigma_m | \chi_n \rangle = 0$ for $n, m = 0, 1, 2, \ldots$ A set whose functions have a norm equal to unity and are mutually orthogonal is called *orthonormal*. Next, (8.37) shows that the set χ_n, σ_n, $n = 0, 1, 2, \ldots$ is complete in $[-\alpha/2, +\alpha/2]$ with respect to any g for which the expansion is allowed. Letting $c_{2n-1} = \langle \chi_n | g \rangle$, $c_{2n} = \langle \sigma_n | g \rangle$, $w_{2n-1} = \chi_n$, $w_{2n} = \sigma_n$, (8.37) takes the even more compact form

$$g = \sum_{m=1}^{\infty} c_m w_m \,, \qquad c_m = \langle w_m | g \rangle \,. \tag{8.38}$$

From the properties of the Fourier series it follows that the set of the σ_n functions alone is complete with respect to any function that is odd in $[-\alpha/2, +\alpha/2]$, hence it is complete with respect to any function over the half interval $[0, +\alpha/2]$. On the other hand, letting $a = \alpha/2$ and comparing with (8.8) show that σ_n (apart from the normalization coefficient) is the eigenfunction of the Schrödinger equation for a particle in a box. In conclusion, the set of eigenfunctions of this equation is complete within $[0, a]$.

One notes the striking resemblance of the first relation in (8.38) with the vector-algebra expression of a vector in terms of its components c_m. The similarity is completed by the second relation in (8.38), that provides each component as the projection of g over w_m. The latter plays the same role as the unit vector in algebra, the difference being that the unit vectors here are functions and that their number is infinite. A further generalization of the same concept is given by (8.32), where the summation index k is continuous.

Expansions like (8.32) or (8.38) hold because $w_k(x)$ and $w_m(x)$ are complete sets, whose completeness is demonstrated in the theory of Fourier's integral or series; such a theory is readily extended to the three-dimensional case, showing that also the three-dimensional counterparts of $w_k(x)$ or $w_m(x)$ form complete sets (in this case

the indices k or m are actually groups of indices, see, e.g., (9.5)). One may wonder whether other complete sets of functions exist, different from those considered in this section; the answer is positive: in fact, completeness is possessed by many other sets of functions,[9] and those of interest in Quantum Mechanics are made of the eigenfunctions of equations like (8.23). A number of examples will be discussed later.

8.4.4 Parseval Theorem

Consider the expansion of a complex function f with respect to a complete and orthonormal set functions w_n,

$$f = \sum_n c_n \, w_n \,, \qquad c_n = \langle w_n | f \rangle \,, \qquad \langle w_n | w_m \rangle = \delta_{nm} \,, \tag{8.39}$$

where the last relation on the right expresses the set's orthonormality. As before, m indicates a single index or a group of indices. The squared norm of f reads

$$||f||^2 = \int_\Omega |f|^2 \, \mathrm{d}\Omega = \left\langle \sum_n c_n \, w_n | \sum_m c_m \, w_m \right\rangle . \tag{8.40}$$

Applying (8.17, 8.18) yields

$$||f||^2 = \sum_n c_n^* \sum_m c_m \langle w_n | w_m \rangle = \sum_n c_n^* \sum_m c_m \, \delta_{nm} = \sum_n |c_n|^2 \,, \tag{8.41}$$

namely, the norm of the function equals the norm of the vector whose components are the expansion's coefficients (*Parseval theorem*). The result applies irrespective of the set that has been chosen for expanding f. The procedure leading to (8.41) must be repeated for the continuous spectrum, where the expansion reads

$$f = \int_\alpha c_\alpha \, w_\alpha \, \mathrm{d}\alpha \,, \qquad c_\alpha = \langle w_\alpha | f \rangle \,. \tag{8.42}$$

Here a difficulty seems to arise, related to expressing the counterpart of the third relation in (8.39). Considering for the sake of simplicity the case where a single index is present, the scalar product $\langle w_\alpha | w_\beta \rangle$ must differ from zero only for $\beta = \alpha$, while it must vanish for $\beta \neq \alpha$ no matter how small the difference $\alpha - \beta$ is. In other terms, for a given value of α such a scalar product vanishes for any β apart from a null set. At the same time, it must provide a finite value when used as a

[9] The completeness of a set of eigenfunctions must be proven on a case-by-case basis.

factor within an integral. An example taken from the case of a free particle shows that the requirements listed above are mutually compatible. In fact, remembering the analysis of Sect. 8.4.3, the scalar product corresponding to the indices α and β reads

$$\langle w_\alpha | w_\beta \rangle = \frac{1}{2\,\pi} \int_{-\infty}^{+\infty} \exp[\mathrm{i}\,(\beta - \alpha)\,x]\,\mathrm{d}x = \delta(\alpha - \beta)\,, \tag{8.43}$$

where the last equality is taken from (C.43). As mentioned in Sect. C.4, such an equality can be used only within an integral. In conclusion,[10]

$$\int_\Omega |f|^2\,\mathrm{d}\Omega = \langle f | f \rangle = \int_{-\infty}^{+\infty} c_\alpha^*\,\mathrm{d}\alpha \int_{-\infty}^{+\infty} c_\beta\,\delta(\alpha - \beta)\,\mathrm{d}\beta = \int_{-\infty}^{+\infty} |c_\alpha|^2\,\mathrm{d}\alpha\,. \tag{8.44}$$

One notes that (8.44) generalizes a theorem of Fourier's analysis that states that the norm of a function equals that of its transform.

8.5 Hamiltonian Operator and Momentum Operator

As mentioned in Sect. 7.5, the form (7.44) of the time-independent Schrödinger equation holds only when the force is conservative. It is readily recast in the more compact form (8.23) by defining the *Hamiltonian operator*

$$\mathscr{H} = -\frac{\hbar^2}{2\,m}\,\nabla^2 + V\,, \tag{8.45}$$

that is, a linear, real operator that gives (7.44) the form

$$\mathscr{H} w = E\,w\,. \tag{8.46}$$

The term used to denote $\mathscr{H}$ stems from the formal similarity of (8.46) with the classical expression $H(\mathbf{p}, \mathbf{q}) = E$ of a particle's total energy in a conservative field, where $H = T + V$ is the Hamiltonian function (Sect. 1.5). By this similarity, the classical kinetic energy $T = p^2/(2m)$ corresponds to the kinetic operator $\mathscr{T} = -\hbar^2/(2\,m)\,\nabla^2$; such a correspondence reads

$$T = \frac{1}{2\,m}\left(p_1^2 + p_2^2 + p_3^2\right) \quad \Longleftrightarrow \quad \mathscr{T} = -\frac{\hbar^2}{2\,m}\left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2}\right). \tag{8.47}$$

[10] The relation (8.44) is given here with reference to the specific example of the free particle's eigenfunctions. For other cases of continuous spectrum the relation $\langle w_\alpha | w_\beta \rangle = \delta(\alpha - \beta)$ is proven on a case-by-case basis.

The units of $\mathscr{T}$ are those of an energy, hence $\hbar^2\,\nabla^2$ has the units of a momentum squared. One notes that to transform T into $\mathscr{T}$ one must replace each component of momentum by a first-order operator as follows:

$$p_i \Leftarrow \hat{p}_i = -\mathrm{i}\,\hbar\,\frac{\partial}{\partial x_i}\,, \tag{8.48}$$

where $\hat{p}_i$ is called *momentum operator*. The correspondence (8.47) would still hold if the minus sign in (8.48) were omitted. However, the minus sign is essential for a correct description of the particle's motion.[11]

From the results of Sect. 8.6.1 one finds that the momentum operator and its three-dimensional form $\hat{\mathbf{p}} = -\mathrm{i}\,\hbar$ grad are Hermitean for square-integrable functions. Their units are those of a momentum. The Hamiltonian operator (8.45) is a real-coefficient, linear combination of ∇^2 and V; combining (8.22) with the findings of Sect. 8.6 shows that (8.45) is Hermitean for square-integrable functions.

The one-dimensional form of the momentum operator yields the eigenvalue equation

$$-\,\mathrm{i}\,\hbar\,\frac{\mathrm{d}v}{\mathrm{d}x} = \tilde{p}\,v\,, \tag{8.49}$$

where $\tilde{p}$ has the units of a momentum. The solution of (8.49) is $v = \text{const} \times \exp(\mathrm{i}\,\tilde{p}/\hbar)$, where $\tilde{p}$ must be real to prevent the solution from diverging. Letting $\text{const} = 1/\sqrt{2\,\pi}$, $k = \tilde{p}/\hbar$ yields $v = v_k(x) = \exp(\mathrm{i}\,k\,x)/\sqrt{2\,\pi}$, showing that the eigenfunctions of the momentum operator form a complete set (compare with (8.31)) and are mutually orthogonal (compare with (8.43)). As $|v_k(x)|^2 = 1/(2\,\pi)$, the eigenfunctions are not square integrable; the spectrum is continuous because the eigenvalue $\hbar\,k$ can be any real number.

8.6 Complements

8.6.1 Examples of Hermitean Operators

A real function V, depending on the spatial coordinates over the domain Ω, and possibly on other variables $\alpha, \beta, \ldots$, may be thought of as a purely multiplicative operator. Such an operator is Hermitean[12]; in fact,

[11] Consider for instance the calculation of the expectation value of the momentum of a free particle based on (10.18). If the minus sign were omitted in (8.48), the direction of momentum would be opposite to that of the propagation of the wave front associated with it.

[12] Some definitions and properties illustrated in this section have an analogue in the theory of matrices outlined in Sect. A.11.

$$\int_\Omega g^* V f \, \mathrm{d}\Omega = \int_\Omega V g^* f \, \mathrm{d}\Omega = \int_\Omega (Vg)^* f \, \mathrm{d}\Omega \,. \tag{8.50}$$

In contrast, an imaginary function $W = \mathrm{i}\, V$, with V real, is not Hermitean because

$$\langle g|Wf\rangle = -\langle W g|f\rangle \,. \tag{8.51}$$

Any operator that fulfills a relation similar to (8.51) is called *anti-Hermitean* or *skew-Hermitean*.

As a second example consider a one-dimensional case defined over a domain Ω belonging to the x axis. It is easily shown that the operator $(\mathrm{i}\,\mathrm{d}/\mathrm{d}x)$ is Hermitean: in fact, integrating by parts and observing that the integrated part vanishes because f and g are square integrable yield

$$\int_\Omega g^* \,\mathrm{i}\, \frac{\mathrm{d}f}{\mathrm{d}x} \, \mathrm{d}\Omega = \left[g^* \,\mathrm{i} f\right]_\Omega - \int_\Omega \mathrm{i}\, \frac{\mathrm{d}g^*}{\mathrm{d}x} f \, \mathrm{d}\Omega = \int_\Omega \left(\mathrm{i}\, \frac{\mathrm{d}g}{\mathrm{d}x}\right)^* f \, \mathrm{d}\Omega \,. \tag{8.52}$$

By the same token one shows that the operator $\mathrm{d}/\mathrm{d}x$ is skew-Hermitean. The three-dimensional generalization of $(\mathrm{i}\,\mathrm{d}/\mathrm{d}x)$ is (i grad). Applying the latter onto the product $g^* f$ yields $g^* \,\mathrm{i}\, \mathrm{grad} f - (\mathrm{i}\ \mathrm{grad}\, g)^* f$. Integrating over Ω with Σ the boundary of Ω and $\mathbf{n}$ the unit vector normal to it yields

$$\int_\Omega g^* \,\mathrm{i}\ \mathrm{grad} f \, \mathrm{d}\Omega - \int_\Omega (\mathrm{i}\ \mathrm{grad}\, g)^* \, f \, \mathrm{d}\Omega = \mathrm{i} \int_\Sigma g^* f \, \mathbf{n} \, \mathrm{d}\Sigma \,. \tag{8.53}$$

The form of the right-hand side of (8.53) is due to (A.25). As f, g vanish over the boundary, it follows $\langle g|\mathrm{i}\ \mathrm{grad} f\rangle = \langle \mathrm{i}\, \mathrm{grad}\, g|f\rangle$, namely, (i grad) is Hermitean.

Another important example, still in the one-dimensional case, is that of the operator $\mathrm{d}^2/\mathrm{d}x^2$. Integrating by parts twice shows that the operator is Hermitean. Its three-dimensional generalization in Cartesian coordinates is ∇^2. Using the second Green theorem (A.25) and remembering that f, g vanish over the boundary provides $\langle g|\nabla^2 f\rangle = \langle \nabla^2 g|f\rangle$, that is, ∇^2 is Hermitean.

8.6.2 A Collection of Operators' Definitions and Properties

A number of definitions and properties of operator algebra are illustrated in this section. The *identity operator* $\mathscr{I}$ is such that $\mathscr{I} f = f$ for all f; the *null operator* $\mathscr{O}$ is such that $\mathscr{O} f = 0$ for all f. The product of two operators, $\mathscr{A}\mathscr{B}$, is an operator whose action on a function is defined as follows: $s = \mathscr{A}\mathscr{B} f$ is equivalent to $g = \mathscr{B} f$, $s = \mathscr{A} g$; in other terms, the operators $\mathscr{A}$ and $\mathscr{B}$ act in a specific order. In general, $\mathscr{B}\mathscr{A} \neq \mathscr{A}\mathscr{B}$. The operators $\mathscr{A}\mathscr{A}$, $\mathscr{A}\mathscr{A}\mathscr{A}$, ... are indicated with $\mathscr{A}^2$, $\mathscr{A}^3$, ...

An operator $\mathscr{A}$ may or may not have an *inverse*, $\mathscr{A}^{-1}$. If the inverse exists, it is unique and has the property $\mathscr{A}^{-1}\mathscr{A} f = f$ for all f. Left multiplying the above by

$\mathscr{A}$ and letting $g = \mathscr{A} f$ yield $\mathscr{A}\mathscr{A}^{-1} g = g$ for all g. The two relations just found can be recast as

$$\mathscr{A}^{-1}\mathscr{A} = \mathscr{A}\mathscr{A}^{-1} = \mathscr{I} \,. \tag{8.54}$$

From (8.54) it follows $(\mathscr{A}^{-1})^{-1} = \mathscr{A}$. If $\mathscr{A}$ and $\mathscr{B}$ have an inverse, letting $\mathscr{C} = \mathscr{B}^{-1}\mathscr{A}^{-1}$ one finds, for all f and using the associative property, the two relations $\mathscr{B}\mathscr{C} f = \mathscr{B}\mathscr{B}^{-1}\mathscr{A}^{-1} f = \mathscr{A}^{-1} f$ and $\mathscr{A}\mathscr{B}\mathscr{C} f = \mathscr{A}\mathscr{A}^{-1} f = f$, namely, $\mathscr{A}\mathscr{B}\mathscr{C} = \mathscr{I}$; in conclusion,

$$(\mathscr{A}\mathscr{B})^{-1} = \mathscr{B}^{-1}\mathscr{A}^{-1} \,. \tag{8.55}$$

From (8.55) one defines the inverse powers of $\mathscr{A}$ as

$$\mathscr{A}^{-2} = (\mathscr{A}^2)^{-1} = (\mathscr{A}\mathscr{A})^{-1} = \mathscr{A}^{-1}\mathscr{A}^{-1} \,, \tag{8.56}$$

and so on. Let $\mathscr{A} v = \lambda v$ be the eigenvalue equation of $\mathscr{A}$. Successive left multiplications by $\mathscr{A}$ yield

$$\mathscr{A}^2 v = \lambda^2 v \,, \quad \mathscr{A}^3 v = \lambda^3 v \,, \quad \ldots \tag{8.57}$$

As a consequence, an operator of the polynomial form

$$P_n(\mathscr{A}) = c_0 \mathscr{A}^n + c_1 \mathscr{A}^{n-1} + c_2 \mathscr{A}^{n-2} + \ldots + c_n \tag{8.58}$$

fulfills the eigenvalue equation

$$P_n(\mathscr{A})\, v = P_n(\lambda)\, v \,, \qquad P_n(\lambda) = c_0 \lambda^n + \ldots + c_n \,. \tag{8.59}$$

By definition, an eigenfunction cannot vanish identically. If $\mathscr{A}$ has an inverse, left-multiplying the eigenvalue equation $\mathscr{A} v = \lambda v$ by $\mathscr{A}^{-1}$ yields $v = \lambda \mathscr{A}^{-1} v \neq 0$, whence $\lambda \neq 0$. Dividing the latter by λ and iterating the procedure show that

$$\mathscr{A}^{-2} v = \lambda^{-2} v \,, \quad \mathscr{A}^{-3} v = \lambda^{-3} v \,, \quad \ldots \tag{8.60}$$

An operator may be defined by a series expansion, if the latter converges:

$$\mathscr{C} = \sigma(\mathscr{A}) = \sum_{k=-\infty}^{+\infty} c_k \mathscr{A}^k \,. \tag{8.61}$$

By way of example,

$$\mathscr{C} = \exp(\mathscr{A}) = \mathscr{I} + \mathscr{A} + \frac{1}{2!}\mathscr{A}^2 + \frac{1}{3!}\mathscr{A}^3 + \ldots \tag{8.62}$$

Given an operator $\mathscr{A}$, its *adjoint* $\mathscr{A}^\dagger$ is defined as in Sect. 8.3.1. Letting $\mathscr{C} = \mathscr{A}^\dagger$, applying the definition of adjoint operator to $\mathscr{C}$, and taking the conjugate of both sides show that $(\mathscr{A}^\dagger)^\dagger = \mathscr{A}$. From the definition of adjoint operator it also follows

$$(\mathscr{A}\mathscr{B})^\dagger = \mathscr{B}^\dagger \mathscr{A}^\dagger \,. \tag{8.63}$$

An operator is *unitary* if its inverse is identical to its adjoint for all f:

$$\mathscr{A}^{-1} f = \mathscr{A}^\dagger f \,. \tag{8.64}$$

Left multiplying (8.64) by $\mathscr{A}$, and left multiplying the result by $\mathscr{A}^\dagger$, yields for a unitary operator

$$\mathscr{A}\mathscr{A}^\dagger = \mathscr{A}^\dagger \mathscr{A} = \mathscr{I} \,. \tag{8.65}$$

The application of a unitary operator to a function f leaves the norm of the latter unchanged. In fact, using definition (8.12), namely, $||f||^2 = \langle f | f \rangle$, and letting $g = \mathscr{A} f$ with $\mathscr{A}$ unitary, yields

$$||g||^2 = \int_\Omega (\mathscr{A} f)^* \, \mathscr{A} f \, \mathrm{d}\Omega = \int_\Omega (\mathscr{A}^\dagger \mathscr{A} f)^* f \, \mathrm{d}\Omega = \int_\Omega f^* f \, \mathrm{d}\Omega = ||f||^2 \,, \tag{8.66}$$

where the second equality holds due to the definition of adjoint operator, and the third one holds because $\mathscr{A}$ is unitary. The inverse also holds true: if the application of $\mathscr{A}$ leaves the function's norm unchanged, that is, if $||\mathscr{A} f|| = ||f||$ for all f, then

$$\int_\Omega (\mathscr{A}^\dagger \mathscr{A} f - f)^* f \, \mathrm{d}\Omega = 0 \,. \tag{8.67}$$

As a consequence, the quantity in parenthesis must vanish, whence the operator is unitary. The product of two unitary operators is unitary:

$$(\mathscr{A}\mathscr{B})^{-1} = \mathscr{B}^{-1}\mathscr{A}^{-1} = \mathscr{B}^\dagger \mathscr{A}^\dagger = (\mathscr{A}\mathscr{B})^\dagger \,, \tag{8.68}$$

where the second equality holds because $\mathscr{A}$ and $\mathscr{B}$ are unitary. The eigenvalues of a unitary operator have the form $\exp(\mathrm{i}\,\nu)$, with ν a real number. Let an eigenvalue equation be $\mathscr{A} v = \lambda v$, with $\mathscr{A}$ unitary. The following hold,

$$\int_\Omega |\mathscr{A} v|^2 \, \mathrm{d}\Omega = |\lambda|^2 \int_\Omega |v|^2 \, \mathrm{d}\Omega \,, \qquad \int_\Omega |\mathscr{A} v|^2 \, \mathrm{d}\Omega = \int_\Omega |v|^2 \, \mathrm{d}\Omega \,, \tag{8.69}$$

the first one because of the eigenvalue equation, the second one because $\mathscr{A}$ is unitary. As an eigenfunction cannot vanish identically, it follows $|\lambda|^2 = 1$ whence $\lambda = \exp(\mathrm{i}\,\nu)$. It is also seen by inspection that if the eigenvalues of an operator have the form $\exp(\mathrm{i}\,\nu)$, with ν a real number, then the operator is unitary.

It has been anticipated above that in general it is $\mathscr{B}\mathscr{A} \neq \mathscr{A}\mathscr{B}$. Two operators $\mathscr{A}$, $\mathscr{B}$ are said to *commute* if

$$\mathscr{B}\mathscr{A}f = \mathscr{A}\mathscr{B}f \tag{8.70}$$

for all f. The *commutator* of $\mathscr{A}$, $\mathscr{B}$ is the operator $\mathscr{C}$ such that

$$\mathrm{i}\,\mathscr{C}f = (\mathscr{A}\mathscr{B} - \mathscr{B}\mathscr{A})f \tag{8.71}$$

for all f. The definition (8.71) is such that if both $\mathscr{A}$ and $\mathscr{B}$ are Hermitean, then $\mathscr{C}$ is Hermitean as well. The commutator of two commuting operators is the null operator. A very important example of non-commuting operators is the pair q, $-\mathrm{i}\,\mathrm{d}/\mathrm{d}q$, where q is any dynamic variable. One finds

$$\mathrm{i}\,\mathscr{C}f = -\mathrm{i}\,q\,\frac{\mathrm{d}f}{\mathrm{d}q} + \mathrm{i}\,\frac{\mathrm{d}(qf)}{\mathrm{d}q} = \mathrm{i}f\,, \tag{8.72}$$

namely, the commutator is in this case the identity operator $\mathscr{I}$.

8.6.3 Examples of Commuting Operators

Operators that contain only spatial coordinates commute; similarly, operators that contain only momentum operators commute. The operators $\mathscr{A}$, $\mathscr{B}$, $\mathscr{C}$ defined in (10.4) commute because they act on different coordinates; note that the definition of $\mathscr{A}$ is such that it may contain both x and $\hat{p}_x = -\mathrm{i}\,\hbar\,\partial/\partial x$, and so on.

As an example of operators containing only momentum operators one may consider the Hamiltonian operator $-(\hbar^2/2\,m)\,\nabla^2$ of a free particle discussed in Sect. 8.2.1 and the momentum operator $-\mathrm{i}\,\hbar\,\nabla$ itself (Sect. 8.5). As for a free particle they commute, a measurement of momentum is compatible in that case with a measurement of energy. Considering a one-dimensional problem, the energy is $E = p^2/(2\,m)$, where the modulus of momentum is given by $p = \hbar\,k$; for a free particle, both energy and momentum are conserved. The eigenfunctions are const $\times \exp(\pm\mathrm{i}\,p\,x/\hbar)$ for both operators.

Remembering (8.72) one concludes that two operators do not commute if one of them contains one coordinate q, and the other one contains the operator $-\mathrm{i}\,\hbar\,\partial/\partial q$ associated with the momentum conjugate to q.

8.6.4 Momentum and Energy of a Free Particle

The eigenfunctions of the momentum operator are the same as those of the Schrödinger equation for a free particle. More specifically, given the sign of $\tilde{p}$,

the solution of (8.49) coincides with either one or the other of the two linearly independent solutions of (8.1). This outcome is coherent with the conclusions reached in Sect. 8.2.1 about the free particle's motion. For a free particle whose momentum is prescribed, the energy is purely kinetic and is prescribed as well, whence the solution of (8.49) must be compatible with that of (8.1). However, prescribing the momentum, both in modulus and direction, for a free particle, provides the additional information that allows one to eliminate one of the two summands from the linear combination (8.2) by setting either c_1 or c_2 to zero. For a given eigenvalue $\tilde{p}$, (8.49) has only one solution (apart from the multiplicative constant) because it is a first-order equation; in contrast, for a given eigenvalue E, the second-order equation (8.1) has two independent solution and its general solution is a linear combination of them.

In a broader sense the momentum operator $\hat{p}_x = -\mathrm{i}\,\hbar\,\mathrm{d}/\mathrm{d}x$ is Hermitean also for functions of the form $v_k(x) = \exp(\mathrm{i}\,k\,x)/\sqrt{2\,\pi}$, which are not square integrable. In fact, remembering (C.43) one finds

$$\langle v_{k'}|\hat{p}_x v_k\rangle = -\frac{\mathrm{i}\,\hbar}{2\,\pi}\int_{-\infty}^{+\infty}\exp(-\mathrm{i}\,k'\,x)\,\frac{\mathrm{d}}{\mathrm{d}x}\exp(\mathrm{i}\,k\,x)\,\mathrm{d}x = \hbar\,k\,\delta(k'-k)\,. \tag{8.73}$$

Similarly it is $\langle \hat{p}_x v_{k'}|v_k\rangle = \hbar\,k'\,\delta(k'-k)$. As mentioned in Sect. C.4, the two equalities just found can be used only within an integral over k or k'. In that case, however, they yield the same result $\hbar\,k$. By the same token one shows that

$$\langle v_{k'}|\hat{p}_x^2 v_k\rangle = \hbar^2\,k^2\,\delta(k'-k)\,, \qquad \langle \hat{p}_x^2 v_{k'}|v_k\rangle = \hbar^2\,(k')^2\,\delta(k'-k)\,, \tag{8.74}$$

hence the Laplacian operator is Hermitean in a broader sense for non-square-integrable functions of the form $v_k(x) = \exp(\mathrm{i}\,k\,x)/\sqrt{2\,\pi}$.

Problems

8.1 The one-dimensional, time-independent Schrödinger equation is a homogeneous equation of the form

$$w'' + q\,w = 0\,, \qquad q = q(x)\,, \tag{8.75}$$

where primes indicate derivatives. In turn, the most general, linear equation of the second order with a nonvanishing coefficient of the highest derivative is

$$f'' + a f' + b f = c\,, \qquad a = a(x)\,, \quad b = b(x)\,, \quad c = c(x)\,. \tag{8.76}$$

Assume that a is differentiable. Show that if the solution of (8.75) is known, then the solution of (8.76) is obtained from the former by simple integrations.

8.2 Consider the one-dimensional, time-independent Schrödinger equation $w'' + q\,w = 0$, where $q(x) = \lambda - g(x)$. Being the above a linear, second-order equation, it is expected to have, at most, two linearly independent solutions for each eigenvalue $\lambda = 2\,m\,E/\hbar^2$. Letting w_1, w_2 be the two solutions corresponding to the same λ, show that the Wronskian of the two solutions is equal to a constant (the definition of Wronskian is in Sect. A.12).

8.3 Given the one-dimensional, time-independent Schrödinger equation $w'' + q\,w = 0$, where $q(x) = \lambda - g(x)$, let w_1, w_2 be two linearly independent solutions corresponding to the eigenvalue $\lambda = 2\,m\,E/\hbar^2$. Show that if the solution w_1 is known, one can obtain the other solution w_2 by simple integrations. Hint: consider an interval I where $w_1 \neq 0$ and let $w_2 = v\,w_1$, with v to be determined.

8.4 Consider the one-dimensional, time-independent Schrödinger equation $w'' + q\,w = 0$, where $q(x) = \lambda - g(x)$ with $\lambda = 2\,m\,E/\hbar^2$ and $g = 2\,m\,V/\hbar^2$. Show that if the eigenvalue E is simple and the potential energy V is even, the eigenfunction corresponding to E is either even or odd.

Chapter 9
Time-Dependent Schrödinger Equation

9.1 Introduction

The time-dependent Schrödinger equation is derived from the superposition principle, in the conservative case first, then in the general case. The derivation of the continuity equation follows, leading to the concept of wave packet and density of probability flux. Then, the wave packet for a free particle is investigated in detail, and the concept of group velocity is introduced. The first complement deals with an application of the semiclassical approximation; through it one explains why an electron belonging to a stationary state emits no power, namely, why the radiative decay predicted by the classical model does not occur. The polar form of the time-dependent Schrödinger equation is then shown, that brings about an interesting similarity with the Hamilton-Jacobi equation of Classical Mechanics. The last complement deals with the Hamiltonian operator of a particle subjected to an electromagnetic field and shows the effect of a gauge transformation on the wave function.

9.2 Superposition Principle

Following de Broglie's line of reasoning one associates the monochromatic wave function $w(\mathbf{r})\,\exp(-\mathrm{i}\,\omega\,t)$ with the motion of a particle with definite and constant energy $E = \hbar\,\omega$. The analogy with the electromagnetic case then suggests that a more general type of wave function—still related to the conservative case—can be expressed as a superposition, that is, a linear combination with constant coefficients, of monochromatic wave functions. This possibility is one of the postulates of de Broglie's theory, and is referred to as *Superposition Principle*. To vest it with a

M. Rudan, *Physics of Semiconductor Devices*, DOI 10.1007/978-3-319-63154-7_9

mathematical form one must distinguish among the different types of spectrum; for the discrete spectrum, indicating with c_n the complex coefficients of the linear combination, the general wave function reads

$$\psi(\mathbf{r}, t) = \sum_n c_n \, w_n \, \exp(-\mathrm{i}\,\omega_n \, t) \,, \tag{9.1}$$

with E_n, w_n the eigenvalues and eigenfunctions of the time-independent Schrödinger equation $\mathscr{H} w_n = E_n \, w_n$, and $\omega_n = E_n/\hbar$. As usual, n stands for a single index or a set of indices. The form of (9.1) is such that the spatial coordinates are separated from the time coordinate; fixing the latter by letting, say, $t = 0$, and remembering that the set of eigenfunctions w_n is complete, yield

$$\psi_{t=0} = \psi(\mathbf{r}, 0) = \sum_n c_n \, w_n \,, \qquad c_n = \langle w_n | \psi_{t=0} \rangle \,. \tag{9.2}$$

The above shows that the coefficients c_n are uniquely determined by the initial condition $\psi_{t=0}$. On the other hand, once the coefficients c_n are known, the whole time evolution of ψ is determined, because the angular frequencies appearing in the time-dependent terms $\exp(-\mathrm{i}\,\omega_n \, t)$ are also known. In other terms, ψ is determined by the initial condition and by the time-independent Hamiltonian operator whence E_n, w_n derive.

An important aspect of (9.1) is that it allows one to construct a wave function of a given form; for such a construction, in fact, it suffices to determine the coefficients by means of the second relation in (9.2). In particular it is possible to obtain a wave function that is square integrable at all times, even if the eigenfunctions w_n are not square integrable themselves. Thanks to this property the wave function (9.1) is localized in space at each instant of time, hence it is suitable for describing the motion of the particle associated with it. Due to the analogy with the electromagnetic case, where the interference of monochromatic waves provides the localization of the field's intensity, a wave function of the form (9.1) is called *wave packet*. Remembering that the wave function provides the probability density $|\psi|^2$ used to identify the position of the particle, one can assume that the wave packet's normalization holds:

$$\int_\Omega |\psi|^2 \, \mathrm{d}^3 r = \sum_n |c_n|^2 = 1 \,, \tag{9.3}$$

where the second equality derives from Parseval's theorem (8.41). From (9.3) it follows that the coefficients are subjected to the constraint $0 \leq |c_n|^2 \leq 1$.

As all possible energies E_n appear in the expression (9.1) of ψ, the wave packet does not describe a motion with a definite energy. Now, assume that an energy measurement is carried out on the particle, and let $t = t_E$ be the instant at which the measurement is completed. During the measurement the Hamiltonian operator of (8.46) does not hold because the particle is interacting with the measuring

apparatus, hence the forces acting on it are different from those whence the potential energy V of (8.45) derives. Instead, for $t > t_E$ the original Schrödinger equation (8.46) is restored, so the expression of ψ is again given by a linear combination of monochromatic waves; however, the coefficients of the combination are expected to be different from those that existed prior to the measurement, due to the perturbation produced by the latter. In particular, the form of ψ for $t > t_E$ must be compatible with the fact that the energy measurement has found a specific value of the energy; this is possible only if the coefficients are set to zero, with the exception of the one corresponding to the energy that is the outcome of the measurement. The latter must be one of the eigenvalues of (8.46) due to the compatibility requirement; if it is, say, E_m, then the form of the wave function for $t > t_E$ is

$$\psi(\mathbf{r}, t) = w_m \exp[-\mathrm{i}\, E_m\, (t - t_E)/\hbar]\,, \tag{9.4}$$

where the only nonvanishing coefficient, c_m, has provisionally been set to unity.

The reasoning leading to (9.4) can be interpreted as follows: the interaction with the measuring apparatus filters out from (9.1) the term corresponding to E_m; as a consequence, the coefficients c_n whose values were previously set by the original ψ are modified by the measurement and become $c_n = \delta_{nm}$ when $n \neq m$. If the filtered eigenfunction w_m is square integrable, then (9.3) holds, whence $\sum_n |c_n|^2 = |c_m|^2 = 1$, $c_m = \exp(\mathrm{i}\, \Phi)$. As the constant phase Φ does not carry any information, it can be set to zero to yield $c_m = 1$. If w_m is not square integrable, the arbitrariness of the multiplicative constant still allows one to set $c_m = 1$.

As the energy measurement forces the particle to belong to a definite energy state (in the example above, E_m), for $t > t_E$ the particle's wave function keeps the monochromatic form (9.4). If, at a later time, a second energy measurement is carried out, the only possible outcome is E_m; as a consequence, after the second measurement is completed, the form of the wave function is still (9.4), whence $|c_m|^2 = 1$. One notes that the condition $|c_m|^2 = 1$ is associated with the certainty that the outcome of the energy measurement is E_m whereas, when the general superposition (9.1) holds, the coefficients fulfill the relations $\sum_n |c_n|^2 = 1$, $0 \leq |c_n|^2 \leq 1$, and the measurement's outcome may be any of the eigenvalues. It is then sensible to interpret $|c_n|^2$ as the *probability* that a measurement of energy finds the result E_n. This interpretation, that has been worked out here with reference to energy, is extended to the other dynamic variables (Sect. 10.2).

When the spectrum is continuous the description of the wave packet is

$$\psi(\mathbf{r}, t) = \iiint_{-\infty}^{+\infty} c_{\mathbf{k}}\, w_{\mathbf{k}} \exp(-\mathrm{i}\, \omega_{\mathbf{k}}\, t)\, \mathrm{d}^3 k\,, \tag{9.5}$$

with $E_{\mathbf{k}}$, $w_{\mathbf{k}}$ the eigenvalues and eigenfunctions of $\mathscr{H} w_{\mathbf{k}} = E_{\mathbf{k}}\, w_{\mathbf{k}}$, and $\omega_{\mathbf{k}} = E_{\mathbf{k}}/\hbar$. Such symbols stand for $E_{\mathbf{k}} = E(\mathbf{k})$, $w_{\mathbf{k}} = w(\mathbf{r}, \mathbf{k})$, and so on, with $\mathbf{k}$ a three-dimensional vector whose components are continuous. The relations corresponding to (9.2, 9.3) are

$$\psi_{t=0} = \psi(\mathbf{r}, 0) = \iiint_{-\infty}^{+\infty} c_{\mathbf{k}}\, w_{\mathbf{k}}\, \mathrm{d}^3 k\,, \qquad c_{\mathbf{k}} = \langle w_{\mathbf{k}} | \psi_{t=0} \rangle\,, \tag{9.6}$$

$$\int_{\Omega} |\psi|^2\, \mathrm{d}^3 r = \iiint_{-\infty}^{+\infty} |c_{\mathbf{k}}|^2\, \mathrm{d}^3 k = 1\,. \tag{9.7}$$

The expression of $\psi_{t=0}$ in (9.6) lends itself to providing an example of a wave function that is square integrable, while the eigenfunctions that build up the superposition are not. Consider, in fact, the relation (C.83) and let $c_k = \sigma\, \exp(-\sigma^2 k^2/2)$, $w_k = \exp(\mathrm{i}\, k\, x)/\sqrt{2\pi}$ in it, with σ a length; in this way (C.83) becomes the one-dimensional case of (9.6) and yields $\psi_{t=0} = \exp[-x^2/(2\,\sigma^2)]$, showing that a square-integrable function like the Gaussian one can be expressed as a combination of the non-square-integrable spatial parts of the plane waves.

The extraction of the probabilities in the continuous-spectrum case accounts for the fact that the $E(\mathbf{k})$ varies continuously with $\mathbf{k}$. To this purpose one takes the elementary volume $\mathrm{d}^3 k$ centered on some $\mathbf{k}$ and considers the product $|c_{\mathbf{k}}|^2\, \mathrm{d}^3 k$. Such a product is given the meaning of infinitesimal probability that the outcome of an energy measurement belongs to the range of $E(\mathbf{k})$ values whose domain is $\mathrm{d}^3 k$ (more comments are made in Sect. 9.7.1).

9.3 Time-Dependent Schrödinger Equation

The Superposition Principle illustrated in Sect. 9.2 prescribes the form of the wave packet in the conservative case. Considering for simplicity a discrete set of eigenfunctions, the time derivative of ψ reads

$$\frac{\partial \psi}{\partial t} = \sum_n c_n\, w_n\, \frac{E_n}{\mathrm{i}\,\hbar}\, \exp(-\mathrm{i} E_n\, t/\hbar)\,. \tag{9.8}$$

Using the time-independent Schrödinger equation $\mathscr{H} w_n = E_n w_n$ transforms the above into

$$\mathrm{i}\,\hbar\, \frac{\partial \psi}{\partial t} = \sum_n c_n\, \mathscr{H} w_n\, \exp(-\mathrm{i}\, E_n\, t/\hbar)\,, \qquad \mathrm{i}\,\hbar\, \frac{\partial \psi}{\partial t} = \mathscr{H} \psi\,. \tag{9.9}$$

The second relation in (9.9) is a linear, homogeneous partial-differential equation, of the second order with respect to the spatial coordinates and of the first order with respect to time, whose solution is the wave function ψ. It is called *time-dependent Schrödinger equation*; as its coefficients are complex, so is ψ. To solve the equation it is necessary to prescribe the initial condition $\psi(\mathbf{r}, t = 0)$ and the boundary conditions. For the latter the same discussion as in Sect. 8.2 applies, because the spatial behavior of ψ is prescribed by the Hamiltonian operator.

The reasoning leading to the second relation in (9.9) is based on the Superposition Principle, namely, once the form of ψ is given, the equation fulfilled by ψ is readily extracted. Such a reasoning is not applicable in the nonconservative cases, because the time-independent equation $\mathscr{H} w_n = E_n w_n$ does not hold then, so the eigenvalues and eigenfunctions upon which the superposition is based are not available. However, another line of reasoning shows that the time-dependent Schrödinger equation holds also for the nonconservative situations [86]. Although in such cases the wave function ψ is not expressible as a superposition of monochromatic waves, it can still be expanded using an orthonormal set. If the set is discrete, $v_n = v_n(\mathbf{r})$, the expansion reads

$$\psi(\mathbf{r}, t) = \sum_n b_n \, v_n(\mathbf{r}) \,, \qquad b_n(t) = \langle v_n | \psi \rangle \,, \tag{9.10}$$

whereas for a continuous set $v_\mathbf{k} = v(\mathbf{r}, \mathbf{k})$ one finds

$$\psi(\mathbf{r}, t) = \iiint_{-\infty}^{+\infty} b_\mathbf{k} \, v_\mathbf{k} \, \mathrm{d}^3 k \,, \qquad b_\mathbf{k}(t) = \langle v_\mathbf{k} | \psi \rangle \,. \tag{9.11}$$

9.4 Continuity Equation and Norm Conservation

Remembering that the square modulus of the wave function provides the localization of the particle, it is of interest to investigate the time evolution of $|\psi|^2$, starting from the time derivative of ψ given by the time-dependent Schrödinger equation (9.9). Here it is assumed that the wave function is normalized to unity and that the Hamiltonian operator is real, $\mathscr{H}^* = \mathscr{H} = -\hbar^2/(2\,m)\,\nabla^2 + V$; a case where the operator is complex is examined in Sect. 9.5. Taking the time derivative of $|\psi|^2$ yields

$$\frac{\partial |\psi|^2}{\partial t} = \psi^* \frac{\partial \psi}{\partial t} + \psi \frac{\partial \psi^*}{\partial t} = \psi^* \frac{\mathscr{H} \psi}{\mathrm{i}\,\hbar} - \psi \frac{\mathscr{H} \psi^*}{\mathrm{i}\,\hbar} \,, \tag{9.12}$$

with $\psi^* \mathscr{H} \psi - \psi \mathscr{H} \psi^* = -\hbar^2/(2m)\,(\psi^* \nabla^2 \psi - \psi \nabla^2 \psi^*)$. Identity (A.17) then yields

$$\frac{\partial |\psi|^2}{\partial t} + \operatorname{div} \mathbf{J}_\psi = 0 \,, \qquad \mathbf{J}_\psi = \frac{\mathrm{i}\,\hbar}{2\,m} \left(\psi \, \operatorname{grad} \psi^* - \psi^* \, \operatorname{grad} \psi \right) . \tag{9.13}$$

The first relation in (9.13) has the form of a continuity equation (compare with (23.3) and (4.23)). As $|\psi|^2$ is the probability density, $\mathbf{J}_\psi$ takes the meaning of *density of the*

probability flux[1]; it is a real quantity because the term in parentheses in the second relation of (9.13) is imaginary.

Relations (9.13) provide a physical explanation of the continuity requirements that were discussed from the mathematical standpoint in Sect. 8.2. Such requirements, namely, the continuity of the wave function and of its first derivatives in space, were introduced in Sect. 8.2 with reference to the solutions of the time-independent Schrödinger equation; however, they hold also for the time-dependent one because the spatial behavior of ψ is prescribed by the Hamiltonian operator. Their physical explanation is that they provide the spatial continuity of the probability density and of the probability-flux density. Integrating (9.13) over a volume Ω' whose surface is Σ' yields

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{\Omega'} |\psi|^2 \, \mathrm{d}\Omega' = - \int_{\Sigma'} \mathbf{J}_\psi \cdot \mathbf{n} \, \mathrm{d}\Sigma' \,, \tag{9.14}$$

with $\mathbf{n}$ the unit vector normal to Σ', oriented in the outward direction. The integral at the left-hand side of (9.14) is the probability of localizing the particle within Ω', that at the right-hand side is the probability flux across Σ' in the outward direction; as a consequence, the meaning of (9.14) is that the time variation of the localization probability within Ω' is the negative probability flux across the surface. If $\Omega' \to \infty$ the surface integral vanishes because ψ is square integrable and, as expected,

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{\infty} |\psi|^2 \, \mathrm{d}\Omega = 0 \,. \tag{9.15}$$

The above is another form of the normalization condition and is also termed *norm-conservation condition*. Note that the integral in (9.15) does not depend on time although ψ does.

The density of the probability flux can be given a different form that uses the momentum operator $\hat{\mathbf{p}} = -\mathrm{i}\,\hbar$ grad introduced in Sect. 8.5; one finds

$$\mathbf{J}_\psi = \frac{1}{2\,m} \left[\psi \, (\hat{\mathbf{p}}\psi)^* + \psi^* \, \hat{\mathbf{p}}\psi \right] = \frac{1}{m} \, \Re \left(\psi^* \, \hat{\mathbf{p}}\psi \right) \,. \tag{9.16}$$

Although this form is used less frequently than (9.13), it makes the analogy with the classical flux density much more intelligible.

When the wave function is of the monochromatic type (9.4), the time-dependent factors cancel each other in (9.13), to yield

$$\operatorname{div} \mathbf{J}_\psi = 0 \,, \qquad \mathbf{J}_\psi = \frac{\mathrm{i}\,\hbar}{2\,m} \, (w \, \operatorname{grad} w^* - w^* \, \operatorname{grad} w) \,. \tag{9.17}$$

If w is real, then $\mathbf{J}_\psi = 0$.

[1] Remembering that $[|\psi|^2] = \mathrm{m}^{-3}$, one finds $[\mathbf{J}_\psi] = \mathrm{m}^{-2}\,\mathrm{s}^{-1}$.

9.5 Hamiltonian Operator of a Charged Particle

The Hamiltonian function of a particle of mass m and charge e, subjected to an electromagnetic field, is given by (1.35), namely,

$$H = \sum_{i=1}^{3} \frac{1}{2m} (p_i - e A_i)^2 + e\varphi \,, \tag{9.18}$$

where the scalar potential φ and the components of the vector potential A_i may depend on the spatial coordinates and time. To find the Hamiltonian operator corresponding to H one could apply the same procedure as in Sect. 8.48, that consists in replacing p_i with $\hat{p}_i = -\mathrm{i}\hbar\,\partial/\partial x_i$. In this case, however, a difficulty arises if A_i depends on the coordinates; in fact, the two equivalent expansions of $(p_i - e A_i)^2$, namely, $p_i^2 + e^2 A_i^2 - 2 p_i e A_i$ and $p_i^2 + e^2 A_i^2 - 2 e A_i p_i$ yield two different operators: the first of them contains the summand $\partial(A_i \psi)/\partial x_i$, the other one contains $A_i\, \partial\psi/\partial x_i$, and neither one is Hermitean. If, instead, one keeps the order of the factors in the expansion, namely, $(p_i - e A_i)^2 = p_i^2 + e^2 A_i^2 - e A_i p_i - p_i e A_i$, the resulting Hamiltonian operator reads $\mathscr{H} = \mathscr{H}_R + \mathrm{i}\,\mathscr{H}_I$, with

$$\mathscr{H}_R = -\frac{\hbar^2}{2m}\nabla^2 + e\varphi + \frac{e^2}{2m}\mathbf{A}\cdot\mathbf{A}\,, \qquad \mathscr{H}_I = \frac{\hbar e}{2m}\sum_{i=1}^{3}\left(A_i\frac{\partial}{\partial x_i} + \frac{\partial}{\partial x_i}A_i\right), \tag{9.19}$$

and is Hermitean (compare with Sect. 10.2). The particle dynamics is determined by the time-dependent Schrödinger equation $\mathrm{i}\hbar\,\partial\psi/\partial t = \mathscr{H}\psi$. The continuity equation fulfilled by ψ is found following the same reasoning as in Sect. 9.4, starting from

$$\frac{\partial|\psi|^2}{\partial t} = \psi^* \frac{(\mathscr{H}_R + \mathrm{i}\,\mathscr{H}_I)\psi}{\mathrm{i}\hbar} - \psi\,\frac{(\mathscr{H}_R - \mathrm{i}\,\mathscr{H}_I)\psi^*}{\mathrm{i}\hbar}\,. \tag{9.20}$$

The terms related to $\mathscr{H}_R$ yield $-\operatorname{div}\Re(\psi^*\,\hat{\mathbf{p}}\,\psi)/m$ as in Sect. 9.4. Those related to $\mathscr{H}_I$ yield $\operatorname{div}(e\,\mathbf{A}\,|\psi|^2)/m$. In conclusion, the continuity equation for the wave function of a charged particle reads

$$\frac{\partial|\psi|^2}{\partial t} + \operatorname{div}\mathbf{J}_\psi = 0\,, \qquad \mathbf{J}_\psi = \frac{1}{m}\,\Re\left[\psi^*\,(\hat{\mathbf{p}} - e\,\mathbf{A})\,\psi\right]. \tag{9.21}$$

It is worth noting that the transformation from the Hamiltonian function (9.18) to the Hamiltonian operator (9.19) produced by replacing p_i with $\hat{p}_i$ is limited to the dynamics of the particle; the electromagnetic field, instead, is still treated through the scalar and vector potentials, and no transformation similar to that used for the particle is carried out. The resulting Hamiltonian operator (9.19) must then be considered as approximate; the term *semiclassical approximation* is in fact used to

indicate the approach based on (9.19), where the electromagnetic field is treated classically whereas Quantum Mechanics is used in the description of the particle's dynamics. The procedure by which the quantum concepts are extended to the electromagnetic field is described in Sect. 12.3.

The semiclassical approximation is useful in several instances, among which there is the calculation of the stationary states of atoms. As shown in Sect. 9.7.2, it explains why the radiative decay predicted in Sect. 5.11.3 using the classical (planetary) model does not actually occur.

9.6 Approximate Form of the Wave Packet for a Free Particle

The energy spectrum of a free particle is continuous, and the wave packet is given by (9.5), with

$$w_{\mathbf{k}}(\mathbf{r}) = \frac{1}{(2\,\pi)^{3/2}}\,\exp(\mathrm{i}\,\mathbf{k}\cdot\mathbf{r})\,, \qquad E_{\mathbf{k}} = \frac{\hbar^2}{2\,m}\,\left(k_1^2 + k_2^2 + k_3^2\right) = \hbar\,\omega_{\mathbf{k}}\,. \tag{9.22}$$

As $-\infty < k_i < +\infty$, here the order of degeneracy of $E_{\mathbf{k}}$ is infinite. Now, remembering that the wave packet is normalized to unity, it follows that $|\psi(\mathbf{r},t)|^2\,\mathrm{d}^3r$ is the infinitesimal probability that at time t the particle is localized within d^3r ; also, from the analysis carried out in Sect. 9.2, the product $|c_{\mathbf{k}}|^2\,\mathrm{d}^3k$ is the infinitesimal probability that the outcome of an energy measurement belongs to the range of $E(\mathbf{k})$ values whose domain is d^3k. Note that $c_{\mathbf{k}}$ does not depend on time.

Considering the example of a Gaussian wave packet given at the end of Sect. 9.2, one can assume that $|\psi(\mathbf{r},t)|^2$ is localized in the $\mathbf{r}$ space and $|c_{\mathbf{k}}|^2$ is localized in the $\mathbf{k}$ space. This means that $|\psi(\mathbf{r},t)|^2$ and $|c_{\mathbf{k}}|^2$ become vanishingly small when $\mathbf{r}$ departs from its average value $\mathbf{r}_0(t)$ and respectively, $\mathbf{k}$ departs from its average value $\mathbf{k}_0$. Such average values are given by[2]

$$\mathbf{r}_0(t) = \iiint_{-\infty}^{+\infty} \mathbf{r}\,|\psi(\mathbf{r},t)|^2\,\mathrm{d}^3r\,, \qquad \mathbf{k}_0 = \iiint_{-\infty}^{+\infty} \mathbf{k}\,|c_{\mathbf{k}}|^2\,\mathrm{d}^3k\,. \tag{9.23}$$

An approximate expression of the wave packet is obtained by observing that, due to the normalization condition (9.7), the main contribution to the second integral in (9.7) is given by the values of $\mathbf{k}$ that are in the vicinity of $\mathbf{k}_0$. From the identity $k_i^2 = (k_{0i} - k_{0i} + k_i)^2 = k_{0i}^2 + 2\,k_{0i}\,(k_i - k_{0i}) + (k_i - k_{0i})^2$ it then follows

$$\omega_{\mathbf{k}} = \frac{\hbar}{2\,m}\,k_0^2 + \frac{\hbar}{m}\,\mathbf{k}_0\cdot(\mathbf{k}-\mathbf{k}_0) + \frac{\hbar}{2\,m}|\mathbf{k}-\mathbf{k}_0|^2 \tag{9.24}$$

[2]Definitions (9.23) provide the correct weighed average of $\mathbf{r}$ and $\mathbf{k}$ thanks to the normalization condition (9.7). A more exhaustive treatment is carried out in Sect. 10.5.

where, for $\mathbf{k}$ close to $\mathbf{k}_0$, one neglects the quadratic term to find

$$\omega_{\mathbf{k}} \simeq \omega_0 + \mathbf{u}\cdot(\mathbf{k}-\mathbf{k}_0)\,, \quad \omega_0 = \frac{\hbar}{2\,m}\,k_0^2\,, \quad \mathbf{u} = \frac{\hbar}{m}\,\mathbf{k}_0 = (\mathrm{grad}_{\mathbf{k}}\omega_{\mathbf{k}})_{\mathbf{k}_0}\,, \tag{9.25}$$

with $\mathbf{u}$ the *group velocity*. The neglect of terms of order higher than the first used to simplify (9.24) could not be applied to $c_{\mathbf{k}}$; in fact, $c_{\mathbf{k}}$ has typically a peak for $\mathbf{k}=\mathbf{k}_0$, so its first derivatives would vanish there. Using (9.25) and letting $\Phi_0 = \mathbf{k}_0\cdot\mathbf{r} - \omega_0\,t$ transform (9.5) into $\psi(\mathbf{r},t) \simeq \exp(\mathrm{i}\,\Phi_0)\,A\,(\mathbf{r}-\mathbf{u}\,t\,;\mathbf{k}_0)$, where the *envelope function* A is defined as

$$A\,(\mathbf{r}-\mathbf{u}\,t\,;\mathbf{k}_0) = \iiint_{-\infty}^{+\infty} \frac{c_{\mathbf{k}}}{(2\,\pi)^{3/2}}\,\exp\left[\mathrm{i}\;(\mathbf{r}-\mathbf{u}\,t)\cdot(\mathbf{k}-\mathbf{k}_0)\right]\mathrm{d}^3k\,. \tag{9.26}$$

Within the limit of validity of (9.25), the envelope function contains the whole information about the particle's localization: $|\psi|^2 = |A|^2$. Also, the dependence of A on $\mathbf{r}$ and t is such that, for any two pairs $(\mathbf{r}_1,t_1)$, $(\mathbf{r}_2,t_2)$ fulfilling $\mathbf{r}_2-\mathbf{u}\,t_2 = \mathbf{r}_1-\mathbf{u}\,t_1$ the form of $A(\mathbf{r}_2,t_2)$ is the same as that of $A(\mathbf{r}_1,t_1)$. In other terms, A moves without distortion in the direction of $\mathbf{u}$, and its velocity is $(\mathbf{r}_2-\mathbf{r}_1)/(t_2-t_1) = \mathbf{u}$. As time evolves, the approximation leading to (9.25) becomes less and less accurate; taking by way of example $t_1 = 0$ as the initial time, and observing that the summands in (9.25) belong to the phase $-\mathrm{i}\,\omega_{\mathbf{k}}\,t$, the approximation holds as long as $\hbar\,|\mathbf{k}-\mathbf{k}_0|^2\,t/(2\,m) \ll 2\pi$.

9.7 Complements

9.7.1 About the Units of the Wave Function

Consider the wave function ψ associated with a single particle. When the wave function is square integrable and normalized to unity, its units are $[\psi] = \mathrm{m}^{-3/2}$ due to $\int |\psi|^2\,\mathrm{d}^3r = 1$. Then, if the eigenvalues of the Hamiltonian operator are discrete, the second equality in (9.3) shows that $|c_n|^2$, c_n are dimensionless and, finally, (9.2) shows that w_n has the same units as ψ. If the eigenvalues are continuous, the second equality in (9.7) shows that $|c_{\mathbf{k}}|^2$ has the dimensions of a volume of the real space, so that $[c_{\mathbf{k}}] = \mathrm{m}^{3/2}$ and, finally, (9.6) shows that $w_{\mathbf{k}}$ has the same units as ψ.

There are situations, however, where units different from those illustrated above are ascribed to the eigenfunctions. One example is that of the eigenfunctions of the form $w_{\mathbf{k}} = \exp(\mathrm{i}\,\mathbf{k}\cdot\mathbf{r})/(2\,\pi)^{3/2}$, worked out in Sect. 8.2.1, which are dimensionless; another example is that of eigenfunctions of the form (10.7), whose units are $[\delta(\mathbf{r}-\mathbf{r}_0)] = \mathrm{m}^{-3}$. When such eigenfunctions occur, the units of the expansion's coefficients must be modified in order to keep the correct units of ψ (compare with the calculation shown in Sect. 14.6).

The considerations carried out here apply to the single-particle case. When the wave function describes a system of two or more particles, its units change accordingly (Sect. 15.2).

9.7.2 An Application of the Semiclassical Approximation

As indicated in Sect. 9.5 for the case of a particle subjected to an electromagnetic field, the semiclassical approximation consists in using the Hamiltonian operator (9.19), which is derived from the Hamiltonian function (9.18) by replacing p_i with $\hat{p}_i$. The electromagnetic field, instead, is still treated through the scalar and vector potentials. Experiments show that the approximation is applicable in several cases of interest. For instance, consider again the problem of the electromagnetic field generated by a single electron, discussed in classical terms in Sect. 5.11.2. If ψ is the wave function (assumed to be square integrable) associated with the electron in the quantum description, the electron's localization is given by $|\psi|^2$. It is then sensible to describe the charge density and the current density produced by the electron as

$$\varrho = -q\,|\psi|^2\,, \qquad \mathbf{J} = -q\,\mathbf{J}_\psi\,, \tag{9.27}$$

where q is the elementary charge and $\mathbf{J}_\psi$ is defined in (9.13). If the electron is in a stationary state, namely, $\psi(\mathbf{r},t) = w(\mathbf{r})\,\exp(-\mathrm{i}\,\omega\,t)$, then ϱ and $\mathbf{J}$ are independent of time. From (4.58, 4.59) it follows that the potentials φ, $\mathbf{A}$ are independent of time as well. As a consequence, the distribution of charge and current density associated with the electron's motion is stationary (compare with (9.17)), which also yields that the acceleration $\dot{u}$ vanishes. From Larmor's formula (5.72) one finally finds that in this situation the electron emits no power; thus, the radiative decay predicted in Sect. 5.11.3 using the classical model does not occur.

9.7.3 Polar Form of the Schrödinger Equation

The time-dependent Schrödinger equation (9.9) is easily split into two real equations by considering the real and imaginary part of ψ. However, in this section the wave function will rather be written in polar form, $\psi = \alpha(\mathbf{r},t)\,\exp[\mathrm{i}\,\beta(\mathbf{r},t)]$, $\alpha \geq 0$, which reminds one of that used to derive the eikonal equation (5.51) of Geometrical Optics. Despite the fact that the resulting relations are nonlinear, the outcome of this procedure is interesting. Considering a Hamiltonian operator of the type $\mathscr{H} = -\hbar^2\,\nabla^2/(2\,m) + V$, replacing the polar expression of ψ in (9.9), and separating the real and imaginary parts yield two coupled, real equations; the first of them reads

$$\frac{\partial \alpha}{\partial t} = -\frac{\hbar}{2\,m}\left(\alpha\,\nabla^2\beta + 2\,\mathrm{grad}\,\alpha\cdot\mathrm{grad}\,\beta\right)\,, \tag{9.28}$$

where the units of α^2 are those of the inverse of a volume. As for the second equation one finds

$$\frac{\partial(\hbar\,\beta)}{\partial t} + \frac{1}{2\,m}|\,\mathrm{grad}(\hbar\beta)|^2 + V + Q = 0\,, \qquad Q = -\frac{\hbar^2}{2\,m}\,\frac{\nabla^2\alpha}{\alpha}\,, \tag{9.29}$$

where the units of $\mathrm{grad}(\hbar\,\beta)$ are those of a momentum. Using the short-hand notation

$$P = \alpha^2 = |\psi|^2\,, \quad S = \hbar\,\beta\,, \quad \mathbf{v}_e = \frac{\mathrm{grad}\,S}{m}\,, \quad H_Q = \frac{1}{2}\,m v_e^2 + Q + V\,, \tag{9.30}$$

and multiplying (9.28) by 2α, transforms (9.28, 9.29) into

$$\frac{\partial P}{\partial t} + \mathrm{div}(P\,\mathbf{v}_e) = 0\,, \qquad \frac{\partial S}{\partial t} + H_Q = 0\,. \tag{9.31}$$

The wave function is assumed to be square integrable, so that $\int_\Omega P\,\mathrm{d}^3r = 1$. It is easily found that the first of (9.31) is the continuity equation (9.13): from the expression (9.16) of the current density one finds in fact

$$\mathbf{J}_\psi = \frac{1}{m}\,\Re\left(\psi^*\,\hat{\mathbf{p}}\psi\right) = \frac{\hbar}{m}\,\Re\left(\alpha^2\;\mathrm{grad}\,\beta - \mathrm{i}\,\alpha\;\mathrm{grad}\,\alpha\right) = P\,\mathbf{v}_e\,. \tag{9.32}$$

The two differential equations (9.31), whose unknowns are α, β, are coupled with each other. The second of them is similar to the Hamilton-Jacobi equation (1.51) of Classical Mechanics, and becomes equal to it in the limit $\hbar \to 0$, that makes Q to vanish and S to become the Hamilton principal function (Sect. 1.7). Note that the limit $Q \to 0$ decouples (9.31) from each other. In the time-independent case (9.31) reduce to $\mathrm{div}(P\,\mathbf{v}_e) = 0$ and $m\,v_e^2/2 + Q + V = E$, coherently with the fact that in this case Hamilton's principal function becomes $S = W - E\,t$, with W the (time-independent) Hamilton characteristic function. Although $\mathbf{v}_e$ plays the role of an average velocity in the continuity equation of (9.31), determining the expectation value needs a further averaging: in fact, taking definition (10.13) of the expectation value and observing that the normalization makes the integral of $\mathrm{grad}\,\alpha^2$ to vanish yield

$$m\,\langle\mathbf{v}_e\rangle = \int_\Omega \psi^*\,\hat{\mathbf{p}}\psi\;\mathrm{d}^3r = \int_\Omega \alpha^2\,\nabla(\hbar\,\beta)\,\mathrm{d}^3r = m\int_\Omega \alpha^2\,\mathbf{v}_e\,\mathrm{d}^3r\,. \tag{9.33}$$

The last relation in (9.30) seems to suggest that Q is a sort of potential energy to be added to V. In fact this is not true, as demonstrated by the calculation of the expectation value of the kinetic energy T,

$$\langle T \rangle = -\frac{\hbar^2}{2m} \int_\Omega \psi^* \nabla^2 \psi \, \mathrm{d}^3 r = \int_\Omega \psi^* \left(\frac{1}{2} m v_e^2 + Q \right) \psi \, \mathrm{d}^3 r \,, \tag{9.34}$$

showing that Q enters the expectation value of T, not V. To better investigate the meaning of Q it is useful to consider alternative expressions of $\langle Q \rangle$, like

$$\langle Q \rangle = \frac{\hbar^2}{2m} \int_\Omega |\nabla \alpha|^2 \, \mathrm{d}^3 r = \frac{1}{2\,m} \left(\langle \hat{\mathbf{p}} \cdot \hat{\mathbf{p}} \rangle - \langle p_e^2 \rangle \right) , \tag{9.35}$$

where $\hat{\mathbf{p}} = -\mathrm{i}\,\hbar$ grad and $\mathbf{p}_e = m\,\mathbf{v}_e$. The derivation of (9.34, 9.35) follows the same pattern as that of (9.28, 9.29). The first form of (9.35) shows that $\langle Q \rangle$ is positive definite irrespective of the shape of α. The second one is the analogue of the definition of dispersion around the average: the analogy with the treatment used in statistical mechanics (compare with (19.79)) suggests that $p_e^2/(2m)$ provides the analogue of the convective part of the kinetic energy, while Q provides the analogue of the thermal part of it [109].

It is interesting to note that the analogy between the Schrödinger equation and a set of a continuity and a Hamilton-Jacobi-like equations had been observed by de Broglie, who introduced the concept of *pilot wave* in [34]. This cost him severe criticism by Pauli at the Fifth Solvay Conference in 1927. He resumed the idea more than 20 years later, stimulated by the papers by Bohm introducing the concept of *quantum potential*, see, e.g., [14]. The most recent paper by de Broglie on the subject is [35], published when the author was 79 years old.

9.7.4 *Effect of a Gauge Transformation on the Wave Function*

The Hamiltonian function (1.35) of a particle of mass m and charge e, subjected to an electromagnetic field, has been derived in Sect. 1.5 and reads $H = \sum_{i=1}^{3} (p_i - e\,A_i)^2/(2\,m) + e\,\varphi$, with p_i the ith component of momentum in Cartesian coordinates, φ the electric potential, and A_i the ith component of the magnetic potential. If a gauge transformation (Sect. 4.5) is carried out, leading to the new potentials

$$\varphi \leftarrow \varphi' = \varphi - \frac{\partial \vartheta}{\partial t} \,, \qquad \mathbf{A} \leftarrow \mathbf{A}' = \mathbf{A} + \mathrm{grad}\, \vartheta \,, \tag{9.36}$$

the resulting Hamiltonian function H' differs from the original one. In other terms, the Hamiltonian function is not gauge invariant. However, the Lorentz force $e\,(\mathbf{E} + \dot{\mathbf{r}} \wedge \mathbf{B})$ is invariant, whence the dynamics of the particle is not affected by a gauge transformation.

Also in the quantum case it turns out that the Hamiltonian operator is not gauge invariant, $\mathscr{H}' \neq \mathscr{H}$. As consequence, the solution of the Schrödinger equation is not gauge invariant either: $\psi' \neq \psi$. However, the particle's dynamics cannot be affected by a gauge transformation because the Lorentz force is invariant. It follows

that if the initial condition is the same, $|\psi'|^2_{t=0} = |\psi|^2_{t=0}$, then it is $|\psi'|^2 = |\psi|^2$ at all times; for this to happen, there must exist a real function σ such that

$$\psi' = \psi \, \exp(\mathrm{i}\,\sigma)\,, \qquad \sigma = \sigma(\mathbf{r}, t)\,. \tag{9.37}$$

From the gauge invariance of $|\psi|^2$ at all times it follows $\partial|\psi'|^2/\partial t = \partial|\psi|^2/\partial t$ whence, from the continuity equation (9.13), one obtains $\mathrm{div}\,\mathbf{J}'_\psi = \mathrm{div}\,\mathbf{J}_\psi$ with $\mathbf{J}_\psi = \Re[\psi^*(\hat{\mathbf{p}} - e\,\mathbf{A})\psi]/m$. Gauge transforming the quantities in brackets yields $(\psi')^*\,\hat{\mathbf{p}}\,\psi' = \psi^*\,\hat{\mathbf{p}}\,\psi + |\psi|^2\,\hbar\,\mathrm{grad}\,\sigma$ and $(\psi')^*\,e\,\mathbf{A}'\,\psi' = |\psi|^2\,e\,\mathbf{A} + |\psi|^2\,e\,\mathrm{grad}\,\vartheta$, whose difference provides

$$\mathbf{J}'_\psi - \mathbf{J}_\psi = \frac{1}{m}\,|\psi|^2\,\mathrm{grad}\,(\hbar\,\sigma - e\,\vartheta)\,. \tag{9.38}$$

In (9.38) it is $|\psi|^2 \neq 0$ and $\mathrm{grad}\,|\psi|^2 \neq 0$; also, $\hbar\,\sigma - e\,\vartheta$ is independent of ψ. It follows that, for $\mathrm{div}(\mathbf{J}'_\psi - \mathbf{J}_\psi) = 0$ to hold, from (A.17) it must be $\mathrm{grad}\,(\hbar\sigma - e\vartheta) = 0$, namely, $\hbar\sigma - e\vartheta$ is an arbitrary function of time only. Setting the latter to zero finally yields the expression for the exponent in (9.37), that reads[3]

$$\sigma = \frac{e}{\hbar}\,\vartheta\,. \tag{9.39}$$

Problems

9.1 Using the one-dimensional form of (9.26) determine the envelope function $A(x - u\,t)$ corresponding to $c_k = \sqrt{\sigma/\sqrt{\pi}}\,\exp(-\sigma^2\,k^2/2)$, with σ a length. Noting that $\int_{-\infty}^{+\infty} |c_k|^2\,\mathrm{d}k = 1$, show that A is normalized to 1 as well.

9.2 Using the envelope function $A(x - u\,t)$ obtained from Prob. 9.1 and the one-dimensional form of definition (9.23), show that $x_0(t) = u\,t$.

[3]The units of ϑ are $[\vartheta] = V\,s$.

Chapter 10
General Methods of Quantum Mechanics

10.1 Introduction

The preceding chapters have provided the introductory information about Quantum Mechanics. Here the general principles of the theory are illustrated, and the methods worked out for the Hamiltonian operator are extended to the operators associated with dynamic variables different from energy. The important concept of separable operator is introduced, and the property of some operators to commute with each other is related to the mutual compatibility of measurements of the corresponding dynamic variables. Then, the concepts of expectation value and uncertainty are introduced, and the Heisenberg uncertainty principle is worked out. This leads in turn to the analysis of the time derivative of the expectation values, showing that the latter fulfill relations identical to those of Classical Mechanics. The form of the minimum-uncertainty wave packet is worked out in the complements.

10.2 General Methods

The discussion carried out in Sect. 9.2 has led to a number of conclusions regarding the eigenvalues of the time-independent Schrödinger equation (7.44). They are:

- The energy of a particle subjected to a conservative force is one of the eigenvalues of the time-independent equation $\mathscr{H} w = E w$, where $\mathscr{H}$ is derived from the corresponding Hamiltonian function by replacing p_i with $-\mathrm{i}\,\hbar\,\partial/\partial x_i$. Any other energy different from an eigenvalue is forbidden.
- The wave function of a conservative case (taking by way of example the discrete-eigenvalue case) is $\psi = \sum_n c_n w_n \exp(-\mathrm{i}\,E_n\,t/\hbar)$. The particle's localization is given by $|\psi|^2$, where it is assumed that ψ is normalized to unity. The probability

M. Rudan, *Physics of Semiconductor Devices*, DOI 10.1007/978-3-319-63154-7_10

that a measurement of energy finds the eigenvalue E_m is $|c_m|^2$; an energy measurement that finds the eigenvalues E_m forces c_n to become δ_{nm}.
- The time evolution of ψ is found by solving the time-dependent Schrödinger equation $\mathrm{i}\hbar\,\partial\psi/\partial t = \mathscr{H}\psi$. The latter holds also in the nonconservative situations; although in such cases the wave function ψ is not expressible as a superposition of monochromatic waves, it can still be expanded using an orthonormal set like in (9.10) or (9.11).

An important issue is now extending the conclusions listed above to the dynamic quantities different from energy (e.g., momentum, angular momentum, and so on). The extension is achieved by analogy, namely, it is assumed that for any dynamic variable one can construct an eigenvalue equation whose solution provides the possible values of the variable itself. This line of reasoning yields the procedures listed below, that are called *general methods of Quantum Mechanics*:

1. Given a dynamic variable A, an operator $\mathscr{A}$ is associated with it. It is found, first, by expressing A in terms of canonical coordinates q_i, p_i (Sect. 1.6), then, by replacing the momentum's components p_i with $\hat{p}_i = -\mathrm{i}\hbar\,\partial/\partial q_i$ in such a way that $\mathscr{A}$ is Hermitean.
2. It is checked whether the eigenvalue equation $\mathscr{A}v = Av$ possesses a complete, orthonormal set of eigenfunction. If the check fails, the operator is not considered; otherwise it is accepted and is called *observable* [96, Chap. V.9]. The eigenvalue equation is subjected to the same boundary or asymptotic conditions as $\mathscr{H}w = Ew$.
3. Let A_n or A_β be the eigenvalues of $\mathscr{A}v = Av$, with n (β) a set of discrete (continuous) indices. Such eigenvalues are the only values that a measure of the dynamic variable A can find.
4. Thanks to completeness, the wave function ψ describing the particle's localization can be written, respectively for discrete or continuous spectra,

$$\psi = \sum_n a_n(t)\, v_n(\mathbf{r})\,, \qquad \psi = \int_\beta a_\beta(t)\, v_\beta(\mathbf{r})\,\mathrm{d}\beta\,, \tag{10.1}$$

   with $a_n = \langle v_n|\psi\rangle$, $a_\beta = \langle v_\beta|\psi\rangle$.
5. If the wave function in (10.1) is normalizable, then $\sum_n |a_n|^2 = 1$, $\int_\beta |a_\beta|^2\,\mathrm{d}\beta = 1$ at all times. For a discrete spectrum, $P_n = |a_n(t_A)|^2$ is the probability that a measurement of A finds the eigenvalue A_n at $t = t_A$. For a continuous spectrum, the infinitesimal probability that at $t = t_A$ the domain of A_β is found in the interval $\mathrm{d}\beta$ around β is $\mathrm{d}P = |a_\beta(t_A)|^2\,\mathrm{d}\beta$.
6. When the measurement is carried out at $t = t_A$ and an eigenvalue, say, A_m, is found, the coefficients of the first expansion in (10.1) are forced by the measurement to become $|a_n(t_A^+)|^2 = \delta_{mn}$, and the wave function at that instant[1]

[1] Measurements are not instantaneous (refer to the discussion in Sect. 9.2). Here it is assumed that the duration of a measurement is much shorter than the time scale of the whole experiment.

becomes $\psi(\mathbf{r}, t_A^+) = v_m(\mathbf{r})$. The time evolution of ψ starting from t_A^+ is prescribed by the time-dependent Schrödinger equation $\mathrm{i}\hbar\,\partial\psi/\partial t = \mathscr{H}\psi$, with $\psi(\mathbf{r}, t_A^+)$ as the initial condition. In this respect there is a neat parallel with Classical Mechanics, where the time evolution of the canonical variables starting from the initial conditions is prescribed by the Hamilton equations (1.42).

According to the general methods listed above, the eigenvalues of $\mathscr{A}$ are the only possible outcome of a measurement of the dynamic variable A. As the eigenvalues represent a physical quantity, they must be real; this makes the requirement that $\mathscr{A}$ must be Hermitean easily understood: if an operator is Hermitean, then its eigenvalues are real (Sect. 8.4.1). The inverse is also true: if the eigenfunctions of $\mathscr{A}$ form a complete set and its eigenvalues are real, then $\mathscr{A}$ is Hermitean. In fact, for any pair of functions f, g, considering the discrete spectrum by way of example, one has

$$\langle g|\mathscr{A}f\rangle - \langle \mathscr{A}g|f\rangle = \sum_n \sum_m g_n^* f_m \left[\langle v_n|\mathscr{A}v_m\rangle - \langle \mathscr{A}v_n|v_m\rangle\right] =$$

$$= \sum_n \sum_m g_n^* f_m \langle v_n|v_m\rangle \left(A_m - A_n^*\right) = \sum_n g_n^* f_n \left(A_n - A_n^*\right) = 0\,, \tag{10.2}$$

which holds for all f, g because the eigenfunctions v_n are mutually orthogonal and the eigenvalues A_n are real.

As indicated at 1 point above, the dynamic variable A is transformed into the operator $\mathscr{A}$ by replacing p_i with $\hat{p}_i$. The operator obtained from such a replacement is not necessarily Hermitean: its hermiticity must be checked on a case-by-case basis. For instance, the dynamic variable $A = x\,p_x$ can be written in equivalent ways as $x\,p_x$, $p_x\,x$, and $(x\,p_x + p_x\,x)/2$. However, their quantum counterparts

$$-\mathrm{i}\,\hbar x\,\frac{\partial}{\partial x}\,, \qquad -\mathrm{i}\,\hbar\,\frac{\partial}{\partial x}\,x\,, \qquad -\mathrm{i}\,\frac{\hbar}{2}\left(x\,\frac{\partial}{\partial x} + \frac{\partial}{\partial x}\,x\right) \tag{10.3}$$

are different from each other, and only the third one is Hermitean (compare with Sect. 9.5).

10.3 Separable Operators

Let $\mathscr{A}$ be an operator acting only on the x coordinate. Similarly, let $\mathscr{B}$ and $\mathscr{C}$ the two operators acting only on y and z, respectively. The eigenvalue equations for the discrete-spectrum case read

$$\mathscr{A}u_k = A_k\,u_k\,, \qquad \mathscr{B}v_m = B_m\,v_m\,, \qquad \mathscr{C}w_n = C_n\,w_n\,, \tag{10.4}$$

where $u_k(x)$, $v_m(y)$, and $w_n(z)$ are three complete and orthonormal sets of eigenfunctions. Given a square-integrable function $f(x, y, z)$, thanks to the completeness of the three sets the following expansion holds:

$$f(x,y,z) = \sum_n a_n(x,y)\, w_n = \sum_n \left[\sum_m b_{mn}(x)\, v_m \right] w_n =$$

$$= \sum_n \left\{ \sum_m \left[\sum_k c_{kmn}\, u_k \right] v_m \right\} w_n = \sum_{kmn} c_{kmn}\, u_k\, v_m\, w_n\,, \tag{10.5}$$

showing that the set made of the products $u_k\, v_m\, w_n$ is complete. Also, for any linear combination of the above operators, with $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ constant vectors, it is

$$(\mathbf{a}\,\mathscr{A} + \mathbf{b}\,\mathscr{B} + \mathbf{c}\,\mathscr{C})\; u_k\, v_m\, w_n = (\mathbf{a}\,A_k + \mathbf{b}\,B_m + \mathbf{c}\,C_n)\; u_k\, v_m\, w_n\,, \tag{10.6}$$

that is, $u_k\, v_m\, w_n$ is an eigenfunction corresponding to eigenvalue $\mathbf{a}\,A_k + \mathbf{b}\,B_m + \mathbf{c}\,C_n$. It is important to add that in (10.4) it is implied that the boundary conditions of $\mathscr{A} u_k = A_k\, u_k$ depend on x alone, those of $\mathscr{B} v_m = B_m\, v_m$ on y alone, and the like for the third equation. In other terms, separability means that at least one set of coordinates exists, such that both the equation and boundary conditions are separable.

As a first example of application of (10.6), consider the classical position of a particle, $\mathbf{r} = x_1\,\mathbf{i}_1 + x_2\,\mathbf{i}_2 + x_3\,\mathbf{i}_3$. Such a dynamic variable does not contain the components of momentum; as a consequence, the operator associated with it is $\mathbf{r}$ itself, and generates the eigenvalue equation $\mathbf{r}\, g(\mathbf{r}) = \mathbf{r}_0\, g(\mathbf{r})$. Separating the latter and considering the eigenvalue equation for x_i, one finds $x_i\, v_i(x_i) = x_{i0}\, v_i(x_i)$, namely, $(x_i - x_{i0})\, v_i(x_i) = 0$ for all $x_i \neq x_{i0}$. It follows $v_i = \delta(x_i - x_{i0})$, whence $\mathbf{r}_0 = x_{10}\,\mathbf{i}_1 + x_{20}\,\mathbf{i}_2 + x_{30}\,\mathbf{i}_3$, and

$$g_{\mathbf{r}_0}(\mathbf{r}) = \delta(x_1 - x_{10})\,\delta(x_2 - x_{20})\,\delta(x_3 - x_{30}) = \delta(\mathbf{r} - \mathbf{r}_0)\,. \tag{10.7}$$

As a second example consider the classical momentum of a particle, $\mathbf{p} = p_1\,\mathbf{i}_1 + p_2\,\mathbf{i}_2 + p_3\,\mathbf{i}_3$. Remembering the discussion of Sect. 8.5 one finds for the operator associated with $\mathbf{p}$,

$$\hat{\mathbf{p}} = -\mathrm{i}\,\hbar \left(\mathbf{i}_1 \frac{\partial}{\partial x_1} + \mathbf{i}_2 \frac{\partial}{\partial x_2} + \mathbf{i}_3 \frac{\partial}{\partial x_3} \right) = -\mathrm{i}\,\hbar\ \mathrm{grad}\,, \tag{10.8}$$

whose eigenvalue equation reads $-\mathrm{i}\,\hbar\ \mathrm{grad} f = \mathbf{p}_0 f$. Separation yields for the ith eigenvalue equation, with $v_i = v_i(x_i)$, the first-order equation $-\mathrm{i}\,\hbar\, \mathrm{d}v_i/\mathrm{d}x_i = p_{i0}\, v_i$ (compare with (8.49)), whence $v_i = (2\,\pi)^{-1/2} \exp(\mathrm{i}\, k_i\, x_i)$, with $k_i = p_{i0}/\hbar$, so that $\mathbf{k} = \mathbf{p}/\hbar = k_1\,\mathbf{i}_1 + k_2\,\mathbf{i}_2 + k_3\,\mathbf{i}_3$, and

$$f_{\mathbf{k}}(\mathbf{r}) = (2\,\pi)^{-3/2} \exp(\mathrm{i}\,\mathbf{k}\cdot\mathbf{r})\,. \tag{10.9}$$

Neither (10.7) nor F (10.9) are square integrable. The indices of the eigenvalues ($\mathbf{r}_0$ in (10.8) and $\mathbf{k}$ in (10.9)) are continuous in both cases. Also, from the results of Sects. C.2 and C.5 one finds that $g_{\mathbf{r}_0}(\mathbf{r}) = g(\mathbf{r}, \mathbf{r}_0)$ is the Fourier transform of $f_{\mathbf{k}}(\mathbf{r}) = f(\mathbf{r}, \mathbf{k})$.

10.4 Eigenfunctions of Commuting Operators

It has been shown in Sect. 10.2 that a measurement of the dynamic variable A at time t_A yields one of the eigenvalues of the equation $\mathscr{A}a = A\,a$. Considering for instance a discrete spectrum, let the eigenvalue be A_m. The initial condition $\psi(\mathbf{r}, t_A^+)$ for the time evolution of the particle's wave function after the measurement is one of the eigenfunctions of $\mathscr{A}$ corresponding to A_m. If a measurement of another dynamic variable B is carried out at a later time t_B, the wave function at $t = t_B$ is forced to become one of the eigenfunctions of $\mathscr{B}b = B\,b$, say, b_k. The latter can in turn be expanded in terms of the complete set derived from $\mathscr{A}$, namely, $b_k = \sum_n \langle a_n | b_k \rangle\, a_n$. As the coefficients of the expansion are in general different from zero, there is a finite probability that a new measurement of A at $t_C > t_B$ finds a value different from A_m. In principle this could be due to the fact that if A is not conserved, its value has evolved, from the outcome A_m of the measurement carried out at $t = t_A$, into something different, as prescribed by the time-dependent Schrödinger equation having $\psi(\mathbf{r}, t_A^+)$ as initial condition.[2] However, the instant t_B of the second measurement can in principle be brought as close to t_A as we please, so that the two measurements can be thought of as simultaneous. As a consequence, the loss of information about the value of A must be ascribed to the second measurement, specifically, to its interference with the wave function, rather than to a natural evolution[3] of the value of A: the gain in information about the eigenvalue of B produces a loss of information about that of A; for this reason, the two measurements are said to be *incompatible*.

From the discussion above one also draws the conclusion that if it were $b_k = a_m$, the two measurements of outcome A_m and B_k would be compatible. This is in itself insufficient for stating that the measurements of A and B are compatible in all cases; for this to happen it is necessary that the whole set of eigenfunctions of A coincides with that of B: in this case, in fact, the condition $b_k = a_m$ is fulfilled no matter what the outcome of the two measurements is.

It would be inconvenient to check the eigenfunctions to ascertain whether two observables A, B are compatible or not. In fact, this is not necessary thanks to the following property: if two operators $\mathscr{A}$ and $\mathscr{B}$ have a common, complete set of

[2] By way of example one may think of A as the position x, that typically evolves in time from the original value $x_A = x(t_A)$ even if the particle is not perturbed.

[3] In any case, the evolution would be predicted exactly by the Schrödinger equation. Besides, the eigenvalue would not change if A were conserved.

eigenfunctions, then they commute, and vice versa (as indicated in Sect. 8.6.2, two operators commute if their commutator (8.71) is the null operator). Still assuming a discrete spectrum, for any eigenfunction v_n it is $\mathscr{A}\mathscr{B}v_n = \mathscr{A}B_n\, v_n = B_n\mathscr{A}\, v_n = B_n A_n\, v_n$. Similarly, $\mathscr{B}\mathscr{A}\, v_n = \mathscr{B}A_n\, v_n = A_n\mathscr{B}v_n = A_n B_n\, v_n$, showing that $\mathscr{A}$ and $\mathscr{B}$ commute for all eigenfunctions. Then, using the completeness of the common set v_n to expand any function f as $f = \sum_n f_n\, v_n$, one finds

$$\mathscr{A}\mathscr{B}f = \sum_n f_n \mathscr{A}\mathscr{B}v_n = \sum_n f_n \mathscr{B}\mathscr{A}\, v_n = \mathscr{B}\mathscr{A}\sum_n f_n\, v_n = \mathscr{B}\mathscr{A}f\,. \tag{10.10}$$

This proves that if two operators have a complete set of eigenfunctions in common, then they commute. Conversely, assume that $\mathscr{A}$ and $\mathscr{B}$ commute and let v_n be an eigenfunction of $\mathscr{A}$; then, $\mathscr{A}\mathscr{B}v_n = \mathscr{B}\mathscr{A}\, v_n$ and $\mathscr{A}\, v_n = A_n\, v_n$. Combining the latter relations yields $\mathscr{A}\mathscr{B}v_n = \mathscr{B}A_n\, v_n$ which, letting $g_n = \mathscr{B}v_n$, is recast as $\mathscr{A}\, g_n = A_n g_n$. In conclusion, both v_n and g_n are eigenfunctions of $\mathscr{A}$ belonging to the same eigenvalue A_n.

If A_n is not degenerate, the eigenfunctions v_n and g_n must be the same function, apart from a multiplicative constant due to the homogeneity of the eigenvalue equation. Let such a constant be B_n; combining $g_n = B_n\, v_n$ with the definition $g_n = \mathscr{B}v_n$ yields $\mathscr{B}v_n = B_n\, v_n$, thus showing that v_n is an eigenfunction of $\mathscr{B}$ as well. The property holds also when A_n is degenerate, although the proof is somewhat more involved [95, Chap. 8-5]. This proves that if two operators commute, then they have a complete set of eigenfunctions in common. Examples are given in Sect. 8.6.3.

10.5 Expectation Value and Uncertainty

The discussion carried out in Sect. 10.2 has led to the conclusion that the wave function ψ describing the particle's localization can be expanded as in (10.1), where v_n or v_β are the eigenfunctions of an Hermitean operator $\mathscr{A}$ that form a complete, orthonormal set. Considering a discrete spectrum first, the coefficients of the expansion are $a_n = \langle v_n|\psi\rangle$; assuming that the wave function is normalizable, it is $\sum_n |a_n|^2 = 1$.

The meaning of the coefficients is that $P_n = |a_n(t)|^2$ is the probability that a measurement of A finds the eigenvalue A_n at time t. From this it follows that the statistical average of the eigenvalues is

$$\langle A\rangle(t) = \sum_n P_n A_n\,. \tag{10.11}$$

The average (10.11) is called *expectation value*.[4] It can be given a different form by observing that $P_n = a_n^*\, a_n = (\sum_m a_m^*\, \delta_{mn})\, a_n$ and that, due to the orthonormality of the eigenfunctions of $\mathscr{A}$, it is $\delta_{mn} = \langle v_m|v_n\rangle$; then,

[4] If the wave function is normalized to a number different from unity, the definition of the expectation value is $\sum_n P_n A_n / \sum_n P_n$, and the other definitions are modified accordingly.

$$\sum_n \left(\sum_m a_m^* \langle v_m | v_n \rangle \right) a_n A_n = \langle \sum_m a_m v_m | \sum_n a_n A_n v_n \rangle = \langle \sum_m a_m v_m | \sum_n a_n \mathscr{A} v_n \rangle \,. \tag{10.12}$$

Combining (10.12) with (10.11) and remembering that $\mathscr{A}$ is Hermitean yields

$$\langle A \rangle = \langle \psi | \mathscr{A} | \psi \rangle \,. \tag{10.13}$$

The same result holds for a continuous spectrum:

$$\langle A \rangle = \int_\alpha P_\alpha A_\alpha \, \mathrm{d}\alpha = \int_\alpha |a_\alpha|^2 A_\alpha \, \mathrm{d}\alpha = \langle \psi | \mathscr{A} | \psi \rangle \,, \tag{10.14}$$

where

$$\int_\alpha |a_\alpha|^2 A_\alpha \, \mathrm{d}\alpha = \int_\alpha \left(\int_\beta a_\beta^* \, \delta(\beta - \alpha) \, \mathrm{d}\beta \right) a_\alpha A_\alpha \, \mathrm{d}\alpha \,, \tag{10.15}$$

and $\int_\beta |a_\beta|^2 \, \mathrm{d}\beta = 1$ at all times. The expectation values of Hermitean operators are real because they are the statistical averages of the eigenvalues, themselves real. Using (8.57) one extends the definition of expectation value to the powers of the eigenvalues; for instance,

$$\langle A^2 \rangle = \langle \psi | \mathscr{A}^2 | \psi \rangle = \int_\Omega \psi^* \, \mathscr{A} \mathscr{A} \psi \, \mathrm{d}\Omega = \langle \mathscr{A} \psi | \mathscr{A} \psi \rangle = ||\mathscr{A} \psi||^2 \geq 0 \,, \tag{10.16}$$

where the hermiticity of $\mathscr{A}$ is exploited. The *variance* of the eigenvalues is given by

$$(\Delta A)^2 = \left\langle (A - \langle A \rangle)^2 \right\rangle = \left\langle A^2 - 2 \langle A \rangle A + \langle A \rangle^2 \right\rangle = \langle A^2 \rangle - \langle A \rangle^2 \,, \tag{10.17}$$

real and nonnegative by construction; as a consequence, $\langle A^2 \rangle \geq \langle A \rangle^2$. The general term used to indicate the positive square root of the variance, $\Delta A = \sqrt{(\Delta A)^2} \geq 0$, is *standard deviation*. When it is used with reference to the statistical properties of the eigenvalues, the standard deviation is called *uncertainty*.

Assume by way of example that the wave function at $t = t_A$ coincides with one of the eigenfunctions of $\mathscr{A}$. With reference to a discrete spectrum (first relation in (10.1)), let $\psi(t_A) = v_m$. From (10.13) and (10.17) it then follows $\langle A \rangle(t_A) = A_m$, $\Delta A(t_A) = 0$. The standard deviation of the eigenvalues is zero in this case, because the measurement of A can only find the eigenvalue A_m. As a second example consider a continuous spectrum in one dimension, and let $\psi(t_A) = \exp(\mathrm{i}\, k\, x)/\sqrt{2\,\pi}$, namely, an eigenfunction of the momentum operator. In this case the wave function is not square integrable, so one must calculate the expectation value as

$$\langle A\rangle(t_A) = \lim_{x_0\to\infty} \frac{\int_{-x_0}^{+x_0} \psi^*(t_A)\, \mathscr{A}\psi(t_A)\, \mathrm{d}x}{\int_{-x_0}^{+x_0} \psi^*(t_A)\, \psi(t_A)\, \mathrm{d}x}\,. \tag{10.18}$$

If one lets $\mathscr{A} = \hat{p} = -\mathrm{i}\,\hbar\,\mathrm{d}/\mathrm{d}x$, the result is $\langle p\rangle(t_A) = \hbar\,k$, $\Delta p(t_A) = 0$. In fact, like in the previous example, the wave function coincides with one of the eigenfunctions of the operator. If, however, one applies another operator to the same wave function, its variance does not necessarily vanish. A remarkable outcome stems from applying $\hat{x} = x$, that is, the operator associated with the dynamic variable canonically conjugate to p: one finds $\langle x\rangle(t_A) = 0$, $\Delta x(t_A) = \infty$.

In conclusion, the examples above show that the term "uncertainty" does not refer to an insufficient precision of the measurements (which in principle can be made as precise as we please), but to the range of eigenvalues that is covered by the form of $\psi(t_A)$. In the last example above all positions x are equally probable because $|\psi(t_A)|^2 = \text{const}$, whence the standard deviation of position diverges.

10.6 Heisenberg Uncertainty Relation

Consider the wave function ψ describing the dynamics of a particle, and let $\mathscr{A}$ and $\mathscr{B}$ be Hermitean operators. A relation exists between the standard deviations of these operators, calculated with the same wave function. Defining the complex functions $f = (\mathscr{A} - \langle A\rangle)\,\psi$ and $g = (\mathscr{B} - \langle B\rangle)\,\psi$ yields

$$||f||^2 = (\Delta A)^2\,, \qquad ||g||^2 = (\Delta B)^2\,, \qquad \langle f|g\rangle - \langle g|f\rangle = \mathrm{i}\,\langle C\rangle\,, \tag{10.19}$$

where the first two relations derive from (10.17) while $\langle C\rangle$ in the third one is the expectation value of the commutator $\mathscr{C} = -\mathrm{i}\,(\mathscr{A}\mathscr{B} - \mathscr{B}\mathscr{A})$. Letting $\mu = \mathrm{i}\,\nu$ in (8.16), with ν real, and using (10.19) provides

$$(\Delta A)^2 + \nu^2\,(\Delta B)^2 - \nu\,\langle C\rangle \geq 0\,, \tag{10.20}$$

namely, a second-degree polynomial in the real parameter ν. In turn, the coefficients of the polynomial are real because they derive from Hermitean operators. For the polynomial to be nonnegative for all ν, the discriminant $\langle C\rangle^2 - 4\,(\Delta A)^2\,(\Delta B)^2$ must be non-positive. The relation between the standard deviations then reads

$$\Delta A\;\Delta B \geq \frac{1}{2}\,|\langle C\rangle|\,. \tag{10.21}$$

The interpretation of this result follows from the discussion carried out at the end of Sect. 10.5. If $\mathscr{A}$ and $\mathscr{B}$ commute, then their commutator is the null operator, whose eigenvalue is zero. As a consequence it is $\Delta A\;\Delta B \geq 0$, namely, the minimum of the product is zero. Remembering the result of Sect. 10.4, when two operators commute

they have a common, complete set of eigenfunctions. If the wave function used for calculating the variance (10.17) is an eigenfunction of $\mathscr{A}$ and $\mathscr{B}$, then both standard deviations ΔA and ΔB vanish and $\Delta A\,\Delta B = 0$, namely, the minimum can in fact be attained. If, instead, $\mathscr{A}$ and $\mathscr{B}$ do not commute, the minimum of the product $\Delta A\,\Delta B$ must be calculated on a case-by-case basis. The most interesting outcome is found when the two operators are associated with conjugate dynamic variables: $\mathscr{A} = q_i$ and $\mathscr{B} = -\mathrm{i}\hbar\,\partial/\partial q_i$. Remembering (8.72) one finds $\mathscr{C} = \hbar\,\mathscr{I}$, $C = \hbar$, $\langle C\rangle = \hbar$, whence

$$\Delta A\,\Delta B \geq \frac{\hbar}{2}\,. \tag{10.22}$$

Inequality (10.22) is also called *Heisenberg principle* or *uncertainty principle*, because it was originally deduced by Heisenberg from heuristic arguments [63].[5] The more formal deduction leading to (10.21) was given shortly after in [77] and [144].

10.7 Time Derivative of the Expectation Value

The expectation value (10.11) of a Hermitean operator is a real function of time. In Classical Mechanics, the generalized coordinates and momenta are also functions of time, whose evolution is given by the Hamilton equations (1.42); the latter express the time derivatives of coordinates and momenta in terms of the Hamiltonian function. Then, for an arbitrary function ϱ of the canonical coordinates, the total derivative with respect of time is expressed through the Poisson bracket as in (1.53). A relation of the same form as (1.53) is found in Quantum Mechanics by calculating the time derivative of the expectation value (10.13). It is assumed that operator $\mathscr{A}$ depends on time, but does not operate on it; as a consequence, the symbol $\partial\mathscr{A}/\partial t$ indicates the operator resulting from differentiating $\mathscr{A}$ with respect to its functional dependence on t. With these premises one finds

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \psi^*\,\mathscr{A}\,\psi\,\mathrm{d}\Omega = \int_\Omega \left(\frac{\partial\psi^*}{\partial t}\,\mathscr{A}\,\psi + \psi^*\,\frac{\partial\mathscr{A}}{\partial t}\,\psi + \psi^*\,\mathscr{A}\,\frac{\partial\psi}{\partial t}\right)\mathrm{d}\Omega\,. \tag{10.23}$$

The time derivative of ψ is obtained from the time-dependent Schrödinger equation (9.9). Considering the case where $\mathscr{H}$ is real yields $\partial\psi/\partial t = -\mathrm{i}\,\mathscr{H}\psi/\hbar$ and $\partial\psi^*/\partial t = \mathrm{i}\,\mathscr{H}\psi^*/\hbar$, whence

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle A\rangle = \int_\Omega \psi^*\,\frac{\partial\mathscr{A}}{\partial t}\,\psi\,\mathrm{d}\Omega + \frac{\mathrm{i}}{\hbar}\int_\Omega \psi^*\,\left(\mathscr{H}\mathscr{A} - \mathscr{A}\mathscr{H}\right)\psi\,\mathrm{d}\Omega\,, \tag{10.24}$$

[5] Namely, (10.22) is a theorem rather than a principle. A similar comment applies to the Pauli principle (Sect. 15.6). The English translation of [63] is in [145].

which has the same structure as (1.53). Other relations similar to those of Sect. 1.8 are also deduced from (10.24). For instance, letting $\mathscr{A} = \mathscr{H}$ yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle H\rangle = \left\langle \frac{\partial \mathscr{H}}{\partial t} \right\rangle , \tag{10.25}$$

similar to (1.44). If $\langle A\rangle$ is a constant of motion, then

$$\int_\Omega \psi^* \frac{\partial \mathscr{A}}{\partial t}\psi \,\mathrm{d}\Omega + \frac{\mathrm{i}}{\hbar}\int_\Omega \psi^* \,(\mathscr{H}\mathscr{A} - \mathscr{A}\mathscr{H})\,\psi \,\mathrm{d}\Omega = 0\,, \tag{10.26}$$

similar to (1.54) while, if $\mathscr{A}$ does not depend on time, (10.24) yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle A\rangle = \frac{\mathrm{i}}{\hbar}\int_\Omega \psi^* \,(\mathscr{H}\mathscr{A} - \mathscr{A}\mathscr{H})\,\psi \,\mathrm{d}\Omega\,, \tag{10.27}$$

similar to (1.55). Finally, if $\mathscr{A}$ does not depend on time and commutes with $\mathscr{H}$, (10.24) yields $\mathrm{d}\langle A\rangle/\mathrm{d}t = 0$, namely, the expectation value $\langle A\rangle$ is a constant of motion.

10.8 Ehrenfest Theorem

An important application of (10.27) is found by replacing $\mathscr{A}$ with either a position operator or a momentum operator. The calculation is shown here with reference to the Hamiltonian operator $\mathscr{H} = -\hbar^2/(2\,m)\nabla^2 + V$, where the potential energy V is independent of time. Letting first $\mathscr{A} = x$ yields

$$(\mathscr{H}\,x - x\,\mathscr{H})\,\psi = \frac{\hbar^2}{2\,m}\left(x\,\frac{\partial^2\psi}{\partial x^2} - \frac{\partial^2 x\,\psi}{\partial x^2}\right) = \frac{\hbar^2}{2m}\left(-2\frac{\partial\psi}{\partial x}\right) \tag{10.28}$$

whence, using $\hat{p}_x = -\mathrm{i}\,\hbar\,\partial/\partial x$, it follows

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle x\rangle = \frac{\mathrm{i}}{\hbar}\int_\Omega \psi^* \frac{\hbar^2}{2\,m}\left(-2\frac{\partial\psi}{\partial x}\right)\mathrm{d}\Omega = \frac{1}{m}\,\langle\psi|\hat{p}_x|\psi\rangle = \frac{\langle p_x\rangle}{m}\,. \tag{10.29}$$

In conclusion, the relation $\mathrm{d}\langle x\rangle/\mathrm{d}t = \langle p_x\rangle/m$ holds, similar to the one found in a classical case when the Hamiltonian function has the form $H = p^2/(2\,m) + V$ (compare with the second relation in (1.33)). Still with $\mathscr{H} = -\hbar^2/(2\,m)\nabla^2 + V$, consider as a second example $\mathscr{A} = \hat{p}_x = -\mathrm{i}\,\hbar\,\partial/\partial x$, to find

$$(\mathscr{H}\,\hat{p}_x - \hat{p}_x\,\mathscr{H})\,\psi = -\mathrm{i}\,\hbar\left(V\,\frac{\partial\psi}{\partial x} - \frac{\partial(V\,\psi)}{\partial x}\right) = \mathrm{i}\,\hbar\,\psi\,\frac{\partial V}{\partial x}\,. \tag{10.30}$$

From this, letting $F_x = -\partial V/\partial x$ be the component of the force along x, it follows

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle p_x\rangle = \frac{\mathrm{i}}{\hbar}\int_\Omega \psi^*\,\mathrm{i}\hbar\,\psi\,\frac{\partial V}{\partial x}\,\mathrm{d}\Omega = \langle F_x\rangle\,, \tag{10.31}$$

also in this case similar to the classical one. Combining (10.29) and (10.31) shows that the expectation values fulfill a relation similar to Newton's law,

$$m\,\frac{\mathrm{d}^2}{\mathrm{d}t^2}\langle x\rangle = \langle F_x\rangle\,. \tag{10.32}$$

This result is called *Ehrenfest theorem*. If the dependence of F_x on position is weak in the region where ψ is significant, the normalization of the wave function yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle p_x\rangle \simeq F_x\int_\Omega \psi^*\,\psi\,\mathrm{d}\Omega = F_x\,. \tag{10.33}$$

In this case, called *Ehrenfest approximation*, the expectation value of position fulfills Newton's law exactly. If, on the contrary, F_x depends strongly on position in the region where ψ is significant (as happens, e.g., when the potential energy has the form of a step or a barrier), then the outcome of the quantum calculation is expected to be different from the classical one (see, e.g., Sects. 11.2 and 11.3).

10.9 Complements

10.9.1 Minimum-Uncertainty Wave Function

It has been shown in Sect. 10.6 that when the two operators are associated with conjugate dynamic variables, $\mathscr{A} = q$ and $\mathscr{B} = -\mathrm{i}\hbar\,\mathrm{d}/\mathrm{d}q$, the relation between their standard deviations is given by (10.22). It is interesting to seek a form of the wave function such that the equality $\Delta A\,\Delta B = \hbar/2$ holds. If it exists, such a form is called *minimum-uncertainty wave function*. To proceed one notes that the equality yields the value $\nu_m = (|\langle C\rangle|/2)/(\Delta B)^2$ corresponding to the minimum of the polynomial (10.20); moreover, such a minimum is zero. On the other hand, imposing the equality in (8.16) after letting $\mu = \mathrm{i}\,\nu_m$ yields the more compact form $||f + \mathrm{i}\,\nu_m\,g||^2 = 0$, equivalent to $f + \mathrm{i}\,\nu_m\,g = 0$. Remembering the definitions given in Sect. 10.6, it is $f = (\mathscr{A} - \langle A\rangle)\,\psi$, $g = (\mathscr{B} - \langle B\rangle)\,\psi$. Now, letting $q_0 = \langle A\rangle$, $p_0 = \hbar\,k_0 = \langle B\rangle$, from the relation $f + \mathrm{i}\,\nu_m\,g = 0$ one obtains the first-order differential equation $\nu_m\,\hbar\,\mathrm{d}\psi/\mathrm{d}q = [\mathrm{i}\,\nu_m\,p_0 - (q - q_0)]\,\psi$, whose solution is

$$\psi(q) = \psi_0\,\exp\left[\mathrm{i}\,k_0\,q - (q - q_0)^2/(2\,\nu_m\,\hbar)\right]\,. \tag{10.34}$$

The normalization condition $\langle\psi|\psi\rangle = 1$ yields $\psi_0 = (\pi\, \nu_m\, \hbar)^{-1/4}$. Using $\langle C\rangle = \hbar$ and combining the expression of ν_m with the equality $\Delta A\, \Delta B = \hbar/2$ provide $\nu_m = 2\,(\Delta A)^2/\hbar$ whence, letting $\Delta q = \Delta A$,

$$\psi(q) = \frac{1}{\sqrt[4]{2\,\pi}\,\sqrt{\Delta q}} \exp\left[\mathrm{i}\, k_0\, q - \frac{(q-q_0)^2}{(2\,\Delta q)^2}\right]. \tag{10.35}$$

The minimum-uncertainty wave function turns out to be proportional to a Gaussian function centered at q_0. The factor $\exp(\mathrm{i}\, k_0\, q)$ disappears from $|\psi|^2$, whose peak value and width are determined by Δq. Note that this calculation leaves the individual values of Δq and $\Delta p = \Delta B$ unspecified.

Problems

10.1 Starting from the wave packet (9.5) describing a free particle, determine the time evolution of its position without resorting to the approximation used in Sect. 9.6.

10.2 Using the results of Prob. 10.1, determine the time evolution of the standard deviation of position.

10.3 Starting from the wave packet (9.5) describing a free particle, determine the time evolution of its momentum without resorting to the approximation used in Sect. 9.6.

10.4 Using the results of Prob. 10.3, determine the time evolution of the standard deviation of momentum.

10.5 Consider a one-dimensional wave function that at some instant of time is given by a minimum-uncertainty packet (10.35) whose polar form is

$$\alpha = \frac{1}{\sqrt[4]{2\,\pi}\,\sqrt{\sigma}} \exp\left[-\frac{(x-x_0)^2}{4\,\sigma^2}\right], \qquad \beta = k_0\, x\,. \tag{10.36}$$

The wave packet is normalized to 1. Using the concepts introduced in Sect. 9.7.3, find the "convective" and "thermal" parts of the expectation value of the kinetic energy.

Part III
Applications of the Schrödinger Equation

Chapter 11
Elementary Cases

11.1 Introduction

The time-independent Schrödinger equation is a linear, second-order equation with a coefficient that depends on position. An analytical solution can be found in a limited number of cases, typically one-dimensional ones. This chapter illustrates some of these cases, starting from the step-like potential energy followed by the potential-energy barrier. In both of them, the coefficient of the Schrödinger equation is approximated with a piecewise-constant function. Despite their simplicity, the step and barrier potential profiles show that the quantum-mechanical treatment may lead to results that differ substantially from the classical ones: a finite probability of transmission may be found where the classical treatment would lead to a reflection only, or vice versa. The transmission and reflection coefficients are defined by considering a plane wave launched towards the step or barrier. It is shown that the definition of the two coefficients can be given also for a barrier of a general form, basing on the formal properties of the second-order linear equations in one dimension. Finally, the case of a finite well is tackled, showing that in the limit of an infinite depth of the well one recovers the results of the particle in a box illustrated in a preceding chapter.

11.2 Step-Like Potential Energy

Consider a one-dimensional, step-like potential energy as shown in Fig. 11.1, with $V = 0$ for $x < 0$ and $V = V_0 > 0$ for $x > 0$. From the general properties of the time-independent Schrödinger equation (Sect. 8.2.3) it follows $E \geq 0$. To proceed, it is convenient to consider the two cases $0 < E < V_0$ and $E > V_0$ separately.

M. Rudan, *Physics of Semiconductor Devices*, DOI 10.1007/978-3-319-63154-7_11

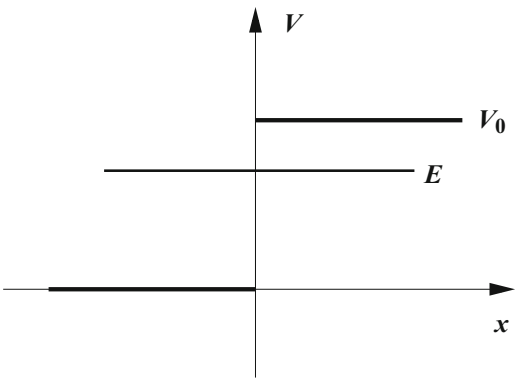


Fig. 11.1 The example of the step-like potential energy analyzed in Sect. 11.2. Only the case $0 \le E \le V_0$ is shown

11.2.1 Case A: $0 < E < V_0$

The Schrödinger equation is split over the two partial domains to yield

$$\begin{cases} x < 0: & -w'' = k^2 w\,, & k = \sqrt{2\,m\,E}/\hbar \\ x > 0: & w'' = \alpha^2 w\,, & \alpha = \sqrt{2\,m\,(V_0 - E)}/\hbar \end{cases} \tag{11.1}$$

where the derivatives are indicated with primes. The solutions on the left and right of the origin are respectively given by

$$w = \begin{cases} w_- = a_1\,\exp(\mathrm{i}\,k\,x) + a_2\,\exp(-\mathrm{i}\,k\,x)\,, & x < 0 \\ w_+ = a_3\,\exp(-\alpha\,x) + a_4\,\exp(\alpha\,x)\,, & x > 0 \end{cases} \tag{11.2}$$

where it must be set $a_4 = 0$ to prevent w_+ from diverging. Using the continuity of w and w' in the origin yields

$$\begin{cases} w_+(0) = w_-(0)\,, & a_1 + a_2 = a_3 \\ w'_+(0) = w'_-(0)\,, & \mathrm{i}\,k\,(a_2 - a_1) = \alpha\,a_3 \end{cases} \tag{11.3}$$

Eliminating a_3 provides the relation $\alpha\,(a_1 + a_2) = \mathrm{i}\,k\,(a_2 - a_1)$ whence

$$\frac{a_2}{a_1} = \frac{\mathrm{i}\,k + \alpha}{\mathrm{i}\,k - \alpha} = \frac{k - \mathrm{i}\,\alpha}{k + \mathrm{i}\,\alpha}\,, \qquad \frac{a_3}{a_1} = 1 + \frac{a_2}{a_1} = \frac{2\,k}{k + \mathrm{i}\,\alpha}\,, \tag{11.4}$$

that determine a_2, a_3 apart from the arbitrary constant a_1. This should be expected as w is not normalizable. From $|a_2/a_1| = |k - \mathrm{i}\,\alpha|/|k + \mathrm{i}\,\alpha| = 1$ one finds $a_2/a_1 = \exp(-\mathrm{i}\,\varphi)$, with $\varphi = 2\,\arctan(\alpha/k)$. The solution of the time-independent Schrödinger equation is then recast as

$$w = \begin{cases} w_- = 2\,a_1\,\exp(-\mathrm{i}\,\varphi/2)\,\cos(k\,x + \varphi/2)\,, & x < 0 \\ w_+ = [2\,k/(k + \mathrm{i}\,\alpha)]\,a_1\,\exp(-\alpha\,x)\,, & x > 0 \end{cases} \tag{11.5}$$

The eigenvalues E are continuous in the range $0 \leq E < V_0$. The monochromatic wave function corresponding to w_k is $\psi_k(x,t) = w(x)\exp(-\mathrm{i}E_k\,t/\hbar)$, with $E_k = \hbar^2\,k^2/(2\,m)$.

The quantity $R = |a_2/a_1|^2$ is called *reflection coefficient* of the monochromatic wave. As shown in Sect. 11.4 from a more general point of view, R is the probability that the wave is reflected by the step. In the case investigated here, where a particle with $0 < E < V_0$ is launched towards a step, it is $a_2/a_1 = \exp(-\mathrm{i}\,\varphi)$ whence $R = 1$.

The solution (11.3) or (11.5) shows that w becomes vanishingly small on the right of the origin as $x \to +\infty$. When a wave packet is built up using a superposition of monochromatic solutions, chosen in such a way that each energy E_k is smaller than V_0, its behavior is similar; as a consequence, the probability of finding the particle on the right of the origin becomes smaller and smaller as the distance from the origin increases. In conclusion, if a particle described by such a packet is launched from $x < 0$ towards the step, the only possible outcome is the same as in a calculation based on Classical Mechanics, namely, a reflection. A difference between the classical and quantum treatment exists though: in the former the reflection occurs at $x = 0$, whereas in the latter the reflection abscissa is not defined. This is better understood by considering a wave packet of standard deviation Δx approaching the origin from $x < 0$. Considering the approximate form of the packet described in Sect. 9.6, the incident envelope has the form $A_i = A(x - u\,t)$, and the localization of the incident particle is described by $|\psi_i|^2 = |A_i|^2$. Due to its finite width, the packet crosses the origin during a time $\Delta t \sim \Delta x/u$ starting, e.g., at $t = 0$. At a later instant $\Delta t + t_0$, where $t_0 \geq 0$ is the time that the wave packet takes to move away from the step to the extent that the interaction with it is practically completed, only the reflected packet exists, described by $|\psi_r|^2 = |A_r|^2$, with $A_r = A[x + u\,(t - t_0)]$. For $0 \leq t \leq \Delta t + t_0$ both incident and reflected packets exist. One could think that the reflection abscissa is given by the product $u\,t_0$; however, t_0 depends on the form of the packet, so the reflection abscissa is not well defined. Before the particle starts interacting with the step, only the incident packet exists and the normalization $\int_{-\infty}^{0} |\psi_i|^2\,\mathrm{d}x = 1$ holds; similarly, after the interaction is completed, only the reflected packet exists and $\int_{-\infty}^{0} |\psi_r|^2\,\mathrm{d}x = 1$. For $0 \leq t \leq \Delta t + t_0$ the normalization is achieved by a superposition of the incident and reflected packets.

11.2.2 Case B: $E > V_0$

Still considering the one-dimensional step of Fig. 11.1, let $E > V_0$. In this case the time-independent Schrödinger equation reads

$$\begin{cases} x < 0: & -w'' = k^2\,w\,, & k = \sqrt{2\,m\,E}/\hbar \\ x > 0: & -w'' = k_1^2\,w\,, & k_1 = \sqrt{2\,m\,(E - V_0)}/\hbar \end{cases} \tag{11.6}$$

whose solution is

$$w = \begin{cases} w_- = a_1 \exp(\mathrm{i}\, k\, x) + a_2 \exp(-\mathrm{i}\, k\, x)\,, & x < 0 \\ w_+ = a_3 \exp(\mathrm{i}\, k_1\, x) + a_4 \exp(-\mathrm{i}\, k_1\, x)\,, & x > 0 \end{cases} \tag{11.7}$$

Remembering the discussion of Sect. 9.6, function w_- in (11.7) describes a superposition of two planar and monochromatic waves, belonging to the $x < 0$ region, that propagate in the forward and backward direction, respectively; a similar meaning holds for w_+ with reference to the $x > 0$ region. Now one assumes that an extra information is available, namely, that the particle was originally launched from $x < 0$ towards the origin; it follows that one must set $a_4 = 0$, because a wave that propagates in the backward direction cannot exist in the region $x > 0$. By the same token one should set $a_1 = 0$ if the particle was launched from $x > 0$ towards the origin. From the continuity of w and w' in the origin it follows

$$\begin{cases} w_+(0) = w_-(0)\,, & a_1 + a_2 = a_3 \\ w'_+(0) = w'_-(0)\,, & k\,(a_1 - a_2) = k_1\, a_3 \end{cases} \tag{11.8}$$

Eliminating a_3 yields $k_1\,(a_1 + a_2) = k\,(a_1 - a_2)$ whence

$$\frac{a_2}{a_1} = \frac{k - k_1}{k + k_1}\,, \qquad \frac{a_3}{a_1} = 1 + \frac{a_2}{a_1} = \frac{2\,k}{k + k_1}\,, \tag{11.9}$$

that determine a_2, a_3 apart from the arbitrary constant a_1. The eigenvalues are continuous in the range $E > V_0$. The monochromatic, time-dependent wave function reads

$$\psi = \begin{cases} \psi_- = w_- \exp(-\mathrm{i}\, E_-\, t/\hbar)\,, & x < 0 \\ \psi_+ = w_+ \exp(-\mathrm{i}\, E_+\, t/\hbar)\,, & x > 0 \end{cases} \tag{11.10}$$

where $E_- = E(k) = \hbar^2\, k^2/(2\,m)$ and $E_+ = E(k_1) = \hbar^2\, k_1^2/(2\,m) + V_0$. Note that $k > k_1 > 0$, namely, the modulus of the particle's momentum in the $x > 0$ region is smaller than in the $x < 0$ region; this is similar to what happens in the classical treatment. On the other hand, from $k > k_1$ it follows $a_2 \neq 0$, showing that a monochromatic plane wave propagating in the backward direction exists in the $x < 0$ region. In other term, a finite probability of reflexion is present, which would be impossible in the classical treatment. As before, the reflection coefficient of the monochromatic wave is defined as $R = |a_2/a_1|^2 < 1$. In turn, the *transmission coefficient* is defined as $T = 1 - R$ whence, from (11.9),

$$R = \left|\frac{a_2}{a_1}\right|^2 = \frac{(k - k_1)^2}{(k + k_1)^2}\,, \qquad T = \frac{4\,k\,k_1}{(k + k_1)^2} = \frac{k_1}{k}\left|\frac{a_3}{a_1}\right|\,. \tag{11.11}$$

Like in the $0 < E < V_0$ case one may build up a wave packet of standard deviation Δx. Still considering a packet approaching the origin from $x < 0$, the

envelope has the form $A_i = A(x - u\,t)$. Before the particle starts interacting with the step, its localization is given by $|\psi_i|^2 = |A_i|^2$. The packet crosses the origin in a time $\Delta t \sim \Delta x/u$ and, using the symbols k_0, k_{10} to indicate the center of the packet in the momentum space for $x < 0$ and $x > 0$, respectively, from the first of (11.11) the reflected packet is described by

$$|\psi_r|^2 = \frac{(k_0 - k_{10})^2}{(k_0 + k_{10})^2}\,|A(x + u\,t)|^2\,. \tag{11.12}$$

It has the same group velocity, hence the same width, as $|\psi_i|^2$. The transmitted packet has the form

$$|\psi_t|^2 = \frac{(2k_0)^2}{(k_0 + k_{10})^2}\,|A(k_0\,x/k_{10} - u\,t)|^2\,, \tag{11.13}$$

and its group velocity is $u_1 = \mathrm{d}x/\mathrm{d}t = k_{10}\,u/k_0 < u$. As all packets cross the origin in the same time interval Δt it follows $\Delta x/u = \Delta x_1/u_1$ whence $\Delta x_1 = (k_{10}/k_0)\,\Delta x < \Delta x$. This result shows that the transmitted packet is slower and narrower than the incident packet (if the incident packet was launched from $x > 0$ towards the origin, the transmitted packet would be faster and broader). From $\int_{-\infty}^{0} |\psi_i|^2\,\mathrm{d}x = 1$ it follows

$$P_r = \int_{-\infty}^{0} |\psi_r|^2\,\mathrm{d}x = \frac{(k_0 - k_{10})^2}{(k_0 + k_{10})^2}\,, \tag{11.14}$$

$$P_t = \frac{k_{10}}{k_0}\int_0^{\infty} |\psi_t|^2\,\mathrm{d}\frac{k_0\,x}{k_{10}} = \frac{k_{10}}{k_0}\,\frac{(2\,k_0)^2}{(k_0 + k_{10})^2} = \frac{4\,k_0\,k_{10}}{(k_0 + k_{10})^2}\,. \tag{11.15}$$

The two numbers P_r, P_t fulfill the relations $0 < P_r, P_t < 1$, $P_r + P_t = 1$, and are the reflection and transmission probabilities of the wave packet. The treatment outlined above for the case $E > V_0$ still holds when $V_0 < 0$, $E > 0$. In particular, if $V_0 < 0$ and $|V_0| \gg E$, like in Fig. 11.2, from (11.16) it turns out $P_r \simeq 1$. This result is quite different from that obtained from a calculation based on Classical Mechanics; in fact, its features are similar to those of the propagation of light across the interface between two media of different refraction index [15, Sects. 1.5,1.6]. The oddity of the result lies, instead, in the fact that term $\sqrt{2\,m}/\hbar$ cancels out in the expressions of P_r and P_t, so that one finds

$$P_r = \frac{(\sqrt{E_0} - \sqrt{E_0 - V_0})^2}{(\sqrt{E_0} + \sqrt{E_0 - V_0})^2}\,, \qquad P_t = \frac{\sqrt{E_0\,(E_0 - V_0)}}{(\sqrt{E_0} + \sqrt{E_0 - V_0})^2}\,, \tag{11.16}$$

with E_0 the total energy corresponding to k_0 and k_{10}. Thus, the classical result $P_r = 0, P_t = 1$ cannot be recovered by making, e.g., m to increase: the discontinuity in the potential energy makes it impossible to apply the Ehrenfest approximation (10.33)

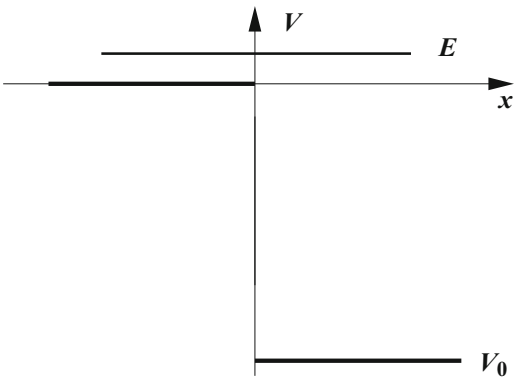


Fig. 11.2 Another example of the step-like potential energy analyzed in Sect. 11.2, with $V_0 < 0$ and $|V_0| \gg E$

no matter what the value of the mass is. The same happens for the monochromatic wave: in fact, using (11.6) and replacing E_0 with E in (11.16) makes the latter equal to (11.11). To recover the classical result it is necessary to consider a potential energy whose asymptotic values 0 and V_0 are connected by a smooth function, and solve the corresponding Schrödinger equation. The classical case is then recovered by letting m increase (Prob. 11.1).

11.3 Energy Barrier

Consider a one-dimensional energy barrier as shown in Fig. 11.3, with $V = V_0 > 0$ for $0 < x < s$ and $V = 0$ elsewhere. From the general properties of the time-independent Schrödinger equation (Sect. 8.2.3) it follows $E \geq 0$. To proceed, it is convenient to consider the two cases $0 < E < V_0$ and $E > V_0$ separately.

11.3.1 Case A: $0 < E < V_0$

The Schrödinger equation is split over the three domains to yield

$$\begin{cases} x < 0 : & -w'' = k^2\, w\,, & k = \sqrt{2\,m\,E}/\hbar \\ 0 < x < s : & w'' = \alpha^2\, w\,, & \alpha = \sqrt{2\,m\,(V_0 - E)}/\hbar \\ s < x : & -w'' = k^2\, w\,, & k = \sqrt{2\,m\,E}/\hbar \end{cases} \tag{11.17}$$

where the derivatives are indicated with primes. The solutions of (11.17) are, respectively,

$$w = \begin{cases} w_- = a_1\, \exp(\mathrm{i}\,k\,x) + a_2\, \exp(-\mathrm{i}\,k\,x)\,, & x < 0 \\ w_B = a_3\, \exp(\alpha\,x) + a_4\, \exp(-\alpha\,x)\,, & 0 < x < s \\ w_+ = a_5\, \exp(\mathrm{i}\,k\,x) + a_6\, \exp(-\mathrm{i}\,k\,x)\,, & s < x \end{cases} \tag{11.18}$$

Using the continuity of w and w' in the origin,

$$\begin{cases} w_-(0) = w_B(0)\,, & a_1 + a_2 = a_3 + a_4 \\ w'_-(0) = w'_B(0)\,, & \mathrm{i}\,k\,(a_1 - a_2) = \alpha\,(a_3 - a_4) \end{cases} \tag{11.19}$$

Solving for a_1, a_2 and letting $\vartheta = 1 + \mathrm{i}\,\alpha/k$ yields $2\,a_1 = \vartheta\,a_4 + \vartheta^*\,a_3$, $2\,a_2 = \vartheta\,a_3 + \vartheta^*\,a_4$. Using the same reasoning as in Sect. 11.2.2, one now assumes that an extra information is available, namely, that the particle was originally launched from $x < 0$ towards the barrier; it follows that one must set $a_6 = 0$ in (11.18), because a wave that propagates in the backward direction cannot exist in the region $x > s$. By the same token one should set $a_1 = 0$ if the particle were launched from $x > s$ towards the barrier. Taking the first case ($a_6 = 0$) and using the continuity of w and w' at $x = s$ provides

$$\begin{cases} w_B(s) = w_+(s)\,, & a_3\,\exp(\alpha\,s) + a_4\,\exp(-\alpha\,s) = a_5\,\exp(\mathrm{i}\,k\,s) \\ w'_B(s) = w'_+(s)\,, & \alpha\,[a_3\,\exp(\alpha\,s) - a_4\,\exp(-\alpha\,s)] = \mathrm{i}\,k\,a_5\,\exp(\mathrm{i}\,k\,s) \end{cases} \tag{11.20}$$

whence $2\,a_3 = a_5\,\sigma\,\exp(\mathrm{i}\,k\,s - \alpha\,s)$, $2\,a_4 = a_5\,\sigma^*\exp(\mathrm{i}\,k\,s + \alpha\,s)$, with $\sigma = 1 + \mathrm{i}\,k/\alpha$. In summary, a_1, a_2 are linear combinations of a_3, a_4; the latter, in turn, are proportional to a_5. It follows that if it were $a_5 = 0$, then it would also be $a_3 = a_4 = 0$ and $a_1 = a_2 = 0$. However, this is impossible because w_- cannot vanish identically; as a consequence it is necessarily $a_5 \neq 0$. This shows that $w_+ = a_5\,\exp(\mathrm{i}\,k\,x)$ differs from zero, namely, that a wave propagating in the forward direction exists for $x > s$.

As the relations involving the coefficients a_i are homogeneous, they determine a_2, a_3, a_4, a_5 apart from the arbitrary constant a_1. This should be expected as w is not normalizable. As shown below, the determination of the ratios between the coefficient does not impose any constraint on the total energy; as a consequence, the eigenvalues $0 \le E < V_0$ are continuous. The ratio a_5/a_1 is found from

$$4\,\frac{a_1}{a_5} = 2\,\frac{a_3}{a_5}\,\vartheta^* + 2\,\frac{a_4}{a_5}\,\vartheta = \sigma\,\exp(\mathrm{i}\,k\,s - \alpha\,s)\,\vartheta^* + \sigma^*\,\exp(\mathrm{i}\,k\,s + \alpha\,s)\,\vartheta\,. \tag{11.21}$$

Letting $\mu = \vartheta\,\sigma^* = 2 + \mathrm{i}\,(\alpha/k - k/\alpha)$ one finds $4\,a_1/a_5\,\exp(-\mathrm{i}\,k\,s) = \mu\,\exp(\alpha\,s) + \mu^*\,\exp(-\alpha\,s)$, one finds $a_1/a_5\,\exp(-\mathrm{i}\,k\,s) = \cosh(\alpha\,s) + \mathrm{i}\,(\alpha/k - k/\alpha)\,\sinh(\alpha\,s)/2$. Using the identity $\cosh^2\zeta - \sinh^2\zeta = 1$ finally yields for the transmission coefficient of the monochromatic wave

$$\frac{1}{T} = \left|\frac{a_1}{a_5}\right|^2 = 1 + \frac{1}{4}\left(\frac{\alpha}{k} + \frac{k}{\alpha}\right)^2\,\sinh^2(\alpha\,s)\,. \tag{11.22}$$

A similar calculation provides the ratio a_2/a_1; it is found

$$R = \left|\frac{a_2}{a_1}\right|^2 = \frac{(\alpha/k + k/\alpha)^2\,\sinh^2(\alpha\,s)/4}{1 + (\alpha/k + k/\alpha)^2\,\sinh^2(\alpha\,s)/4} = 1 - \left|\frac{a_5}{a_1}\right|^2 = 1 - T\,. \tag{11.23}$$

In the classical treatment, if a particle with $0 < E < V_0$ is launched from the left towards the barrier, it is reflected at $x = 0$; similarly, it is reflected at $x = s$ if it is launched from the right. In the quantum treatment it is $T > 0$: in contrast to the classical case, the particle can cross the barrier. For a given width s of the barrier and total energy E of the particle, the transmission coefficient T decreases when V_0 increases; for E and V_0 fixed, T becomes proportional to $\exp(-2\alpha s)$ when s increases to the extent that $\alpha\, s \gg 1$. Finally, $T \to 1$ as $s \to 0$; this was expected because the potential energy becomes equal to zero everywhere.[1]

The interaction of the particle with the barrier is better understood by considering a wave packet approaching the origin from $x < 0$. The incident envelope has the form $A_i = A(x - u\,t)$. Before reaching the origin the particle's localization is described by $|\psi_i|^2 = |A_i|^2$. After the interaction with the barrier is completed, both the reflected and transmitted packet exist, that move in opposite directions with the same velocity. Letting $P_r = \int_{-\infty}^{0} |\psi_r|^2 \,\mathrm{d}x$, $P_t = \int_{s}^{\infty} |\psi_t|^2 \,\mathrm{d}x$, and observing that $\int_{-\infty}^{0} |\psi_i|^2 \,\mathrm{d}x = 1$, it follows that the two numbers P_r P_t fulfill the relations $0 < P_r, P_t < 1$, $P_r + P_t = 1$, and are the reflection and transmission probabilities of the wave packet. In summary, the solution of the Schrödinger equation for the energy barrier shows that a particle with $0 < E < V_0$ has a finite probability of crossing the barrier, which would be impossible in the classical treatment. The same result holds when the form of the barrier is more complicate than the rectangular one (Sect. 11.4). The phenomenon is also called *tunnel effect*.

11.3.2 Case B: $0 < V_0 < E$

Still considering the one-dimensional barrier of Fig. 11.3 let $0 < V_0 < E$. The Schrödinger equation over the three domains reads

$$\begin{cases} x < 0: & -w'' = k^2\, w\,, & k = \sqrt{2\,m\,E}/\hbar \\ 0 < x < s: & -w'' = k_1^2\, w\,, & k_1 = \sqrt{2\,m\,(E - V_0)}/\hbar \\ s < x: & -w'' = k^2\, w\,, & k = \sqrt{2\,m\,E}/\hbar \end{cases} \tag{11.24}$$

where the derivatives are indicated with primes. The solutions of (11.24) are, respectively,

$$w = \begin{cases} w_- = a_1\, \exp(\mathrm{i}\,k\,x) + a_2\, \exp(-\mathrm{i}\,k\,x)\,, & x < 0 \\ w_B = a_3\, \exp(\mathrm{i}\,k_1\,x) + a_4\, \exp(-\mathrm{i}\,k_1\,x)\,, & 0 < x < s \\ w_+ = a_5\, \exp(\mathrm{i}\,k\,x) + a_6\, \exp(-\mathrm{i}\,k\,x)\,, & s < x \end{cases} \tag{11.25}$$

[1] If the potential energy were different on the two sides of the barrier, namely, $V = V_0 > 0$ for $0 < x < s$, $V = V_L$ for $x < 0$, and $V = V_R \neq V_L$ for $x > s$, with $V_0 > V_L, V_R$, the limit $s \to 0$ would yield the case discussed in Sect. 11.2 (compare also with Sects. 11.4 and 17.9.4).

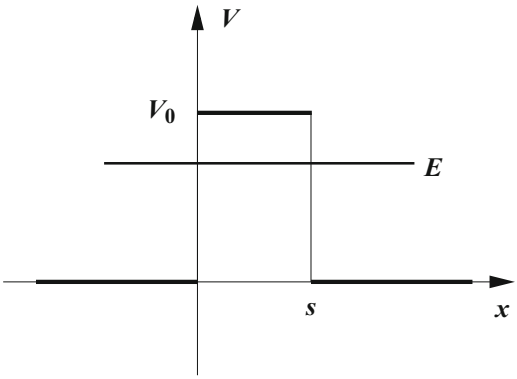


Fig. 11.3 The example of the one-dimensional energy barrier analyzed in Sect. 11.3. Only the case $0 \leq E \leq V_0$ is shown

As in Sect. 11.3.2 one assumes that the particle was originally launched from $x < 0$, so that $a_6 = 0$. The calculation follows the same line as in Sect. 11.3.2, and yields that w_- and w_+ are the same as in (11.24), whereas w_B is found by replacing α with $\mathrm{i}\, k_1$ there. The determination of the ratios between the coefficient a_i does not impose any constraint on the total energy; as a consequence, the eigenvalues $E > V_0$ are continuous. Using $\cosh(\mathrm{i}\,\zeta) = \cos(\zeta)$, $\sinh(\mathrm{i}\,\zeta) = \mathrm{i}\,\sin(\zeta)$ then yields

$$\left|\frac{a_1}{a_5}\right|^2 = 1 + \frac{1}{4}\left(\frac{k}{k_1} - \frac{k_1}{k}\right)^2 \sin^2(k_1\, s) \tag{11.26}$$

where, from (11.24), $k_1/k = \sqrt{1 - V_0/E}$. Similarly,

$$\left|\frac{a_2}{a_1}\right|^2 = \frac{(k/k_1 - k_1/k)^2\, \sin^2(k_1\, s)/4}{1 + (k/k_1 - k/k_1)^2\, \sin^2(k_1\, s)/4} = 1 - \left|\frac{a_5}{a_1}\right|^2 . \tag{11.27}$$

In the classical treatment, if a particle with $0 < V_0 < E$ is launched towards the barrier, it is always transmitted. In the quantum treatment it may be $R > 0$: in contrast to the classical case, the particle can be reflected.[2] The barrier is transparent ($R = 0$) for $k_1\, s = i\,\pi$, with i any integer. Letting $\lambda_1 = 2\,\pi/k_1$ yields $s = i\,\lambda_1/2$, which is equivalent to the optical-resonance condition in a sequence of media of refractive indices n_1, n_2, n_1 [15, Sect. 1.6].

[2]The reflection at the barrier for $k_1\, s \neq i\,\pi$ explains why the experimental value of the Richardson constant A is lower than the theoretical one. Such a constant appears in the expression $J_s = A\, T^2\, \exp[-E_W/(k_B\, T)]$ of the vacuum-tube characteristics [31]. This is one of the cases where the effect is evidenced in macroscopic-scale experiments. Still considering the vacuum tubes, another experimental evidence of the tunnel effect is the lack of saturation of the forward current-voltage characteristic at increasing bias.

11.4 Energy Barrier of a General Form

The interaction of a particle with an energy barrier has been discussed in Sect. 11.3 with reference to the simple case of Fig. 11.3. Here it is extended to the case of a barrier of a general form—still considering the one-dimensional, time-independent Schrödinger equation for a particle of mass m, to the purpose of calculating the transmission coefficient T. The equation reads

$$\frac{\mathrm{d}^2 w}{\mathrm{d}x^2} + q\, w = 0\,, \qquad q(x) = \frac{2\,m}{\hbar^2}\,(E - V)\,, \tag{11.28}$$

where the potential energy $V(x)$ is defined as follows:

$$V = V_L = \text{const.}\,, \quad x < 0\,; \qquad V = V_R = \text{const.}\,, \quad 0 < s < x\,. \tag{11.29}$$

In the interval $0 \le x \le s$ the potential energy is left unspecified, with the only provision that its form is such that (11.28) is solvable. It will also be assumed that $E > V_L, V_R$; as a consequence, the total energy is not quantized and all values of E larger than V_L and V_R are allowed. For a given E the time-dependent wave function takes the form $\psi(x,t) = w(x)\,\exp(-\mathrm{i}\,E\,t/\hbar)$, where

$$w(x) = a_1\,\exp(\mathrm{i}\,k_L\,x) + a_2\,\exp(-\mathrm{i}\,k_L\,x)\,, \qquad x < 0\,, \tag{11.30}$$

$$w(x) = a_5\,\exp(\mathrm{i}\,k_R\,x) + a_6\,\exp(-\mathrm{i}\,k_R\,x)\,, \qquad s < x\,. \tag{11.31}$$

The real parameters $k_L, k_R > 0$ are given by $k_L = \sqrt{2\,m\,(E - V_L)}/\hbar$ and, respectively, $k_R = \sqrt{2\,m\,(E - V_R)}/\hbar$. Like in the examples of Sect. 11.3 it is assumed that the particle is launched from $-\infty$, so that the plane wave corresponding to the term multiplied by a_6 in (11.31) does not exist. As a consequence one lets $a_6 = 0$, whereas a_1, a_2, a_5 are left undetermined. In the interval $0 \le x \le s$ the general solution of (11.28) is

$$w(x) = a_3\,u(x) + a_4\,v(x)\,, \qquad 0 \le x \le s\,, \tag{11.32}$$

where u, v are two linearly independent solutions. The continuity equation for the wave function (9.13), becomes in this case $\mathrm{d}J_\psi/\mathrm{d}x = 0$, namely, $J_\psi = \text{const.}$ In turn, the density of the probability flux reads

$$J_\psi = \frac{\mathrm{i}\,\hbar}{2\,m}\left(w\,\frac{\mathrm{d}w^*}{\mathrm{d}x} - w^*\,\frac{\mathrm{d}w}{\mathrm{d}x}\right). \tag{11.33}$$

Applying (11.33) to (11.30) and (11.31) yields, respectively,

$$J_\psi = \frac{\hbar\,k_L}{m}\left(|a_1|^2 - |a_2|^2\right), \quad x < 0\,; \qquad J_\psi = \frac{\hbar\,k_R}{m}\,|a_5|^2\,, \quad s < x\,. \tag{11.34}$$

As J_ψ is a constant one may equate the two expressions in (11.34) to obtain

$$\left|\frac{a_2}{a_1}\right|^2 + \frac{k_R}{k_L}\left|\frac{a_5}{a_1}\right|^2 = 1\,. \tag{11.35}$$

The division by $|a_1|^2$ leading to (11.35) is allowed because, by hypothesis, the particle is launched from $-\infty$, so that $a_1 \neq 0$. From (11.35) one defines the reflection and transmission coefficients

$$R = \left|\frac{a_2}{a_1}\right|^2\,, \qquad T = \frac{k_R}{k_L}\left|\frac{a_5}{a_1}\right|^2\,. \tag{11.36}$$

Given E, V_L, V_R, and a_1, the transmission and reflection coefficients depend on the form of the potential energy in the interval $0 \leq x \leq s$. By way of example, if $V_L = 0$ and the potential energy in the interval $0 \leq x \leq s$ is equal to some constant $V_B \geq 0$, then R is expected to vary from 0 to 1 as V_B varies from 0 to $+\infty$. On the other hand, R and T cannot vary independently from each other because of the relation $R + T = 1$. As a consequence, it suffices to consider only one coefficient, say, T. From the discussion above it follows that the coefficient T depends on the shape of the potential energy that exists within the interval $0 \leq x \leq s$, namely, it is a functional of V: $0 \leq T = T[V] \leq 1$. One may also note that the relation $R + T = 1$ derives only from the constancy of J_ψ due to the one-dimensional, steady-state condition. In other terms, the relation $R + T = 1$ does not depend on the form of the potential energy within the interval $0 \leq x \leq s$. It must then reflect some invariance property intrinsic to the solution of the problem. In fact, the invariance is that of the density of the probability flux J_ψ, which is intrinsic to the form of the Schrödinger equation and leads to the relation (11.35).

For the sake of generality one provisionally considers a slightly more general equation than (11.28), built by the linear operator

$$\mathscr{L} = \frac{\mathrm{d}^2}{\mathrm{d}x^2} + p(x)\,\frac{\mathrm{d}}{\mathrm{d}x} + q(x)\,, \tag{11.37}$$

where the functions p and q are real. If u is a solution of the differential equation $\mathscr{L}w = 0$ in the interval $0 \leq x \leq s$, let $P(x)$ be any function such that $p = \mathrm{d}P/\mathrm{d}x$, and define

$$v(x) = u(x)\int_a^x \frac{\exp[-P(\xi)]}{u^2(\xi)}\,\mathrm{d}\xi\,, \tag{11.38}$$

where a, x belong to the same interval. It is found by inspection that $\mathscr{L}v = 0$ in the interval $0 \leq x \leq s$, namely, v is also a solution. Moreover, the Wronskian of u and v (Sect. A.12) reads

$$W(x) = u\,v' - u'\,v = \exp(-P)\,. \tag{11.39}$$

As the Wronskian never vanishes, u and v are linearly independent. This shows that for any solution u of the differential equation $\mathscr{L}w = 0$ in a given interval, (11.38) provides another solution which is linearly independent from u (see also Prob. 8.3). As the differential equation is linear and of the second order, the general solution is then given by a linear combination of u and v.

Being the equation $\mathscr{L}w = 0$ homogeneous, the solution u may be replaced by λu, with $\lambda \neq 0$ a constant. In this case v must be replaced by v/λ due to (11.38). It follows that the Wronskian (11.39) is invariant under scaling of the solutions. Another consequence of the homogeneity of $\mathscr{L}w = 0$ is that the dimensions of w may be chosen arbitrarily. The same holds for the dimensions of u. Once the latter have been chosen, the dimensions of v follow from (11.38); in fact, the product uv has the dimensions of a length. From (11.32) it then follows that the products $a_3\,u$ and $a_4\,v$ have the same dimensions.

The linear independency allows one to choose for u and v the two fundamental solutions, namely, those having the properties [66, Sect. 5.2]

$$u(0) = 1\,, \qquad u'(0) = 0\,, \qquad v(0) = 0\,, \qquad v'(0) = 1\,, \tag{11.40}$$

so that the Wronskian W equals 1 everywhere. Then, letting $a_6 = 0$ in (11.31) and prescribing the continuity of the solution and its derivative at $x = 0$ and $x = s$ yields, from (11.32),

$$a_3 = a_1 + a_2\,, \qquad a_4 = \mathrm{i}\,k_L\,(a_1 - a_2)\,, \tag{11.41}$$

$$a_5 \exp(jk_R s) = a_3\,u_s + a_4\,v_s\,, \qquad \mathrm{i}\,k_R\,a_5\,\exp(\mathrm{i}\,k_R\,s) = a_3\,u'_s + a_4\,v'_s\,, \tag{11.42}$$

where suffix s indicates that the functions are calculated at $x = s$. Eliminating $a_5\,\exp(\mathrm{i}\,k_R\,s)$ yields $(\mathrm{i}\,k_R\,u_s - u'_s)\,a_3 = (v'_s - \mathrm{i}\,k_R\,v_s)\,a_4$ whence, from (11.41),

$$\frac{a_2}{a_1} = \frac{k_L k_R v_s + u'_s + j\,(k_L\,v'_s - k_R\,u_s)}{k_L k_R v_s - u'_s + j\,(k_L\,v'_s + k_R\,u_s)} = \frac{A + jB}{C + jD}\,. \tag{11.43}$$

In conclusion, $T = 1 - |a_2/a_1|^2 = (C^2 - A^2 + D^2 - B^2)/(C^2 + D^2)$. Using $W = 1$ transforms the numerator into $C^2 - A^2 + D^2 - B^2 = 4\,k_L\,k_R\,(u_s\,v'_s - v_s\,u'_s) = 4\,k_L\,k_R$. In turn, the denominator reads $C^2 + D^2 = 2\,k_L\,k_R + (k_R\,u_s)^2 + (u'_s)^2 + (k_L\,v'_s)^2 + (k_L\,k_R\,v_s)^2$, whence

$$T = \frac{4\,k_L\,k_R}{2\,k_L\,k_R + (k_R\,u_s)^2 + (u'_s)^2 + (k_L\,v'_s)^2 + (k_L\,k_R\,v_s)^2}\,. \tag{11.44}$$

The expression of the transmission coefficient may be recast as $T = 1/(1+F)$, with

$$F = \frac{1}{4\,k_L\,k_R}\left[\left(u'_s\right)^2 + (k_R\,u_s)^2\right] + \frac{k_L}{4\,k_R}\left[\left(v'_s\right)^2 + (k_R\,v_s)^2\right] - \frac{1}{2}\,. \tag{11.45}$$

It is easily shown that $F > 0$; in fact, this condition is equivalent to $(k_R\, u_s - k_L\, v'_s)^2 + (u'_s + k_L\, k_R\, v_s)^2 > 0$. In conclusion, (11.44) is the expression of the transmission coefficient across a barrier of any form, with no approximation [110]. To calculate T it is necessary to determine the four quantities u_s, u'_s, v_s, and v'_s. Actually only three of them suffice thanks to the condition $u_s\, v'_s - u'_s\, v_s = 1$.

Repeating the calculation of this section after letting $a_1 = 0$, $a_6 \neq 0$ provides the known result that, for a given barrier, the transmission probability for a particle of energy E is the same whether the particle is launched from the left or from the right. The property holds also in the relativistic case [98].

11.5 Energy Well

Taking a one-dimensional case, let $V = V_0 < 0$ for $0 < x < s$ and $V = 0$ elsewhere. From the general properties of the time-independent Schrödinger equation it follows $E \geq V_0$. The case $E > 0$ is treated in the same way as that of Sect. 11.3.2 and leads to similar results. The case $V_0 < E < 0$, shown in Fig. 11.4, is instead different from those investigated above: the total energy is quantized and the wave function is square integrable. The Schrödinger equation over the three domains reads

$$\begin{cases} x < 0: & w'' = \alpha^2\, w\,, & \alpha = \sqrt{-2\,m\,E}/\hbar \\ 0 < x < s: & -w'' = k^2\, w\,, & k = \sqrt{2\,m\,(E - V_0)}/\hbar \\ s < x: & w'' = \alpha^2\, w\,, & \alpha = \sqrt{-2\,m\,E}/\hbar \end{cases} \tag{11.46}$$

where the derivatives are indicated with primes. The solutions of (11.46) are, respectively,

$$w = \begin{cases} w_- = a_1\, \exp(\alpha\, x) + a_5\, \exp(-\alpha\, x)\,, & x < 0 \\ w_W = a_2\, \exp(\mathrm{i}\, k\, x) + a_3\, \exp(-\mathrm{i}\, k\, x)\,, & 0 < x < s \\ w_+ = a_4\, \exp(-\alpha\, x) + a_6\, \exp(\alpha\, x)\,, & s < x \end{cases} \tag{11.47}$$

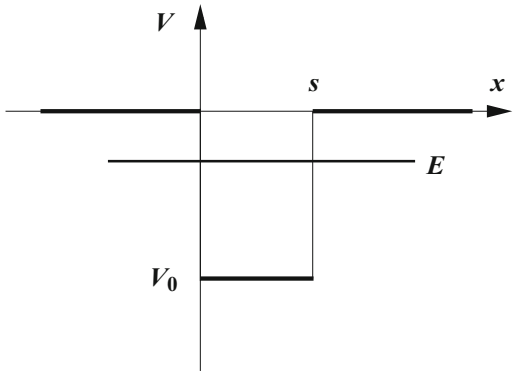


Fig. 11.4 The example of the one-dimensional energy well analyzed in Sect. 11.5. Only the case $V_0 < E < 0$ is shown

where it must be set $a_5 = a_6 = 0$ to prevent w_- and w_+ from diverging. Using the continuity of w and w' in the origin,

$$\begin{cases} w_-(0) = w_W(0)\,, & a_1 = a_2 + a_3 \\ w'_-(0) = w'_W(0)\,, & \alpha\, a_1 = \mathrm{i}\, k\, (a_2 - a_3) \end{cases} \tag{11.48}$$

Solving for a_2, a_3 and letting $\vartheta = 1 + \mathrm{i}\,\alpha/k$ yields $2\,a_2 = \vartheta^*\, a_1$, $2\,a_3 = \vartheta\, a_1$ whence $a_2/a_3 = \vartheta^*/\vartheta$. Then, using the continuity of w and w' at $x = s$ yields

$$\begin{cases} w_W(s) = w_+(s)\,, & a_4\, \exp(-\alpha\, s) = a_2\, \exp(\mathrm{i}\, k\, s) + a_3\, \exp(-\mathrm{i}\, k\, s) \\ w'_W(s) = w'_+(s)\,, & -\alpha\, a_4\, \exp(-\alpha\, s) = \mathrm{i}\, k\, [a_2\, \exp(\mathrm{i}\, k\, s) - a_3\, \exp(-\mathrm{i}\, k\, s)] \end{cases}$$

Solving for a_2, a_3 yields $2\,a_2 = a_4\, \vartheta\, \exp(-\alpha\, s - \mathrm{i}\, k\, s)$, $2\,a_3 = a_4\, \vartheta^*\, \exp(-\alpha\, s + \mathrm{i}\, k\, s)$, whence $a_2/a_3 = (\vartheta/\vartheta^*)\, \exp(-2\,\mathrm{i}\, k\, s)$. In summary, a_2, a_3 are proportional to a_1 and to a_4. It follows that if it were $a_1 = 0$ or $a_4 = 0$, then it would also be $a_2 = a_3 = 0$. However, this is impossible because w cannot vanish identically; as a consequence it is necessarily $a_1 \neq 0$, $a_4 \neq 0$. This shows that, in contrast to the classical case, the particle penetrates the boundaries of the well. The relations found so far determine two different expressions for a_2/a_3. For them to be compatible, the equality $\vartheta^2\, \exp(-\mathrm{i}\, k\, s) = (\vartheta^*)^2\, \exp(\mathrm{i}\, k\, s)$ must hold, which represents the condition of a vanishing determinant of the 4×4, homogeneous algebraic system whose unknowns are a_1, a_2, a_3, and a_4. Using $\vartheta = 1 + \mathrm{i}\,\alpha/k$, the equality is recast as

$$\left(1 - \frac{\alpha^2}{k^2}\right)\, \sin(k\, s) = 2\, \frac{\alpha}{k}\, \cos(k\, s)\,, \qquad \frac{k^2 - \alpha^2}{2\,\alpha\, k} = \cot(k\, s)\,. \tag{11.49}$$

Finally, replacing the expressions of α and k provides the transcendental equation

$$\frac{E - V_0/2}{\sqrt{-E\,(E - V_0)}} = \cot\left(s\, \frac{\sqrt{2\,m}}{\hbar}\, \sqrt{E - V_0}\right)\,, \qquad V_0 < E < 0\,, \tag{11.50}$$

in the unknown E, whose roots fulfill the compatibility condition. As a consequence, such roots are the eigenvalues of E. Given m, s, and V_0, let $n \geq 1$ be an integer such that

$$(n - 1)\, \pi < s\, \sqrt{-(2\,m/\hbar^2)\, V_0} \leq n\pi\,. \tag{11.51}$$

Such an integer always exists and indicates the number of branches of $\cot(k\, s)$ that belong (partially or completely) to the interval $V_0 < E < 0$. In such an interval the left-hand side of (11.50) increases monotonically from $-\infty$ to $+\infty$; as a consequence, (11.50) has n roots $V_0 < E_1\,, E_2\,, \ldots\,, E_n < 0$. An example with five roots is shown in Fig. 11.5; the corresponding calculation is carried out in Prob. 11.2. When an eigenvalue, say E_i, is introduced into (11.46), it provides α_i, k_i, and $\vartheta_i = 1 + \mathrm{i}\,\alpha_i/k_i$; in conclusion,

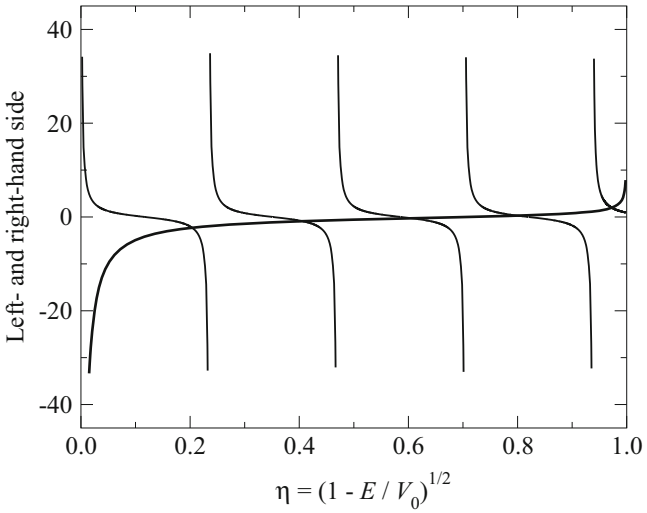


Fig. 11.5 Graphic solution of (11.50) using the auxiliary variable η. The solutions $\eta_1, \ldots, \eta_5$ are the intercepts of the left-hand side (*thicker line*) with the branches of the right-hand side. The data are given in Prob. 11.2

$$a_{i2} = \frac{1}{2}\,\vartheta_i^*\,a_{i1}\,, \qquad a_{i3} = \frac{1}{2}\,\vartheta_i\,a_{i1}\,, \qquad a_{i4} = \frac{\vartheta_i^*}{\vartheta_i}\,\exp(\alpha_i\,s + \mathrm{i}\,k_i\,s)\,a_{i1}\,. \tag{11.52}$$

The ith eigenfunction w_i can thus be expressed, from (11.50), in terms of a_{i1} alone. The latter, in turn, is found from the normalization condition $\int_{-\infty}^{+\infty} |w_i|^2\,\mathrm{d}x = 1$.

The case of the box treated in Sect. 8.2.2 is obtained by letting $V_0 \to -\infty$ here; at the same time one lets $E \to -\infty$ in such a way that the difference $E - V_0$ is kept finite. In this way, w_- and w_+ in (11.47) vanish identically and yield the boundary conditions for w_W used in Sect. 8.2.2.

Problems

11.1 Consider a smooth potential energy described by

$$V(x) = \frac{V_0}{1 + \exp(-x/a)}\,, \tag{11.53}$$

with $V_0 > 0, a > 0$ (Fig. 11.6). The limit of $V(x)$ for $a \to 0$ yields the discontinuous step of Sect. 11.2. Considering a monochromatic wave with $E > V_0$ launched from the left towards the barrier, the reflection coefficient is found to be [55, Sect. 2.2]

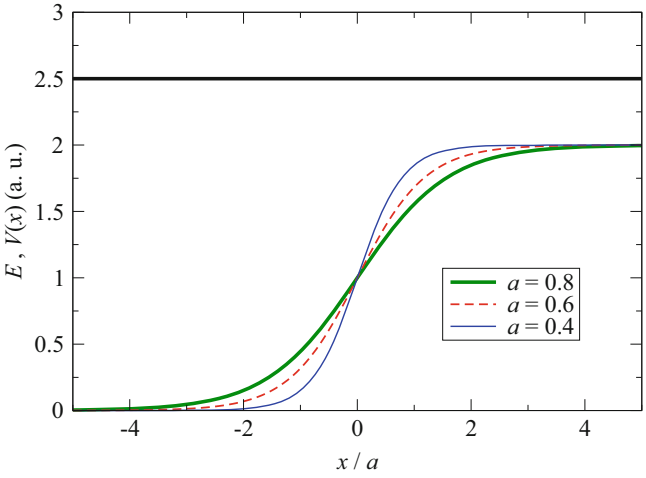


Fig. 11.6 The smooth potential energy considered in Prob. 11.1, with $V_0 = 2$ and $E = 2.5$ (arbitrary units)

$$R(a) = \frac{\sinh^2[\pi\, a\, \sqrt{2\,m}\,(\sqrt{E} - \sqrt{E - V_0})/\hbar]}{\sinh^2[\pi\, a\, \sqrt{2\,m}\,(\sqrt{E} + \sqrt{E - V_0})/\hbar]}\,. \tag{11.54}$$

Discuss the limiting cases $a \to 0$ and $m \to \infty$ and compare the results with those found in Sect. 11.2.

11.2 Find the eigenvalues of the Schrödinger equation for an energy well like that of Fig. 11.5, having a width $s = 15$ Å $= 1.5 \times 10^{-9}$ m and a depth[3] $-V_0 = 3$ eV $\simeq 4.81 \times 10^{-19}$ J. Use $m \simeq 9.11 \times 10^{-31}$ kg, $\hbar \simeq 1.05 \times 10^{-34}$ J s.

[3]The *electron Volt* (eV) is a unit of energy obtained by multiplying 1 J by a number equal to the modulus of the electron charge expressed in C (Table D.1).

Chapter 12
Cases Related to the Linear Harmonic Oscillator

12.1 Introduction

The chapter is devoted to the solution of the Schrödinger equation for the linear harmonic oscillator, and to a number of important application of the results. The importance of the problem has already been outlined in the sections devoted to the classical treatment: its results can in fact be applied, with little or no modification, to mechanical situations where the positional force acting on the particle can be replaced with a first-order expansion, or to more complicate systems whose degrees of freedom can be separated into a set of Hamiltonian functions of the linear-harmonic-oscillator type. Such systems are not necessarily mechanical: for instance, the energy of the electromagnetic field *in vacuo* is amenable to such a separation, leading to the formal justification of the concept of photon. Similarly, a system of particles near a mechanical-equilibrium point can be separated in the same manner, providing the justification of the concept of phonon. An interesting property of the Fourier transform of the Schrödinger equation for the linear harmonic oscillator is shown in the complements.

12.2 Linear Harmonic Oscillator

An example of paramount importance is that of the linear harmonic oscillator, whose classical treatment is given in Sect. 3.3. The potential energy is shown in Fig. 12.1. From the general properties of the time-independent Schrödinger equation (Sect. 8.2.3) it follows $E \geq 0$. Also, the time-independent wave function w is expected to be square integrable. The Hamiltonian operator is found by replacing p with $\hat{p} = -\mathrm{i}\,\hbar\,\mathrm{d}/\mathrm{d}x$ in (3.1), yielding the Schrödinger equation

$$-\frac{\hbar^2}{2\,m}\,\frac{\mathrm{d}^2 w}{\mathrm{d}x^2} + \frac{1}{2}\,m\,\omega^2\,x^2\,w = E\,w\,, \qquad \omega = \sqrt{c/m}\,, \tag{12.1}$$

M. Rudan, *Physics of Semiconductor Devices*, DOI 10.1007/978-3-319-63154-7_12

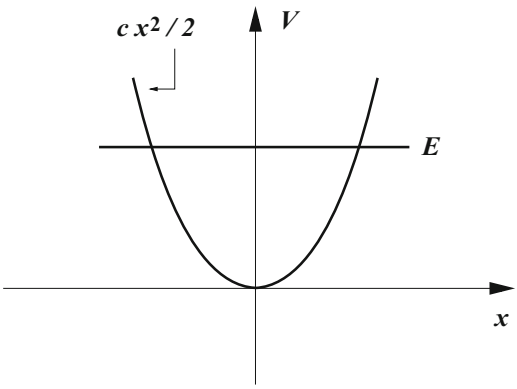


Fig. 12.1 The potential energy of the linear harmonic oscillator (Sect. 12.2)

with c the elastic constant. The equation is conveniently recast in normalized form by defining the dimensionless quantities $\epsilon = E/(\hbar\,\omega)$, $\xi = (m\,\omega/\hbar)^{1/2}\,x$:

$$\mathscr{H}'w = \epsilon\,w\,, \qquad \mathscr{H}' = \frac{1}{2}\left(\xi^2 - \frac{\mathrm{d}^2}{\mathrm{d}\xi^2}\right)\,. \tag{12.2}$$

The eigenvalues and eigenfunctions of (12.2) are found by means of the factorization method illustrated in Sect. 13.3. To begin, one defines the first-order operator

$$\hat{a} = \frac{1}{\sqrt{2}}\left(\xi + \frac{\mathrm{d}}{\mathrm{d}\xi}\right)\,, \tag{12.3}$$

whence, for all f and g vanishing at infinity, an integration by parts yields

$$\int_{-\infty}^{+\infty} g^*\,\hat{a}f\,\mathrm{d}\xi = \int_{-\infty}^{+\infty} \left(\hat{a}^\dagger g\right)^*\,f\,\mathrm{d}\xi\,, \tag{12.4}$$

with $\hat{a}^\dagger = (\xi - \mathrm{d}/\mathrm{d}\xi)/\sqrt{2}$ the adjoint operator of $\hat{a}$. As $\hat{a}^\dagger \neq \hat{a}$, $\hat{a}$ is not Hermitean. Also, $\hat{a}$ and $\hat{a}^\dagger$ do not commute; using the definitions of $\hat{a}$ and $\hat{a}^\dagger$ yields

$$\hat{a}\hat{a}^\dagger - \hat{a}^\dagger\hat{a} = \mathscr{I}\,, \qquad \hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a} = 2\,\mathscr{H}'\,, \tag{12.5}$$

with $\mathscr{I}$ the identity operator. From (12.5), after defining the operator $\mathscr{N} = \hat{a}^\dagger\hat{a}$, called *number operator*, one finds $2\,\mathscr{H}' = (\mathscr{I} + \hat{a}^\dagger\hat{a}) + \hat{a}^\dagger\hat{a} = 2\,\mathscr{N} + \mathscr{I}$, whence

$$\mathscr{H}'w = \mathscr{N}w + \mathscr{I}w/2 = \epsilon\,w \qquad \mathscr{N}w = \left(\epsilon - \frac{1}{2}\right)w = \nu w\,, \tag{12.6}$$

where $\nu = \epsilon - 1/2$. The second relation in (12.6) shows that $\mathscr{H}'$ and $\mathscr{N}$ have the same eigenfunctions, while their eigenvalues differ by $1/2$. Using the same relation

and observing that $\hat{a}$ and $\hat{a}^\dagger$ are real, one finds

$$\int_{-\infty}^{+\infty} w^* \, \hat{a}^\dagger \hat{a} w \, \mathrm{d}\xi = \int_{-\infty}^{+\infty} |\hat{a} w|^2 \, \mathrm{d}\xi = \nu \int_{-\infty}^{+\infty} |w|^2 \, \mathrm{d}\xi \,, \tag{12.7}$$

showing that $\mathscr{N} = \hat{a}^\dagger \hat{a}$, in contrast to $\hat{a}$ and $\hat{a}^\dagger$, is Hermitean, hence its eigenvalues ν are real. Moreover, they are nonnegative: in fact, the integral of $|w|^2$ in (12.7) is strictly positive because an eigenfunction cannot vanish identically; in turn, the integral of $|\hat{a} w|^2$ is strictly positive if $\hat{w} \neq 0$, whereas it vanishes if $\hat{w} = 0$ identically. As a consequence it is $\nu = 0$ if and only if $\hat{a} w = 0$, otherwise it is $\nu > 0$. Another relation is found by left multiplying by $\hat{a}^\dagger$ the first relation in (12.5), to find $\mathscr{N} \hat{a}^\dagger - \hat{a}^\dagger \mathscr{N} = \hat{a}^\dagger$, whence a new operator is defined: $\mathscr{N} \hat{a}^\dagger = \hat{a}^\dagger \, (\mathscr{N} + \mathscr{I})$. The action of the latter upon an eigenfunction w of $\mathscr{N}$ results in

$$\mathscr{N} \hat{a}^\dagger w = \hat{a}^\dagger \, (\mathscr{N} + \mathscr{I}) \, w = \hat{a}^\dagger \nu \, w + \hat{a}^\dagger w = (\nu + 1) \, \hat{a}^\dagger w \,. \tag{12.8}$$

In other terms, if ν is an eigenvalue of $\mathscr{N}$ and w is an eigenfunction belonging to ν, then $\nu + 1$ is also an eigenvalue of $\mathscr{N}$, with eigenfunction $\hat{a}^\dagger w$. This reasoning can be repeated indefinitely: $\mathscr{N} \hat{a}^\dagger \hat{a}^\dagger w = (\nu + 2) \, \hat{a}^\dagger \hat{a}^\dagger w$, ...; its conclusion is that if a number $\nu \geq 0$ is known to be an eigenvalue of $\mathscr{N}$, then all numbers $\nu + 1$, $\nu + 2, \ldots$ belonging to the unlimited ladder[1] beginning with ν are also eigenvalues of $\mathscr{N}$. Also, if w is an eigenfunction belonging to ν, the eigenfunctions belonging to $\nu + 1$, $\nu + 2, \ldots$ are $\hat{a}^\dagger w$, $\hat{a}^\dagger \hat{a}^\dagger w, \ldots$.

One may argue that the same kind of reasoning should be applicable in the backward direction as well, to check whether $\nu - 1$, $\nu - 2$ are eigenvalues of $\mathscr{N}$. In fact, right multiplying by $\hat{a}$ the first relation in (12.5) yields $\hat{a} \mathscr{N} - \mathscr{N} \hat{a} = \hat{a}$, whence $\hat{a} \, (\mathscr{N} - \mathscr{I}) = \mathscr{N} \hat{a}$. The action of $\mathscr{N} \hat{a}$ upon an eigenfunction w of $\mathscr{N}$ results in

$$\mathscr{N} \hat{a} w = \hat{a} \, (\mathscr{N} - \mathscr{I}) \, w = \hat{a} \nu \, w - \hat{a} w = (\nu - 1) \, \hat{a} w \,. \tag{12.9}$$

Remembering that the eigenvalues of $\mathscr{N}$ are nonnegative, (12.9) shows that if ν is an eigenvalue of $\mathscr{N}$ and $\nu \geq 1$, than $\nu - 1$ is also an eigenvalue of $\mathscr{N}$, to which $\hat{a} w$ belongs. By induction, $\mathscr{N} \hat{a} \hat{a} w = (\nu - 2) \, \hat{a} \hat{a} w$, and so on. However, the backward process cannot be repeated indefinitely because one of the numbers $\nu - 2$, $\nu - 3, \ldots$ will eventually be found to be negative: this result contradicts (12.7), that shows that the eigenvalues cannot be negative. The contradiction is due to the implicit assumption that the eigenvalue ν can be any real number, and is readily eliminated by specifying that $\nu = n = 0, 1, 2, \ldots$; consider in fact the eigenvalue $n = 1$ and let w be an eigenfunction belonging to it. Applying (12.9) shows that $w_0 = \hat{a} w$ is an eigenfunction of $\mathscr{N}$ belonging to $n = 0$, namely, $\mathscr{N} w_0 = 0$. The next step in the backward process would yield $\hat{a} w_0$ which, however, is not an eigenfunction because it vanishes identically: this is easily found by combining $\langle w_0 | \mathscr{N} w_0 \rangle = ||\hat{a} w_0||^2$

[1] The term "ladder" is introduced in Sect. 13.3.

with $\mathscr{N} w_0 = 0$. In other terms, the backward process comes automatically to an end when the eigenvalue $n = 0$ is reached.

The above reasoning shows that only the numbers $n = 0, 1, 2, \ldots$ are eigenvalues of $\mathscr{N}$. It also provides an easy method to calculate the eigenfunctions, that starts from the result $\hat{a} w_0 = 0$ just found. Such a relation is a differential equation of the form

$$\frac{1}{\sqrt{2}} \left(\xi + \frac{\mathrm{d}}{\mathrm{d}\xi} \right) w_0 = 0 \,, \qquad \frac{\mathrm{d}w_0}{w_0} = -\xi \, \mathrm{d}\xi \,, \qquad w_0 = c_0 \, \exp\left(-\frac{\xi^2}{2} \right) . \tag{12.10}$$

The normalization constant is found from (C.27), which yields $c_0 = \pi^{-1/4}$. The eigenfunctions corresponding to $n = 1, 2, \ldots$ are found recursively with $w_1 = \hat{a}^\dagger w_0$, $w_2 = \hat{a}^\dagger w_1 = \hat{a}^\dagger \hat{a}^\dagger w_0$, ... For example,

$$w_1 = \frac{1}{\sqrt{2}} \left(\xi w_0 - \frac{\mathrm{d}w_0}{\mathrm{d}\xi} \right) = \frac{w_0}{\sqrt{2}} \, 2 \, \xi \,. \tag{12.11}$$

From this construction it is easily found that the eigenvalues are not degenerate, and that w_n is even (odd) if n is even (odd).[2] Also, it can be shown that w_n has the form [96, Chap. XII.7]

$$w_n(\xi) = \left(n! \, 2^n \, \sqrt{\pi} \right)^{-1/2} \exp(-\xi^2/2) \, H_n(\xi) \,, \tag{12.12}$$

where H_n is the nth *Hermite polynomial*

$$H_n(\xi) = (-1)^n \, \exp(\xi^2) \, \frac{\mathrm{d}^n}{\mathrm{d}\xi^n} \, \exp(-\xi^2) \,. \tag{12.13}$$

The eigenfunctions of the linear harmonic oscillator form a real, orthonormal set:

$$\int_{-\infty}^{+\infty} w_n w_m \, \mathrm{d}x = \delta_{nm} \,. \tag{12.14}$$

By remembering that $\nu = \epsilon - 1/2 = n$ and $\epsilon = E/(\hbar\omega)$, one finds for the energy E of the linear harmonic oscillator the eigenvalues

$$E_n = \left(n + \frac{1}{2} \right) \hbar \, \omega \,, \qquad n = 0, 1, 2, \ldots \tag{12.15}$$

[2] The eigenfunction's nature of being either even or odd is related to a general property of the one-dimensional Schrödinger equation, specifically: if the eigenvalue E is simple and the potential energy V is even, the eigenfunction corresponding to E is either even or odd (Prob. 8.4).

In conclusion, the energy of the linear harmonic oscillator is the sum of the minimum energy $E_0 = \hbar\,\omega/2 > 0$, also called *zero-point energy*, and of an integer number of elementary quanta of energy $\hbar\omega$. The paramount importance of the example of the linear harmonic oscillator has already been emphasized in the classical treatment of Sect. 3.13.1. Examples of application of the quantization of the linear harmonic oscillator are given in Sects. 12.3, 12.4, and 12.5.

The Hermite polynomials (12.13) fulfill the recursive relation [59]

$$H_{n+1} - 2\,\xi\,H_n + 2\,n\,H_{n-1} = 0\,, \qquad n = 1, 2, \ldots\,, \tag{12.16}$$

which is useful to determine some properties of the $\hat{a}$, $\hat{a}^\dagger$ operators. For instance, combining the definition of $\hat{a}^\dagger$ with (12.12) one obtains

$$\hat{a}^\dagger w_n = \frac{1/\sqrt{2}}{(n!\,2^n\,\sqrt{\pi})^{1/2}}\,\exp(-\xi^2/2)\left(2\,\xi\,H_n - \frac{\mathrm{d}H_n}{\mathrm{d}\xi}\right). \tag{12.17}$$

On the other hand, one finds from (12.13) that $\mathrm{d}H_n/\mathrm{d}\xi = 2\,\xi\,H_n - H_{n+1}$. Replacing the derivative in (12.17) yields

$$\hat{a}^\dagger w_n = \frac{1/\sqrt{2}}{(n!\,2^n\,\sqrt{\pi})^{1/2}}\,\exp(-\xi^2/2)\,H_{n+1} = \sqrt{n+1}\,w_{n+1}\,. \tag{12.18}$$

This result shows that, apart from the multiplicative constant, $\hat{a}^\dagger$ transforms the state of index n into that of index $n+1$. Due to (12.15), the transformation corresponds to an increase $E_{n+1} - E_n = \hbar\,\omega$ in the total energy of the oscillator. Since its action "creates" a quantum of energy, $\hat{a}^\dagger$ is called *creation operator*. Using the same procedure, combined with the recursive relation (12.16), one finds

$$\hat{a} w_n = \sqrt{n}\,w_{n-1}\,. \tag{12.19}$$

Due to the above, $\hat{a}$ is called *destruction operator* or *annihilation operator*. Note that, due to (12.18, 12.19), the successive application of $\hat{a}$ and $\hat{a}^\dagger$ to w_n is equivalent to $\mathscr{N} w_n = n\,w_n$.

12.3 Quantization of the Electromagnetic Field's Energy

The energy of the electromagnetic field within a finite volume V free of charge has been calculated in Sect. 5.6 in terms of the field's modes. In such a calculation the shape of V was chosen as that of a box whose sides d_1, d_2, d_3 are aligned with the coordinate axes and start from the origin (Fig. 5.1), so that $V = d_1\,d_2\,d_3$. The energy reads

$$W_{\mathrm{em}} = \sum_{\mathbf{k}\sigma} W_{\mathbf{k}\sigma}\,, \qquad W_{\mathbf{k}\sigma} = 2\varepsilon_0 V\,\omega^2\,s_{\mathbf{k}\sigma}\,s^*_{\mathbf{k}\sigma}\,, \tag{12.20}$$

where ε_0 is the vacuum permittivity, $\mathbf{k} = \sum_{i=1}^{3} 2\pi n_i \mathbf{i}_i/d_i$, $n_i = 0, \pm 1, \pm 2, \ldots$, $\sigma = 1, 2$, and $\omega = c\,|\mathbf{k}|$, with c the speed of light and $\mathbf{i}_i$ the unit vector of the ith axis. Finally, $s_{\mathbf{k}\sigma}(t)$ is one of the two components of the complex vector defined by (5.29). The energy $W_{\mathbf{k}\sigma}$ of the degree of freedom $\mathbf{k}, \sigma$ is written in Hamiltonian form by introducing the canonical coordinates $q_{\mathbf{k}\sigma}, p_{\mathbf{k}\sigma}$ such that

$$2\sqrt{\varepsilon_0 V}\,\omega\, s_{\mathbf{k}\sigma} = \omega\, q_{\mathbf{k}\sigma} + \mathrm{i}\, p_{\mathbf{k}\sigma}\,. \tag{12.21}$$

Using (12.21) transforms (12.20) into

$$W_{\mathbf{k}\sigma} = \frac{1}{2}\,(\omega\, q_{\mathbf{k}\sigma} + \mathrm{i}\, p_{\mathbf{k}\sigma})\,(\omega\, q_{\mathbf{k}\sigma} - \mathrm{i}\, p_{\mathbf{k}\sigma}) = \frac{1}{2}\,\left(p_{\mathbf{k}\sigma}^2 + \omega^2\, q_{\mathbf{k}\sigma}^2\right)\,. \tag{12.22}$$

The Hamiltonian operator is obtained by replacing $q_{\mathbf{k}\sigma}$ with $\hat{q}_{\mathbf{k}\sigma} = q_{\mathbf{k}\sigma}$ and $p_{\mathbf{k}\sigma}$ with $\hat{p}_{\mathbf{k}\sigma} = -\mathrm{i}\,\hbar\,\mathrm{d}/\mathrm{d}q_{\mathbf{k}\sigma}$ in the last expression on the right of (12.22):

$$\mathscr{H}^0_{\mathbf{k}\sigma} = -\frac{\hbar^2}{2}\,\frac{\mathrm{d}^2}{\mathrm{d}q_{\mathbf{k}\sigma}^2} + \frac{\omega^2}{2}\,q_{\mathbf{k}\sigma}^2\,. \tag{12.23}$$

It should be noted, however, that if the intermediate expression in (12.22), instead of that on the right, is used for the replacement of the classical coordinates with the corresponding operators, a different result is obtained. In fact, $\hat{q}_{\mathbf{k}\sigma}$ and $\hat{p}_{\mathbf{k}\sigma}$ do not commute so that, from (8.72), one obtains

$$\mathscr{H}^-_{\mathbf{k}\sigma} = \mathscr{H}^0_{\mathbf{k}\sigma} - \frac{1}{2}\,\hbar\,\omega\,. \tag{12.24}$$

A third form of the Hamiltonian operator, different from the two above, is obtained by exchanging the order of factors in the intermediate expression in (12.22) prior to the replacement of the coordinates with operators:

$$\mathscr{H}^+_{\mathbf{k}\sigma} = \mathscr{H}^0_{\mathbf{k}\sigma} + \frac{1}{2}\,\hbar\,\omega\,. \tag{12.25}$$

To proceed one considers $\mathscr{H}^0_{\mathbf{k}\sigma}$ first. The Schrödinger equation generated by it,

$$\mathscr{H}^0_{\mathbf{k}\sigma} w^0_{\mathbf{k}\sigma} = E^0_{\mathbf{k}\sigma}\, w^0_{\mathbf{k}\sigma} \tag{12.26}$$

is the equation (12.2) of the linear harmonic oscillator, whose eigenvalues (12.15) read $E^0(n_{\mathbf{k}\sigma}) = (n_{\mathbf{k}\sigma} + 1/2)\,\hbar\,\omega$, with $n_{\mathbf{k}\sigma} = 0, 1, 2\ldots$, and are nondegenerate. The second form (12.24) of the Hamiltonian operator generates the Schrödinger equation $\mathscr{H}^-_{\mathbf{k}\sigma} w^-_{\mathbf{k}\sigma} = E^-_{\mathbf{k}\sigma}\, w^-_{\mathbf{k}\sigma}$, namely,

$$\mathscr{H}^0_{\mathbf{k}\sigma}\, w^-_{\mathbf{k}\sigma} = \left(E^-_{\mathbf{k}\sigma} + \frac{1}{2}\,\hbar\,\omega\right)\, w^-_{\mathbf{k}\sigma}\,, \tag{12.27}$$

again the equation for the linear harmonic oscillator, whose operator is identical to that of (12.26). As a consequence, its eigenvalues are $E^-(n_{\mathbf{k}\sigma}) + \hbar\,\omega/2 = (n_{\mathbf{k}\sigma} + 1/2)\,\hbar\,\omega$, whence $E^-(n_{\mathbf{k}\sigma}) = n_{\mathbf{k}\sigma}\,\hbar\,\omega$. By the same token the eigenvalues of (12.25) are found to be $E^+(n_{\mathbf{k}\sigma}) = (n_{\mathbf{k}\sigma} + 1)\,\hbar\,\omega$. From (12.20), the energy of the electromagnetic field is the sum of the energy of the single modes; the three cases considered above then yield:

$$W^0_{\mathrm{em}} = \sum_{\mathbf{k}\sigma} \left(n_{\mathbf{k}\sigma} + \frac{1}{2} \right) \hbar\,\omega\,, \qquad W^-_{\mathrm{em}} = \sum_{\mathbf{k}\sigma} n_{\mathbf{k}\sigma}\,\hbar\omega\,, \qquad W^+_{\mathrm{em}} = \sum_{\mathbf{k}\sigma} (n_{\mathbf{k}\sigma} + 1)\,\hbar\,\omega\,,$$

with $\omega = c\,|\mathbf{k}|$. In the expression of W^0_{em} and W^+_{em} the sum over $\mathbf{k}$ of the terms $\hbar\,\omega/2$ and, respectively, $\hbar\,\omega$, diverges. This is not acceptable because the total energy within V is finite. On the contrary, for the expression of W^-_{em} to converge it is sufficient that $n_{\mathbf{k}\sigma}$ vanishes from some $|\mathbf{k}|$ on; the correct Hamiltonian is thus $\mathscr{H}^-_{\mathbf{k}\sigma}$. Grouping for each $\mathbf{k}$ the summands corresponding to $\sigma = 1$ and $\sigma = 2$, and letting $n_{\mathbf{k}} = n_{\mathbf{k}1} + n_{\mathbf{k}2}$, provides

$$W^-_{\mathrm{em}} = \sum_{\mathbf{k}} n_{\mathbf{k}}\,\hbar\,\omega\,. \tag{12.28}$$

In conclusion, the energy of each mode of oscillation is a multiple (0 included) of the elementary quantum of energy $\hbar\,\omega(\mathbf{k})$. This result provides the formal justification of the concept of *photon*. The integer $n_{\mathbf{k}\sigma}$ is the *occupation number* of the pair $\mathbf{k}, \sigma$, whereas $n_{\mathbf{k}}$ is the number of photons[3] of the mode corresponding to $\mathbf{k}$. Like in the classical treatment, the energy of the electromagnetic field is the sum of the energies of each mode of oscillation.

12.4 Quantization of the Electromagnetic Field's Momentum

The momentum of the electromagnetic field within a finite volume V free of charge has been calculated in Sect. 5.7 in terms of the field's modes. Premises and symbols here are the same as in Sect. 12.3. The momentum reads

$$\int_V \frac{\mathbf{S}}{c^2}\,\mathrm{d}V = 2\,\varepsilon_0\,V \sum_{\mathbf{k}} \omega\,\mathbf{s}_{\mathbf{k}} \cdot \mathbf{s}^*_{\mathbf{k}}\,\mathbf{k} = \sum_{\mathbf{k}\sigma} \frac{1}{c}\,W_{\mathbf{k}\sigma}\,\frac{\mathbf{k}}{k}\,, \tag{12.29}$$

[3] To complete the description of the photon it is necessary to work out also the quantum expression of its momentum. This is done in Sect. 12.4. The concept of photon was introduced by Einstein in 1905 [45] (the English translation of [45] is in [133]). The quantization procedure shown here is given in [41].

with $\mathbf{S}$ the Poynting vector and $k = |\mathbf{k}|$. For each pair $\mathbf{k}, \sigma$, the same quantization procedure used for the energy in Sect. 12.3 is applicable here and yields the operator $\mathscr{H}^-_{\mathbf{k}\sigma}\, \mathbf{k}/(c\,k)$. As the latter differs by a constant vector from the Hamiltonian operator (12.24) corresponding to the same pair $\mathbf{k}, \sigma$, the eigenvalues of the ith component of momentum turn out to be

$$\frac{1}{c}\, n_{\mathbf{k}\sigma}\, \hbar\, \omega\, \frac{k_i}{k} = n_{\mathbf{k}\sigma}\, \hbar k_i\,, \qquad n_{\mathbf{k}\sigma} = 0, 1, 2, \ldots \tag{12.30}$$

In conclusion, the eigenvalues of momentum corresponding to $\mathbf{k}, \sigma$ are $n_{\mathbf{k}\sigma}\, \hbar\mathbf{k}$. Letting as in Sect. 12.3 $n_{\mathbf{k}} = n_{\mathbf{k}1} + n_{\mathbf{k}2}$, the momentum of the electromagnetic field is expressed in quantum terms as $\sum_{\mathbf{k}} n_{\mathbf{k}}\, \hbar\, \mathbf{k}$. This result shows that the momentum of each mode of oscillation is a multiple (0 included) of the elementary quantum of momentum $\hbar\, \mathbf{k}$, and completes the formal justification of the concept of *photon* started in Sect. 12.3. Each photon has energy and momentum, given by $\hbar\, \omega$ and $\hbar\, \mathbf{k}$ respectively. Like in the classical treatment, the momentum of the electromagnetic field is the sum of the momenta of each mode of oscillation.

12.5 Quantization of a Diagonalized Hamiltonian Function

A system of particles near an equilibrium point has been investigated in Sects. 3.9 and 3.10. The analysis led to the separation of the Hamiltonian function that reads

$$H_a - V_{a0} = \sum_{\sigma=1}^{3N} H_\sigma\,, \qquad H_\sigma = \frac{1}{2}\, \dot{b}_\sigma^2 + \frac{1}{2}\, \omega_\sigma^2\, b_\sigma^2\,. \tag{12.31}$$

In (12.31), $3N$ is the number of degrees of freedom, b_σ the normal coordinate of index σ, ω_σ the angular frequency corresponding to b_σ, and V_{a0} the minimum of the system's potential energy. Apart from the constant V_{a0}, the Hamiltonian function H_a is given by a sum of terms, each associated with a single degree of freedom. In turn, each summand H_σ is identical to the Hamiltonian function of a linear harmonic oscillator with $m = 1$. As a consequence, the quantum operator corresponding to (12.31) takes the form

$$\mathscr{T}_a + V_a = \sum_{\sigma=1}^{3N} \mathscr{H}_\sigma + V_{a0}\,, \qquad \mathscr{H}_\sigma = -\frac{\hbar^2}{2}\, \frac{\partial^2}{\partial b_\sigma^2} + \frac{1}{2}\, \omega_\sigma^2\, b_\sigma^2\,, \tag{12.32}$$

and generates the eigenvalue equation $(\mathscr{T}_a + V_a)\, v = E\, v$, where

$$\left(\sum_{\sigma=1}^{3N} \mathscr{H}_\sigma \right) v = E'\, v\,, \qquad E' = E - V_{a0}\,. \tag{12.33}$$

Being $\mathscr{H}_a$ the sum of operators acting on individual degrees of freedom, the Schrödinger equation is separable (Sect. 10.3) and splits into $3N$ equations of the form

$$\mathscr{H}_\sigma \, v_{\sigma\,\zeta(\sigma)} = E_{\sigma\zeta(\sigma)} \, v_{\sigma\,\zeta(\sigma)} \,, \qquad E' = \sum_{\sigma=1}^{3N} E_{\sigma\zeta(\sigma)} \,, \tag{12.34}$$

where $\sigma = 1, 2, \ldots, 3N$ refers to the degrees of freedom, whereas $\zeta(\sigma) = 0, 1, 2, \ldots$ counts the set of eigenvalue indices corresponding to a given σ. Remembering the solution of the Schrödinger equation for linear harmonic oscillator (Sect. 12.2), the energy of the individual degree of freedom is

$$E_{\sigma\zeta(\sigma)} = \left[\zeta(\sigma) + \frac{1}{2}\right] \hbar\omega_\sigma \,, \qquad \zeta(\sigma) = 0, 1, 2, \ldots \,. \tag{12.35}$$

The total energy of the system then reads

$$E = V_{a0} + \sum_{\sigma=1}^{3N} \left[\zeta(\sigma) + \frac{1}{2}\right] \hbar\,\omega_\sigma \,. \tag{12.36}$$

As indicated in Sect. 3.10, the oscillation of the normal coordinate of index σ is called *mode* of the vibrating system. The classical expression of the energy associated with each mode has the same form as that of a mode of the electromagnetic field (compare (12.31) with (12.22)). By analogy with the electromagnetic case, a particle of energy $\hbar\,\omega_\sigma$, called *phonon*, is introduced in the quantum description, and the energy of the mode is ascribed to the set of phonons belonging to the mode. The integers $\zeta(\sigma)$ are the *occupation numbers* of the normal modes of oscillation.[4]

Note that the quadratic form (12.31) of the total energy H_σ of each degree of freedom of the oscillating system was derived directly, in contrast with that of the electromagnetic field where the product of two linear forms was involved (compare with (12.22)). For this reason, the discussion about three possible forms of the Hamiltonian operator, carried out in Sect. 12.3, is not necessary here. The total energy (12.36) does not diverge because the number of degrees of freedom is finite.

The standard way of describing the interaction of an electron with the vibrating nuclei is using the quantum-mechanical, first-order perturbation theory (Sect. 17.8).

[4]The equilibrium distribution of the phonons' occupation numbers $\zeta(\sigma)$ is the Bose-Einstein statistics (Sect. 15.8.2).

12.6 Complements

12.6.1 Comments About the Linear Harmonic Oscillator

The normalized form (12.2) of the Schrödinger equation for the linear harmonic oscillator is $-\mathrm{d}^2 w_n/\mathrm{d}\xi^2 + \xi^2\, w_n = 2\,\epsilon_n\, w_n$. Due to the exponential decay at infinity, the Fourier transform (C.16) of the eigenfunction exists. Let

$$u_n(\eta) = \mathscr{F} w_n = \frac{1}{\sqrt{2\,\pi}} \int_{-\infty}^{+\infty} w_n(\xi)\, \exp(-\mathrm{i}\,\eta\,\xi)\, \mathrm{d}\xi\,. \tag{12.37}$$

Thanks to the property (C.22) of the Fourier transform it is $\mathscr{F}\,\mathrm{d}^2 w_n/\mathrm{d}\xi^2 = -\eta^2\, u_n$, $\mathscr{F}\,\xi^2\, w_n = -\mathrm{d}^2 u_n/\mathrm{d}\eta^2$. Fourier transforming (12.2) thus yields

$$\frac{1}{2}\,\eta^2\, u_n - \frac{1}{2}\,\frac{\mathrm{d}^2 u_n}{\mathrm{d}\eta^2} = \epsilon_n\, u_n\,, \tag{12.38}$$

namely, an equation identical to (12.2), having the same eigenvalue. As ϵ_n is not degenerate, it follows $u_n \propto w_n$, namely, the eigenfunctions of the linear harmonic oscillator are equal to their own Fourier transforms apart from a multiplicative constant at most.[5]

[5]Compare with (C.83), where the property is demonstrated for the Gaussian function; the latter, apart from scaling factors, coincides with the eigenfunction of the linear harmonic oscillator belonging to the eigenvalue corresponding to $n = 0$.

Chapter 13
Other Examples of the Schrödinger Equation

13.1 Introduction

A number of properties of the one-dimensional, time-independent Schrödinger equation can be worked out without specifying the form of the coefficient. To this purpose one examines the two fundamental solutions, which are real because the coefficient is such. One finds that the fundamental solutions do not have multiple zeros and do not vanish at the same point; more precisely, the zeros of the first and second fundamental solution separate each other. It is also demonstrated that the character of the fundamental solutions within an interval is oscillatory or non-oscillatory depending on the sign of the equation's coefficient in such an interval. After completing this analysis, the chapter examines an important and elegant solution method, consisting in factorizing the operator. The theory is worked out for the case of localized states, corresponding to discrete eigenvalues. The procedure by which the eigenfunctions' normalization is embedded into the solution scheme is also shown. The chapter continues with the analysis of the solution of a Schrödinger equation whose coefficient is periodic; this issue finds important applications in the case of periodic structures like, e.g., crystals. Finally, the solution of the Schrödinger equation for a particle subjected to a central force is worked out; the equation is separated and the angular part is solved first, followed by the radial part whose potential energy is specified in the Coulomb case. The first complements deal with the operator associated with the angular momentum and to the solution of the angular and radial equations by means of the factorization method. The last complement generalizes the solution method for the one-dimensional Schrödinger equation in which the potential energy is replaced with a piecewise-constant function, leading to the concept of transmission matrix.

M. Rudan, *Physics of Semiconductor Devices*, DOI 10.1007/978-3-319-63154-7_13

13.2 Properties of the One-Dimensional Schrödinger Equation

In the general expression (11.44) for the transmission coefficient, the fundamental solutions u, v appear in the denominator. It is then necessary to investigate the zeros of the solutions of (11.28). Due to the $u(0) = 1$ prescription, the possible zeros of u belong to the interval $0 < x \leq s$, while those of v belong to the interval $0 \leq x \leq s$.

If one or more zero exist, they cannot be multiple. In fact, if u had a multiple zero at x_m it would be $u(x_m) = 0$, $u'(x_m) = 0$, hence $u = 0$ would be a solution of (11.28) compatible with such conditions. In fact, because of the uniqueness of the solution, $u = 0$ would be the only solution. Observing that u is continuous, this would contradict the condition $u(0) \neq 0$. Similarly, if v had a multiple zero it would vanish identically. Remembering that the derivative of the solution is continuous, this would contradict the condition $v'(0) = 1$ of (11.40). Another property is that u and v cannot vanish at the same point. This is apparent from the relation $W(x) = u\,v' - u'\,v = 1$ demonstrated in Sect. A.12. For the same reason, u' and v' cannot vanish at the same point.

If one of the solutions, say u, has more than one zero in $0 < x \leq s$, then the following property holds: between two consecutive zeros of u there is one and only one zero of v. Let x_L, x_R be two consecutive zeros of u, with $0 < x_L < x_R \leq s$. The property is demonstrated by showing, first, that a contradiction would arise if there were no zeros of v between x_L and x_R (that is to say, at least one zero must exist there) and, second, that if a zero of v exists between x_L and x_R, it must be unique [127]. To proceed one considers the function u/v in the interval $x_L \leq x \leq x_R$. By definition u/v vanishes at the boundaries of such an interval while, as shown above, v cannot vanish at the boundaries. If one assumes that there are no zeros of v inside the interval, then u/v exists everywhere in the interval, and is also everywhere continuous with a continuous first derivative because u and v are solutions of the second-order differential equation (11.28). As u/v vanishes at x_L and x_R, its derivative must vanish at least once in the open interval $x_L < x < x_R$. However, this is impossible because $\mathrm{d}(u/v)/\mathrm{d}x = -W/v^2 = -1/v^2 \neq 0$. This shows that v must have at least one zero between x_L and x_R. Such a zero is also unique because, if v had two zeros in $x_L < x < x_R$, then by the same reasoning u would have one zero between them, so x_L and x_R would not be consecutive. The property may be restated as *the zeros of two real linearly independent solutions of a second-order linear differential equation separate each other*. The property does not hold for complex solutions.

So far the properties demonstrated in this section did not consider the sign of the coefficient $q(x) = 2\,m\,(E - V)/\hbar^2$ of (11.28). The coefficient separates the interval $0 \leq x \leq s$ into subintervals where q is alternatively positive or negative. If q is continuous the extrema of the subintervals are the zeros of q, otherwise they may be discontinuity points of q. In either case the behavior of the solution u within each subinterval depends on the sign of q there. To show this, consider the function $\mathrm{d}(u\,u')/\mathrm{d}x = (u')^2 - q\,u^2$, where the expression at the right-hand side has been found

by means of (11.28). If $q \leq 0$ in the subinterval, then $\mathrm{d}(u\,u')/\mathrm{d}x$ is nonnegative. It follows that $u\,u'$ is a nondecreasing function in the subinterval, hence it has one zero at most. Remembering that u and u' cannot vanish at the same point, one of the following holds: *(i)* neither u nor u' vanishes in the subinterval, *(ii)* either u or u' vanishes once in the subinterval. For a given interval a function is called *non-oscillatory* if its derivative vanishes at most once. It follows that the solution u is non-oscillatory in those subintervals where $q \leq 0$. The case $V = V_0 > E > 0$ in the interval $0 < x < s$, considered in Sect. 11.3.1, is of this type.

The investigation about the behavior of the solutions may be extended to the case of an infinite domain, at least for typical forms of the potential energy V [96, Chap. III.9]. Consider for instance the domain $0 \leq x < \infty$, and assume that for $x > 0$ the potential energy is such that the constraint $-q(x) \geq \alpha^2 > 0$ is fulfilled. Without further specifying the form of q, it is possible to analyze the asymptotic behavior of the solutions by comparison with those of the simple case $-q = \text{const} = \alpha^2$. The fundamental solutions[1] of the simple case are $\tilde{u} = \cosh(\alpha\,x)$ and $\tilde{v} = \sinh(\alpha\,x)/\alpha$, neither of which is acceptable from a physical point of view because both diverge as $x \to \infty$. However, two linear combinations of the fundamental solutions, namely, $\tilde{u} - \alpha\,\tilde{v} = \exp(-\alpha\,x)$ and $\tilde{u} + \alpha\,\tilde{v} = \exp(\alpha\,x)$, are also linearly independent solutions of $w'' = \alpha^2\,w$, the first of which is physically acceptable.

Turning now to the general case $-q(x) \geq \alpha^2 > 0$, and letting u, v be its fundamental solutions, one finds that u is by construction positive and concave upward for $x = 0$; thus, it increases when x departs from 0, thus making the second derivative $-q\,u$ even more positive. In conclusion, u remains positive for $x > 0$; moreover, it keeps increasing at least as fast as $\exp(\alpha\,x)$. One can prove the last statement by the following procedure: multiply by $\tilde{u}$ both sides of $u'' = -q\,u$, then multiply by u both sides of $\tilde{u}'' = \alpha^2\,\tilde{u}$, and subtract the resulting expressions from each other. Following the same steps as in Prob. 8.2, and remembering that here $\alpha^2 + q$ is non-positive, one finds for the Wronskian $W(u, \tilde{u}) = u\,\alpha\,\sinh(\alpha\,x) - u'\,\cosh(\alpha\,x)$ the relation[2]

$$\frac{\mathrm{d}W(u,\tilde{u})}{\mathrm{d}x} = (\alpha^2 + q)\,u\,\tilde{u} \leq 0\,. \tag{13.1}$$

Integrating (13.1) from 0 to x and observing that $W(u, \tilde{u})$ vanishes at $x = 0$, one finds $W(u, \tilde{u}) \leq 0$; from this, remembering that u does not vanish, it follows

$$\frac{u'}{u} = \frac{\mathrm{d}}{\mathrm{d}x}\log u \geq \alpha\,\frac{\sinh(\alpha\,x)}{\cosh(\alpha\,x)} = \frac{\mathrm{d}}{\mathrm{d}x}\log\cosh(\alpha\,x) \tag{13.2}$$

whence, integrating both sides from 0 to x and taking the exponentials, $u \geq \cosh(\alpha\,x)$; moreover, (13.2) shows that asymptotically it is $u' \geq \alpha\,u$. In a similar

[1] The fundamental solutions are defined by (11.40).

[2] The definition of Wronskian is in Sect. A.12.

manner one finds that, by construction, the second fundamental solution v has a positive derivative and is concave upward for $x = 0$; thus, it remains positive and keeps increasing for $x > 0$. Following the same steps as for u yields $W(v, \tilde{v}) = v \cosh(\alpha x) - v' \sinh(\alpha x)/\alpha \leq 0$, whence

$$\frac{v'}{v} = \frac{\mathrm{d}}{\mathrm{d}x} \log v \geq \alpha \frac{\cosh(\alpha x)}{\sinh(\alpha x)} = \frac{\mathrm{d}}{\mathrm{d}x} \log \sinh(\alpha x) \,. \tag{13.3}$$

Integrating both sides from 0 to x and taking the exponentials yield $v \geq \sinh(\alpha x)$; also, (13.3) shows that asymptotically it is $v' \geq \alpha v$. As a final consideration, the Wronskian of the two linearly independent solutions u, v is equal to a constant (Prob. 8.2); using the values of the fundamental solutions and their derivatives at $x = 0$ yields $W(u, v) = 1$.

Having discussed some properties of the fundamental solutions, one may seek a linear combination of u, v that vanishes at infinity. To this purpose, one observes that the ratios u/v and u'/v' have the following property:

$$\frac{u}{v} - \frac{u'}{v'} = \frac{W(u, v)}{v\, v'} = \frac{1}{v\, v'} \,; \tag{13.4}$$

next, using again $W(u, v) = 1$ along with $u'' = -q\, u$ and $v'' = -q\, v$, one finds

$$\frac{\mathrm{d}}{\mathrm{d}x} \left(\frac{u}{v} \right) = -\frac{1}{v^2} < 0 \,, \qquad \frac{\mathrm{d}}{\mathrm{d}x} \left(\frac{u'}{v'} \right) = -\frac{q}{(v')^2} > 0 \,. \tag{13.5}$$

Due to (13.5), u/v is a decreasing function while u'/v' is an increasing function; as their difference (13.4) vanishes asymptotically, the two ratios have a common positive limit λ as $x \to \infty$. The following inequalities then hold, $u'/v' < \lambda < u/v$ or, equivalently, $u'/v' - \lambda < 0 < u/v - \lambda$. Using (13.4), the latter form can further be refined to yield

$$-\frac{1}{v\, v'} = \frac{u'}{v'} - \frac{u}{v} < \frac{u'}{v'} - \lambda < 0 < \frac{u}{v} - \lambda < \frac{u}{v} - \frac{u'}{v'} = \frac{1}{v\, v'} \,, \tag{13.6}$$

which in turn splits as

$$-\frac{1}{v} < u' - \lambda\, v' < 0 \,, \qquad 0 < u - \lambda\, v < \frac{1}{v'} \,. \tag{13.7}$$

Due to the asymptotic behaviors discussed above, the denominators v, v' in (13.7) are such that $v \geq \sinh(\alpha x)$, $v' \geq \alpha v \geq \alpha \sinh(\alpha x)$; as a consequence, the linear combination $u - \lambda v$ of the two fundamental solutions is always positive, and vanishes at infinity at least as fast as $1/v'$, *a fortiori* at least as fast as $1/\sinh(\alpha x) \sim \exp(-\alpha x)$. At the same time the derivative $u' - \lambda v'$ of the linear combination is always negative, and vanishes at infinity in the same manner.

Finally, the linear combination found above is the only solution that vanishes at infinity; in fact, consider another linear combination, say, $w = u + \sigma\, v$ with $\sigma \neq -\lambda$: the new solution can be recast in the form $w = u - \lambda\, v + (\sigma + \lambda)\, v$. While the term $u - \lambda\, v$ vanishes asymptotically, the remaining term diverges with the same strength as v because $\sigma + \lambda \neq 0$.

13.3 Localized States—Operator's Factorization

It may happen that the form of the potential energy V in the interval $0 \leq x \leq s$ is such that V has one or more negative minima (Fig. 13.1). In this case negative eigenvalues of E may exist, giving rise to localized states. To treat this case one must preliminarily observe that the eigenfunctions do not vanish identically outside the interval $0 \leq x \leq s$, because the minima of V are finite. As a consequence, it is convenient to replace the above interval with $x_1 \leq x \leq x_2$, where the two boundaries may be brought (independently from each other) to $-\infty$ or $+\infty$, respectively. Letting

$$\lambda = \frac{2\,m}{\hbar^2}\,E\,, \qquad r_l(x) = -\frac{2\,m}{\hbar^2}\,V\,, \tag{13.8}$$

the Schrödinger equation (11.28) becomes

$$w'' + r_l\, w + \lambda\, w = 0\,. \tag{13.9}$$

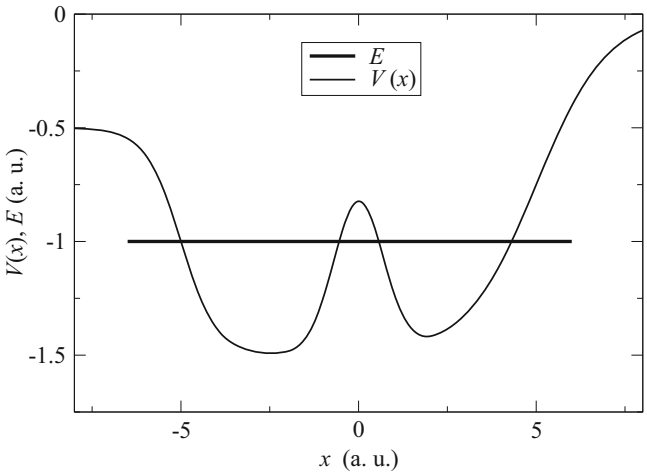


Fig. 13.1 Form of the potential energy that gives rise to localized states (Sect. 13.3). Only one state E is shown

The integer index $l = 0, 1, \ldots$ is attached to $r(x)$ for convenience. In fact, in typical applications the form of the potential energy in (13.8) may be prescribed by a previous separation procedure in which the index is involved. As will be shown after the analysis of the solution procedure, index l may eventually be disposed of. For the time being, the solutions of (13.9) must be considered as dependent on the eigenvalue λ and on the index l, namely, $w = w_{\lambda l}(x)$. Also, for a given pair λ, l the eigenfunctions are nondegenerate due to the normalization condition. As a consequence, two eigenfunctions belonging to the same pair differ by a multiplicative constant at most.

13.3.1 Factorization Scheme

A possible method for solving (13.9) is expanding w into a power series, replacing the series into (13.9), collecting the terms of equal power, and letting their coefficients vanish. This provides a recurrence relation involving the coefficients of the series; then, the latter is suitably truncated to obtain a square-integrable solution. Another method for solving (13.9), that greatly reduces the calculation effort and brings about a procedure of supreme elegance, is the operator's factorization. The conditions that make the factorization possible are illustrated in [67] and are briefly reported here. In essence they amount to finding a function $g_l(x)$ and a parameter L_l such that (13.9) may be recast as

$$\left(g_{l+1} + \frac{\mathrm{d}}{\mathrm{d}x}\right)\left(g_{l+1} - \frac{\mathrm{d}}{\mathrm{d}x}\right) w_{\lambda l} = (\lambda - L_{l+1})\, w_{\lambda l}\,, \tag{13.10}$$

$$\left(g_l - \frac{\mathrm{d}}{\mathrm{d}x}\right)\left(g_l + \frac{\mathrm{d}}{\mathrm{d}x}\right) w_{\lambda l} = (\lambda - L_l)\, w_{\lambda l}\,. \tag{13.11}$$

Note that both (13.10) and (13.11) must be identical to (13.9). An additional constraint is that, for a given integer n, it must be $L_{n+1} > L_{l+1}$, $l = 0, 1, \ldots, n-1$. To proceed, the boundary conditions $w_{\lambda l}(x_1) = w_{\lambda l}(x_2) = 0$ will be assumed. If one or both boundaries are at infinity, the condition $\int |w_{\lambda l}|^2\, \mathrm{d}x < \infty$ will also be assumed. Now, imposing that (13.10) is identical to (13.9) yields $g_{l+1}^2 + g_{l+1}' + L_{l+1} = -r_l$ whence, letting $l \leftarrow l - 1$,

$$g_l^2 + g_l' + L_l = -r_{l-1}\,. \tag{13.12}$$

Similarly, imposing that (13.11) is identical to (13.9) leads to

$$g_l^2 - g_l' + L_l = -r_l\,. \tag{13.13}$$

Adding (13.13) to (13.12) and subtracting (13.13) from (13.12) yield, respectively,

$$g_l^2 + L_l = -\frac{1}{2}\,(r_{l-1} + r_l)\,, \qquad g_l' = -\frac{1}{2}\,(r_{l-1} - r_l)\,. \tag{13.14}$$

Differentiating the first relation of (13.14) with respect to x and replacing g_l' from the second one provides

$$g_l = \frac{1}{2}\,\frac{r_{l-1}' + r_l'}{-2\,g_l'} = \frac{1}{2}\,\frac{r_{l-1}' + r_l'}{r_{l-1} - r_l}\,. \tag{13.15}$$

Finally, replacing (13.15) into the first relation of (13.14),

$$L_l = -\frac{1}{2}\,(r_{l-1} + r_l) - \frac{1}{4}\left(\frac{r_{l-1}' + r_l'}{r_{l-1} - r_l}\right)^2\,. \tag{13.16}$$

In conclusion, the factorization is possible if L_l given by (13.16) is independent of x. In this case, g_l is given by (13.15). As r_l is real, both L_l and g_l are real as well.

13.3.2 First-Order Operators

If the factorization (13.10,13.11) succeeds, it is useful to define the first-order, real operators

$$\mathscr{A}_l^+ = g_l + \frac{\mathrm{d}}{\mathrm{d}x}\,, \qquad \mathscr{A}_l^- = g_l - \frac{\mathrm{d}}{\mathrm{d}x}\,, \tag{13.17}$$

so that (13.10,13.11) are rewritten as

$$\mathscr{A}_{l+1}^+ \mathscr{A}_{l+1}^- w_{\lambda l} = (\lambda - L_{l+1})\, w_{\lambda l}\,, \qquad \mathscr{A}_l^- \mathscr{A}_l^+ w_{\lambda l} = (\lambda - L_l)\, w_{\lambda l}\,. \tag{13.18}$$

The two operators (13.17) are mutually adjoint. In fact, for any pair of functions f_1, f_2 fulfilling the same boundary conditions as $w_{\lambda l}$ one finds

$$\int_{x_1}^{x_2} f_1^*\, \mathscr{A}_l^+ f_2\, \mathrm{d}x = \left[f_1^*\, f_2\right]_{x_1}^{x_2} + \int_{x_1}^{x_2} (\mathscr{A}_l^- f_1)^*\, f_2\, \mathrm{d}x\,, \tag{13.19}$$

where the integrated part vanishes due to the boundary conditions. From the above result one finds a property of the eigenvalue λ. In fact, multiplying (13.18) by $w_{\lambda l}^*$ and integrating, one finds

$$\int_{x_1}^{x_2} |\mathscr{A}_{l+1}^- w_{\lambda l}|^2\, \mathrm{d}x = (\lambda - L_{l+1}) \int_{x_1}^{x_2} |w_{\lambda l}|^2\, \mathrm{d}x\,, \tag{13.20}$$

where (13.19) has been used after letting $f_1 = w_{\lambda l}, f_2 = \mathscr{A}_{l+1}^- w_{\lambda l}$. From (13.20) it follows that if $w_{\lambda l}$ is an eigenfunction, that is, if $w_{\lambda l}$ does not vanish identically, then the case $\lambda < L_{l+1}$ is impossible. In fact, the integral at the right-hand side of (13.20) is strictly positive, while that at the left-hand side is nonnegative. There remain two possibilities, namely:

1. $\mathscr{A}_{l+1}^- w_{\lambda l}$ does not vanish identically, whence both integrals of (13.20) are strictly positive. It follows that $\lambda > L_{l+1}$. Also, as will be shown later, $\mathscr{A}_{l+1}^- w_{\lambda l}$ is another eigenfunction of (13.9). The opposite is also true, namely, $\lambda > L_{l+1}$ implies that $\mathscr{A}_{l+1}^- w_{\lambda l}$ does not vanish identically.
2. $\mathscr{A}_{l+1}^- w_{\lambda l}$ vanishes identically, whence $\lambda = L_{l+1}$. The opposite is also true, namely, $\lambda = L_{l+1}$ implies that $\mathscr{A}_{l+1}^- w_{\lambda l}$ vanishes identically. In this case $\mathscr{A}_{l+1}^- w_{\lambda l}$ is not an eigenfunction of (13.9).

The discussion above allows one to identify the eigenvalue λ. In fact, there must be a value of the index l, say $l = n$, such that the equality $\lambda = L_{n+1}$ holds. It follows that the eigenvalue is identified by the index n, $\lambda = \lambda_n$.

As mentioned before the condition $L_{n+1} > L_{l+1}$, $l = 0, 1, \ldots, n-1$ holds. As a consequence, the eigenfunction corresponding to the pair λ_n, l may be indicated with w_{nl} instead of $w_{\lambda l}$. In particular, the eigenfunction corresponding to $l = n$ is w_{nn}. As shown in case 2 above, such an eigenfunction corresponds to the equality $\lambda_n = L_{n+1}$ which, in turn, implies the condition $\mathscr{A}_{n+1}^- w_{nn} = 0$. Remembering the second relation of (13.17), such a condition yields the first-order equation $(g_{n+1} - \mathrm{d}/\mathrm{d}x)\, w_{nn} = 0$, whose solution is real and reads

$$w_{nn} = c_{nn} \exp\left[\int_{x_1}^{x} g_{n+1}(\xi)\,\mathrm{d}\xi\right], \quad \frac{1}{c_{nn}^2} = \int_{x_1}^{x_2} \exp\left[\int_{x_1}^{x} 2\, g_{n+1}(\xi)\,\mathrm{d}\xi\right]\mathrm{d}x, \tag{13.21}$$

with $c_{nn} = \sqrt{c_{nn}^2} > 0$.

13.3.3 The Eigenfunctions Corresponding to $l < n$

The result given in (13.21) shows that if the factorization is achieved, the eigenfunction corresponding to $l = n$ is found by solving a first-order equation. It remains to determine the eigenfunctions corresponding to $l = 0, 1, \ldots, n-1$. For this, left multiplying the first relation in (13.18) by $\mathscr{A}_{l+1}^-$, letting $l \leftarrow l+1$ in the second relation of (13.18), and remembering that $\lambda = L_{n+1}$, one finds

$$\mathscr{A}_{l+1}^- \mathscr{A}_{l+1}^+ \mathscr{A}_{l+1}^- w_{nl} = (L_{n+1} - L_{l+1})\, \mathscr{A}_{l+1}^- w_{nl}\,, \tag{13.22}$$

$$\mathscr{A}_{l+1}^- \mathscr{A}_{l+1}^+ w_{\lambda,l+1} = (L_{n+1} - L_{l+1})\, w_{n,l+1}\,. \tag{13.23}$$

The above shows that both $w_{n,l+1}$ and $\mathscr{A}_{l+1}^{-} w_{nl}$ are eigenfunctions of the operator $\mathscr{A}_{l+1}^{-} \mathscr{A}_{l+1}^{+}$ belonging to the same eigenvalue. As the eigenfunctions are nondegenerate, it must be

$$\mathscr{A}_{l+1}^{-} w_{nl} = \text{const} \times w_{n,l+1} \,, \tag{13.24}$$

where the constant may be determined by imposing the normalization condition. The result shows that if w_{n0} is known, one may calculate a sequence of eigenfunctions belonging to $\lambda_n = L_{n+1}$ (apart from the normalization constant) by successive applications of first-order operators: $\mathscr{A}_1^{-} w_{n0} = \text{const} \times w_{n1}$, $\mathscr{A}_2^{-} w_{n1} = \text{const} \times w_{n2}, \ldots$. The process stops for $l = n$ because, as shown earlier, $\mathscr{A}_{n+1}^{-} w_{nn} = 0$ is not an eigenfunction anymore, so any further application of the operator beyond $l = n$ provides a sequence of zeros. In a similar manner, left multiplying the second relation in (13.18) by $\mathscr{A}_l^{+}$ and letting $l \leftarrow l - 1$ in the first relation of (13.18), one finds

$$\mathscr{A}_l^{+} \mathscr{A}_l^{-} w_{n,l-1} = (L_{n+1} - L_l)\, w_{n,l-1} \,, \tag{13.25}$$

$$\mathscr{A}_l^{+} \mathscr{A}_l^{-} \mathscr{A}_l^{+} w_{nl} = (L_{n+1} - L_l)\, \mathscr{A}_l^{+} w_{nl} \,. \tag{13.26}$$

From the above one finds that both $w_{n,l-1}$ and $\mathscr{A}_l^{+} w_{nl}$ are eigenfunctions of the operator $\mathscr{A}_l^{+} \mathscr{A}_l^{-}$ belonging to the same eigenvalue, whence

$$\mathscr{A}_l^{+} w_{nl} = \text{const} \times w_{n,l-1} \,. \tag{13.27}$$

The result shows that if w_{nn} is known, one may calculate a sequence of eigenfunctions belonging to $\lambda_n = L_{n+1}$ (apart from the normalization constant) by successive applications of first-order operators: $\mathscr{A}_n^{+} w_{nn} = \text{const} \times w_{n,n-1}$, $\mathscr{A}_{n-1}^{+} w_{n,n-1} = \text{const} \times w_{n,n-2}, \ldots$. The process stops for $l = 0$ which, by hypothesis, is the minimum of l. The derivation also shows that since w_{nn} and the operators are real, all the eigenfunctions found using the factorization method are real as well.

13.3.4 Normalization

The results of this section may be summarized as follows: (13.24) shows that the application of the first-order operator $\mathscr{A}_{l+1}^{-}$ to an eigenfunction of indices n, l provides an eigenfunction of indices $n, l + 1$. Similarly, (13.27) shows that the application of the first-order operator $\mathscr{A}_l^{+}$ to an eigenfunction of indices n, l provides an eigenfunction of indices $n, l - 1$. These results may be described as a process of going up or down along a ladder characterized by an index $n \geq 0$, whose steps are numbered by a second index $l = 0, 1, \ldots, n$. It follows that by applying two suitably chosen operators one may go up and down (or down and up) one step in the ladder and return to the same eigenfunction apart from a multiplicative constant.

This is indeed true, as shown by (13.18), that also indicate that the multiplicative constant to be used at the end of the two steps starting from w_{nl} is $L_{n+1} - L_l$ when the operators' index is l. It follows that the constants in (13.24,13.27) must be chosen as $\sqrt{L_{n+1} - L_{l+1}}$ and $\sqrt{L_{n+1} - L_l}$, respectively. This provides a method for achieving the normalization of the eigenfunctions, starting from an eigenfunction w_{nn} normalized to unity as in (13.21). For this one defines the auxiliary, mutually adjoint operators

$$\mathscr{B}_{nl}^{+} = \frac{\mathscr{A}_l^{+}}{\sqrt{L_{n+1} - L_l}}\,, \qquad \mathscr{B}_{nl}^{-} = \frac{\mathscr{A}_l^{-}}{\sqrt{L_{n+1} - L_l}}\,, \tag{13.28}$$

so that (13.18) becomes

$$\mathscr{B}_{n,l+1}^{+}\mathscr{B}_{n,l+1}^{-} w_{nl} = w_{nl}\,, \qquad \mathscr{B}_{nl}^{-}\mathscr{B}_{nl}^{+} w_{nl} = w_{nl}\,. \tag{13.29}$$

Thanks to the auxiliary operators the multiplicative constant at the end of the two steps becomes unity. Remembering that the eigenfunctions and operators are real, multiplying both of (13.29) by w_{nl} and integrating yield

$$\int_{x_1}^{x_2} (\mathscr{B}_{n,l+1}^{-} w_{nl})^2 \,\mathrm{d}x = \int_{x_1}^{x_2} (\mathscr{B}_{nl}^{+} w_{nl})^2 = \int_{x_1}^{x_2} w_{nl}^2 \,\mathrm{d}x\,, \tag{13.30}$$

On the other hand, replacing the constant in (13.24), (13.27) with $\sqrt{L_{n+1} - L_{l+1}}$, $\sqrt{L_{n+1} - L_l}$, respectively, one derives

$$\mathscr{B}_{n,l+1}^{-} w_{nl} = w_{n,l+1}\,, \qquad \mathscr{B}_{nl}^{+} w_{nl} = w_{n,l-1}\,. \tag{13.31}$$

Comparing (13.31) with (13.30) shows that if one of the eigenfunctions of the ladder is normalized to unity, all the others have the same normalization. In particular, if w_{nn} is normalized to unity as in (13.21), then the whole ladder of normalized eigenfunction is found by repeatedly applying the same procedure:

$$\mathscr{B}_{nn}^{+} w_{nn} = w_{n,n-1}\,, \quad \mathscr{B}_{n,n-1}^{+} w_{n,n-1} = w_{n,n-2}\,, \quad \ldots\,, \quad \mathscr{B}_{n1}^{+} w_{n1} = w_{n0}\,. \tag{13.32}$$

13.4 Schrödinger Equation with a Periodic Coefficient

An important case of (11.28) occurs when the coefficient q is periodic [49], with a period that will be denoted with $2\,\omega$. The independent variable (not necessarily a Cartesian one) will be denoted with z:

$$w''(z) + q(z)\, w(z) = 0\,, \qquad q(z + 2\,\omega) = q(z)\,, \tag{13.33}$$

where primes indicate derivatives. Here the variable z is considered real; the theory, however, can be extended to the case of a complex variable. Let $u(z)$, $v(z)$ be fundamental solutions (Sect. 11.4), with $u(0) = 1$, $u'(0) = 0$, $v(0) = 0$, $v'(0) = 1$. As (13.33) holds for any z, it holds in particular for $z + 2\omega$. From the periodicity of q it follows

$$w''(z + 2\omega) + q(z)\, w(z + 2\omega) = 0\,, \tag{13.34}$$

namely, $w(z+2\omega)$ is also a solution. Similarly, $u(z+2\omega)$, $v(z+2\omega)$ are solutions. As the equation has only two independent solutions it must be

$$u(z+2\omega) = a_{11}\, u(z) + a_{12}\, v(z)\,, \qquad v(z+2\omega) = a_{21}\, u(z) + a_{22}\, v(z)\,, \tag{13.35}$$

with a_{ij} suitable constants. The values of the latter are readily related to those of u, v by letting $z = 0$ and using the initial conditions: $u(2\omega) = a_{11}$, $u'(2\omega) = a_{12}$, $v(2\omega) = a_{21}$, $v'(2\omega) = a_{22}$. As the Wronskian of u, v equals unity it follows $a_{11}a_{22} - a_{12}a_{21} = 1$. One now seeks a constant s such that

$$u(z + 2\omega) = s\, u(z)\,, \qquad v(z + 2\omega) = s\, v(z)\,. \tag{13.36}$$

This is equivalent to diagonalizing (13.35), namely, s must fulfill for any z the following relations:

$$(a_{11} - s)\, u(z) + a_{12}\, v(z) = 0\,, \qquad a_{21}\, u(z) + (a_{22} - s)\, v(z) = 0\,. \tag{13.37}$$

Equating to zero the determinant of the coefficients in (13.37) yields

$$s = \frac{a_0}{2} \pm \sqrt{\frac{a_0^2}{4} - 1}\,, \qquad a_0 = a_{11} + a_{22}\,. \tag{13.38}$$

If $a_0 = 2$ the two solutions of (13.38) are real and take the common value $s^- = s^+ = 1$. In this case, as shown by (13.36), the functions u, v are periodic with period 2ω. Similarly, if $a_0 = -2$ the two solutions of (13.38) are real and take the common value $s^- = s^+ = -1$. In this case, as shown by (13.36), the functions u, v are periodic with period 4ω, whereas their moduli $|u|$, $|v|$ are periodic with period 2ω. As the moduli $|u|$, $|v|$ do not diverge as further and further periods are added to z, the case $a_0 = \pm 2$ is stable. If $a_0 > 2$ the two solutions of (13.38) are real and range over the intervals $0 < s^- < 1$ and, respectively, $s^+ > 1$. In particular it is $s^- \to 0$ for $a_0 \to \infty$. If $a_0 < -2$ the two solutions of (13.38) are real and range over the intervals $s^+ < -1$ and, respectively, $-1 < s^- < 0$. In particular it is $s^- \to 0$ for $a_0 \to -\infty$. From the relations (13.36) it follows that the moduli of the solutions corresponding to s^+ diverge: in fact one has $u(z + n\,2\omega) = (s^+)^n\, u(z)$ and $v(z + n\,2\omega) = (s^+)^n\, v(z)$, so the case $|a_0| > 2$ is unstable. When $|a_0| < 2$ the

two solutions are complex, namely, $s^{\pm} = \exp(\pm \mathrm{i}\,\mu)$ with $\tan^2 \mu = 4/a_0^2 - 1$. As the modulus of the solutions is unity, the case $|a_0| < 2$ is stable.

The discussion about stability may seem incomplete because the possible cases depend on the value of $a_0 = a_{11} + a_{22} = u(2\,\omega) + v'(2\,\omega)$, which depends on the fundamental solutions that are yet to be found. On the other hand, the analysis of stability must eventually reduce to considering only the properties of the coefficient of (13.33). In fact, it can be shown that if $q(z)$ is positive for all values of z and the absolute value of $2\,\omega \int_0^{2\omega} q(z)\,\mathrm{d}z$ is not larger than 4, then $|a_0| < 2$ [92]. Multiplying both sides of the first relation in (13.36) by $\exp[-\alpha\,(z + 2\,\omega)]$, with α an undetermined constant, yields

$$\tilde{u}(z + 2\,\omega) = s\,\exp(-2\,\omega\alpha)\,\tilde{u}(z)\,, \qquad \tilde{u}(z) = \exp(-\alpha\,z)\,u(z)\,. \tag{13.39}$$

A similar treatment is applied to $v(z)$, to yield another auxiliary function $\tilde{v}$. Now one exploits the undetermined constant to impose that the auxiliary functions be periodic of period $2\,\omega$: for this one lets $s\exp(-2\,\omega\,\alpha) = 1$, whence

$$\alpha = \frac{\log s}{2\,\omega}\,. \tag{13.40}$$

The constant α defined by (13.40) is termed *characteristic exponent* or *Floquet exponent*. Three of the cases listed in the discussion above about stability, namely, $s^{\pm} = \exp(\pm \mathrm{i}\,\mu)$, $s^{\pm} = 1$, and $s^{\pm} = -1$ lead now to the single expression $\alpha^{\pm} = \pm \mathrm{i}\,\mu/(2\,\omega)$, with $0 \le \mu \le \pi$, showing that α is purely imaginary. The cases $0 < s^- < 1$ and $s^+ > 1$ lead to real values of α (negative and positive, respectively) and, finally, the cases $s^+ < -1$ and $-1 < s^- < 0$ lead to complex values of α. Considering the stable cases only, one transforms (13.33) by replacing $w(z)$ with, e.g., $\tilde{u}(z)\,\exp[\pm \mathrm{i}\,\mu\,z/(2\,\omega)]$ to find

$$\tilde{u}''(z) \pm 2\,\mathrm{i}\,\frac{\mu}{2\,\omega}\,\tilde{u}'(z) + \left[q(z) - \frac{\mu^2}{4\,\omega^2}\right]\tilde{u}(z) = 0\,. \tag{13.41}$$

The coefficients and the unknown function of (13.41) are periodic functions of period $2\,\omega$. As a consequence it suffices to solve the equation within the single period, say, $0 \le z \le 2\,\omega$. An example is given in Sect. 17.9.4; a different approach leading to the generalization to three dimensions is shown in Sect. 17.6.

13.5 Schrödinger Equation for a Central Force

In the investigation about the properties of atoms it is important to analyze the dynamics of particle subjected to a central force in the case where the motion is limited. The treatment based on the concepts of Classical Mechanics is given in Sects. 3.4 (where the general properties are illustrated), 3.7 (for the two-particle

interaction), 3.8 (for a Coulomb field in the repulsive case), and 3.13.6 (for a Coulomb field in the attractive case). To proceed one considers a particle of mass[3] m_0 acted upon by a force deriving from a potential energy of the central type, $V = V(r)$, and expresses the time-independent Schrödinger equation $-\hbar^2/(2\,m_0)\,\nabla^2 w + V(r)\,w = E\,w$ in spherical coordinates r, ϑ, φ (Sect. B.1). Remembering the transformation (B.25) of the ∇^2 operator one obtains

$$\frac{1}{r}\,\frac{\partial^2(r\,w)}{\partial r^2} + \frac{1}{r^2}\,\hat{\Omega} w + \frac{2\,m_0}{\hbar^2}\,[E - V(r)]\,w = 0\,, \tag{13.42}$$

where operator $\hat{\Omega}$ is defined as

$$\hat{\Omega} = \frac{1}{\sin^2\vartheta}\left[\sin\vartheta\,\frac{\partial}{\partial\vartheta}\left(\sin\vartheta\,\frac{\partial}{\partial\vartheta}\right) + \frac{\partial^2}{\partial\varphi^2}\right]. \tag{13.43}$$

The r coordinate is separated by letting $w = \varrho(r)\,Y(\vartheta,\varphi)$ in (13.42) and dividing both sides by w/r^2:

$$r^2\left[\frac{1}{r\,\varrho}\,\frac{\mathrm{d}^2(r\,\varrho)}{\mathrm{d}r^2} + \frac{2\,m_0}{\hbar^2}\,(E - V)\right] = -\frac{1}{Y}\,\hat{\Omega} Y\,. \tag{13.44}$$

Each side of (13.44) must equal the same dimensionless constant, say, c, whence the original Schrödinger equation separates into the pair

$$\hat{\Omega} Y = -c\,Y\,, \quad \left[-\frac{\hbar^2}{2\,m_0}\,\frac{\mathrm{d}^2}{\mathrm{d}r^2} + V_e(r)\right] r\,\varrho = E\,r\,\varrho\,, \quad V_e = V + \frac{c\,\hbar^2}{2\,m_0\,r^2}\,. \tag{13.45}$$

The first equation in (13.45), called *angular equation*, does not depend on any parameter specific to the problem in hand. As a consequence, its eigenvalues c and eigenfunctions Y can be calculated once and for all. Being the equation's domain two dimensional, the eigenfunctions Y are expected to depend onto two indices, say, l, m. After the angular equation is solved, inserting each eigenvalue c into the second equation of (13.45), called *radial equation*, provides the eigenvalues and eigenfunctions of the latter. For the radial equation, the solution depends on the form of $V(r)$. It is also worth noting the similarity of V_e with its classical counterpart (3.5), that reads

$$V_e = V + \frac{M^2}{2\,m_0\,r^2}\,, \qquad M^2 = \text{const}\,. \tag{13.46}$$

[3]To avoid confusion with the azimuthal quantum number m, the particle's mass is indicated with m_0 in the sections dealing with the angular momentum in the quantum case.

To tackle the solution of the angular equation $\hat{\Omega} Y = -c\, Y$ one associates an operator $\mathscr{L}_x$, $\mathscr{L}_y$, $\mathscr{L}_z$ with each component of the classical angular momentum $\mathbf{M} = \mathbf{r} \wedge \mathbf{p}$, and another operator $\mathscr{L}^2$ to its square modulus M^2. The procedure, illustrated in Sect. 13.6.1, shows that the three operators $\mathscr{L}_x$, $\mathscr{L}_y$, $\mathscr{L}_z$ do not commute with each other, whereas $\mathscr{L}^2$ commutes with each of them. Also, it is found that $\mathscr{L}^2$ is proportional to $\hat{\Omega}$, specifically, $\mathscr{L}^2 = -\hbar^2\, \hat{\Omega}$. In conclusion, the Schrödinger equation in the case of a central force reads

$$\mathscr{H} w = E\, w\,, \qquad \mathscr{H} = -\frac{\hbar^2}{2\, m_0\, r}\, \frac{\partial^2}{\partial r^2}\, r + \frac{\mathscr{L}^2}{2\, m_0\, r^2} + V(r)\,. \tag{13.47}$$

As the r coordinate does not appear in $\mathscr{L}^2$, the latter commutes with $\mathscr{H}$; moreover, $\mathscr{L}^2$ does not depend on time. As a consequence, its expectation value is a constant of motion (Sect. 10.7). Similarly, $\mathscr{L}_z$ commutes with $\mathscr{L}^2$ and does not contain r or t, so it commutes with $\mathscr{H}$ as well and its expectation value is also a constant of motion. As $\mathscr{H}$, $\mathscr{L}^2$, and $\mathscr{L}_z$ commute with each other, they have a common set of eigenfunctions.

13.5.1 Angular Part of the Equation

The conservation of the expectation values of $\mathscr{L}^2$ and $\mathscr{L}_z$ is the counterpart of the classical result of the conservation of M^2 and M_z (Sect. 2.8). In contrast, the expectation values of $\mathscr{L}_x$ and $\mathscr{L}_y$ are not constants of motion. To determine the eigenfunctions w of $\mathscr{H}$, $\mathscr{L}^2$ and $\mathscr{L}_z$ it is convenient to solve the eigenvalue equation for $\mathscr{L}_z$ first:

$$\mathscr{L}_z w = L_z\, w\,, \qquad -\mathrm{i}\,\hbar\, \frac{\partial w}{\partial \varphi} = L_z\, w\,, \qquad w = v(r, \vartheta)\, \exp(\mathrm{i}\, L_z\, \varphi/\hbar)\,, \tag{13.48}$$

with v yet undetermined. For an arbitrary value of L_z, the exponential part of w is a multi-valued function of φ. This is not acceptable because w should not vary when φ is changed by integer multiples of 2π. A single-valued function is achieved by letting $L_z/\hbar = m$, with m an integer. In conclusion, the eigenvalues and eigenfunctions of $\mathscr{L}_z$ are

$$L_z = m\,\hbar\,, \qquad w = v(r, \vartheta)\, \exp(\mathrm{i}\, m\, \varphi)\,. \tag{13.49}$$

Combining (13.49) with (13.43) provides

$$\hat{\Omega} w = \frac{\exp(\mathrm{i}\, m\, \varphi)}{\sin^2 \vartheta} \left[\sin\vartheta\, \frac{\partial}{\partial \vartheta} \left(\sin\vartheta\, \frac{\partial}{\partial \vartheta} \right) - m^2 \right] v(r, \vartheta)\,. \tag{13.50}$$

This result shows that in (13.47) the factor $\exp(\mathrm{i}\, m\, \varphi)$ cancels out, so that $\hat{\Omega}$ actually involves the angular coordinate ϑ only. This suggests to seek the function $v(r, \vartheta)$ by separation. Remembering that w was originally separated as $w = \varrho(r)\, Y(\vartheta, \varphi)$ one finds

$$v(r, \vartheta) = \varrho(r)\, P(\vartheta)\,, \qquad Y(\vartheta, \varphi) = P(\vartheta)\, \exp(\mathrm{i}\, m\, \varphi)\,. \tag{13.51}$$

As the separation transforms $\hat{\Omega} Y = -c\, Y$ into $\hat{\Omega} P = -c\, P$, the equation to be solved for a given integer m reduces to

$$\frac{1}{\sin^2 \vartheta} \left[\sin \vartheta \, \frac{\mathrm{d}}{\mathrm{d}\vartheta} \left(\sin \vartheta \, \frac{\mathrm{d}}{\mathrm{d}\vartheta} \right) - m^2 \right] P = -c\, P\,. \tag{13.52}$$

From $\mathscr{L}^2 = -\hbar^2\, \hat{\Omega}$, it follows that the eigenvalue of $\mathscr{L}^2$ is $\lambda = \hbar^2\, c$. The eigenvalues c of (13.52) are found by the factorization method described in Sect. 13.3.1; they have the form $c = l\,(l+1)$, with l a nonnegative integer, called *orbital* (or *total*) *angular momentum quantum number*. For a given l, the allowed values of m, called *azimuthal* (or *magnetic*) *quantum number*, are the $2\,l+1$ integers $-l, \ldots, 0, \ldots, l$. The details of the eigenvalue calculation are given in Sect. 13.6.2.

The factorization method provides also the eigenfunctions of $\mathscr{L}^2$ and $\mathscr{L}_z$, that are called *spherical harmonics* and, as expected, depend on the two indices l, m. The details of the calculation of the eigenfunctions Y_l^m are given in Sect. 13.52. The lowest-order spherical harmonics are shown in Table 13.1. As the eigenfunctions fulfill the equations

$$\mathscr{L}^2 Y_l^m = \hbar^2\, l\,(l+1)\, Y_l^m\,, \qquad \mathscr{L}_z Y_l^m = \hbar\, m\, Y_l^m\,, \tag{13.53}$$

the only possible results of a measurement of M^2 are $\hbar^2\, l\,(l+1)$, with $l = 0, 1, 2, \ldots$ and, for a given l, the only possible results of a measurement of M_z are $\hbar\, m$, with $m = -l, \ldots, -1, 0, 1, \ldots, l$. It follows that the only possible results of a measurement of M are $\hbar\, \sqrt{l\,(l+1)}$. For any $l > 0$ it is $\max(|M_z|) = \hbar\, l < \hbar\, \sqrt{l\,(l+1)} = M$; as a consequence, for $l > 0$ the angular momentum $\mathbf{M}$ lies on a cone centered on the z axis. The half amplitude $\alpha = \arccos[m/\sqrt{l(l+1)}]$ of such a cone is strictly positive, showing that the other two components M_x, M_y cannot vanish together when $M \neq 0$. A geometrical construction describing the relation between $\mathbf{M}$ and M_z is given in Fig. 13.2. The groups of states corresponding to the orbital quantum

Table 13.1 The lowest-order spherical harmonics

Y_l^m	Form of the function
Y_0^0	$1/\sqrt{4\,\pi}$
Y_1^{-1}	$\sqrt{3/(8\,\pi)}\, \sin\vartheta\, \exp(-\mathrm{i}\,\varphi)$
Y_1^0	$-\sqrt{3/(4\,\pi)}\, \cos\vartheta$
Y_1^1	$-\sqrt{3/(8\,\pi)}\, \sin\vartheta\, \exp(\mathrm{i}\,\varphi)$

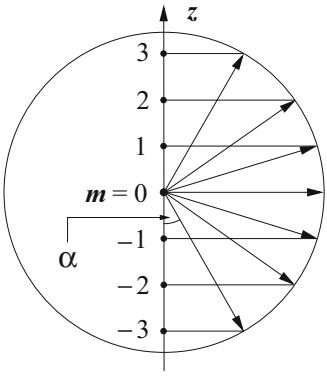


Fig. 13.2 Geometrical construction showing the relation between **M** and M_z. The $l = 3$ is case considered, whence one finds $m = -3, \ldots, 0, \ldots, 3$ and $\sqrt{l\,(l+1)} \simeq 3.46$

Table 13.2 Symbols and names for the states corresponding to $l = 0, 1, 2, 3$

l	m	Symbol	Name
0	0	s	Sharp
1	−1, 0, 1	p	Principal
2	−2, −1, 0, 1, 2	d	Diffuse
3	−3, −2, −1, 0, 1, 2, 3	f	Fundamental

numbers $l = 0, 1, 2, 3$ are denoted with special symbols and names that originate from spectroscopy [13, Chap. 15] and are listed in Table 13.2. For $l \geq 4$ the symbols continue in alphabetical order ("g," "h," ...), while no special names are used.

13.5.2 Radial Part of the Equation in the Coulomb Case

To solve the radial part of the Schrödinger equation (second and third relation in (13.45)) one uses the eigenvalue $c = l\,(l+1)$ to find

$$\left[-\frac{\hbar^2}{2\,m_0}\,\frac{\mathrm{d}^2}{\mathrm{d}r^2} + V_e(r)\right] r\,\varrho(r) = E\,r\,\varrho(r)\,, \qquad V_e = V + \frac{\hbar^2\,l\,(l+1)}{2\,m_0\,r^2}\,. \tag{13.54}$$

As anticipated above, the solution of (13.54) depends on the form of $V(r)$. Of particular interest is the Coulomb potential (3.31) that is considered here in the attractive case

$$V(r) = -\frac{Z\,q^2}{4\,\pi\,\varepsilon_0\,r}\,, \tag{13.55}$$

with ε_0 the vacuum permittivity, $q > 0$ the elementary electric charge, and $Z\,q$ the charge whence the central force originates. This form of the potential energy is typical of the case of an electron belonging to a hydrogen or hydrogen-like atom. As usual, the arbitrary constant inherent in the definition of the potential energy is

such that $\lim_{r\to\infty} V = 0$. As a consequence, the electron is bound if $E < 0$ (in other terms, according to the definition given in Sect. 3.13.6, the classical motion is limited). The eigenvalues E of (13.54,13.55) are found by the factorization method described in Sect. 13.3.1; they have the form

$$E = E_n = -\frac{m_0}{2\hbar^2}\left(\frac{Z q^2}{4\pi\,\varepsilon_0}\right)^2 \frac{1}{n^2}\,, \tag{13.56}$$

where n is an integer, called *principal quantum number*, fulfilling the relation $n \geq l+1$. The details of the eigenvalue calculation are given in Sect. 13.6.2. As $l \geq 0$, the minimum value of n is 1. For a given n the possible values of the orbital quantum number are $l = 0, 1, \ldots, n-1$; also, as found earlier, for each l the possible values of the azimuthal quantum number are $m = -l, \ldots, 0, \ldots, l$. It follows that for a given n the number of different pairs l, m is

$$\sum_{l=0}^{n-1}(2\,l+1) = n^2\,, \tag{13.57}$$

namely, each eigenvalue E_n of the energy corresponds to n^2 possible combinations[4] of the eigenvalues of M and M_z. As for the angular part of the equation, the factorization method provides also the eigenfunctions of (13.54); the details are given in Sect. (13.6.5).

13.6 Complements

13.6.1 Operators Associated with Angular Momentum

Consider the classical angular momentum $\mathbf{M} = \mathbf{r}\wedge\mathbf{p}$ (Sect. 2.6), whose components in rectangular coordinates are given by (2.38), namely,

$$M_x = y\,p_z - z\,p_y\,, \qquad M_y = z\,p_x - x\,p_z\,, \qquad M_z = x\,p_y - y\,p_x\,. \tag{13.58}$$

The operators corresponding to (13.58) are

$$\begin{cases} \mathscr{L}_x = -\mathrm{i}\,\hbar\,(y\,\partial/\partial z - z\,\partial/\partial y) \\ \mathscr{L}_y = -\mathrm{i}\,\hbar\,(z\,\partial/\partial x - x\,\partial/\partial z) \\ \mathscr{L}_z = -\mathrm{i}\,\hbar\,(x\,\partial/\partial y - y\,\partial/\partial x) \end{cases} \tag{13.59}$$

[4] The actual degree of degeneracy of E_n is $2\,n^2$, where factor 2 is due to spin (Sect. 15.5.1).

It is easily found that $\mathscr{L}_x$, $\mathscr{L}_y$, $\mathscr{L}_z$ are Hermitean and fulfill the relations

$$\begin{cases} \mathscr{L}_x \mathscr{L}_y - \mathscr{L}_y \mathscr{L}_x = \mathrm{i}\hbar \mathscr{L}_z \\ \mathscr{L}_y \mathscr{L}_z - \mathscr{L}_z \mathscr{L}_y = \mathrm{i}\hbar \mathscr{L}_x \\ \mathscr{L}_z \mathscr{L}_x - \mathscr{L}_x \mathscr{L}_z = \mathrm{i}\hbar \mathscr{L}_y \end{cases} \tag{13.60}$$

namely, $\mathscr{L}_x$, $\mathscr{L}_y$, $\mathscr{L}_z$ do not commute with each other. Left multiplying the third relation in (13.60) by $\mathscr{L}_x$ and the second one by $\mathscr{L}_y$ provide, respectively,

$$\mathscr{L}_x^2 \mathscr{L}_z = \mathscr{L}_x \mathscr{L}_z \mathscr{L}_x - \mathrm{i}\hbar \mathscr{L}_x \mathscr{L}_y \,, \qquad \mathscr{L}_y^2 \mathscr{L}_z = \mathscr{L}_y \mathscr{L}_z \mathscr{L}_y + \mathrm{i}\hbar \mathscr{L}_y \mathscr{L}_x \,. \tag{13.61}$$

Similarly, right multiplying the third relation in (13.60) by $\mathscr{L}_x$ and the second one by $\mathscr{L}_y$,

$$\mathscr{L}_z \mathscr{L}_x^2 = \mathscr{L}_x \mathscr{L}_z \mathscr{L}_x + \mathrm{i}\hbar \mathscr{L}_y \mathscr{L}_x \,, \qquad \mathscr{L}_z \mathscr{L}_y^2 = \mathscr{L}_y \mathscr{L}_z \mathscr{L}_y - \mathrm{i}\hbar \mathscr{L}_x \mathscr{L}_y \,. \tag{13.62}$$

The operator associated with $M^2 = M_x^2 + M_y^2 + M_z^2$ is $\mathscr{L}^2 = \mathscr{L}_x^2 + \mathscr{L}_y^2 + \mathscr{L}_z^2$ whence, using (13.61,13.62),

$$\mathscr{L}^2 \mathscr{L}_z - \mathscr{L}_z \mathscr{L}^2 = (\mathscr{L}_x^2 + \mathscr{L}_y^2) \mathscr{L}_z - \mathscr{L}_z (\mathscr{L}_x^2 + \mathscr{L}_y^2) = 0 \,. \tag{13.63}$$

Similarly,

$$\mathscr{L}^2 \mathscr{L}_x - \mathscr{L}_x \mathscr{L}^2 = 0 \,, \qquad \mathscr{L}^2 \mathscr{L}_y - \mathscr{L}_y \mathscr{L}^2 = 0 \,. \tag{13.64}$$

In conclusion, the components $\mathscr{L}_x$, $\mathscr{L}_y$, $\mathscr{L}_z$ do not commute with each other, while the square modulus of the angular momentum commutes with any single component of it. To check whether $\mathscr{L}^2$ or any of the components $\mathscr{L}_x$, $\mathscr{L}_y$, $\mathscr{L}_z$ commute with the Hamiltonian operator of a central force, it is necessary to express all operators in spherical coordinates. To this purpose, using $\mathscr{L}_z$ by way of example, one finds

$$-\frac{\mathscr{L}_z}{\mathrm{i}\hbar} = x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} = r \sin\vartheta \cos\varphi \frac{\partial}{\partial y} - r \sin\vartheta \sin\varphi \frac{\partial}{\partial x} \,. \tag{13.65}$$

The partial derivatives $\partial/\partial x$ and $\partial/\partial y$ in terms of the spherical coordinates are extracted from (B.4); in particular, the first one reads $\partial/\partial x = \sin\vartheta \cos\varphi \, \partial/\partial r + (1/r) \cos\vartheta \cos\varphi \, \partial/\partial\vartheta - (1/r) (\sin\varphi/\sin\vartheta) \, \partial/\partial\varphi$, while the expression of $\partial/\partial y$ is obtained from that of $\partial/\partial x$ by replacing $\cos\varphi$ with $\sin\varphi$ and $\sin\varphi$ with $-\cos\varphi$. When such expressions of the partial derivatives are used within (13.65), several terms cancel out to finally yield the relation

$$\mathscr{L}_z = -\mathrm{i}\hbar \frac{\partial}{\partial\varphi} \tag{13.66}$$

which, consistently with the classical one, $M_z = p_\varphi$ (Sect. 3.4), shows that the operator associated with the z component of the angular momentum is conjugate to the generalized coordinate φ. The quantum relation can thus be derived directly from the classical one by letting $\mathscr{L}_z = \hat{p}_\varphi = -\mathrm{i}\,\hbar\,\partial/\partial\varphi$. As already noted in Sect. 2.8, the remaining components of M_x, M_y are not conjugate momenta. The expression of $\mathscr{L}_x$ in spherical coordinates reads

$$\mathscr{L}_x = \mathrm{i}\,\hbar \left(\sin\varphi\, \frac{\partial}{\partial\vartheta} + \frac{\cos\vartheta}{\sin\vartheta} \cos\varphi\, \frac{\partial}{\partial\varphi} \right) , \tag{13.67}$$

while that of $\mathscr{L}_y$ is obtained from (13.67) by replacing $\cos\varphi$ with $\sin\varphi$ and $\sin\varphi$ with $-\cos\varphi$. Combining the above findings, one calculates the expression of $\mathscr{L}^2 = \mathscr{L}_x^2 + \mathscr{L}_y^2 + \mathscr{L}_z^2$, that turns out to be

$$\mathscr{L}^2 = -\frac{\hbar^2}{\sin^2\vartheta} \left[\sin\vartheta\, \frac{\partial}{\partial\vartheta} \left(\sin\vartheta\, \frac{\partial}{\partial\vartheta} \right) + \frac{\partial^2}{\partial\varphi^2} \right] = -\hbar^2\, \hat{\Omega}\,. \tag{13.68}$$

13.6.2 Eigenvalues of the Angular Equation

The solution of the angular equation is found by the factorization method described in Sect. 13.3.1. Remembering that $\mathscr{L}_z$ and $\mathscr{L}^2$ commute, the whole eigenfunction $Y = P(\vartheta)\, \exp(\mathrm{i}\, m\, \varphi)$, introduced in Sect. 13.5 and common to both operators, will be used here. The following hold:

$$\mathscr{L}_z Y = L_z\, Y\,, \qquad \mathscr{L}^2 Y = \lambda\, Y\,, \tag{13.69}$$

with $L_z = m\,\hbar$, m an integer. Applying the operator $\mathscr{L}_x \pm \mathrm{i}\,\mathscr{L}_y$ to the first equation in (13.69) yields $\left(\mathscr{L}_x \pm \mathrm{i}\,\mathscr{L}_y\right)\, \mathscr{L}_z Y = m\,\hbar\, \left(\mathscr{L}_x \pm \mathrm{i}\,\mathscr{L}_y\right)\, Y$, where the upper (lower) signs hold together. Due to the commutation rules (13.60) the left-hand side of the above transforms into

$$\left(\mathscr{L}_z\mathscr{L}_x - \mathrm{i}\,\hbar\,\mathscr{L}_y\right) Y \pm \mathrm{i}\, \left(\mathscr{L}_z\mathscr{L}_y + \mathrm{i}\,\hbar\,\mathscr{L}_x\right) Y = \mathscr{L}_z \left(\mathscr{L}_x \pm \mathrm{i}\,\mathscr{L}_y\right) Y \mp \hbar\, \left(\mathscr{L}_x \pm \mathrm{i}\,\mathscr{L}_y\right) Y\,,$$

whence the first eigenvalue equation in (13.69) becomes

$$\mathscr{L}_z\, \left(\mathscr{L}_x \pm \mathrm{i}\,\mathscr{L}_y\right) Y = (m \pm 1)\,\hbar\, \left(\mathscr{L}_x \pm \mathrm{i}\,\mathscr{L}_y\right) Y\,. \tag{13.70}$$

Iterating the above reasoning shows that if Y is an eigenfunction of $\mathscr{L}_z$ belonging to the eigenvalue $m\,\hbar$, then $(\mathscr{L}_x + \mathrm{i}\,\mathscr{L}_y)Y$, $(\mathscr{L}_x + \mathrm{i}\,\mathscr{L}_y)^2 Y, \ldots$ are also eigenfunctions of $\mathscr{L}_z$ which belong, respectively, to $(m+1)\,\hbar$, $(m+2)\,\hbar, \ldots$, and so on. Similarly, $(\mathscr{L}_x - \mathrm{i}\,\mathscr{L}_y)Y$, $(\mathscr{L}_x - \mathrm{i}\,\mathscr{L}_y)^2 Y, \ldots$ are also eigenfunctions of $\mathscr{L}_z$ belonging, respectively, to $(m-1)\,\hbar$, $(m-2)\,\hbar, \ldots$, and so on. At the same time, due to the commutativity of $\mathscr{L}^2$ with $\mathscr{L}_x$ and $\mathscr{L}_y$, it is

$$\left(\mathscr{L}_x \pm \mathrm{i}\,\mathscr{L}_y\right) \mathscr{L}^2 Y = \mathscr{L}^2 \left(\mathscr{L}_x \pm \mathrm{i}\,\mathscr{L}_y\right) Y = \lambda \left(\mathscr{L}_x \pm \mathrm{i}\,\mathscr{L}_y\right) Y\,, \tag{13.71}$$

showing that $(\mathscr{L}_x \pm \mathrm{i}\,\mathscr{L}_y)Y$ is also an eigenfunction of $\mathscr{L}^2$, belonging to the same eigenvalue as Y. By induction, $(\mathscr{L}_x \pm \mathrm{i}\,\mathscr{L}_y)^2 Y, \ldots$ are also eigenfunctions of $\mathscr{L}^2$, belonging to the same eigenvalue as Y. To summarize, if $Y = P(\vartheta)\,\exp(\mathrm{i}\,m\,\varphi)$ is an eigenfunction common to operators $\mathscr{L}_z$ and $\mathscr{L}^2$, belonging to the eigenvalues $L_z = m\,\hbar$ and λ, respectively, then,

1. $(\mathscr{L}_x + \mathrm{i}\,\mathscr{L}_y)Y$ is another eigenfunction of $\mathscr{L}^2$ still belonging to λ, and is also an eigenfunction of $\mathscr{L}_z$ belonging to $(m+1)\,\hbar$. Similarly, $(\mathscr{L}_x + \mathrm{i}\,\mathscr{L}_y)^2 Y$ is still another eigenfunction of $\mathscr{L}^2$ belonging to λ, and is also an eigenfunction of $\mathscr{L}_z$ belonging to $(m+2)\,\hbar$, and so on.
2. $(\mathscr{L}_x - \mathrm{i}\,\mathscr{L}_y)Y$ is another eigenfunction of $\mathscr{L}^2$ still belonging to λ, and is also an eigenfunction of $\mathscr{L}_z$ belonging to $(m-1)\,\hbar$. Similarly, $(\mathscr{L}_x - \mathrm{i}\,\mathscr{L}_y)^2 Y$ is still another eigenfunction of $\mathscr{L}^2$ belonging to λ, and is also an eigenfunction of $\mathscr{L}_z$ belonging to $(m-2)\,\hbar$, and so on.

By this reasoning, starting from a given pair λ, Y it seems possible to construct as many degenerate eigenfunctions of $\mathscr{L}^2$ as we please. This, however, leads to unbounded eigenvalues of $\mathscr{L}_z$, which are not admissible as shown below. As a consequence, the procedure depicted here can be applied only a finite number of times. To demonstrate that the eigenvalues of $\mathscr{L}_z$ are bounded one starts from the relation $\mathscr{L}^2 = \mathscr{L}_x^2 + \mathscr{L}_y^2 + \mathscr{L}_z^2$ and from a given pair λ, Y. As Y is also an eigenfunction $\mathscr{L}_z$ belonging to, say, $m\,\hbar$, an application of $\mathscr{L}^2$ to Y followed by a left scalar multiplication by Y^* yields, thanks to (13.69),

$$\lambda - m^2\hbar^2 = \frac{\langle \mathscr{L}_x Y | \mathscr{L}_x Y\rangle + \langle \mathscr{L}_y Y | \mathscr{L}_y Y\rangle}{\langle Y | Y\rangle} \geq 0\,, \tag{13.72}$$

where the hermiticity of $\mathscr{L}_x$, $\mathscr{L}_y$ has been exploited. Inequality (13.72) provides the upper bound for $|m|$. To find the acceptable values of m one defines $m^+ = \max(m)$, and lets Y^+ be an eigenfunction of $\mathscr{L}^2$ and $\mathscr{L}_z$ belonging to λ and $m^+\,\hbar$, respectively. From (13.70) one obtains $\mathscr{L}_z\,(\mathscr{L}_x + \mathrm{i}\,\mathscr{L}_y)Y^+ = (m^+ + 1)\,\hbar\,(\mathscr{L}_x + \mathrm{i}\,\mathscr{L}_y)Y^+$ but, as the eigenvalue $(m^+ + 1)\hbar$ is not acceptable, it must be $(\mathscr{L}_x + \mathrm{i}\,\mathscr{L}_y)Y^+ = 0$. Similarly, letting $m^- = \min(m)$, and letting Y^- be an eigenfunction of $\mathscr{L}^2$ and $\mathscr{L}_z$ belonging to λ and $m^-\,\hbar$, respectively, it must be $(\mathscr{L}_x - \mathrm{i}\,\mathscr{L}_y)Y^- = 0$. Due to the commutation rules it is $(\mathscr{L}_x - \mathrm{i}\,\mathscr{L}_y)\,(\mathscr{L}_x + \mathrm{i}\,\mathscr{L}_y) = \mathscr{L}_x^2 + \mathscr{L}_y^2 - \hbar\,\mathscr{L}_z$, whence

$$\mathscr{L}^2 = (\mathscr{L}_x - \mathrm{i}\,\mathscr{L}_y)\,(\mathscr{L}_x + \mathrm{i}\,\mathscr{L}_y) + \hbar\,\mathscr{L}_z + \mathscr{L}_z^2\,. \tag{13.73}$$

Application of (13.73) to Y^+ and Y^- yields

$$\begin{cases} \mathscr{L}^2 Y^+ = \left(\mathscr{L}_z^2 + \hbar\,\mathscr{L}_z\right) Y^+ = \hbar^2\,m^+\,\left(m^+ + 1\right)\,Y^+ \\ \mathscr{L}^2 Y^- = \left(\mathscr{L}_z^2 - \hbar\,\mathscr{L}_z\right) Y^- = \hbar^2\,m^-\,(m^- - 1)\,Y^- \end{cases} \tag{13.74}$$

By construction, Y^+ and Y^- belong to the same eigenvalue of $\mathscr{L}^2$; as a consequence it must be $m^+\left(m^+ + 1\right) = m^-\,(m^- - 1)$. A possible integer solution of the above is $m^- = m^+ + 1$ which, however, is not acceptable because $m^+ = \max(m)$. The only acceptable solution left is $m^- = -m^+$. In conclusion, letting $l = m^+$ (so that $m^- = -l$) and using (13.74), the eigenvalues of $\mathscr{L}^2$ take the form

$$\lambda = \hbar^2\, l\,(l+1)\,, \tag{13.75}$$

with l a nonnegative integer. For a given l, the allowed values of m are the $2\,l + 1$ integers $-l, \ldots, 0, \ldots, l$.

13.6.3 Eigenfunctions of the Angular Equation

Due to the findings illustrated in Sect. 13.6.2, the eigenfunctions of $\hat{\Omega}$, whose form is $Y(\vartheta, \varphi) = P(\vartheta)\,\exp(\mathrm{i}\,m\,\varphi)$, depend on the two indices l, m and, for this reason, will be indicated with Y_l^m. In particular, the eigenfunction Y^+ introduced in Sect. 13.6.2, which belongs to $l = \max(m)$ and fulfills the equation $(\mathscr{L}_x + \mathrm{i}\,\mathscr{L}_y)Y = 0$, will be indicated with Y_l^l. Similarly, as P depends on l and may depend on m as well, it will be indicated with P_l^m. The eigenfunction Y_l^l is readily found by solving the first-order equation $(\mathscr{L}_x + \mathrm{i}\,\mathscr{L}_y)Y_l^l = 0$, where operator $\mathscr{L}_x$ is expressed in terms of φ, ϑ through (13.67), and $\mathscr{L}_y$ is obtained from (13.67) by replacing $\cos\varphi$ with $\sin\varphi$ and $\sin\varphi$ with $-\cos\varphi$. After eliminating the factor $\hbar\,\exp[\mathrm{i}\,(l+1)\,\varphi]$ one finds a differential equation for P_l^l, that reads

$$\frac{\mathrm{d}P_l^l}{\mathrm{d}\vartheta} - l\,\frac{\cos\vartheta}{\sin\vartheta}\,P_l^l = 0\,, \qquad \frac{1}{P_l^l}\,\frac{\mathrm{d}P_l^l}{\mathrm{d}\vartheta} = \frac{l}{\sin\vartheta}\,\frac{\mathrm{d}\sin\vartheta}{\mathrm{d}\vartheta}\,. \tag{13.76}$$

In conclusion it is found, with a an arbitrary constant,

$$P_l^l = a\,(\sin\vartheta)^l\,, \qquad Y_l^l = a\,\exp(\mathrm{i}\,l\,\varphi)\,(\sin\vartheta)^l\,. \tag{13.77}$$

Then, remembering the discussion of Sect. 13.6.2, the remaining $2\,l$ eigenfunctions $Y_l^{l-1}, \ldots, Y_l^0, \ldots, Y_l^{-l}$ are found by successive applications of

$$\mathscr{L}_x - \mathrm{i}\,\mathscr{L}_y = -\hbar\,\exp(-\mathrm{i}\,\varphi)\left(\frac{\partial}{\partial\vartheta} - \mathrm{i}\,\frac{\cos\vartheta}{\sin\vartheta}\,\frac{\partial}{\partial\varphi}\right), \tag{13.78}$$

with the help of the auxiliary relations $-\mathrm{i}\,\partial Y_l^l/\partial\varphi = l\,Y_l^l$ and

$$\left(\frac{\partial}{\partial\vartheta} + l\,\frac{\cos\vartheta}{\sin\vartheta}\right) Y_l^l = \frac{\partial[(\sin\vartheta)^l\,Y_l^l]/\partial\vartheta}{(\sin\vartheta)^l} = a\,\frac{\exp(\mathrm{i}\,l\,\varphi)}{(\sin\vartheta)^l}\,\frac{\mathrm{d}}{\mathrm{d}\vartheta}\,(\sin\vartheta)^{2l}\,. \tag{13.79}$$

In fact, combining $\mathscr{L}_z\left[\left(\mathscr{L}_x - \mathrm{i}\,\mathscr{L}_y\right) Y_l^m\right] = (m-1)\,\hbar\left[\left(\mathscr{L}_x - \mathrm{i}\,\mathscr{L}_y\right) Y_l^m\right]$ with $\mathscr{L}_z\, Y_l^m = m\hbar\, Y_l^m$ provides the recursive relation $Y_l^{m-1} = \left(\mathscr{L}_x - \mathrm{i}\,\mathscr{L}_y\right) Y_l^m$. In particular, letting $m = l, m = l-1, \ldots$ yields

$$Y_l^{l-1} = \left(\mathscr{L}_x - \mathrm{i}\,\mathscr{L}_y\right) Y_l^l\,, \qquad Y_l^{l-2} = \left(\mathscr{L}_x - \mathrm{i}\,\mathscr{L}_y\right) Y_l^{l-1}\,, \qquad \ldots \tag{13.80}$$

where Y_l^l is given by (13.77) while

$$\left(\mathscr{L}_x - \mathrm{i}\,\mathscr{L}_y\right) Y_l^l = Y_l^{l-1} = a\,\frac{\exp[\mathrm{i}\,(l-1)\,\varphi]}{(\sin\vartheta)^{l-1}}\,\frac{-\hbar}{\sin\vartheta}\,\frac{\mathrm{d}}{\mathrm{d}\vartheta}\,(\sin\vartheta)^{2l}\,. \tag{13.81}$$

The denominator $(\sin\vartheta)^l$ in the above has been split into two parts for the sake of convenience. The next functions are found from

$$Y_l^{l-s-1} = \left(\mathscr{L}_x - \mathrm{i}\,\mathscr{L}_y\right) Y_l^{l-s} = \frac{-\hbar}{\exp(\mathrm{i}\,\varphi)}\left[\frac{\partial}{\partial\vartheta} + (l-s)\,\frac{\cos\vartheta}{\sin\vartheta}\right] Y_l^{l-s} =$$

$$= \frac{\exp(-\mathrm{i}\,\varphi)}{(\sin\vartheta)^{l-s-1}}\,\frac{-\hbar}{\sin\vartheta}\,\frac{\partial}{\partial\vartheta}\left[(\sin\vartheta)^{l-s}\, Y_l^{l-s}\right], \tag{13.82}$$

where the product $(\sin\vartheta)^{l-s}\, Y_l^{l-s}$ is taken from the previously calculated expression of Y_l^{l-s}. Iterating the procedure yields

$$Y_l^{l-s} = a\,\frac{\exp[\mathrm{i}\,(l-s)\varphi]}{(\sin\vartheta)^{l-s}}\,\underbrace{\frac{-\hbar}{\sin\vartheta}\,\frac{\mathrm{d}}{\mathrm{d}\vartheta}\cdots\frac{-\hbar}{\sin\vartheta}\,\frac{\mathrm{d}}{\mathrm{d}\vartheta}}_{s\text{ times}}\,(\sin\vartheta)^{2l}\,. \tag{13.83}$$

As Y_l^{l-s} is a solution of the linear, homogeneous equations $\mathscr{L}_z Y = L_z\, Y$ and $\mathscr{L}^2 Y = \lambda\, Y$, the constant $a\,\hbar^s$ that builds up in the derivation can be dropped. Letting $m = l-s$ one finds

$$Y_l^m = c_{lm}\,\frac{\exp(\mathrm{i}\,m\,\varphi)}{(\sin\vartheta)^m}\,\underbrace{\frac{-1}{\sin\vartheta}\,\frac{\mathrm{d}}{\mathrm{d}\vartheta}\cdots\frac{-1}{\sin\vartheta}\,\frac{\mathrm{d}}{\mathrm{d}\vartheta}}_{l-m\text{ times}}\,(\sin\vartheta)^{2l}\,, \tag{13.84}$$

where the coefficient c_{lm} has been added for normalization purposes. One may recast (13.84) in a more compact form by letting $\zeta = \cos\vartheta$, whence $-1 \le \zeta \le 1$ and $\mathrm{d}\zeta = -\sin\vartheta\,\mathrm{d}\vartheta$. As a consequence,

$$Y_l^m = c_{lm}\,\frac{\exp(\mathrm{i}\,m\,\varphi)}{(1-\zeta^2)^{m/2}}\,\frac{\mathrm{d}^{l-m}}{\mathrm{d}\zeta^{l-m}}\,(1-\zeta^2)^l\,. \tag{13.85}$$

The eigenfunctions Y_l^m are square integrable and mutually orthogonal [96, App. B.10]. To examine some of their properties it is convenient to introduce some special functions; to begin with, the *associate Legendre functions* are defined in the interval $-1 \le \zeta \le 1$ by

$$P_l^m(\zeta) = \frac{(-1)^m}{2^l\, l\,!}\,(1-\zeta^2)^{m/2}\,\frac{\mathrm{d}^{l+m}}{\mathrm{d}\zeta^{l+m}}(\zeta^2-1)^l\,, \tag{13.86}$$

with $l = 0, 1, \ldots$ and, for each l, $m = -l, \ldots, -1, 0, 1, \ldots, l$. As $(1-\zeta^2)^{m/2}$ and $(\zeta^2-1)^l$ are even functions of ζ, P_l^m is even (odd) if $l+m$ is even (odd): $P_l^m(-\zeta) = (-1)^{l+m}\,P_l^m(\zeta)$. Furthermore, it is

$$P_l^m(\zeta) = (-1)^m\,\frac{(l+m)\,!}{(l-m)\,!}\,P_l^{-m}\,. \tag{13.87}$$

Replacing m with $-m$ in (13.86) shows that Y_l^m is proportional to P_l^{-m} which, in turn, is proportional to P_l^m due to (13.87). In conclusion, using $\zeta = \cos\vartheta$,

$$Y_l^m(\vartheta,\varphi) = c_{lm}\,\exp(\mathrm{i}\,m\,\varphi)\,P_l^m(\cos\vartheta)\,, \tag{13.88}$$

with $0 \le \vartheta \le \pi$ and $0 \le \varphi \le 2\pi$. As for the indices it is $l = 0, 1, \ldots$ and $m = -l, \ldots, -1, 0, 1, \ldots, l$. The functions Y_l^m defined by (13.88) are called *spherical harmonics*. Combining the definition of Y_l^m with the properties of P_l^m shown above yields $Y_l^{-m} = (-1)^m\,(Y_l^m)^*$. Note that Y_l^m and Y_l^{-m} are linearly independent, whereas P_l^m and P_l^{-m} are not. Letting

$$c_{lm} = \left[\frac{(2l+1)\,(l-m)\,!}{4\,\pi\,(l+m)\,!}\right]^{1/2}\,, \tag{13.89}$$

the set made of the spherical harmonics is orthonormal, namely,

$$\int_0^{2\,\pi}\int_0^{\pi}\left(Y_\lambda^\mu\right)^*\,Y_l^m\,\mathrm{d}\vartheta\,\mathrm{d}\varphi = \begin{cases} 1\,, & \lambda = l \text{ and } \mu = m \\ 0\,, & \text{otherwise} \end{cases} \tag{13.90}$$

and complete (Sect. 8.4.3), namely,

$$F(\vartheta,\varphi) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} a_{lm}\,Y_l^m(\vartheta,\varphi)\,, \qquad a_{lm} = \int_0^{2\pi}\int_0^{\pi}\left(Y_l^m\right)^*\,F\,\mathrm{d}\vartheta\,\mathrm{d}\varphi\,, \tag{13.91}$$

where F is a sufficiently regular function of the angles. The inner sum of (13.91),

$$Y_l(\vartheta,\varphi) = \sum_{m=-l}^{l} a_{lm}\,Y_l^m(\vartheta,\varphi) \tag{13.92}$$

is also called *general spherical harmonic of order l*, whereas the special case $m = 0$ of the associate Legendre function,

$$P_l^0(\zeta) = \frac{1}{2^l\, l!}\, \frac{\mathrm{d}^l}{\mathrm{d}\zeta^l}(\zeta^2 - 1)^l\,, \tag{13.93}$$

is a polynomial of degree l called *Legendre polynomial*.

13.6.4 Eigenvalues of the Radial Equation—Coulomb Case

The case $E < 0$ of the radial equation (13.54) is considered here, corresponding to a limited motion. As a consequence, the eigenvectors are expected to depend on a discrete index, and the eigenfunctions are expected to be square integrable. Calculating the derivative and multiplying both sides of (13.54) by $-2\,m_0/(\hbar^2\, r)$ yield

$$\frac{\mathrm{d}^2\varrho}{\mathrm{d}r^2} + \frac{2}{r}\,\frac{\mathrm{d}\varrho}{\mathrm{d}r} - \frac{2\,m_0}{\hbar^2}\,V_e\,\varrho + \frac{2\,m_0}{\hbar^2}\,E\,\varrho = 0\,. \tag{13.94}$$

To proceed one scales the independent variable by multiplying both sides of (13.94) by a^2, where a is a length. The term involving V_e becomes

$$-\,a^2\,\frac{2\,m_0}{\hbar^2}\,V_e\,\varrho = \left[\frac{2\,m_0\,Z\,q^2\,a}{4\,\pi\,\varepsilon_0\,\hbar^2\,(r/a)} - \frac{l\,(l+1)}{(r/a)^2}\right]\varrho\,, \tag{13.95}$$

where both fractions in brackets are dimensionless. As a may be chosen arbitrarily, it is convenient to select for it a value that makes the first fraction equal to $2/(r/a)$, namely,

$$a = \frac{4\,\pi\,\varepsilon_0\,\hbar^2}{m_0\,Z\,q^2}\,. \tag{13.96}$$

As a consequence, the term involving E becomes

$$a^2\,\frac{2\,m_0}{\hbar^2}\,E\,\varrho = \lambda\,\varrho\,, \qquad \lambda = \left(\frac{4\,\pi\,\varepsilon_0}{Z\,q^2}\right)^2\,\frac{2\,\hbar^2}{m_0}\,E\,. \tag{13.97}$$

Adopting the dimensionless variable $x = r/a$ and using the relations $a^2\,\mathrm{d}^2\varrho/\mathrm{d}r^2 = \mathrm{d}^2\varrho/\mathrm{d}x^2$, $(2\,a^2/r)\,\mathrm{d}\varrho/\mathrm{d}r = (2/x)\,\mathrm{d}\rho/\mathrm{d}x$ yields the radial equation in scaled form,

$$\frac{\mathrm{d}^2\varrho}{\mathrm{d}x^2} + \frac{2}{x}\,\frac{\mathrm{d}\varrho}{\mathrm{d}x} + \left[\frac{2}{x} - \frac{l\,(l+1)}{x^2}\right]\varrho + \lambda\,\varrho = 0\,. \tag{13.98}$$

The range of the independent variable is $0 \le x < \infty$, so the equation has a double pole in the origin: it is necessary to select solutions that vanish in the origin in such a way as to absorb the pole (more on this in Sect. 13.6.5). The replacement $\varrho = \sigma/x$ gives (13.98) the simpler form

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}x^2} + \left[\frac{2}{x} - \frac{l\,(l+1)}{x^2}\right]\sigma + \lambda\,\sigma = 0\,, \tag{13.99}$$

which is identical to (13.9). The factorization of (13.99) is then accomplished following the scheme shown in Sect. 13.3, and is based upon the function $g_l = l/x - 1/l$; in this case operators (13.17) and parameter (13.16) read, respectively,

$$\mathscr{A}_l^+ = \frac{l}{x} - \frac{1}{l} + \frac{\mathrm{d}}{\mathrm{d}x}\,, \qquad \mathscr{A}_l^- = \frac{l}{x} - \frac{1}{l} - \frac{\mathrm{d}}{\mathrm{d}x}\,, \qquad L_l = -1/l^2\,. \tag{13.100}$$

The latter depends only on l and fulfills the relation $L_{n+1} > L_{l+1}$, $l = 0, 1, \ldots, n-1$. As a consequence, remembering the second relation in (13.97), the eigenvalues[5] $\lambda = \lambda_n = L_{n+1}$ of (13.99) and those of (13.94) are, respectively,

$$\lambda_n = -\frac{1}{(n+1)^2}\,, \qquad E_n = -\left(\frac{Z\,q^2}{4\,\pi\,\varepsilon_0}\right)^2 \frac{m_0/(2\,\hbar^2)}{(n+1)^2}\,, \qquad n = 0, 1, \ldots\,. \tag{13.101}$$

13.6.5 *Eigenfunctions of the Radial Equation—Coulomb Case*

The eigenfunction corresponding to $l = n$ is found by applying (13.21). As the eigenfunction is indicated here with σ_{nn}, one must solve $\mathscr{A}_{n+1}^- \sigma_{nn} = 0$, namely, using (13.100), $[(n+1)/x - 1/(n+1) - \mathrm{d}/\mathrm{d}x]\,\sigma_{nn} = 0$, whose solution is

$$\sigma_{nn} = c_{nn}\, x^{n+1}\, \exp\left(-\frac{x}{n+1}\right)\,, \qquad n = 0, 1, \ldots\,, \tag{13.102}$$

which vanishes both in the origin and at infinity, thus fulfilling the requirements stated in Sect. 13.6.4. The eigenfunction (13.102) is also square integrable with, from (C.95,C.97),

$$\frac{1}{c_{nn}^2} = \int_0^\infty x^{2n+2} \exp\left(-\frac{2\,x}{n+1}\right) \mathrm{d}x = (2\,n+2)\,! \left(\frac{n+1}{2}\right)^{2n+3}\,. \tag{13.103}$$

[5] In (13.56) a more convenient notation is used, obtained from (13.101) through the replacements $n+1 \leftarrow n' \leftarrow n$, with $n' = 1, 2, \ldots$.

Combining the first definition in (13.28) with the third relation in (13.100), the auxiliary operator $\mathscr{B}_{nl}^{+}$ reads

$$\mathscr{B}_{nl}^{+} = \frac{l\,(n+1)}{\sqrt{(n+1-l)(n+1+l)}}\,\mathscr{A}_{l}^{+}\,. \tag{13.104}$$

Then, from the second of (13.31), the normalized eigenfunctions corresponding to $l < n$ are found recursively from

$$\sigma_{n,n-1} = \mathscr{B}_{nn}^{+}\sigma_{nn}\,, \qquad \sigma_{n,n-2} = \mathscr{B}_{n,n-1}^{+}\sigma_{n,n-1}\,, \qquad \ldots \tag{13.105}$$

The last eigenfunction found by the recursive procedure is $\sigma_{n0} = \mathscr{B}_{n1}^{+}\sigma_{n1}$ as expected. In fact, a further iteration would not yield an eigenfunction because $\mathscr{B}_{n0}^{+} = 0$.

The eigenfunction of (13.94) corresponding to the lowest total energy $E_{\min} = E(n=0)$ is found by combining (13.102,13.103) with $\sigma = \varrho/x$ and $x = r/a$, thus yielding $\varrho(r) = (1/2)\,\exp(-r/a)$. There is only one spherical harmonic compatible with this energy eigenvalue, specifically, $Y_0^0 = 1/\sqrt{4\,\pi}$ (Table 13.1). Thus, the product $w(E_{\min}) = (c/2)\,\exp(-r/a)/\sqrt{4\,\pi}$, with c a normalization constant, is the eigenfunction of the Schrödinger equation (13.42) corresponding to the lowest total energy. The normalization constant is necessary because ϱ is obtained from scaling another function σ, originally normalized to unity. Taking the Jacobian determinant $J = r^2\,\sin\vartheta$ from (B.3) one finds

$$\frac{1}{c^2} = \frac{1}{16\,\pi}\int_0^{\infty}\int_0^{\pi}\int_0^{2\,\pi}\exp\left(-\frac{2\,r}{a}\right)r^2\,\sin\vartheta\,\mathrm{d}r\,\mathrm{d}\vartheta\,\mathrm{d}\varphi = \frac{a^3}{16}\,, \tag{13.106}$$

whence $w(E_{\min}) = \exp(-r/a)/\sqrt{\pi\,a^3}$.

13.6.6 Transmission Matrix

The one-dimensional, time-independent Schrödinger equation (11.28) is solvable analytically in a limited number of cases, some of which have been illustrated in the sections above. When the analytical solution is not known one must resort to approximate methods; an example is given here, with reference to a finite domain $0 \le x \le s$. The latter is tessellated by selecting N points $x_1 < x_2 < \ldots < x_N$, internal to the domain, called *nodes*. The boundaries of the domain are indicated with $0 = x_0 < x_1$ and $s = x_{N+1} > x_N$. The segment bounded by x_i and x_{i+1} is indicated with h_{i+1} and is called *element*. The same symbol indicates the length of the element, $h_{i+1} = x_{i+1} - x_i$. Finally, a subdomain Ω_i, called *cell*, is associated with each node. For the internal nodes $x_1, \ldots, x_N$ the cell is bounded by $x_i - h_i/2$ and $x_i + h_{i+1}/2$. The same symbol is used to indicate also the cell length, $\Omega_i =$

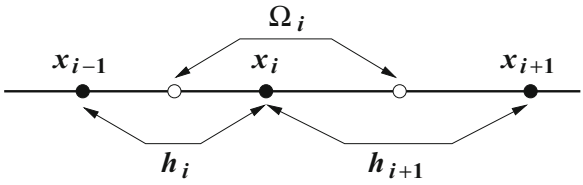


Fig. 13.3 Illustration of the concepts of node, element, and cell (Sect. 13.6.6)

$(h_i + h_{i+1})/2$ (Fig. 13.3). The left boundary x_0 is associated with the cell Ω_0 of length $h_1/2$ placed on the right of x_0, while the right boundary x_{N+1} is associated with the cell Ω_{N+1} of length $h_{N+1}/2$ placed on the left of x_{N+1}.

The approximation methods that are applicable to a given tessellation are numerous. The method depicted in this section replaces the coefficient $q(x)$ of (11.28) over each element h_i with an approximating function $q_i(x)$ such that the solution $w_i(x)$ to (11.28) over h_i can be found analytically. The approximating functions q_i may differ from an element to another, thus yielding different analytical solutions. Then, the continuity of the analytical solutions and their derivatives is imposed at each node; finally, the same continuity is imposed at the boundaries 0 and s, where the form of the wave function is supposed to be known.

To proceed, consider an internal node $i = 1, 2, \ldots, N$ and the two elements h_i, h_{i+1} adjacent to it. The solutions w_i, w_{i+1} over the two elements are expressed in terms of the fundamental solutions u, v (compare with Sect. 11.4):

$$w_i(x) = a_i^u u_i(x) + a_i^v v_i(x)\,, \qquad w_{i+1}(x) = a_{i+1}^u u_{i+1}(x) + a_{i+1}^v v_{i+1}(x)\,, \tag{13.107}$$

with a_i^u, a_i^v, a_{i+1}^u, a_{i+1}^v undetermined constants. The fundamental solutions fulfill the boundary conditions

$$u_{i+1}(x_i) = 1\,, \qquad u'_{i+1}(x_i) = 0\,, \qquad v_{i+1}(x_i) = 0\,, \qquad v'_{i+1}(x_i) = 1\,, \tag{13.108}$$

$i = 1, 2, \ldots, N$, where primes indicate derivatives. Imposing the continuity of w, w' at x_i yields

$$a_{i+1}^u = a_i^u u_i(x_i) + a_i^v v_i(x_i)\,, \qquad a_{i+1}^v = a_i^u u'_i(x_i) + a_i^v v'_i(x_i)\,. \tag{13.109}$$

Letting

$$\mathbf{a}_i = \begin{bmatrix} a_i^u \\ a_i^v \end{bmatrix}\,, \qquad \mathbf{N}_i = \begin{bmatrix} u_i(x_i) & v_i(x_i) \\ u'_i(x_i) & v'_i(x_i) \end{bmatrix}\,, \tag{13.110}$$

the relations (13.109) take the form

$$\mathbf{a}_{i+1} = \mathbf{N}_i\, \mathbf{a}_i\,. \tag{13.111}$$

Matrix $\mathbf{N}_i$ is known by construction, and provides the link between the unknown vectors $\mathbf{a}_i$ and $\mathbf{a}_{i+1}$. Vector $\mathbf{a}_i$ belongs to element h_i only, whereas matrix $\mathbf{N}_i$ belongs to element h_i (due to u_i, v_i) and also to node x_i (because u_i, v_i are calculated at x_i). Iterating (13.111) yields

$$\mathbf{a}_{N+1} = \mathbf{N}_I \mathbf{a}_1 \,, \qquad \mathbf{N}_I = \mathbf{N}_N \mathbf{N}_{N-1} \ldots \mathbf{N}_2 \mathbf{N}_1 \,. \tag{13.112}$$

Remembering the discussion in Sect. A.12 one finds $\det \mathbf{N}_i = W = 1$, whence $\det \mathbf{N}_I = \det \mathbf{N}_N \ldots \det \mathbf{N}_1 = 1$. Now it is necessary to link the solution over h_1 with that over $x < 0$, is given by (11.30). Although the two functions $\exp(\pm \mathrm{i}\, k_L\, x)$ in the latter are not fundamental solutions, it is convenient to keep the form (11.30) because a_2/a_1 provides the information about the reflection coefficient directly. Letting

$$\mathbf{a}_L = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} , \qquad \mathbf{N}_L = \begin{bmatrix} 1 & 1 \\ \mathrm{i}\, k_L & -\mathrm{i}\, k_L \end{bmatrix} , \tag{13.113}$$

the continuity of w and w' at $x = 0$ yields $\mathbf{a}_1 = \mathbf{N}_L\, \mathbf{a}_L$. Similarly, it is necessary to link the solution over h_{N+1} with that over $x > 0$, which is given by (11.31). Again, the two functions $\exp(\pm \mathrm{i}\, k_R\, x)$ in the latter are not fundamental solutions; however, they are kept here because a_5/a_1 provides the information about the transmission coefficient directly. Letting

$$\mathbf{a}_R = \begin{bmatrix} a_5 \\ a_6 \end{bmatrix} , \qquad \mathbf{N}_R = \begin{bmatrix} \exp(\mathrm{i}\, k_R\, s) & \exp(-\mathrm{i}\, k_R\, s) \\ \mathrm{i}\, k_R\, \exp(\mathrm{i}\, k_R\, s) & -\mathrm{i}\, k_R\, \exp(-\mathrm{i}\, k_R\, s) \end{bmatrix} , \tag{13.114}$$

the continuity of w and w' at $x = s$ yields $\mathbf{N}_R\, \mathbf{a}_R = \mathbf{N}_{N+1}\, \mathbf{a}_{N+1}$, with $\det \mathbf{N}_{N+1} = 1$, $\det \mathbf{N}_L = -2\,\mathrm{i}\, k_L$, $\det \mathbf{N}_R = -2\,\mathrm{i}\, k_R$. Combining the relations found so far,

$$\mathbf{a}_R = \mathbf{N}\, \mathbf{a}_L \,, \qquad \mathbf{N} = \mathbf{N}_R^{-1}\, \mathbf{N}_{N+1}\, \mathbf{N}_I\, \mathbf{N}_L \,, \tag{13.115}$$

where

$$\mathbf{N}_R^{-1} = \begin{bmatrix} \exp(-\mathrm{i}\, k_R\, s)/2 & \exp(-\mathrm{i}\, k_R\, s)/(2\,\mathrm{i}\, k_R) \\ \exp(\mathrm{i}\, k_R\, s)/2 & \mathrm{i}\, \exp(\mathrm{i}\, k_R\, s)/(2\, k_R) \end{bmatrix} , \qquad \det \mathbf{N}_R^{-1} = -\frac{1}{2\,\mathrm{i}\, k_R} \,. \tag{13.116}$$

whence $\det \mathbf{N} = \det \mathbf{N}_R^{-1} \det \mathbf{N}_{N+1} \det \mathbf{N}_I \det \mathbf{N}_L = k_L/k_R$. Matrix $\mathbf{N}$ (also called *transmission matrix* in [24]) provides the link between $\mathbf{a}_L$ and $\mathbf{a}_R$. Splitting the first relation of (13.115) into its components gives

$$a_5 = N_{11}\, a_1 + N_{12}\, a_2 \,, \qquad a_6 = N_{21}\, a_1 + N_{22}\, a_2 \,. \tag{13.117}$$

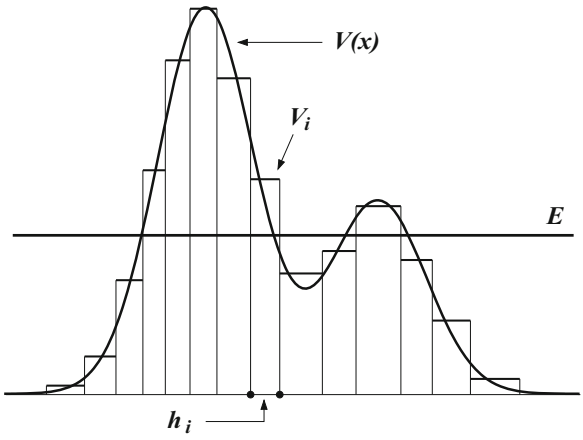


Fig. 13.4 Example of a potential energy $V(x)$ replaced with a piecewise-constant function V_i (Sect. 13.6.6)

If the particle is launched, e.g., from $-\infty$ one lets $a_6 = 0$, whence

$$\frac{a_2}{a_1} = -\frac{N_{21}}{N_{22}}\,, \qquad \frac{a_5}{a_1} = \frac{\det \mathbf{N}}{N_{22}} = \frac{k_L/k_R}{N_{22}}\,. \tag{13.118}$$

Combining (13.118) with (11.35) yields[6] $|N_{21}|^2 + k_L/k_R = |N_{22}|^2$.

The derivation of the transmission matrix has been carried out here without specifying the form of the fundamental solutions u_i, v_i over the corresponding element h_i. In the practical cases, to easily find an analytical solution over each element one approximates the coefficient $q(x)$ of (11.28) with a constant, $q_i = \text{const}$ in h_i; this is equivalent to replacing the potential energy $V(x)$ with a piecewise-constant function V_i (Fig. 13.4). Depending on the sign of $q_i = 2\,m\,(E - V_i)/\hbar^2$ the possible cases for u_i, v_i are:

$$\begin{cases} q_i = -\alpha_i^2 < 0 & u_i = \cosh[\alpha_i\,(x - x_{i-1})] & v_i = \sinh[\alpha_i\,(x - x_{i-1})]/\alpha_i \\ q_i = k_i^2 > 0 & u_i = \cos[k_i\,(x - x_{i-1})] & v_i = \sin[k_i\,(x - x_{i-1})]/k \\ q_i = 0 & u_i = 1 & v_i = x - x_{i-1} \end{cases} \tag{13.119}$$

with α_i, k_i real. As the potential energy is replaced with a piecewise-constant function, the accuracy of the approximation is not very high. Other methods are more accurate; in particular, it can be shown that the Numerov process provides a higher-order accuracy without a significant increase in the computational cost [19], [20]. The details of the method are illustrated in Sect. A.13.3 and its application to the solution of the time-independent Schrödinger equation is shown in Prob. 13.3.

[6] Within a numerical solution of the Schrödinger equation, the relation $|N_{21}|^2 + k_L/k_R = |N_{22}|^2$ may be exploited as a check for the quality of the approximation.

Problems

13.1 Letting $Z = 1$ in (13.56) one finds the expression of the energy levels of the hydrogen atom in a bound state, consistently with that obtained from the Bohr hypothesis (Sect. 7.4.4). Use the same equation to calculate the minimum energy that must be given to the electron to extract it from the hydrogen atom (*ionization energy*).

13.2 With reference to the hydrogen atom, calculate the expectation value of the radius r corresponding to the eigenfunction $w(E_{\min}) = \exp(-r/a)/\sqrt{\pi\, a^3}$ found in Sect. 13.6.5.

13.3 Apply the Numerov process illustrated in Sect. A.13.3 to discretize the time-independent Schrödinger equation (11.28).

Chapter 14
Time-Dependent Perturbation Theory

14.1 Introduction

In many physical problems it is necessary to consider the collision of a particle with another particle or system of particles. The treatment based on Classical Mechanics is given in Sects. 3.5,3.6 with reference to the motion's asymptotic conditions, without considering the form of the interaction, while Sect. 3.8 shows a detailed treatment of the Coulomb interaction. Here the approach based on Quantum Mechanics is shown, dealing with the following problem: a particle in a conservative motion enters at $t = 0$ an interaction with another particle or system of particles; such an interaction has a finite duration t_P, at the end of which the particle is in a conservative motion again. The perturbation produced by the interaction, which is described by a suitable Hamiltonian operator, may change the total energy of the particle; the analysis carried out here, called *time-dependent perturbation theory*, allows one to calculate such an energy change. The other particle or system, with which the particle under consideration interacts, is left unspecified. However, it is implied that the larger system, made of the particle under consideration and the entity with which it interacts, is isolated, so that the total energy is conserved: if the particle's energy increases due to the interaction, then such an energy is absorbed from the other entity, or vice versa. As in Classical Mechanics, other dynamic properties of an isolated system are conserved; an example of momentum conservation is given in Sect. 14.8.3.

The discussion is carried out first for the case where the eigenvalues prior and after the perturbation are discrete and nondegenerate. Starting from the general solution of the perturbed problem, a first-order approximation is applied, which holds for small perturbations, and the probability per unit time of the transition from a given initial state to another state is found. The analysis is repeated for the degenerate case (still for discrete eigenvalues) and, finally, for the situation where both the initial and final state belong to a continuous set. The last section shows the calculation of the perturbation matrix for a screened Coulomb perturbation. The complements deal with the important problems of a perturbation constant in

M. Rudan, *Physics of Semiconductor Devices*, DOI 10.1007/978-3-319-63154-7_14

time and a harmonic perturbation; a short discussion follows about the Fermi golden rule, the transitions from discrete to continuous levels, and the general form of the harmonic perturbation.

14.2 Discrete Eigenvalues

Let $\mathscr{H}$ be the Hamiltonian operator that describes the dynamics of the particle when the perturbation is absent. Such an operator is assumed to be conservative, namely, $\mathscr{H} = -\hbar^2 \nabla^2/(2\,m) + V(\mathbf{r})$. When the perturbation is present, namely, for $0 \le t \le t_P$, the Hamiltonian operator is referred to as $\mathscr{H}'$. The two operators, respectively called *unperturbed Hamiltonian* and *perturbed Hamiltonian*, are Hermitean, so their difference $\delta\mathscr{H} = \mathscr{H}' - \mathscr{H}$, called *perturbation Hamiltonian*, is Hermitean as well. Also, it is assumed that $\mathscr{H}$ and $\mathscr{H}'$ are real, and that $\delta\mathscr{H}$ does not act on time; however, it is $\delta\mathscr{H} = \delta\mathscr{H}(t)$ because the perturbation is present only when $0 \le t \le t_P$.

For $t < 0$ and $t > t_P$ the wave function is unperturbed; remembering the concepts introduced in Sect. 9.2 and considering the case of discrete eigenvalues, it reads

$$\psi = \sum_n c_n\, w_n(\mathbf{r})\, \exp(-\mathrm{i}\, E_n\, t/\hbar)\,, \tag{14.1}$$

with w_n the solutions of $\mathscr{H} w_n = E_n\, w_n$. As usual, n stands for a single index or a set of indices. For the sake of simplicity, here it is assumed provisionally that the energy eigenvalues are not degenerate, so that a one-to-one correspondence exists between E_n and w_n; this assumption will be removed later (Sect. 14.5). The wave function is assumed to be square integrable and normalized to unity, whence $|c_n|^2$ is the probability that the particle under consideration is found in the nth state, and $\langle w_s | w_n \rangle = \delta_{sn}$. As noted above, expansion (14.1) holds for both $t < 0$ and $t > t_P$. However, prior to the perturbation and as a consequence of a measurement, the additional information about the energy state is available; assuming that the outcome of the measurement was the rth state, it follows (Sect. 9.2) that for $t < 0$ it is $c_n = 0$ when $n \neq r$, and[1]

$$|c_r|^2 = 1\,, \qquad c_r = 1\,, \qquad \psi = w_r(\mathbf{r})\, \exp(-\mathrm{i}\, E_r\, t/\hbar)\,. \tag{14.2}$$

When $t \to 0$ the last relation in (14.2) yields $\psi(\mathbf{r}, t = 0) = w_r$ which, by continuity, provides the initial condition for the time-dependent Schrödinger equation to be solved in the interval $0 \le t \le t_P$; such an equation reads

[1] The phase of c_r is irrelevant and is set to zero.

$$(\mathscr{H} + \delta\mathscr{H})\ \psi = \mathrm{i}\,\hbar\,\frac{\partial\psi}{\partial t}\,. \tag{14.3}$$

Thanks to the completeness of the w_ns one expands the wave function as $\psi = \sum_n b_n(t)\,w_n(\mathbf{r})$, where functions b_n are unknown (compare with (9.10)). However, it is convenient to transfer the role of unknowns to a new set of functions $a_n(t) = b_n(t)\,\exp(\mathrm{i}\,E_n\,t/\hbar)$, so that the expansion reads

$$\psi = \sum_n a_n(t)\,w_n(\mathbf{r})\,\exp(-\mathrm{i}\,E_n\,t/\hbar)\,. \tag{14.4}$$

By this token, the initial condition yields $a_n(0) = \delta_{nr}$. Replacing (14.4) in (14.3),

$$\sum_n a_n\,\exp(-\mathrm{i}\,E_n\,t/\hbar)\ (\mathscr{H}w_n + \delta\mathscr{H}w_n) =$$

$$= \mathrm{i}\,\hbar\,\sum_n w_n\,\exp(-\mathrm{i}\,E_n\,t/\hbar)\ (\mathrm{d}a_n/\mathrm{d}t - \mathrm{i}\,E_n\,a_n/\hbar)\ , \tag{14.5}$$

where the first and last terms cancel out due to $\mathscr{H}w_n = E_n\,w_n$. The next step is a scalar multiplication of the remaining terms by one of the eigenfunctions of the unperturbed problem, say, w_s. The sum $\mathrm{i}\,\hbar\,\sum_n w_n\,\exp(-\mathrm{i}\,E_n\,t/\hbar)\,\mathrm{d}a_n/\mathrm{d}t$ at the right-hand side of (14.5) transforms, due to $\langle w_s|w_n\rangle = \delta_{sn}$, into the single term $\mathrm{i}\,\hbar\,\exp(-\mathrm{i}\,E_s\,t/\hbar)\,\mathrm{d}a_s/\mathrm{d}t$. In conclusion, the time-dependent Schrödinger equation (14.3) becomes a set of infinite, coupled linear equations in the unknowns a_s:

$$\frac{\mathrm{d}a_s}{\mathrm{d}t} = \frac{1}{\mathrm{i}\,\hbar}\,\sum_n a_n\,h_{ns}\,\exp(-\mathrm{i}\,\omega_{ns}\,t)\,, \qquad a_s(0) = \delta_{sr}\,, \tag{14.6}$$

with

$$h_{ns}(t) = \int_\Omega w_s^*\,\delta\mathscr{H}w_n\,\mathrm{d}^3r\,, \qquad \omega_{ns} = (E_n - E_s)/\hbar\,. \tag{14.7}$$

The coefficients of (14.6) embed the eigenvalues and eigenfunctions of the unperturbed problem. Due to its form, the set of elements $h_{ns}(t)$ is called *perturbation matrix*; remembering that $\delta\mathscr{H}$ is real, the definition (14.7) of the elements shows that $h_{sn} = h_{ns}^*$.

14.3 First-Order Perturbation

The differential equations (14.6) are readily transformed into a set of integral equations by integrating from $t = 0$ to $t \le t_P$ and using the initial condition:

$$a_s = \delta_{sr} + \frac{1}{\mathrm{i}\,\hbar}\,\int_0^t \sum_n a_n(t')\,h_{ns}(t')\,\exp(-\mathrm{i}\,\omega_{ns}\,t')\,\mathrm{d}t'\,. \tag{14.8}$$

As mentioned above, the solution of the Schrödinger equation for $t > t_P$ is (14.1), with $|c_n|^2$ the probability that a measurement carried out at $t = t_P$ yields the eigenvalue E_n. The coefficients c_n are found by observing that after the solution of (14.6) or (14.8) is calculated, the time-continuity of ψ and the uniqueness of the expansion yield $c_n = a_n(t_P)$. It follows that the probability that at $t = t_P$ an energy measurement finds the eigenvalue E_s is $|a_s(t_P)|^2$. On the other hand, the energy state prior to the perturbation was assumed to be E_r, and the functions $a_n(t)$ inherit this assumption through the initial condition $a_n(0) = \delta_{nr}$; as a consequence, the quantity $|a_s(t_P)|^2 = |b_s(t_P)|^2$ can be thought of as the probability that the perturbation brings the particle from the initial state E_r to the final state E_s: for this reason, $P_{rs} = |a_s(t_P)|^2$ is called *transition probability from state r to state s*. Thanks to the normalization condition it is $\sum_s P_{rs} = \int_\Omega |\psi(t_P)|^2 \, \mathrm{d}^3 r = 1$; the term of equal indices, P_{rr}, is the probability that the perturbation leaves the particle's state unchanged.

The two forms (14.6) and (14.8) are equivalent to each other; however, the second one is better suited for an iterative-solution procedure, that reads

$$a_s^{(k+1)} = \delta_{sr} + \frac{1}{\mathrm{i}\,\hbar} \int_0^t \sum_n a_n^{(k)} \, h_{ns} \, \exp(-\mathrm{i}\,\omega_{ns}\, t') \, \mathrm{d}t' \,, \tag{14.9}$$

where $a_n^{(k)}(t)$ is the kth iterate. The iterations are brought to an end when $||a_s^{(k+1)} - a_s^{(k)}|| < \epsilon$, where the bars indicate a suitable norm and ϵ is a small positive constant. To start the procedure it is necessary to choose the iterate of order zero, $a_n^{(0)}(t)$, which is typically done by letting $a_n^{(0)}(t) = a_n^{(0)}(0) = \delta_{nr}$; in other terms, the initial iterate of a_n is selected as a constant equal to the initial condition of a_n. Replacing this value into the integral of (14.9) yields the first-order iterate

$$a_r^{(1)} = 1 + \frac{1}{\mathrm{i}\,\hbar} \int_0^t h_{rr} \, \mathrm{d}t' \,, \qquad a_s^{(1)} = \frac{1}{\mathrm{i}\,\hbar} \int_0^t h_{rs} \, \exp(-\mathrm{i}\,\omega_{rs}\, t') \, \mathrm{d}t' \,, \tag{14.10}$$

$s \neq r$. If the perturbation is small enough, the first-order iterate is already close to the solution, so that $a_r \simeq a_r^{(1)}$, $a_s \simeq a_s^{(1)}$. This case happens when the norm of the integrals in (14.10) is much smaller than unity; it follows $P_{rr} \simeq 1$, $P_{rs} \ll 1$. The approximate solution thus found is called *first-order solution*, or *first-order perturbation*. Note that, as h_{rr} is real, the iterate $a_r^{(1)}$ is a complex number whose real part equals unity; as a consequence it is $|a_r^{(1)}|^2 > 1$. This nonphysical result is due to the approximation.

14.4 Comments

Considering the case where the initial and final states are different, and observing that the entries of the perturbation matrix vanish for $t < 0$ and $t > t_P$, one can calculate $a_s(t_P)$ by replacing the integration limits 0 and t_P with $-\infty$ and $+\infty$, respectively. This shows that $a_s^{(1)}$ is proportional to the Fourier transform (C.16) of h_{rs} evaluated at $\omega_{rs} = (E_r - E_s)/\hbar$,

$$a_s^{(1)} = \frac{1}{\mathrm{i}\,\hbar} \int_{-\infty}^{+\infty} h_{rs} \exp(-\mathrm{i}\,\omega_{rs}\,t')\,\mathrm{d}t' = \frac{\sqrt{2\,\pi}}{\mathrm{i}\,\hbar}\ \mathscr{F}h_{rs}|_{\omega=\omega_{rs}}\ . \tag{14.11}$$

In conclusion, the first-order solution of the time-dependent Schrödinger equation (14.3) yields the following probability of a transition from state r to state s:

$$P_{rs} = \frac{2\,\pi}{\hbar^2}\ |\mathscr{F}h_{rs}|^2\ . \tag{14.12}$$

The units of h_{ns} are those of an energy. It follows that the units of $\mathscr{F}h_{rs}$ are those of an action, and P_{rs} is dimensionless, as expected. Some important consequences derive from (14.12):

1. It may happen that for a given perturbation Hamiltonian $\delta\mathscr{H}$ the eigenfunctions w_r, w_s ($s \neq r$) are such that $h_{rs} = 0$. In this case $\delta\mathscr{H}$ is not able to induce the transition from state r to state s: the transition is *forbidden*. Basing on this observation one can determine the pairs of indices for which the transitions are permitted, thus providing the so-called *transition rules* or *selection rules*. For this analysis it is sufficient to consider the symmetry of the integrand in the definition (14.7) of h_{rs}, without the need of calculating the integral itself.
2. By exchanging r and s one finds $h_{sr} = h_{rs}^*$, while ω_{rs} becomes $-\omega_{rs}$. From (14.11) it follows that $\mathscr{F}h_{sr} = (\mathscr{F}h_{rs})^*$, whence $P_{sr} = P_{rs}$: for a given perturbation Hamiltonian $\delta\mathscr{H}$ the probability of the $r \to s$ and $s \to r$ transitions are the same.

The transition from an energy state to a different one entails a change in the total energy of the particle under consideration. Such a change is due to the interaction with another particle or system of particles whence $\delta\mathscr{H}$ originates. Examples are given in Sects. 14.8.1 and 14.8.2.

The replacement of the integration limits 0 and t_P with $-\infty$ and $+\infty$, carried out above, has the advantage of making the presence of the Fourier transform clearer; however, remembering that the duration t_P of the perturbation is finite, one observes that the probability P_{rs} is a function of t_P proper. From this, the *probability per unit time* of the transition is defined as

$$\dot{P}_{rs} = \frac{\mathrm{d}P_{rs}}{\mathrm{d}t_P}\ . \tag{14.13}$$

14.5 Degenerate Energy Levels

In Sect. 14.2 nondegenerate energy levels have been assumed for the sake of simplicity. The case of degenerate levels is considered here, still assuming that the indices are discrete. By way of example, let each energy value E_n correspond to a set $w_{n1}, \ldots, w_{n\gamma}, \ldots$ of linearly independent, mutually orthogonal eigenfunctions. An example of this is given by the eigenvalues (13.56) of the Schrödinger equation for a central force of the Coulomb type, whose degree of degeneracy in the spinless case is given by (13.57). Expression (14.1) of the unperturbed wave function, that holds for $t < 0$ and $t > t_P$, becomes in this case

$$\psi = \sum_{n\gamma} c_{n\gamma}\, w_{n\gamma}(\mathbf{r})\, \exp(-\mathrm{i}\, E_n\, t/\hbar)\,, \tag{14.14}$$

with $w_{n\gamma}$ the solutions of $\mathscr{H} w_{n\gamma} = E_n\, w_{n\gamma}$. As before, the wave function is assumed to be square integrable and normalized to unity, whence $|c_{n\gamma}|^2$ is the probability that the particle under consideration is found in the state labeled by n, γ, and $\langle w_{s\beta} | w_{n\gamma} \rangle = \delta_{sn}\, \delta_{\beta\gamma}$. Prior to the perturbation and as a consequence of measurements, the additional information about the energy state is available, along with that of the observable associated with index γ, whose measurement is compatible with that of energy (compare with Sect. 10.4); assuming that the outcome of the measurements was the state labeled r, α, it follows that for $t < 0$ it is $c_{r\alpha} = 1$, $\psi = w_{r\alpha}(\mathbf{r})\, \exp(-\mathrm{i}\, E_r\, t/\hbar)$, while all other coefficients vanish. As a consequence, the initial condition for the time-dependent Schrödinger equation to be solved in the interval $0 \leq t \leq t_P$ is $\psi(\mathbf{r}, t = 0) = w_{r\alpha}$. Following the same reasoning as in Sect. 14.2 shows that in such an interval the expansion $\psi = \sum_{n\gamma} a_{n\gamma}(t)\, w_{n\gamma}(\mathbf{r})\, \exp(-\mathrm{i}\, E_n\, t/\hbar)$ holds, and the time-dependent Schrödinger equation transforms into the set of infinite, coupled linear equations

$$\frac{\mathrm{d} a_{s\beta}}{\mathrm{d}t} = \frac{1}{\mathrm{i}\,\hbar} \sum_{n\gamma} a_{n\gamma}\, h_{ns}^{\gamma\beta}\, \exp(-\mathrm{i}\, \omega_{ns}\, t)\,, \qquad a_{s\beta}(0) = \delta_{sr}\, \delta_{\beta\alpha}\,, \tag{14.15}$$

with

$$h_{ns}^{\gamma\beta}(t) = \int_\Omega w_{s\beta}^*\, \delta\mathscr{H} w_{n\gamma} \mathrm{d}^3 r\,, \qquad \omega_{ns} = (E_n - E_s)/\hbar\,. \tag{14.16}$$

The first-order perturbative solution of (14.15) is obtained following the same path as in Sect. 14.3. Within this approximation, and considering a final state s, β different from the initial one, the probability of a $r, \alpha \to s, \beta$ transition induced by the perturbation is

$$P_{rs}^{\alpha\beta} = \frac{1}{\hbar^2} \left| \int_0^{t_P} h_{rs}^{\alpha\beta}(t)\, \exp(-\mathrm{i}\, \omega_{ns}\, t)\, \mathrm{d}t \right|^2 . \tag{14.17}$$

Thanks to the normalization condition it is $\sum_{s\beta} P_{rs}^{\alpha\beta} = \int_\Omega |\psi(t_P)|^2 \, \mathrm{d}^3 r = 1$, which can be expressed as

$$\sum_s P_{rs}^{\alpha} = 1 \,, \qquad P_{rs}^{\alpha} = \sum_\beta P_{rs}^{\alpha\beta} \,. \tag{14.18}$$

This shows that the inner sum P_{rs}^{α} is the probability that the perturbation induces a transition from the initial state r, α to any final state whose energy is $E_s \neq E_r$.

14.6 Continuous Energy Levels

When the spectrum is continuous, the wave packet describing a particle in a three-dimensional space and in a conservative case is given by (9.5), namely

$$\psi(\mathbf{r}, t) = \iiint_{-\infty}^{+\infty} c_{\mathbf{k}} \, w_{\mathbf{k}}(\mathbf{r}) \, \exp(-\mathrm{i} \, E_{\mathbf{k}} \, t/\hbar) \, \mathrm{d}^3 k \,, \tag{14.19}$$

with $E_{\mathbf{k}}$, $w_{\mathbf{k}}$ the eigenvalues and eigenfunctions of $\mathscr{H} w_{\mathbf{k}} = E_{\mathbf{k}} \, w_{\mathbf{k}}$, and $\mathbf{k}$ a three-dimensional vector whose components are continuous. If the wave function is square integrable and normalized to unity, (9.7) holds:

$$\int_\Omega |\psi|^2 \, \mathrm{d}^3 r = \iiint_{-\infty}^{+\infty} |c_{\mathbf{k}}|^2 \, \mathrm{d}^3 k = 1 \,. \tag{14.20}$$

Remembering the discussion of Sects. 9.2, 9.6, and 10.2, the product $|\psi(\mathbf{r}, t)|^2 \, \mathrm{d}^3 r$ is the infinitesimal probability that at time t the particle is localized within $\mathrm{d}^3 r$ around $\mathbf{r}$, and the product $|c_{\mathbf{k}}|^2 \, \mathrm{d}^3 k$ is the infinitesimal probability that the outcome of an energy measurement belongs to the range of $E(\mathbf{k})$ values whose domain is $\mathrm{d}^3 k$.

To proceed one assumes that the unperturbed Hamiltonian operator is that of a free particle, $\mathscr{H} = -(\hbar^2/2\,m)\,\nabla^2$; it follows that the wave function and energy corresponding to a wave vector $\mathbf{k}$ read (Sect. 9.6)

$$w_{\mathbf{k}}(\mathbf{r}) = \frac{1}{(2\,\pi)^{3/2}} \, \exp(\mathrm{i} \, \mathbf{k} \cdot \mathbf{r}) \,, \qquad E_{\mathbf{k}} = \frac{\hbar^2}{2\,m} \left(k_1^2 + k_2^2 + k_3^2\right) = \hbar \, \omega_{\mathbf{k}} \,. \tag{14.21}$$

Thanks to the completeness of the eigenfunctions (14.21), during perturbation the wave function is given by

$$\psi(\mathbf{r}, t) = \iiint_{-\infty}^{+\infty} a_{\mathbf{k}}(t) \, w_{\mathbf{k}}(\mathbf{r}) \, \exp(-\mathrm{i} \, E_{\mathbf{k}} \, t/\hbar) \, \mathrm{d}^3 k \,. \tag{14.22}$$

Due to (14.20), the units of $|c_{\mathbf{k}}|^2$ and, consequently, of $|a_{\mathbf{k}}|^2$, are those of a volume. The same reasoning as in Sect. 14.2 yields in this case

$$\int \frac{\mathrm{d}a_{\mathbf{k}}}{\mathrm{d}t}\, w_{\mathbf{k}}(\mathbf{r})\, \exp\left(-\mathrm{i}\, E_{\mathbf{k}}\, t/\hbar\right)\, \mathrm{d}^3k = \int \frac{a_{\mathbf{k}}}{\mathrm{i}\,\hbar}\, \delta\mathscr{H} w_{\mathbf{k}}(\mathbf{r})\, \exp\left(-\mathrm{i}\, E_{\mathbf{k}}\, t/\hbar\right)\, \mathrm{d}^3k \tag{14.23}$$

(for the sake of simplicity, the symbol of triple integral over $\mathbf{k}$ or $\mathbf{r}$ is replaced with $\int$ in (14.23) and in the relations below). Considering a state $\mathbf{g}$, a scalar multiplication of (14.23) by the corresponding eigenfunction $w_{\mathbf{g}}$ is carried out; performing the integration over $\mathbf{r}$ first, yields at the right-hand side the entry of the perturbation matrix of labels $\mathbf{k}$ and $\mathbf{g}$:

$$h_{\mathbf{kg}}(t) = \frac{1}{(2\,\pi)^3} \int \exp(-\mathrm{i}\,\mathbf{g}\cdot\mathbf{r})\, \delta\mathscr{H} \exp(\mathrm{i}\,\mathbf{k}\cdot\mathbf{r})\, \mathrm{d}^3r\,, \tag{14.24}$$

where the units of $h_{\mathbf{kg}}(t)$ are those of an energy times a volume. At the left-hand side of (14.23), still performing the integration over $\mathbf{r}$ first, and using (C.56), provides

$$\int \frac{\mathrm{d}a_{\mathbf{k}}}{\mathrm{d}t}\, \delta(\mathbf{k}-\mathbf{g})\, \exp\left(-\mathrm{i}\, E_{\mathbf{k}}\, t/\hbar\right)\, \mathrm{d}^3k = \frac{\mathrm{d}a_{\mathbf{g}}}{\mathrm{d}t}\, \exp\left(-\mathrm{i}\, E_{\mathbf{g}}\, t/\hbar\right) \tag{14.25}$$

which, combined with (14.24) and (14.23), yields

$$\frac{\mathrm{d}a_{\mathbf{g}}}{\mathrm{d}t} = \frac{1}{\mathrm{i}\,\hbar} \int a_{\mathbf{k}}\, h_{\mathbf{kg}}\, \exp\left[-\mathrm{i}\,(E_{\mathbf{k}} - E_{\mathbf{g}})\, t/\hbar\right]\, \mathrm{d}^3k\,, \tag{14.26}$$

the analogue of (14.6) and (14.15). However, a difference with respect to the discrete case exists, because a_n in (14.6) and $a_{n\gamma}$ in (14.15) are dimensionless quantities, whereas $a_{\mathbf{k}}$ in (14.26) is not. As a consequence, when the first-order perturbation method is used, and $a_{\mathbf{k}}$ within the integral of (14.26) is replaced with the initial condition, its expression contains one or more parameters whose values enter the final result. Given these premises, choose for the initial condition, e.g., a Gaussian function centered on some vector $\mathbf{b} \neq \mathbf{g}$,

$$a_{\mathbf{k}}(0) = \pi^{-3/4}\, \lambda^{3/2}\, \exp(-\lambda^2\, |\mathbf{k}-\mathbf{b}|^2/2)\,, \tag{14.27}$$

with $\lambda > 0$ a length. Inserting (14.27) into (14.22) yields the initial condition for the wave function (compare with (C.83)),

$$\psi(\mathbf{r}, 0) = \pi^{-3/4}\, \lambda^{-3/2}\, \exp[-r^2/(2\,\lambda^2) + \mathrm{i}\,\mathbf{b}\cdot\mathbf{r}]\,. \tag{14.28}$$

Both (14.27) and (14.28) are square integrable and normalized to unity for any positive λ; when the latter becomes large, $\psi(\mathbf{r}, 0)$ becomes more and more similar to a plane wave, while the peak of $a_{\mathbf{k}}(0)$ around $\mathbf{b}$ becomes narrower and higher. Assuming that the $\mathbf{k}$-dependence of $h_{\mathbf{kg}}$ is weaker than that of $a_{\mathbf{k}}(0)$, one replaces $\mathbf{k}$

with $\mathbf{b}$ in $h_{\mathbf{kg}}$ and $E_{\mathbf{k}}$, so that in (14.26) only the integral of $a_{\mathbf{k}}(0)$ is left, which yields $(2\sqrt{\pi}/\lambda)^{3/2}$. Completing the calculation as in Sects. 14.4 and 14.5, and remembering that $\mathbf{g} \neq \mathbf{b}$, provides

$$a_{\mathbf{g}}(t_P) \simeq \frac{(2\sqrt{\pi}/\lambda)^{3/2}}{\mathrm{i}\hbar} \int_0^{t_P} h_{\mathbf{bg}}(t)\, \exp\left(-\mathrm{i}\,\omega_{\mathbf{bg}}\, t\right) \mathrm{d}t\,, \qquad \omega_{\mathbf{bg}} = \frac{E_{\mathbf{b}} - E_{\mathbf{g}}}{\hbar}\,. \tag{14.29}$$

The product $\mathrm{d}P_{\mathbf{bg}} = |a_{\mathbf{g}}(t_P)|^2\, \mathrm{d}^3 g$ is the infinitesimal probability that at the end of the perturbation the outcome of an energy measurement belongs to the range of $E(\mathbf{g})$ values whose domain is $\mathrm{d}^3 g$.

Typical applications of (14.26) are encountered in the cases where the perturbation matrix is independent of time, $h_{\mathbf{bg}} = h^{(0)}_{\mathbf{bg}} = \text{const} \neq 0$. A calculation similar to that of Sect. 14.8.1 yields in this case

$$\mathrm{d}P_{\mathbf{bg}} = \frac{8\,\pi^{3/2}\,|h^{(0)}_{\mathbf{bg}}|^2}{\lambda^3\,\hbar^2} f(\omega_{\mathbf{bg}})\, \mathrm{d}^3 g\,, \qquad f(\omega_{\mathbf{bg}}) = \left[\frac{\sin(\omega_{\mathbf{bg}}\, t_P/2)}{\omega_{\mathbf{bg}}/2}\right]^2. \tag{14.30}$$

In place of the domain $\mathrm{d}^3 g$ one can consider the corresponding range of energy $\mathrm{d}E_{\mathbf{g}}$; for this, one profits by the concept of density of states introduced in Sect. B.5. Here the calculation is simple because the $E = E(\mathbf{g})$ relation is given by (14.21) so that, by the same calculation leading to (B.34), one obtains $\mathrm{d}^3 g = (1/2)\, \sin\vartheta\, \mathrm{d}\vartheta\, \mathrm{d}\varphi\, (2\,m/\hbar^2)^{3/2}\, \sqrt{E_{\mathbf{g}}}\, \mathrm{d}E_{\mathbf{g}}$. Considering that the initial state $\mathbf{b}$ and the duration t_P are prescribed, the factor $f(\omega_{\mathbf{bg}})$ in (14.30) depends only on $E_{\mathbf{g}}$, while $h^{(0)}_{\mathbf{bg}}$ may depend on the angles ϑ, φ (compare with Prob. 14.1). Integrating (14.30) over the angles and letting

$$H^{(0)}_{\mathbf{b}}(E_{\mathbf{g}}) = \int_0^{\pi} \int_0^{2\,\pi} |h^{(0)}_{\mathbf{bg}}|^2\, \sin\vartheta\, \mathrm{d}\vartheta\, \mathrm{d}\varphi\,, \tag{14.31}$$

yield the infinitesimal probability that a perturbation constant in time induces a transition from the initial condition (14.28) to a final state whose energy belongs to the range $\mathrm{d}E_{\mathbf{g}}$:

$$\mathrm{d}P_{\mathbf{b}} = \int_0^{\pi} \int_0^{2\,\pi} \mathrm{d}P_{\mathbf{bg}}\, \sin\vartheta\, \mathrm{d}\vartheta\, \mathrm{d}\varphi = \left(\frac{2\,\pi\,m}{\hbar^2}\right)^{3/2} \frac{4 f(\omega_{\mathbf{bg}})\, H^{(0)}_{\mathbf{b}}}{\lambda^3\,\hbar^2}\, \sqrt{E_{\mathbf{g}}}\, \mathrm{d}E_{\mathbf{g}}\,. \tag{14.32}$$

14.7 Screened Coulomb Perturbation

An important case of perturbation is that of a charged particle deflected by another charged particle fixed in the origin. The perturbation Hamiltonian is independent of time and, *in vacuo*, takes the form (3.31) of the Coulomb potential energy.[2] Though,

[2] This case is the quantum analogue of that treated in classical terms in Sect. 3.8.

the perturbation matrix (14.24) calculated using (3.31) diverges. Such an outcome is explained by observing that as the vanishing behavior of the Coulomb potential energy away from the origin is weak, the particle is actually subjected to it at large distances from the origin; as a consequence, using the free particle's eigenfunctions $\exp(\mathrm{i}\,\mathbf{k}\cdot\mathbf{r})/(2\,\pi)^{3/2}$ as solutions of the unperturbed Schrödinger equation is too strong an approximation. A more appropriate approach adopts for the perturbation Hamiltonian the *screened Coulomb potential energy*

$$\delta\mathscr{H} = \frac{A}{4\,\pi\,r}\,\exp(-q_c\,r)\,, \qquad r > 0\,, \qquad A = \frac{\kappa\,Z\,e^2}{\varepsilon_0}\,, \tag{14.33}$$

with $e > 0$ the elementary electric charge, Z a positive integer, ε_0 the vacuum permittivity, $q_c > 0$ the *inverse screening length*, and, finally, $\kappa = 1\,(-1)$ in the repulsive (attractive) case. The asymptotic vanishing of (14.33) is much stronger than that of the pure Coulomb case, and the resulting matrix elements are finite, as shown below. Although the choice of a screened potential energy is not realistic *in vacuo*, an expression like (14.33) is more appropriate than (3.31) when a solid material is considered (Sect. 20.5).

To calculate (14.24) one lets $\mathbf{q} = \mathbf{k} - \mathbf{g}$ and chooses a Cartesian reference such that $\mathbf{q}$ is aligned with the z axis: turning to spherical coordinates (B.1) transforms $\mathrm{d}^3 r$ into $r^2\,\sin\vartheta\,\mathrm{d}\vartheta\,\mathrm{d}\varphi\,\mathrm{d}r$ and $\mathbf{q}\cdot\mathbf{r}$ into $q\,r\,\cos\vartheta$. Letting $\mu = \cos\vartheta$, and observing that the integration over φ yields a $2\,\pi$ factor, one gets

$$h^{(0)}_{\mathbf{kg}} = \frac{A/2}{(2\,\pi)^3}\int_0^\infty \left[\int_{-1}^{+1} \exp(-\mathrm{i}\,q\,r\,\mu)\,\mathrm{d}\mu\right] \exp(-q_c\,r)\,r\,\mathrm{d}r = \frac{A/(2\,\pi)^3}{q_c^2 + q^2}\,. \tag{14.34}$$

14.8 Complements

14.8.1 Perturbation Constant in Time

The simplest example of the time-dependent perturbation theory of Sect. 14.1 occurs when the matrix elements h_{rs} are constant in time during the perturbation. In this case one lets $h_{rs} = h^{(0)}_{rs} = \text{const} \neq 0$ for $0 \leq t \leq t_P$, and $h_{rs} = 0$ elsewhere. From (14.10) it follows

$$h^{(0)}_{rs}\int_0^{t_P} \exp(-\mathrm{i}\,\omega_{rs}\,t)\,\mathrm{d}t = h^{(0)}_{rs}\,\exp(-\mathrm{i}\,\omega_{rs}\,t_P/2)\,\frac{\sin(\omega_{rs}\,t_P/2)}{\omega_{rs}/2}\,, \tag{14.35}$$

$$P_{rs} = \frac{|h^{(0)}_{rs}|^2}{\hbar^2} f(\omega_{rs})\,, \qquad f(\omega_{rs}) = \left[\frac{\sin(\omega_{rs}\,t_P/2)}{\omega_{rs}/2}\right]^2\,. \tag{14.36}$$

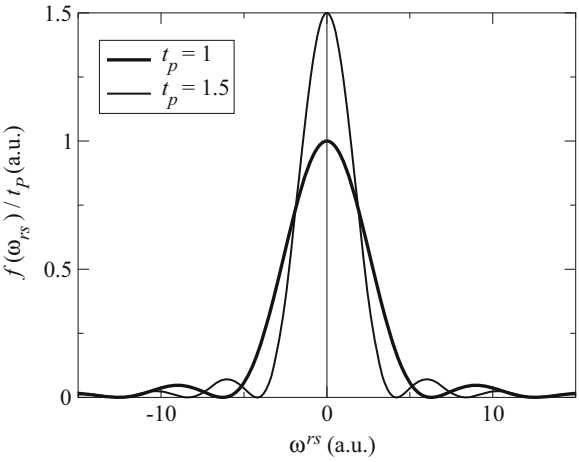


Fig. 14.1 Form of $f(\omega_{rs})/t_P$, with f given by the second expression in (14.36), for different values of t_P (in arbitrary units)

The form of $f(\omega_{rs})/t_P$ is shown in Fig. 14.1 in arbitrary units. The zeros of f nearest to $\omega_{rs} = 0$ are $\omega^+ = 2\,\pi/t_P$ and $\omega^- = -2\,\pi/t_P$, which provides the width $\omega^+ - \omega^- = 4\,\pi/t_P$ of the peak; in turn, the height of the peak is $f(\omega_{rs} = 0) = t_P^2$, thus indicating that the area of the peak is proportional to t_P. In fact, from (C.15) one obtains

$$\int_{-\infty}^{+\infty} f(\omega_{rs})\,\mathrm{d}\omega_{rs} = 2\,\pi\,t_P\,. \tag{14.37}$$

The form of $f(\omega_{rs})/t_P$ suggests that if t_P is sufficiently large, such a ratio may be approximated with a Dirac delta (Sect. C.4), namely, $f(\omega_{rs}) \approx 2\,\pi\,t_P\,\delta(\omega_{rs})$, where the coefficient $2\,\pi\,t_P$ is chosen for consistency with (14.37). To this purpose one also notes that due to the smallness of $\hbar$, the modulus of $\omega_{rs} = (E_r - E_s)/\hbar$ is very large (whence $f(\omega_{rs})$ is very small) unless $E_s = E_r$. Using the approximate form within the probability's definition (14.36) yields

$$P_{rs} \approx 2\,\pi\,\frac{|h_{rs}^{(0)}|^2}{\hbar^2}\,t_P\,\delta(\omega_{rs}) = 2\,\pi\,\frac{|h_{rs}^{(0)}|^2}{\hbar}\,t_P\,\delta(E_r - E_s)\,. \tag{14.38}$$

As expected, P_{rs} is invariant when r and s are interchanged (compare with (C.55) and comments therein). Differentiating (14.38) with respect to t_P yields the probability[3] per unit time of the transition from state r to state s:

$$\dot{P}_{rs} \approx 2\,\pi\,\frac{|h_{rs}^{(0)}|^2}{\hbar^2}\,\delta(\omega_{rs}) = 2\,\pi\,\frac{|h_{rs}^{(0)}|^2}{\hbar}\,\delta(E_r - E_s)\,. \tag{14.39}$$

[3]The expressions in terms of energy in (14.38,14.39) are obtained from $\delta(\omega)\,\mathrm{d}\omega = \delta(E)\,\mathrm{d}E = \delta(E)\,\mathrm{d}\hbar\omega$. Compare with the comments about the dimension of Dirac's δ made in Sect. C.5.

This shows that the particle's energy is approximately conserved when the perturbation lasts for a long time. The result is intuitive, because in the limit $t_P \to \infty$ a constant perturbation is equivalent to a shift in the potential energy, which makes the Hamiltonian operator conservative at all times. On the other hand, the conservation of energy does not imply the conservation of other dynamic quantities like, e.g., momentum (compare with the analysis of the two-particle collision carried out in classical terms in Sect. 3.6 and in quantum terms in Sects. 14.6 and 17.8).

14.8.2 Harmonic Perturbation

Another important example is that of the harmonic perturbation at an angular frequency $\omega_0 > 0$: in this case the matrix elements read $h_{rs} = h_{rs}^{(0)} \cos(\omega_0 t)$ for $0 \le t \le t_P$, $h_{rs}^{(0)} = \text{const} \neq 0$, and $h_{rs} = 0$ elsewhere. From (14.10) it follows

$$\int_0^{t_P} h_{rs} \exp(-\mathrm{i}\,\omega_{rs}\, t)\,\mathrm{d}t = \frac{h_{rs}^{(0)}}{2}\left[\exp\left(\frac{S\,t_P}{2\,\mathrm{i}}\right)\sigma(S) + \exp\left(\frac{D\,t_P}{2\,\mathrm{i}}\right)\sigma(D)\right], \tag{14.40}$$

with $S = \omega_{rs} + \omega_0$, $D = \omega_{rs} - \omega_0$, $\sigma(\eta) = \sin(\eta\, t_P/2)/(\eta/2)$. Comparing the definition of σ with (14.36) shows that $f = \sigma^2$, whence

$$P_{rs} = \frac{|h_{rs}^{(0)}|^2}{4\,\hbar^2} F(\omega_{rs})\,, \qquad F = f(S) + f(D) + 2\,\sigma(S)\,\sigma(D)\,\cos(\omega_0\, t_P)\,. \tag{14.41}$$

The form of $F(\omega_{rs})/t_P$ is shown in Fig. 14.2 in arbitrary units. The largest peaks correspond to $\omega_{rs} = \omega_0$ and $\omega_{rs} = -\omega_0$; if $t_P \gg 1/\omega_0$ holds, the two peaks are practically separate whence, using as in Sect. 14.8.1 the approximation $f \simeq 2\,\pi\, t_P\, \delta$, one finds

$$P_{rs} \simeq 2\,\pi\,\frac{|h_{rs}^{(0)}|^2}{4\,\hbar^2}\,[\delta(\omega_{rs} + \omega_0) + \delta(\omega_{rs} - \omega_0)]\; t_P\,. \tag{14.42}$$

Remembering that $\omega_{rs} = (E_r - E_s)/\hbar$, the transition probability per unit time due to the harmonic perturbation is finally found to be

$$\dot{P}_{rs} = 2\,\pi\,\frac{|h_{rs}^{(0)}|^2}{4\,\hbar}\,[\delta(E_r - E_s + \hbar\,\omega_0) + \delta(E_r - E_s - \hbar\,\omega_0)]\,. \tag{14.43}$$

When r and s are interchanged, the two summands between brackets replace each other. As $\omega_0 \neq 0$, the arguments of δ in (14.42) or (14.43) cannot vanish simultaneously. If the first vanishes it is $E_s = E_r + \hbar\,\omega_0 > E_r$, namely, the final energy E_s is larger than the initial one: the particle acquires the quantum of energy $\hbar\,\omega_0$ from the perturbing entity (*absorption*). If the second argument vanishes it

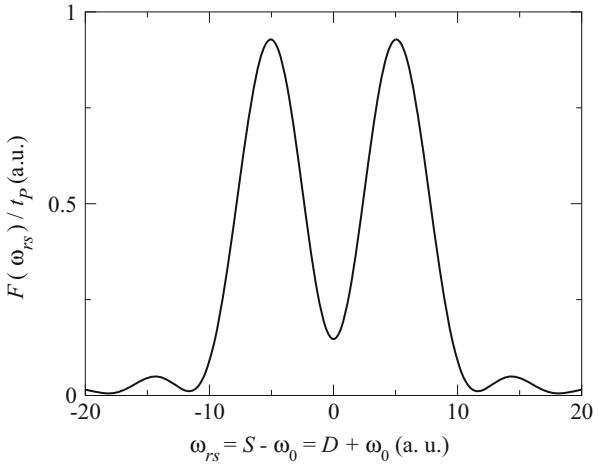


Fig. 14.2 Form of $F(\omega_{rs})/t_P$, with F given by the second expression in (14.41), with $t_P = 1$, $\omega_0 = 5$ (in arbitrary units)

is $E_s = E_r - \hbar\omega_0 < E_r$: the particle releases the energy quantum to the perturbing entity (*emission*). For example, the energy can be absorbed from, or emitted towards, an electromagnetic field; in this case the particle interacts with the mode at angular frequency ω_0 by absorbing or emitting a photon (Sect. 12.28). The total energy of the particle and field is conserved in both cases.[4] The same description applies to the interaction with a vibrational field; in this case the particle absorbs or emits a phonon (Sects. 12.5 and 17.8).

14.8.3 Fermi's Golden Rule

Expression (14.32) gives the infinitesimal probability that a perturbation constant in time induces a transition from the initial state **b** to a final state whose energy belongs to the range $\mathrm{d}E_{\mathbf{g}}$; it is an example of a more general expression denoted with *Fermi's Golden Rule*. Remembering from Sect. 14.8.1 that for a sufficiently large value of t_P, it is $f(\omega_{\mathbf{bg}}) \approx 2\,\pi\,t_P\,\delta(\omega_{\mathbf{bg}})$, one finds from (14.32)

$$\mathrm{d}P_{\mathbf{b}} \approx \left(\frac{2\,\pi\,m}{\hbar^2}\right)^{3/2} \frac{8\,\pi\,t_P\,\delta(E_{\mathbf{b}} - E_{\mathbf{g}})\,H^{(0)}_{\mathbf{b}}}{\lambda^3\,\hbar}\,\sqrt{E_{\mathbf{g}}}\,\mathrm{d}E_{\mathbf{g}}\,. \tag{14.44}$$

Dividing (14.44) by t_P provides the infinitesimal probability per unit time. Factor $\delta(E_{\mathbf{b}} - E_{\mathbf{g}})$ entails the conservation of energy; as the unperturbed Hamiltonian

[4] The spatial dependence of the field is embedded in $h^{(0)}_{rs}$.

operator, upon which the derivation of (14.44) is based, is that of a free particle, the relation $E_{\mathbf{g}} = E_{\mathbf{b}}$ combined with the second relation in (14.21) implies $g^2 = b^2$, namely, the modulus of momentum is also conserved. The result is the same as in the case of the classical treatment of a particle's collision with another particle having a much larger mass (Sect. 3.6).

14.8.4 Transitions from Discrete to Continuous Levels

The transition probability is calculated following a reasoning similar to that of Sect. 14.5 also in the case where the initial state is labeled by a discrete index and the final state belongs to a continuous set. A physical situation where such a transition may occur is that of a particle initially trapped within a well: with reference to Sect. 11.5, the energy levels are discrete if $E < 0$, whereas they are continuous for $E > 0$, namely, the spectrum is mixed (Sect. 8.4). A particle whose initial state belongs to the discrete set may absorb from the perturbation an amount of energy sufficient for reaching the continuous set of states, thus leaving the well. As the final energy belongs to a continuous set, the outcome of the calculation is the expression of an infinitesimal probability like in Sect. 14.6.

14.8.5 Harmonic Perturbation—General Case

The general case of harmonic perturbation derives from a perturbation Hamiltonian of the form (compare with the example of Sect. 14.8.2)

$$\delta\mathscr{H} = \mathscr{A}\, \exp(\mathrm{i}\,\omega_{\mathbf{q}}\, t) + \mathscr{A}^{\dagger}\, \exp(-\mathrm{i}\,\omega_{\mathbf{q}}\, t)\,, \qquad \omega_{\mathbf{q}} > 0\,, \tag{14.45}$$

in the time interval $0 \leq t \leq t_P$, where $\mathscr{A}$ is a Hermitean, time-independent operator having the dimensions of an energy, such that $\mathscr{A} = 0$ for $t < 0$ and $t > t_P$; the angular frequency depends on a three-dimensional vector $\mathbf{q}$ that is provisionally left unspecified. From the definition (14.45) it follows that $\delta\mathscr{H}$ is Hermitean. Defining the complex constants

$$a_{\mathbf{k}\,\mathbf{k}'} = \int_{\Omega} w^{*}_{\mathbf{k}'}\, \mathscr{A} w_{\mathbf{k}}\, \mathrm{d}^3 r\,, \qquad a^{*}_{\mathbf{k}'\,\mathbf{k}} = \int_{\Omega} w^{*}_{\mathbf{k}'}\, \mathscr{A}^{\dagger} w_{\mathbf{k}}\, \mathrm{d}^3 r\,, \tag{14.46}$$

with $\mathbf{k}$, $\mathbf{k}'$ three-dimensional indices, and applying (14.7), one finds the elements of the perturbation matrix, $h_{\mathbf{k}\,\mathbf{k}'}(t) = a_{\mathbf{k}\,\mathbf{k}'}\, \exp(\mathrm{i}\,\omega_{\mathbf{q}}\, t) + a^{*}_{\mathbf{k}'\,\mathbf{k}}\, \exp(-\mathrm{i}\,\omega_{\mathbf{q}}\, t)$. In (14.46), Ω is the domain where the eigenfunctions are defined. The Fourier transform of $h_{\mathbf{k}\,\mathbf{k}'}$ with respect to time is found to be

$$\mathscr{F} h_{\mathbf{k}\mathbf{k}'} = a_{\mathbf{k}\mathbf{k}'} \exp\left(\frac{D\,t_P}{2\,\mathrm{i}}\right) \sigma(D) + a^*_{\mathbf{k}'\mathbf{k}} \exp\left(\frac{S\,t_P}{2\,\mathrm{i}}\right) \sigma(S)\,, \tag{14.47}$$

with $S = \omega_{\mathbf{k}\mathbf{k}'} + \omega_{\mathbf{q}}$, $D = \omega_{\mathbf{k}\mathbf{k}'} - \omega_{\mathbf{q}}$, and $\sigma(\eta) = \sin(\eta\,t_P/2)/(\eta/2)$. Letting $f = \sigma^2$ and $a_{\mathbf{k}\mathbf{k}'} = c_{\mathbf{k}\mathbf{k}'} \exp(\mathrm{i}\,\varphi_{\mathbf{k}\mathbf{k}'})$, with $c_{\mathbf{k}\mathbf{k}'}$, $\varphi_{\mathbf{k}\mathbf{k}'}$ real quantities, the expression of $P_{\mathbf{k}\mathbf{k}'} = |\mathscr{F} h_{\mathbf{k}\mathbf{k}'}|^2/\hbar^2$ reads

$$P_{\mathbf{k}\mathbf{k}'} = \frac{1}{\hbar^2}\left[|a_{\mathbf{k}\mathbf{k}'}|^2 f(D) + |a_{\mathbf{k}'\mathbf{k}}|^2 f(S) + 2\,c_{\mathbf{k}\mathbf{k}'}\,\sigma(S)\,\sigma(D)\,\cos(\omega_{\mathbf{q}}\,t_P + \varphi_{\mathbf{k}\mathbf{k}'})\right]. \tag{14.48}$$

The discussion of the result is similar to that of Sect. 14.8.2: for a sufficiently large value of t_P it is $f \simeq 2\,\pi\,t_P\,\delta$; also, for $\omega_{\mathbf{k}\mathbf{k}'} = \omega_{\mathbf{q}}$ the quantity in brackets in (14.48) becomes

$$t_P^2\left[|a_{\mathbf{k}\mathbf{k}'}|^2 + |a_{\mathbf{k}'\mathbf{k}}|^2\,\frac{\sin^2(\omega_{\mathbf{q}}\,t_P)}{(\omega_{\mathbf{q}}\,t_P)^2} + 2\,c_{\mathbf{k}\mathbf{k}'}\,\frac{\sin(\omega_{\mathbf{q}}\,t_P)}{\omega_{\mathbf{q}}\,t_P}\,\cos(\omega_{\mathbf{q}}\,t_P + \varphi_{\mathbf{k}\mathbf{k}'})\right]. \tag{14.49}$$

The case $\omega_{\mathbf{k}\mathbf{k}'} = -\omega_{\mathbf{q}}$ yields a similar result with $|a_{\mathbf{k}\mathbf{k}'}|^2$ and $|a_{\mathbf{k}'\mathbf{k}}|^2$ interchanged. If $\omega_{\mathbf{q}} \gg 1/t_P$ holds, the two peaks of $P_{\mathbf{k}\mathbf{k}'}$ at $\omega_{\mathbf{q}}$ and $-\omega_{\mathbf{q}}$ are practically separated (compare with Fig. 14.2) whence, using $f \simeq 2\,\pi\,t_P\,\delta$, one finds

$$P_{\mathbf{k}\mathbf{k}'} \simeq \frac{2\,\pi}{\hbar^2}\left[|a_{\mathbf{k}\mathbf{k}'}|^2\,\delta(\omega_{\mathbf{k}\mathbf{k}'} - \omega_{\mathbf{q}}) + |a_{\mathbf{k}'\mathbf{k}}|^2\,\delta(\omega_{\mathbf{k}\mathbf{k}'} + \omega_{\mathbf{q}})\right] t_P\,. \tag{14.50}$$

Using $\hbar\,\omega_{\mathbf{k}\mathbf{k}'} = E_{\mathbf{k}} - E_{\mathbf{k}'}$ and taking the derivative of (14.50) with respect to t_P yield the probability per unit time of the $\mathbf{k} \rightarrow \mathbf{k}'$ transition:

$$\dot{P}_{\mathbf{k}\mathbf{k}'} = \frac{2\,\pi}{\hbar}\left[|a_{\mathbf{k}\mathbf{k}'}|^2\,\delta(E_{\mathbf{k}} - E_{\mathbf{k}'} - \hbar\,\omega_{\mathbf{q}}) + |a_{\mathbf{k}'\mathbf{k}}|^2\,\delta(E_{\mathbf{k}} - E_{\mathbf{k}'} + \hbar\,\omega_{\mathbf{q}})\right]. \tag{14.51}$$

Like in the case of (14.43), the arguments of the δ symbols in (14.51) cannot vanish simultaneously because $\omega_{\mathbf{q}} \neq 0$. If the first argument vanishes it is $E_{\mathbf{k}'} = E_{\mathbf{k}} - \hbar\,\omega_{\mathbf{q}} < E_{\mathbf{k}}$, namely, the particle undergoing the transition emits the quantum of energy $\hbar\,\omega_{\mathbf{k}}$. If the second argument vanishes it is $E_{\mathbf{k}'} = E_{\mathbf{k}} + \hbar\,\omega_{\mathbf{q}} > E_{\mathbf{k}}$, namely, the particle undergoing the transition absorbs the quantum of energy $\hbar\,\omega_{\mathbf{q}}$. This results may be viewed in two ways: if the particle is considered on its own, the result of the perturbation is that the particle energy is either increased by the amount $\hbar\,\omega_{\mathbf{k}}$ (*absorption*) or decreased by the same amount (*emission*). If, instead, the system made of the particle and the perturbing agent is viewed as a whole, the result of the perturbation is the exchange of the energy $\hbar\,\omega_{\mathbf{q}}$ between the two parts of the system, whose total energy is conserved.

The result established by (14.50) and (14.51) is based upon the obvious assumption that states $E_{\mathbf{k}}$ and $E_{\mathbf{k}'}$ exist, namely, that they are eigenvalues of the unperturbed Hamiltonian of the particle under investigation. It follows that an energy transition of the particle may occur only if, within its energy spectrum, there exist two states

fulfilling the relation $E_{\mathbf{k}} - E_{\mathbf{k}'} = |\hbar\,\omega_{\mathbf{q}}|$. If they do not exist, the arguments of the δ symbols in (14.50) or (14.51) never vanish and no transition occurs. In other terms, the particle is not able to exchange energy with the perturbing agent. This apparently surprising result is due to the limit $\tau_P \to \infty$. In fact, in the limit the frequency spectrum of the perturbation becomes monochromatic and only the energy $\hbar\,\omega_{\mathbf{q}}$ becomes available. Without the limit $\tau_P \to \infty$ the transition probability per unit time is found by taking the derivative of (14.48) with respect to t_P. The resulting spectrum extends over all frequencies; correspondingly, there are much more pairs of states $E_{\mathbf{k}}$ and $E_{\mathbf{k}'}$ between which the particle may undergo a transition due to the interaction with the perturbing agent, albeit with different probabilities. The analysis above applies in the same manner to the results of Sects. 14.8.1 and 14.8.2.

14.8.6 Spatially Oscillating Harmonic Perturbation

An interesting type of perturbation is that whose spatial part oscillates. In this case, operator $\mathscr{A}$ of (14.45) is simply multiplicative, $\mathscr{A} = W(\mathbf{r})$, so that

$$\delta\mathscr{H} = W\,\exp(\mathrm{i}\,\omega_{\mathbf{q}}\,t) + W^*\,\exp(-\mathrm{i}\,\omega_{\mathbf{q}}\,t)\,, \tag{14.52}$$

and the integrals to be calculated take the form (compare with (14.46)),

$$I = \int_{\Omega} w^*_{\mathbf{k}'}\,W(\mathbf{r})\,w_{\mathbf{k}}\,\mathrm{d}^3 r\,, \tag{14.53}$$

where the indices at the left-hand side are omitted for simplicity. Some results of this section are used in the description of the interaction between an electron and the lattice, given in Sect. 17.8; to proceed it is then necessary to anticipate some concepts that will be illustrated in more detail in Sect. 17.5.1. When the Hamiltonian operator is periodic, the eigenfunctions of the unperturbed problem have the form of *Bloch functions*, so that the expansion of the wave function takes the form (17.51). In the following, the Bloch functions belonging to a given branch of the dispersion relation are considered, so that the sum over index i in (17.51) is omitted, and indices $\mathbf{k}$, $\mathbf{k}'$ only are retained. In turn, Ω in (14.53) is the crystal volume. Dropping the branch index and letting $\mathbf{l}$ be any translation vector of the crystal (Sect. 17.2) give the Bloch functions the form

$$w_{\mathbf{k}} = u_{\mathbf{k}}(\mathbf{r})\,\exp(\mathrm{i}\,\mathbf{k}\cdot\mathbf{r})\,, \qquad u_{\mathbf{k}}(\mathbf{r}+\mathbf{l}) = u_{\mathbf{k}}(\mathbf{r})\,. \tag{14.54}$$

Due to the general properties of the Schrödinger equation, $w_{\mathbf{k}}$ fulfills the orthogonality condition $\int_{\Omega} w^*_{\mathbf{k}'}\,w_{\mathbf{k}}\,\mathrm{d}^3 r = \delta[\mathbf{k}-\mathbf{k}']$, where the right-hand side is the three-dimensional generalization of the Kronecker symbol (compare with (C.130)). Using the Bloch functions (14.54) transforms (14.53) into

$$I = \int_{\Omega} u^*_{\mathbf{k}'}\, u_{\mathbf{k}}\, W(\mathbf{r})\, \exp\left[-\mathrm{i}\, \left(\mathbf{k}' - \mathbf{k}\right) \cdot \mathbf{r}\right]\, \mathrm{d}^3 r\,, \tag{14.55}$$

showing that for the eigenfunctions considered here the perturbation-matrix entry is proportional to the spatial Fourier transform of $u^*_{\mathbf{k}'}\, u_{\mathbf{k}}\, W(\mathbf{r})$. The expression (14.55) is rather complicate due to the necessity of explicitly calculating the periodic part $u_{\mathbf{k}}(\mathbf{r})$ of the Bloch function for each crystal. The calculation becomes easier when W has a special form, as shown in the example below. Prior to specifying any form of W, one observes that the crystal volume Ω is the union of the volumes of the lattice cells, whose number is N_c. It follows $\Omega = \Omega_0 \cup \Omega_1 \cup \ldots \cup \Omega_{N_c-1}$, whence

$$I = \sum_{m=0}^{N_c-1} I_m\,, \qquad I_m = \int_{\Omega_m} w^*_{\mathbf{k}'}(\mathbf{r})\, W(\mathbf{r})\, w_{\mathbf{k}}(\mathbf{r})\, \mathrm{d}^3 r\,. \tag{14.56}$$

Letting Ω_0 denote the cell containing the origin of the reference, it is

$$I_m = \int_{\Omega_0} w^*_{\mathbf{k}'}(\mathbf{r} + \mathbf{l}_m)\, W(\mathbf{r} + \mathbf{l}_m)\, w_{\mathbf{k}}(\mathbf{r} + \mathbf{l}_m)\, \mathrm{d}^3 r\,, \tag{14.57}$$

where $\mathbf{l}_m$ is the direct-lattice vector of the mth cell (that is, the vector connecting the cell associated with the origin with the cell of index m), and $w_{\mathbf{k}} = u_{\mathbf{k}}(\mathbf{r})\, \exp(\mathrm{i}\, \mathbf{k} \cdot \mathbf{r})$. Due to the periodicity of $u_{\mathbf{k}}$ it is

$$I_m = \exp[\mathrm{i}\, (\mathbf{k} - \mathbf{k}') \cdot \mathbf{l}_m] \int_{\Omega_0} u^*_{\mathbf{k}'}(\mathbf{r})\, u_{\mathbf{k}}(\mathbf{r})\, W(\mathbf{r} + \mathbf{l}_m)\, \exp[\mathrm{i}\, (\mathbf{k} - \mathbf{k}') \cdot \mathbf{r}]\, \mathrm{d}^3 r\,. \tag{14.58}$$

An interesting case occurs when W has the form of the spatial part of a plane wave, $W(\mathbf{r}) = W_0\, \exp(\mathrm{i}\, \mathbf{k}_d \cdot \mathbf{r})$, with $\mathbf{k}_d$ a vector of the scaled reciprocal lattice; in this case, (14.58) becomes

$$I_m = W_0\, \exp[\mathrm{i}\, (\mathbf{k} - \mathbf{k}' + \mathbf{k}_d) \cdot \mathbf{l}_m] \int_{\Omega_0} u^*_{\mathbf{k}'}(\mathbf{r})\, u_{\mathbf{k}}(\mathbf{r})\, \exp[\mathrm{i}\, (\mathbf{k} - \mathbf{k}' + \mathbf{k}_d) \cdot \mathbf{r}]\, \mathrm{d}^3 r\,. \tag{14.59}$$

Relation Between Initial and Final State

In (14.59), the dependence of I_m on m is actually on the three integer indices[5] of $\mathbf{l}_m = m_1\, \mathbf{a}_1 + m_2\, \mathbf{a}_2 + m_3\, \mathbf{a}_3$ whence, letting

$$\mathbf{k} - \mathbf{k}' + \mathbf{k}_d = 2\,\pi\, \gamma_1\, \mathbf{b}_1 + 2\,\pi\, \gamma_2\, \mathbf{b}_2 + 2\,\pi\, \gamma_3\, \mathbf{b}_3 \tag{14.60}$$

[5] Here and in the following, $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ are the characteristic vectors of the Bravais lattice (Sect. 17.2), and $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ are the characteristic vectors of the reciprocal lattice (Sect. 17.3).

and using $\mathbf{a}_i \cdot \mathbf{b}_j = \delta_{ij}$ with δ_{ij} the Kronecker symbol (A.18), one finds, for the exponential factor multiplying the integral in (14.59), the expression $\exp[\mathrm{i}\,(\mathbf{k}-\mathbf{k}'+\mathbf{k}_d)\cdot\mathbf{l}_m] = \exp[2\,\pi\,\mathrm{i}\,(\gamma_1\,m_1+\gamma_2\,m_2+\gamma_3\,m_3)]$. It follows that (14.59) is recast as

$$I_m = W_0\,I_0\,\exp[2\,\pi\,\mathrm{i}\,(\gamma_1\,m_1+\gamma_2\,m_2+\gamma_3\,m_3)]\,, \tag{14.61}$$

where only the last factor depends on m. This expedites the conclusion of the analysis; in fact, to calculate the transition probability one adds up the cells' contributions, as prescribed by the first relation in (14.56), and takes the square modulus of the result. It is shown in Prob. 14.3 that if N_c is large (as is in the practical cases), it follows

$$\left|\sum_{m=0}^{N_c-1}\exp[\mathrm{i}\,(\mathbf{k}-\mathbf{k}'+\mathbf{k}_d)\cdot\mathbf{l}_m]\right|^2 \simeq N_c^2\sum_{\gamma}\delta[\mathbf{k}-\mathbf{k}'+\mathbf{k}_d-\mathbf{g}_\gamma]\,, \tag{14.62}$$

with $\mathbf{g}_\gamma$ any translation vector of the scaled reciprocal lattice. As shown in Prob. 14.3, the left-hand side of (14.62) is expressible as the product of three functions of the form $\sin^2(\pi\,\gamma_s\,N_s)/\sin^2(\pi\,\gamma_s)$, with γ_s any integer. Figure 14.3 shows a plot of one of them: the function is invariant when any integer is added to γ_s; also, it is appreciably different from zero only in the small regions around integer values of γ_s. Note that the value used for N_s in the figure is rather small; more realistic values of N_s provide even more pronounced peaks[6]: in other terms, the function is practically zero everywhere, except when γ_s is an integer; in this case the value of the function is N_s^2. Thus, the product of three such functions vanishes everywhere except at the points where γ_1, γ_2, γ_3 are simultaneously integer, in which

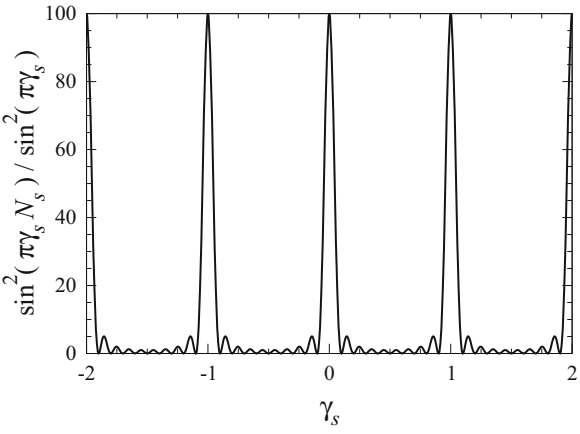


Fig. 14.3 Plot of $\sin^2(\pi\,\gamma_s\,N_s)/(\pi\,\gamma_s)^2$ with $N_s = 10$. The function is appreciable only around integer values of γ_s, even when the value of N_s is rather small

[6] In a cube of material with an atomic density of 6.4×10^{27} m^{-3}, the number of atoms per unit length in each direction is 4,000 μm^{-1}. The value of N_s is similar.

case the function equals $(N_1 N_2 N_3)^2 = N_c^2$ and $\mathbf{k} - \mathbf{k}' + \mathbf{k}_d$ equals any translation vector $\mathbf{g}_\gamma$ of the scaled reciprocal lattice. In conclusion, (14.62) holds (with the sum of the Kronecker symbol ranging over the three integers γ_1, γ_2, γ_3), and the vectors $\mathbf{k}$, $\mathbf{k}'$, and $\mathbf{k}_d$ fulfill the constraint

$$\mathbf{k} - \mathbf{k}' + \mathbf{k}_d = \mathbf{g}_\gamma = 2\,\pi\,\gamma_1\,\mathbf{b}_1 + 2\,\pi\,\gamma_2\,\mathbf{b}_2 + 2\,\pi\,\gamma_3\,\mathbf{b}_3\,. \tag{14.63}$$

As a consequence, the product $\mathbf{g}_\gamma \cdot \mathbf{l}_m/(2\,\pi) = m_1\,\gamma_1 + m_2\,\gamma_2 + m_3\,\gamma_3$ is an integer, thus making the exponential factor multiplying the integral in (14.59) equal to unity. Thanks to the above findings, the integral in (14.59) is given the more compact form

$$Y_{\mathbf{k}\,\mathbf{k}'}(\mathbf{g}_\gamma) = \int_{\Omega_0} u^*_{\mathbf{k}'}(\mathbf{r})\,u_{\mathbf{k}}(\mathbf{r})\,\exp(\mathrm{i}\,\mathbf{g}_\gamma \cdot \mathbf{r})\,\mathrm{d}^3 r\,, \tag{14.64}$$

whence the dependence on m disappears, and

$$\begin{aligned} I_m &= W_0\,Y_{\mathbf{k}\,\mathbf{k}'}(\mathbf{g}_\gamma)\,, \qquad & \mathbf{k} + \mathbf{k}_d &= \mathbf{k}' + \mathbf{g}_\gamma \\ I_m &= 0\,, & \mathbf{k} + \mathbf{k}_d &\neq \mathbf{k}' + \mathbf{g}_\gamma\,. \end{aligned} \tag{14.65}$$

The above are easily adapted to the case where W^* is used instead of W. In fact, one replaces W_0 with W_0^* and $\mathbf{k}_d$ with $-\mathbf{k}_d$. Defining the *overlap factor*

$$G_{\mathbf{k}\,\mathbf{k}'}(\mathbf{g}_\gamma) = |Y_{\mathbf{k}\,\mathbf{k}'}(\mathbf{g}_\gamma)|^2 \tag{14.66}$$

one also finds

$$\begin{aligned} |I|^2 &= |W_0|^2 N_c^2\,G_{\mathbf{k}\,\mathbf{k}'}(\mathbf{g}_\gamma)\,, \qquad & \mathbf{k} \pm \mathbf{k}_d &= \mathbf{k}' + \mathbf{g}_\gamma \\ |I|^2 &= 0\,, & \mathbf{k} \pm \mathbf{k}_d &\neq \mathbf{k}' + \mathbf{g}_\gamma\,, \end{aligned} \tag{14.67}$$

with $\mathbf{g}_\gamma$ any translation vector of the scaled reciprocal lattice. The positive (negative) sign holds when W (W^*) is used.

Selection Rules

It has been found in Sect. 14.8.6 that when the spatial part of the perturbation has a plane-wave form, the connection among the state of the electron before ($\mathbf{k}$) and after ($\mathbf{k}'$) the perturbation, and the wave vector $\mathbf{k}_d$ of the perturbing wave, is given by (14.67). One in fact assumes that $\mathbf{k}$ and $\mathbf{k}_d$ are given, whereas $\mathbf{k}'$ is to be found from (14.67). Therefore, the relation is a selection rule for the possible final states; as $\mathbf{g}_\gamma$ is an arbitrary translation vector, many final states are possible. The number of allowed transitions is strongly reduced by the constraint that $\mathbf{k}$, $\mathbf{k}'$, and $\mathbf{k}_d$ belong to the first Brillouin zone (Sect. 17.4). In this case (compare with (17.37)),

Table 14.1 Possible coefficients of $\mathbf{g}_\gamma$ when $\mathbf{k} + \mathbf{k}_d = \mathbf{k}' + \mathbf{g}_\gamma$ in (14.67)

γ_1	0	−1	0	0	0	−1	−1	−1
γ_2	0	0	−1	0	−1	0	−1	−1
γ_3	0	0	0	−1	−1	−1	0	−1

$$\mathbf{k} = \sum_{s=1}^{3} \frac{n_s}{N_s} 2\pi \mathbf{b}_s , \qquad n_s = 0, 1, \ldots, N_s - 1 . \tag{14.68}$$

A similar expression holds for $\mathbf{k}'$, with n_s replaced by another integer n'_s, and for $\mathbf{k}_d$, with n_s replaced by d_s, still an integer. A prismatic cell is assumed with one vertex coinciding with the origin; it follows that the coefficients of the selection rule are related by

$$n_s \pm d_s = n'_s + \gamma_s N_s , \tag{14.69}$$

with γ_s an arbitrary integer and $0 \le n_s,\ n'_s,\ d_s \le N_s - 1$. When the positive sign holds in (14.69), the range of possible values of the left-hand side is $0 \le n_s + d_s \le 2N_s - 2$. If the values of n_s, d_s are such that $0 \le n_s + d_s \le N_s - 1$, then the selection rule for the sth axis is fulfilled with $\gamma_s = 0$ and $n'_s = n_s + d_s$. If the values of n_s, d_s, instead, are such that $N_s \le n_s + d_s \le 2N_s - 2$, then the selection rule for the sth axis is fulfilled with $\gamma_s = -1$ and $n'_s = n_s + d_s - N_s$. Note that the outcome $\gamma_s = 0$ or $\gamma_s = -1$ occurs independently for each index s. As a consequence, the eight combinations shown in Table 14.1 are possible.

When the negative sign holds in (14.69), the range of possible values of the left-hand side is $1 - N_s \le n_s - d_s \le N_s - 1$. If the values of n_s, d_s are such that $0 \le n_s - d_s \le N_s - 1$, then the selection rule for the sth axis is fulfilled with $\gamma_s = 0$ and $n'_s = n_s - d_s$. If the values of n_s, d_s, instead, are such that $1 - N_s \le n_s - d_s \le -1$, then the selection rule for the sth axis is fulfilled with $\gamma_s = 1$ and $n'_s = n_s - d_s + N_s$. As a consequence, the eight combinations shown in Table 14.2 are possible. In all cases the values of γ_s and n'_s are unique once n_s and d_s are given. In other terms, for a given pair of vectors $\mathbf{k}$, $\mathbf{k}_d$ of the first Brillouin zone, the only $\mathbf{g}_\gamma$ vectors that make $\mathbf{k}'$ to belong to the first Brillouin zone as well are those listed in Table 14.1 or Table 14.2. If vector $\mathbf{g}_\gamma$ is null, then the transition from $\mathbf{k}$ to $\mathbf{k}'$ is called an *N-process* ("N" stands for "normal"). If $\mathbf{g}_\gamma$ is different from zero, then the transition is called a *U-process* ("U" stands for *Umklapp*, meaning "flip over"). Note that a process may be either N or U depending on how the scaled reciprocal space is divided into cells.[7]

The selection rules may also be considered from another viewpoint, namely, that of prescribing the initial state $\mathbf{k}$ and the final state $\mathbf{k}'$ of the transition, and determining the pairs $\mathbf{k}_d$, $\mathbf{g}_\gamma$ that make the transition possible. As before, the

[7] The term *Umklapp* derives from the German verb *umklappen*. R. Peierls in his autobiography [102, p. 43] states the he was the originator of the term in 1929, when he was studying crystal lattices.

Table 14.2 Possible coefficients of $\mathbf{g}_\gamma$ when $\mathbf{k} - \mathbf{k}_d = \mathbf{k}' + \mathbf{g}_\gamma$ in (14.67)

γ_1	0	1	0	0	0	1	1	1
γ_2	0	0	1	0	1	0	1	1
γ_3	0	0	0	1	1	1	0	1

constraint that $\mathbf{k}$, $\mathbf{k}'$, and $\mathbf{k}_d$ belong to the first Brillouin zone limits the choices of $\mathbf{g}_\gamma$ according to Table 14.1 or Table 14.2.

Like in Sects. 14.8.1, 14.8.2, and 14.8.5, the analysis is concluded by calculating the Fourier transform of the perturbation-matrix entry, in order to determine the relation between the initial and final energy (to be coupled with the relations between $\mathbf{k}$ and $\mathbf{k}'$ found in this section) and, from this, the transition probability from the initial to the final state. This part of the analysis is carried out in Sect. 17.8, with reference to the important application to the interaction of an electron with a periodic lattice.

Problems

14.1 Using (14.31) and (14.34), find $H^{(0)}_{\mathbf{b}}$ for the screened Coulomb perturbation. Assume for simplicity that the condition $g = b$ holds (Sect. 14.8.3).

14.2 Analyze the behavior of (14.51) upon exchange of the indices $\mathbf{k}$ and $\mathbf{k}'$.

14.3 Prove the limit shown in (14.62).

14.4 Prove that when N_c is large, it is $\sum_{m=0}^{N_c-1} \exp[\mathrm{i}\,(\mathbf{k}-\mathbf{k}'+\mathbf{k}_d)\cdot\mathbf{l}_m] \simeq N_c \sum_\gamma \delta[\mathbf{k}-\mathbf{k}'+\mathbf{k}_d-\mathbf{g}_\gamma]$, without resorting to the square modulus like in Prob. 14.3.

14.5 Calculate the overlap factor (14.66) when the potential energy V vanishes within Ω.

Part IV
Systems of Interacting Particles: Quantum Statistics

Chapter 15
Many-Particle Systems

15.1 Introduction

This chapter illustrates the properties of many-particle systems. The quantum-mechanical description of the latter is obtained by solving the time-dependent Schrödinger equation. After commenting the simplifications that occur when the Hamiltonian operator is separable, the important issue of the symmetry or anti-symmetry of the wave function is introduced for the purpose of illustrating the peculiar properties possessed by the systems of identical particles. Then, the concept of spin and the exclusion principle are introduced. After a general discussion, the above concepts are applied to the important case of a conservative system, and further properties related to the separability of the Hamiltonian operator are worked out. The remaining part of the chapter is devoted to the derivation of the equilibrium statistics in the quantum case (Fermi-Dirac and Bose-Einstein statistics). The connection between the microscopic statistical concepts and the macroscopic thermodynamic properties is illustrated in the complements, where two important examples of calculation of the density of states are also given.

15.2 Wave Function of a Many-Particle System

The quantum-mechanical concepts outlined in Parts II and III dealt with wave functions ψ describing a single particle. In such a case, if ψ is normalized to unity, the product $|\psi(\mathbf{r}, t)|^2 \, \mathrm{d}^3 r$ is the infinitesimal probability that at time t the particle's position belongs to the elementary volume $\mathrm{d}^3 r = \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z$ centered on $\mathbf{r}$; specifically, the x coordinate belongs to $\mathrm{d}x$, and so on. It is now necessary to extend the treatment to the case of many-particle systems. This is readily accomplished by considering, first, a system made of two particles: the wave function ψ describing such a system depends on two sets of coordinates, $\mathbf{r}_1$, $\mathbf{r}_2$, and time. The first set

M. Rudan, *Physics of Semiconductor Devices*, DOI 10.1007/978-3-319-63154-7_15

labels one of the particles, the second set labels the other particle. Assume that ψ is normalized to unity,

$$\int |\psi(\mathbf{r}_1, \mathbf{r}_2, t)|^2 \, \mathrm{d}^3 r_1 \, \mathrm{d}^3 r_2 = 1 \,, \tag{15.1}$$

where $\int$ is a short-hand notation for a six-fold integral over $\mathrm{d}x_1 \ldots \mathrm{d}z_2$. Then, the product $|\psi(\mathbf{r}_1, \mathbf{r}_2, t)|^2 \, \mathrm{d}^3 r_1 \, \mathrm{d}^3 r_2$ is the infinitesimal probability that at time t, set $\mathbf{r}_1$ belongs to $\mathrm{d}^3 r_1$ and set $\mathbf{r}_2$ belongs to $\mathrm{d}^3 r_2$. The wave function in (15.1) is the solution of the time-dependent Schrödinger equation

$$\mathrm{i}\hbar \frac{\partial \psi}{\partial t} = \mathscr{H}\psi \,, \qquad \psi(\mathbf{r}_1, \mathbf{r}_2, 0) = \psi_0(\mathbf{r}_1, \mathbf{r}_2) \,, \tag{15.2}$$

where the initial condition ψ_0 is prescribed. In turn, the Hamiltonian operator in (15.2) is derived, following the procedure illustrated in Sect. 10.2, from the Hamiltonian function that describes the two-particle system in the classical case. Considering by way of example a case where the forces acting on the two particles derive from a potential energy $V = V(\mathbf{r}_1, \mathbf{r}_2, t)$, the Hamiltonian function and the Hamiltonian operator read, respectively,

$$H = \frac{p_1^2}{2\,m_1} + \frac{p_2^2}{2\,m_2} + V \,, \qquad \mathscr{H} = -\frac{\hbar^2}{2\,m_1}\nabla_1^2 - \frac{\hbar^2}{2\,m_2}\nabla_2^2 + V \,, \tag{15.3}$$

where m_1, m_2 are the particles' masses, while $p_1^2 = p_{x1}^2 + p_{y1}^2 + p_{z1}^2$, $\nabla_1^2 = \partial^2/\partial x_1^2 + \partial^2/\partial y_1^2 + \partial^2/\partial z_1^2$, and the same for label 2.

It may happen that the Hamiltonian operator is separable with respect to the two sets $\mathbf{r}_1$, $\mathbf{r}_2$, namely, $\mathscr{H} = \mathscr{H}_1 + \mathscr{H}_2$ such that $\mathscr{H}_1$ does not contain any component of $\mathbf{r}_2$ and $\mathscr{H}_2$ does not contain any component[1] of $\mathbf{r}_1$. Also, let $\psi_1 = \psi_1(\mathbf{r}_1, t)$, $\psi_2 = \psi_2(\mathbf{r}_2, t)$ be solutions, respectively, of

$$\mathrm{i}\hbar \frac{\partial \psi_1}{\partial t} = \mathscr{H}_1 \psi_1 \,, \qquad \mathrm{i}\hbar \frac{\partial \psi_2}{\partial t} = \mathscr{H}_2 \psi_2 \,, \tag{15.4}$$

with the initial conditions $\psi_{10} = \psi_1(\mathbf{r}_1, 0)$, $\psi_{20} = \psi_2(\mathbf{r}_2, 0)$. Letting $\psi = \psi_1 \, \psi_2$ and using (15.4) yield

$$\mathrm{i}\hbar \frac{\partial \psi}{\partial t} - \mathscr{H}\psi = \psi_2 \left(\mathrm{i}\hbar \frac{\partial \psi_1}{\partial t} - \mathscr{H}_1 \psi_1 \right) + \psi_1 \left(\mathrm{i}\hbar \frac{\partial \psi_2}{\partial t} - \mathscr{H}_2 \psi_2 \right) = 0 \,, \tag{15.5}$$

showing that $\psi_1 \, \psi_2$ solves the Schrödinger equation for the particles' system, with $\psi_{10} \, \psi_{20}$ as initial condition. The concepts introduced in this section are readily

[1] By way of example, (15.3) is separable if $V = V_1(\mathbf{r}_1, t) + V_2(\mathbf{r}_2, t)$.

extended to the case of larger systems. Letting $N > 2$ be the number of particles, and still assuming that the system's wave function is normalized to unity,

$$\int |\psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N, t)|^2 \, \mathrm{d}^3 r_1 \, \mathrm{d}^3 r_2 \ldots \mathrm{d}^3 r_N = 1 \,, \tag{15.6}$$

the product $|\psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N, t)|^2 \, \mathrm{d}^3 r_1 \, \mathrm{d}^3 r_2 \ldots \mathrm{d}^3 r_N$ is the infinitesimal probability that at time t, set $\mathbf{r}_i$ belongs to $\mathrm{d}^3 r_i$, with $i = 1, 2, \ldots, N$. If the Hamiltonian operator is separable, the solution of the time-dependent Schrödinger equation of the system has the form $\psi = \psi_1(\mathbf{r}_1, t) \ldots \psi_N(\mathbf{r}_N, t)$. From (15.6) one also notes that the units of ψ depend on the number of particles involved; specifically, in (15.6) it is $[\psi] = \mathrm{cm}^{-3N/2}$ (compare with the discussion of Sect. 9.7.1).

15.3 Symmetry of Functions and Operators

The Hamiltonian operator and the wave function describing a many-particle system contain sets of coordinates like $\mathbf{r}_1$, $\mathbf{r}_2$ .... It is important to introduce a number of properties related to the exchange of two such sets within the operator or the wave function. The problem is tackled first in a rather abstract way; the applications to specific cases of interest are shown in Sect. 15.6.

Consider a function $f = f(q_1, q_2, \ldots, q_n)$, where q_k represents a group of coordinates.[2] Let $\mathscr{S}_{ij}$ be an operator such that [96, Chap. XIV.3]

$$\mathscr{S}_{ij} f(q_1, \ldots, q_i, \ldots, q_j, \ldots, q_n) = f(q_1, \ldots, q_j, \ldots, q_i, \ldots, q_n) \,, \tag{15.7}$$

namely, $\mathscr{S}_{ij}$ exchanges the names of the ith and jth group, leaving the rest unchanged. From the definition it follows $\mathscr{S}_{ij}^2 = \mathscr{S}_{ij} \mathscr{S}_{ij} = \mathscr{I}$. Now, let λ be an eigenvalue of $\mathscr{S}_{ij}$, and w an eigenfunction corresponding to it: $\mathscr{S}_{ij} w = \lambda w$. The following relations hold together:

$$\mathscr{S}_{ij}^2 w = w \,, \qquad \mathscr{S}_{ij}^2 w = \lambda^2 w \,, \tag{15.8}$$

the first due to the general property shown before, the second to the definition of λ and w. As a consequence, $\lambda = \pm 1$, namely, $\mathscr{S}_{ij}$ has two eigenvalues. As their modulus equals unity, $\mathscr{S}_{ij}$ is unitary (Sect. 8.6.2), namely, $\mathscr{S}_{ij}^{-1} = \mathscr{S}_{ij}^{\dagger}$.

The properties of the operator's eigenfunctions are found by letting $w^s = \mathscr{S}_{ij} w$, so that w^s is the function that results from exchanging the names of the ith and jth group of coordinates. Depending on the eigenvalue, two cases are possible: the first one is $\lambda = +1$, whence $\mathscr{S}_{ij} w = +1 \times w$ and $\mathscr{S}_{ij} w = w^s$ hold together, so that $w^s = w$; the second case is $\lambda = -1$, whence $\mathscr{S}_{ij} w = -1 \times w$ and $\mathscr{S}_{ij} w = w^s$ hold

[2] A "group" of coordinates may also consist of a single coordinate.

together, so that $w^s = -w$. A function such that $w^s = w$ is called *symmetric* with respect to indices ij, while a function such that $w^s = -w$ is called *antisymmetric* with respect to indices ij. In conclusion,

- all symmetric functions are eigenfunctions of $\mathscr{S}_{ij}$ belonging to $\lambda = +1$;
- all antisymmetric functions are eigenfunctions of $\mathscr{S}_{ij}$ belonging to $\lambda = -1$.

The set of eigenfunctions of $\mathscr{S}_{ij}$ is complete; in fact, for any function f it is

$$f = \frac{1}{2}(f + \mathscr{S}_{ij} f) + \frac{1}{2}(f - \mathscr{S}_{ij} f), \tag{15.9}$$

where the first term at the right-hand side is symmetric and the second one is antisymmetric, so that both terms at the right-hand side are eigenfunctions of $\mathscr{S}_{ij}$. This shows that any function is expressible as a linear combination of eigenfunction of $\mathscr{S}_{ij}$.

Only a specific pair ij of coordinate groups has been considered so far. On the other hand, it may happen that a function is symmetric (antisymmetric) with respect to all pairs of indices; in this case it is called *symmetric* (*antisymmetric*) with no further specification.

The definitions above extend to operators. For instance, an operator $\mathscr{A}$ is symmetric with respect to ij if $\mathscr{A}^s = \mathscr{S}_{ij}\mathscr{A} = \mathscr{A}$; it is symmetric without further specification if $\mathscr{A}^s = \mathscr{A}$ for any pair ij. Given a function f and an operator $\mathscr{A}$, and letting let $f^s = \mathscr{S}_{ij} f$, it is for all f,

$$(\mathscr{A} f)^s = \mathscr{S}_{ij}\mathscr{A} f = \mathscr{S}_{ij}\mathscr{A}\mathscr{S}_{ij}^{-1}\mathscr{S}_{ij} f = \mathscr{A}^s f^s . \tag{15.10}$$

If $\mathscr{A}$ is symmetric, replacing $\mathscr{A}^s = \mathscr{A}$ in (15.10) shows that $\mathscr{S}_{ij}\mathscr{A} f = \mathscr{A}\mathscr{S}_{ij} f$ namely, $\mathscr{S}_{ij}$ commutes with all symmetric operators.

The operator whose symmetry properties are of interest is typically the Hamiltonian one. Considering for instance a system of N particles interacting with each other through Coulomb interactions *in vacuo*, one has

$$\mathscr{H} = -\sum_{k=1}^{N} \frac{\hbar^2}{2 m_k} \nabla_k^2 + \frac{1}{2} \sum_{k=1}^{N} \sum_{\substack{s=1 \\ s \neq k}}^{N} \frac{e_k e_s}{4 \pi \varepsilon_0 |\mathbf{r}_k - \mathbf{r}_s|} \tag{15.11}$$

with ε_0 the vacuum permittivity, $\nabla_k^2 = \partial^2/\partial x_k^2 + \partial^2/\partial y_k^2 + \partial^2/\partial z_k^2$, and m_k, e_k the mass and charge of the kth particle, respectively. In general, this operator has no particular symmetry property; however, it is symmetric with respect to the groups of coordinates x_k, y_k, z_k when the particles are identical to each other ($m_1 = m_2 = \ldots = m_N$, $e_1 = e_2 = \ldots = e_N$).

15.4 Conservation of Symmetry in Time

Consider a wave function ψ expanded into a complete set of orthonormal functions w_k. Using as in Sect. 15.3 the symbols q_i for the groups of coordinates, one has

$$\psi(q_1, \ldots, q_n, t) = \sum_k a_k(t)\, w_k(q_1, \ldots, q_n)\,. \tag{15.12}$$

The wave function is assumed to be normalized to unity, so that from Parseval theorem (8.41) it follows

$$\langle \psi | \psi \rangle = \sum_k |a_k|^2 = 1 \tag{15.13}$$

at all times. Now assume that ψ is the solution of a Schrödinger equation deriving from a symmetric Hamiltonian operator, and that ψ is itself symmetric at some instant t' with respect to the pair ij. As the functions w_k of (15.12) are linearly independent, the symmetry of ψ entails that of w_k for all k. As a consequence, for the pair ij, w_k is an eigenfunction of the operator $\mathscr{S}_{ij}$ corresponding to $\lambda = 1$. Combining (15.13) with the definition (10.13) of the expectation value of the eigenvalues yields $\langle \lambda \rangle = 1$ at $t = t'$. In turn, due to symmetry, $\mathscr{H}$ commutes with $\mathscr{S}_{ij}$; this yields, for the time derivative (10.27) of the average value of the eigenvalues of $\mathscr{S}_{ij}$,

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle \lambda \rangle = -\mathrm{i}\,\hbar \int \psi^* \left(\mathscr{H}\mathscr{S}_{ij} - \mathscr{S}_{ij}\mathscr{H} \right) \psi \, \mathrm{d}q_1 \ldots \mathrm{d}q_n = 0\,, \tag{15.14}$$

namely, $\langle \lambda \rangle$ is conserved in time. The above calculation can be summarized as follows:

- Given a symmetric Hamiltonian $\mathscr{H}$, select a pair of indices ij. Due to commutativity, a complete set of eigenfunctions w_k of $\mathscr{H}$ exists, that belongs also to operator $\mathscr{S}_{ij}$.
- The eigenfunctions w_k of $\mathscr{H}$ can thus be separated into two sets, made of symmetric and antisymmetric functions, respectively.
- Let $\psi(q_1, \ldots, t) = \sum_k a_k(t) w_k(q_1, \ldots)$ be the wave function of the system described by $\mathscr{H}$, and let ψ be symmetric at some instant t'. It follows that the nonvanishing coefficients $a_k(t')$ in the expansion of $\psi(q_1, \ldots, q_n, t')$ are only those multiplying the symmetric eigenfunctions.

As $\langle \lambda \rangle = \int \psi^* \mathscr{S}_{ij} \psi \, \mathrm{d}q_1 \ldots q_n = 1$ at all times, the expansion of ψ is made in terms of the symmetric w_ks at all times. Hence, ψ is always symmetric with respect to the groups of coordinates of indices ij. The above reasoning can be repeated for all pairs of indices for which ψ is symmetric. Note that, in order to repeat the reasoning for different pairs of indices, say, ij and jk, one needs not assume that the

corresponding operators $\mathscr{S}_{ij}$, $\mathscr{S}_{jk}$ commute with each other (in fact, they typically do not commute). The analysis holds equally for the case where ψ, at time t', is antisymmetric with respect to two indices. In such case, it remains antisymmetric at all times with respect to them.

15.5 Identical Particles

It is interesting to ascertain whether the mathematical properties related to symmetry or antisymmetry, briefly discussed in Sects. 15.3 and 15.4, correspond to some physical property. This is indeed so, and is especially important when a system of identical particles is considered.

Take for instance a system of two identical particles interacting with each other.[3] In classical terms, the two identical objects that form the system can always be made distinguishable from each other, without disturbing their motion. The typical example is that of two identical billiard balls that are made distinguishable by a different coloring; although the latter has no influence on the balls' dynamics, it allows one to distinguish them from each other irrespective of the number of collisions they undergo. As a consequence, the conjugate variables describing the motion of each particle (e.g., position and momentum) are exactly known at each instant of time. By way of example consider the collision of two identical, charged particles schematically illustrated in Fig. 15.1: it is assumed that the particles are initially far apart, with equal moduli of the initial velocities; the initial velocity of the particle labeled 1 is described by the upper-left arrow visible in cases *a* and *b* of the figure, while the initial velocity of the particle labeled 2 is described by the lower-right arrow. As they come closer, the particles repel each other due to the

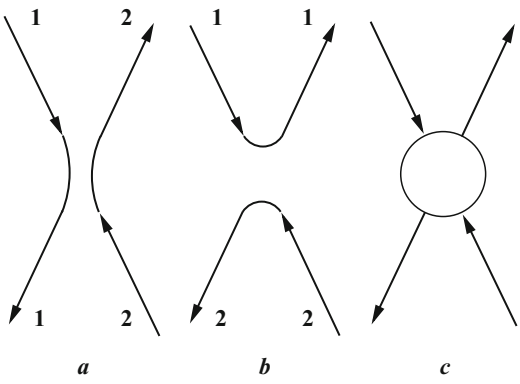


Fig. 15.1 Schematic description of a system made of two identical particles

[3]The reasoning outlined here does not apply to systems where the particles are different, i.e., they can be distinguished in $|\psi|^2$, e.g., by the mass or electric charge.

Coulomb interaction, and their classical motion is described as in Sects. 3.5–3.8. If the initial velocities were exactly aligned (this case is not shown in the figure), the two particles would bounce back along the same direction; if, however, the initial velocity of particle 1 was slightly misaligned to the left, the collision would yield case *a*, whereas it would yield case *b* for a right misalignment. Even if the misalignment is made as small as we please, either case *a* or *b* occurs, and the two possible outcomes are distinguishable from each other.

In the quantum-mechanical description, instead, it is not possible to track each particle separately, because the information about the system's dynamics derives from the wave function; when the particles come closer, the norm of the wave function in (15.1) is significantly different from zero in a finite region of space, which is schematically indicated by the circle of case *c* in Fig. 15.1. Due to the Heisenberg principle (10.22), it is impossible to determine the position and momentum of each particle at the same time with arbitrary precision. It follows that, for identical particles, the collision is described as an event where two particles enter the circle and two particles eventually leave it, without the possibility of distinguishing between cases *a* and *b*: the two cases must in fact be counted as one, and the wave function describing the system must be consistent with it. This requires $|\psi|^2$ be symmetric with respect to the groups $\mathbf{r}_1, \mathbf{r}_2$:

$$|\psi^s|^2 = |\mathscr{S}_{12}\psi|^2 = |\psi|^2 \,, \tag{15.15}$$

which implies $\psi^s = \exp(\mathrm{i}\alpha)\,\psi$ with α a real constant. On the other hand, remembering the first equation in (15.8),

$$\mathscr{S}_{12}^2\psi = \exp(2\,\mathrm{i}\alpha)\,\psi \,, \qquad \mathscr{S}_{12}^2\psi = \psi \,, \tag{15.16}$$

whence $\exp(j\alpha) = \pm 1$. In conclusion, when the two particles of the system are indistinguishable from each other, the system's wave function is either symmetric or antisymmetric. One may argue that when the system is made of more than two identical particles, its wave function could be symmetric with respect to some pairs of indices and antisymmetric with respect to other pairs. However, this is not possible, as the simple case of three identical particles shows [86, Sect. 26]. Assume that ψ is symmetric with respect to $\mathbf{r}_1, \mathbf{r}_2$ and antisymmetric with respect to $\mathbf{r}_1, \mathbf{r}_3$; it follows

$$\psi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, t) = \psi(\mathbf{r}_2, \mathbf{r}_1, \mathbf{r}_3, t) = -\psi(\mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_1, t) = -\psi(\mathbf{r}_1, \mathbf{r}_3, \mathbf{r}_2, t) = \tag{15.17}$$

$$= \psi(\mathbf{r}_3, \mathbf{r}_1, \mathbf{r}_2, t) = \psi(\mathbf{r}_3, \mathbf{r}_2, \mathbf{r}_1, t) = -\psi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, t) \,, \tag{15.18}$$

that is, a contradiction. In other terms, ψ can either be symmetric with respect to all the identical particles of the system, or antisymmetric with respect to all of them. This result has a far-reaching consequence, namely, the class of all particles is made of two subclasses: the first one collects the types of particles whose systems are

described by symmetric wave functions; they are called *Bose particles* or *bosons*. The second subclass collects the types of particles whose systems are described by antisymmetric wave functions; they are called *Fermi particles* or *fermions*. This applies to all known particles, including composite ones (e.g., atoms) and particles describing collective motions (e.g., phonons, photons).[4]

15.5.1 Spin

The properties discussed so far in this section bring about an interesting question: consider a single particle, e.g., an electron. Although electrons are fermions,[5] here the concept of symmetry or antisymmetry does not apply because the wave function of a single electron contains only one group of coordinates. Then add a second electron, so that a system of two identical particles is formed: how do these electrons "know" that, when paired, the wave function of their system must be antisymmetric? It is reasonable to assume that each particle in itself must possess a property that makes it to behave like a fermion or a boson within a system of particles identical to it. Such a property, called *spin*, does in fact exist; as its existence can be proven only within the frame of the relativistic quantum theory [42, 43], [124, Sect. 15], which is beyond the scope if this book, only a brief illustration of spin's properties of interest will be given.

In contrast with the other dynamic quantities considered so far, there is no classical counterpart of spin. Therefore, the latter cannot be derived from the expression of a dynamic variable by replacing conjugate coordinates with suitable operators. It can be shown that the eigenvalues of spin are derived in a manner similar to that of angular momentum: this leads, like in Sect. 13.5.1, to determining the square modulus of spin, σ^2, and its component along one of the coordinate axes, say, σ_z. Their values are given by expressions similar to (13.53), specifically,

$$S^2 = \hbar^2 \, s\,(s+1)\,, \qquad S_z = \hbar \, s_z\,. \tag{15.19}$$

The important difference with (13.53) is that s, instead of being a nonnegative integer, is a nonnegative half integer: $s = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots$; in turn, s_z can take the $2\,s + 1$ values $-s, -s+1, \ldots, s-1, s$.

The introduction of spin must be accounted for in the expression of the wave function: the latter, in the case of a single particle, must be indicated with $\psi(\mathbf{r}, s_z, t)$, and its normalization to unity, if existing, is expressed by

[4] The names "bosons," "fermions" of the two subclasses have this origin: when a system of identical particles is in thermodynamic equilibrium, the particles' energy follows a statistical distribution whose expression is named after Bose and Einstein (Sect. 15.8.2) and, respectively, Fermi and Dirac (Sect. 15.8.1).

[5] It must be noted, however, that in condensed-matter physics two electrons or other fermions may bind together at low temperatures to form a so-called Cooper pair, which turns out to have an integer spin, namely, it is a composite boson [30].

$$\sum_{s_z} \int_{\Omega} |\psi(\mathbf{r}, s_z, t)|^2 \, \mathrm{d}^3 r = 1 \,. \tag{15.20}$$

If (15.20) holds, the product $|\psi(\mathbf{r}, s_z, t)|^2 \, \mathrm{d}^3 r$ is the probability that at time t the particle is in the elementary volume $\mathrm{d}^3 r$ centered on $\mathbf{r}$, and the component of its spin along the z axis is $S_z = \hbar \, s_z$.

The connection between spin and boson-like or fermion-like behavior is the following: the quantum number s is integer for bosons, half integer for fermions [101]. It is then meaningful to use the terms "boson" or "fermion" for an individual particle. All known fermions have $s = 1/2$, whence $2\,s + 1 = 2$. It follows that for fermions the z-component of spin has two possible values, $\hbar/2$ (*spin up*) and $-\hbar/2$ (*spin down*). As anticipated above, electrons are fermions. Photons are bosons with $s = 1$.

The similarity between the expressions of the quantum numbers for spin and those of the angular momentum (Eqs. (13.53) cited above) is the origin of the qualitative visualization of spin in classical terms: spin is described as an intrinsic angular momentum of the particle, as if the particle was a sphere spinning on its axis.

15.6 Pauli Exclusion Principle

Consider a system of identical particles, so that its Hamiltonian operator $\mathscr{H}$ is symmetric with respect to each pair of particle labels ij. Its wave function, in turn, is either symmetric or antisymmetric depending on the nature of the particles forming the system. It may happen that when solving the Schrödinger equation of the system, a solution is found, say φ, that does not possess the necessary symmetry properties. One can then exploit a relation like (15.9) to construct from φ another solution which is either symmetric or antisymmetric. For this, one must remember that φ depends on the groups of coordinates $(\mathbf{r}_1, s_{z1})$, $(\mathbf{r}_2, s_{z2}), \ldots$ which, for the sake of conciseness, will be indicated with the symbol $\mathbf{q}_i = (\mathbf{r}_i, s_{zi})$. Remembering from Sect. 15.3 that the Hamiltonian operator, due to its symmetries, commutes with any operator $\mathscr{S}_{ij}$, one finds

$$\mathscr{S}_{ij} \left(\mathscr{H} - \mathrm{i}\,\hbar \, \frac{\partial}{\partial t} \right) \varphi = \left(\mathscr{H} - \mathrm{i}\,\hbar \, \frac{\partial}{\partial t} \right) \varphi^s \,, \tag{15.21}$$

where $\mathscr{S}_{ij}$ exchanges $\mathbf{q}_i$ with $\mathbf{q}_j$. The parenthesis on the left-hand side of (15.21) is zero because φ solves the Schrödinger equation; it follows that φ^s is also a solution. Due to the linearity of the Schrödinger equation, the two functions

$$\varphi + \varphi^s \,, \qquad \varphi - \varphi^s \,, \tag{15.22}$$

are solutions of the Schrödinger equation, which are also symmetric and, respectively, antisymmetric with respect to the pair ij. The procedure is easily generalized to obtain a wave function that is symmetric or antisymmetric with respect to all pairs of indices. To this purpose, one considers the $N!$ permutations of the system's N particles; let ν be an index representing the order of each permutation with respect to the fundamental one ($1 \leq \nu \leq N!$), and $\mathscr{S}_\nu$ the operator that achieves the νth permutation of the particles' coordinates within φ. The functions

$$\psi = a \sum_\nu \mathscr{S}_\nu \varphi \,, \qquad \psi = b \sum_\nu (-1)^\nu \mathscr{S}_\nu \varphi \tag{15.23}$$

are solutions of the Schrödinger equation, which are also symmetric and, respectively, antisymmetric with respect to all particles' permutations. Symbols a and b denote two constants, that can be used to normalize ψ if φ is normalizable. The above constructions can be worked out at any instant, as the symmetry or antisymmetry of the wave function is conserved in time (Sect. 15.4). The second relation in (15.23) lends itself to an interesting derivation. Considering for simplicity the case $N = 2$ one finds

$$\psi = b \left[\varphi(\mathbf{r}_1, s_{z1}, \mathbf{r}_2, s_{z2}, t) - \varphi(\mathbf{r}_2, s_{z2}, \mathbf{r}_1, s_{z2}, t)\right] , \tag{15.24}$$

If it were $\mathbf{r}_2 = \mathbf{r}_1$, $s_{z2} = s_{z1}$, the wave function (15.24) would vanish, which is not acceptable. The same unphysical results are found by letting $\mathbf{r}_j = \mathbf{r}_i$, $s_{zj} = s_{zi}$ in the second expression in (15.23). The conclusion is that in a system of identical fermions, two (or more) particles with the same spin cannot occupy the same position; this finding derives solely from the antisymmetry of the wave function for a system of identical fermions, and is called *Pauli principle* or *exclusion principle*.[6] As shown in Sect. 15.7 it can be restated in different forms depending on the system under consideration. No similar restriction applies to systems of identical bosons, as the form of the first relation in (15.23) shows.

15.7 Conservative Systems of Particles

An important example of a system of N interacting particles occurs when the forces are conservative. To begin, the general case of nonidentical particles is considered. The Hamiltonian function and the corresponding Hamiltonian operator read, respectively:

[6] Like the Heisenberg principle illustrated in Sect. 10.6, that of Pauli was originally deduced from heuristic arguments. The analysis of this section shows in fact that it is a theorem rather than a principle.

$$H = \sum_{i=1}^{N} \frac{p_i^2}{2m_i} + V\,, \qquad \mathscr{H} = -\sum_{i=1}^{N} \frac{\hbar^2}{2\,m_i}\nabla_i^2 + V\,, \tag{15.25}$$

where the symbols are the same as in (15.3). Here the potential energy depends only on the spatial coordinates, $V = V(\mathbf{r}_1, \ldots, \mathbf{r}_N)$. If the system is in a state of definite and constant energy E_S, its wave function reads

$$\varphi = W\,\exp(-\mathrm{i}\,E_S\,t/\hbar)\,, \qquad W = W(\mathbf{q}_1, \ldots \mathbf{q}_N)\,, \tag{15.26}$$

where E_S is an eigenvalue of $\mathscr{H}W = E\,W$. Extending to this case the definition of Sect. 15.2, the system is separable if $V = \sum_i V_i(\mathbf{r}_i)$, which gives the Hamiltonian operator the form

$$\mathscr{H} = \sum_{i=1}^{N} \mathscr{H}_i\,, \qquad \mathscr{H}_i = -\frac{\hbar^2}{2\,m_i}\nabla_i^2 + V_i(\mathbf{r}_i)\,. \tag{15.27}$$

Assuming that the eigenvalues are discrete, the ith Hamiltonian yields the single-particle equations

$$\mathscr{H}_i w_{n(i)} = E_{n(i)} w_{n(i)}\,, \tag{15.28}$$

where index $n(i)$ denotes the nth eigenvalue of the ith particle. From the general properties of operators (Sect. 10.3) it follows that each eigenfunction of the whole system is the product of eigenfunctions like $w_{n(i)}$,

$$W = w_{n(1)}(\mathbf{q}_1)\, w_{n(2)}(\mathbf{q}_2)\, \ldots\, w_{n(N)}(\mathbf{q}_N)\,, \tag{15.29}$$

and the eigenvalue of $\mathscr{H}$ is the sum of eigenvalues like $E_{n(i)}$:

$$E_S = E_{n(1)} + E_{n(2)} + \ldots + E_{n(N)}\,. \tag{15.30}$$

If the particles are identical, $m_1 = m_2 = \ldots = m_N$, then the single-particle Hamiltonian operators $\mathscr{H}_i$ become identical to each other; as a consequence, each eigenvalue equation like (15.28) produces the same set of eigenvalues and eigenfunctions. It follows that all $N!$ permutations of the indices $1, 2, \ldots, N$ in (15.30) leave the total energy unchanged. On the other hand, as for any pair of groups $\mathbf{q}_r$, $\mathbf{q}_s$ it is

$$w_{n(r)}(\mathbf{q}_r)\, w_{n(s)}(\mathbf{q}_s) \neq w_{n(s)}(\mathbf{q}_r)\, w_{n(r)}(\mathbf{q}_s)\,, \tag{15.31}$$

the total eigenfunction is changed by a permutation of the coordinate indices. Thus, to E_S there correspond $N!$ eigenfunctions w, namely, the eigenvalues of $\mathscr{H}W = E\,W$ for a system of identical particles are $N!$-fold degenerate.

As noted in Sect. 15.6, the solution (15.26) of the system's Schrödinger equation is not necessarily symmetric or antisymmetric. A solution with the correct symmetry property is found from (15.23) and has the form

$$\psi = a \exp(-\mathrm{i}\, E_S\, t/\hbar) \sum_{\nu} \mathscr{S}_\nu w_{n(1)}(\mathbf{q}_1) \dots w_{n(N)}(\mathbf{q}_N) \tag{15.32}$$

in the symmetric case, and

$$\psi = b \exp(-\mathrm{i}\, E_S\, t/\hbar) \sum_{\nu} (-1)^\nu \mathscr{S}_\nu w_{n(1)}(\mathbf{q}_1) \dots w_{n(N)}(\mathbf{q}_N) \tag{15.33}$$

in the antisymmetric one. It is worth specifying that $\mathscr{S}_\nu$ acts on the coordinate groups $\mathbf{q}_1, \dots, \mathbf{q}_N$, not on the indices $n(1), \dots, n(N)$.

When the wave function has the form (15.32) or (15.33), and the eigenfunctions $w_{n(1)}, \dots, w_{n(N)}$ are normalized to unity, the constants a and b are readily found. Considering the symmetric case with $N = 2$, the wave function reads

$$\psi = a \exp(-\mathrm{i}\, E_S\, t/\hbar) \left[w_{n(1)}(\mathbf{q}_1)\, w_{n(2)}(\mathbf{q}_2) + w_{n(2)}(\mathbf{q}_1)\, w_{n(1)}(\mathbf{q}_2) \right], \tag{15.34}$$

where it is assumed that the single-particle eigenfunctions are normalized to unity:

$$\sum_{s_{z1}} \int |w_{n(i)}(\mathbf{r}_1, s_{z1})|^2 \, \mathrm{d}^3 r_1 = 1\,, \qquad \sum_{s_{z2}} \int |w_{n(i)}(\mathbf{r}_2, s_{z2})|^2 \, \mathrm{d}^3 r_2 = 1\,. \tag{15.35}$$

In (15.35) it is $i = 1, 2$, and the indication of the domain of $\mathbf{r}_1$, $\mathbf{r}_2$ is omitted. As $n(2) \neq n(1)$, the pairs of eigenfunctions with such indices are mutually orthogonal (Sect. 8.4.1), whence

$$\sum_{s_{z1}} \sum_{s_{z2}} \iint |\psi(\mathbf{r}_1, s_{z1}, \mathbf{r}_2, s_{z2})|^2 \, \mathrm{d}^3 r_1 \, \mathrm{d}^3 r_2 = 2\,|a|^2\,. \tag{15.36}$$

Imposing the normalization of ψ to unity, and observing that the phase factor in a is irrelevant, yield $a = 1/\sqrt{2}$. The treatment of (15.33) is identical and yields the same result for b. By the same token, one finds $a = b = 1/\sqrt{N!}$ in the N-particle case. Still with reference to the antisymmetric wave function (15.33), one notes that its spatial part can be recast as a determinant,

$$\sum_{\nu} (-1)^\nu \mathscr{S}_\nu w_{n(1)}(\mathbf{q}_1) \dots w_{n(N)}(\mathbf{q}_N) = \begin{bmatrix} w_{n(1)}(\mathbf{q}_1) & \cdots & w_{n(N)}(\mathbf{q}_1) \\ \vdots & \ddots & \vdots \\ w_{n(1)}(\mathbf{q}_N) & \cdots & w_{n(N)}(\mathbf{q}_N) \end{bmatrix}, \tag{15.37}$$

that is called *Slater determinant*. A transposition of two particles involves the exchange of the corresponding coordinate sets, but not of the eigenfunction indices;

this is equivalent to exchanging two rows of the determinant, whence the change of sign. Also, if two or more particles belonged to the same state (including spin), two or more columns of the Slater determinant would be equal to each other and the wave function (15.33) would vanish. This is another form of the proof of Pauli's exclusion principle.

15.8 Equilibrium Statistics in the Quantum Case

This section illustrates the quantum-mechanical treatment of a system of particles in a condition of macroscopic equilibrium. The approach is the same as that outlined in Sect. 6.3 for the classical case; however, the constraints to which the particles are subjected are different. Here the term "particle" is used in a broader meaning, incorporating, e.g., also the case of photons (Sect. 12.3) and phonons (Sect. 16.6). As in Sect. 6.3 one considers a conservative system of identical particles, having a total energy E_S, enclosed in a stationary container of volume Ω. The conservation of the total energy introduces a first constraint, identical to (6.10):

$$F_E(N_1, N_2, \ldots) = 0\,, \qquad F_E = E_S - \sum_i N_i E_i\,. \tag{15.38}$$

The constraint identical to (6.9),

$$F_N(N_1, N_2, \ldots) = 0\,, \qquad F_N = N - \sum_i N_i\,, \tag{15.39}$$

describing the conservation of the total number of particles, may, instead, be fulfilled or not depending on the type of particles. For instance, a system of photons does not fulfill it: in fact, a photon may be absorbed by the container's wall and, say, two photons may be emitted by it, such that the energy of the emitted photons equals that of the absorbed one. In this way, constraint (15.38) applies, whereas constraint (15.39) does not. Another difference from the classical treatment is that, as remarked in Sect. 15.5, identical quantum particles belonging to a system are not distinguishable from each other; as a consequence, the method of counting their placement into the cells of the phase space (illustrated in Sect. 6.3 for the classical case) is different here. A further distinction must be made between the cases of systems made of fermions, to which the exclusion principle applies, and systems made of bosons, to which it does not apply. To appreciate the strong differences that are introduced by the constraints due to indistinguishability and exclusion, the example below is of help.

Consider a system made of three identical particles, whose total energy E_S is, in some units η, equal to $3\,\eta$. For simplicity, the particles are considered spinless

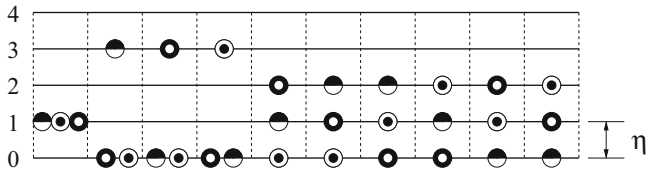


Fig. 15.2 Placement of three identical particles into equally spaced energy states. The particles' total energy equals three energy units η. Different graphic symbols are used for the particles to make the classical treatment clearer

and, instead of the phase space, a space made of energy cells[7] is used, where the energies of the particles are assumed to be quantized starting from a ground level. As a consequence, the energies allowed to the particles are 0, η, $2\,\eta$, $3\,\eta, \ldots$; the energy levels are shown in Fig. 15.2, where the particles are drawn with different graphics to make them provisionally distinguishable. If the system is considered as classical, it is easily found that there are ten possible ways of placing the three particles into the available states in such a way as to fulfill the energy constraint $E_S = 3\,\eta$. One way is to place all three particles in the η state. Another choice is to place one particle in the $3\,\eta$ state and the remaining two in the ground state; this provides three different ways, as there are three distinct possibilities to choose the particle to be placed in the $3\,\eta$ state. The last choice is to place one particle in the ground state, another particle in the η state, and the remaining one in the $2\,\eta$ state; this yields 3! ways as shown in the figure. If, instead, the particles are bosons, the three former combinations with energies $(0, 0, 3\eta)$ reduce to a single one because of the particles' indistinguishability; by the same token, the six former combinations with energies $(0, \eta, 2\eta)$ reduce to a single one as well. This gives a total of three ways for bosons, in contrast with the ten ways of the classical particles. Finally, if the particles are fermions, the combinations (η, η, η) and $(0, 0, 3\eta)$ must be excluded because the exclusion principle forbids one or more fermions to occupy the same state; as a consequence, the only way left for fermions is $(0, \eta, 2\eta)$ that fulfills both indistinguishability and exclusion.

Coming back to the general case, consider a system in thermodynamic equilibrium at temperature T, with E_S the system's total energy. As in the classical case outlined in Sect. 6.3, the system is considered *dilute*, namely such that the mutual interaction among the particles, albeit necessary for the existence of the system, is weak enough to assume that the energy of the interaction among the particles is negligible within the Hamiltonian operator. It follows that the latter is separable, and the expressions found in Sect. 15.7 are applicable. As the particles are identical, the single-particle eigenvalues $E_{n(i)}$ and eigenfunctions $w_{n(i)}(\mathbf{q}_i)$ obtained

[7]The use of energy intervals does not entail a loss of generality, as the subsequent treatment of the general case will show.

by solving (15.28) are the same for all particles. Also, the indices[8] $n(i)$ are those of an energy, so that the procedure of placing the particles into the available states can be carried out directly in the energy space. To account for spin, states corresponding to the same energy and different spin are to be considered as distinct. The minimum eigenvalue of (15.28) is fixed by the form of the potential energy within the Hamiltonian operator. Given these premises, the energy axis is divided into equal intervals of length ΔE; as in the classical case, the partitioning has the advantage that the set of intervals is countable. The intervals are numbered starting from the one containing the minimum eigenvalue mentioned above, and their size ΔE is such that each of them contains a number of eigenvalues of (15.28). Let g_r be the number of eigenvalues within the rth interval, $r = 1, 2 \ldots$, and N_r the number of particles whose eigenvalues belong to the same interval; if the size ΔE is taken small, one can approximate the energy of the N_r particles with the product $N_r E_r$, where E_r is the energy at the interval's center. Following the same procedure as in Sect. 6.4, one then constructs the function

$$F(N_1, N_2, \ldots, \alpha, \beta) = \log W + \alpha F_N + \beta F_E \,, \tag{15.40}$$

where α, β are the Lagrange multipliers, respectively, related to the total number of particles and total energy of the system, and the form of W depends on the type of particles, as shown below. If the constraint on the total number of particles is not applicable, one lets $\alpha = 0$. Using the numbers $N_1, N_2, \ldots$ as continuous variables, taking the derivative of F with respect to N_r, and equating it to zero yield

$$\frac{\partial}{\partial N_r} \log W = \alpha + \beta E_r \,. \tag{15.41}$$

On the other hand, it is $W = W_1 W_2 \ldots, W_r \ldots$, where W_r is the number of ways in which N_r particles can be placed into the g_r states of the rth interval, subjected to the constraints of the type of particles under consideration. As in the left-hand side of (15.41) only the rth summand depends on N_r, the relation to be worked out is eventually

$$\frac{\partial}{\partial N_r} \log W_r = \alpha + \beta E_r \,. \tag{15.42}$$

The expression of W_r depends on the type of particles; it is given in Sect. 15.8.1 for the case of fermions and in Sect. 15.8.2 for that of bosons.

[8] Here the eigenvalues of the Hamiltonian operator are discrete because the system is enclosed in a container, hence the wave function is normalizable. As usual, the notation $n(i)$ stands for a group of indices.

15.8.1 Fermi-Dirac Statistics

For a system of fermions it is $N_r \leq g_r$ due to the exclusion principle. To calculate the number of ways of placing N_r particles into g_r states one provisionally assumes that the particles are distinguishable. There are g_r possibilities to place the first particle; after the latter has been placed, there remain $g_r - 1$ states due to the exclusion principle, hence the different ways of placing the first two (distinct) particles are $g_r\,(g_r - 1)$. The process continues until all N_r particles are used up, thus leading to $g_r\,(g_r - 1)\,\ldots\,(g_r - N_r + 1)$ ways of placing them. On the other hand, the particles are not distinct; as a consequence, after the placement is completed, any of the $N_r!$ permutations of the particles corresponds to the same placement. In conclusion, the product above must be divided by $N_r!$, thus leading to

$$W_r = \frac{g_r\,(g_r - 1)\,\ldots\,(g_r - N_r + 1)}{N_r!} = \frac{g_r!}{N_r!\,(g_r - N_r)!} = \binom{g_r}{N_r}. \tag{15.43}$$

Using the same procedure as in Sect. 6.4 yields

$$\frac{\mathrm{d}\log W_r}{\mathrm{d}N_r} = \frac{\mathrm{d}\log[(g_r - N_r)!]}{\mathrm{d}(-N_r)} - \frac{\mathrm{d}\log(N_r!)}{\mathrm{d}N_r} \simeq \log(g_r - N_r) - \log(N_r) \tag{15.44}$$

which, combined with (15.42), provides $N_r = g_r/[\exp(\alpha + \beta E_r) + 1]$. For convenience, the total number of particles is indicated here with N_S instead of N; the constraints then read

$$\sum_{r=1}^{U} \frac{g_r}{\exp(\alpha + \beta\,E_r) + 1} = N_S\,, \qquad \sum_{r=1}^{U} \frac{g_r E_r}{\exp(\alpha + \beta\,E_r) + 1} = E_S\,, \tag{15.45}$$

by which the two Lagrange multipliers α and β are determined. The sums are carried out up to a maximum energy E_U; in fact, being the total energy of the system prescribed, there exists a maximum energy that the single particle cannot exceed. As outlined in Sect. 15.9.1, a comparison with the findings of Thermodynamics shows that in full analogy with the classical case treated in Sect. 6.4, it is

$$\beta = \frac{1}{k_B T}\,, \tag{15.46}$$

where $k_B \simeq 1.38 \times 10^{-23}\ \mathrm{J\,K^{-1}}$ is Boltzmann's constant and T the system temperature. As a consequence it is $\alpha = \alpha(T)$. Defining the *Fermi energy* or *Fermi level* $E_F(T) = -k_B\,T\,\alpha(T)$, the expression of N_r becomes

$$N_r = \frac{g_r}{\exp\left[(E_r - E_F)\,/\,(k_B\,T)\right] + 1}\,. \tag{15.47}$$

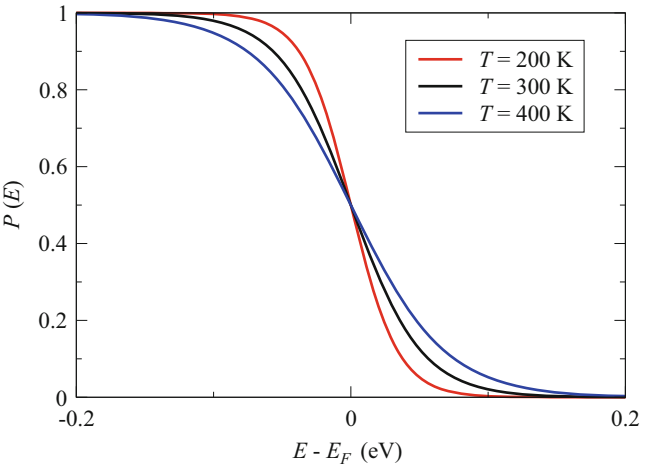


Fig. 15.3 The Fermi-Dirac statistics as a function of energy for different values of the system's temperature. For simplicity the temperature dependence of the Fermi level E_F is not considered

In general the number of energy levels is large and their separation small so that, instead of using the number g_r of states at energy E_r, one disposes of the index and considers the number $g(E)\,\mathrm{d}E$ of states in the infinitesimal interval $\mathrm{d}E$ around E. Thus, $g(E)$ indicates the number of states per unit energy, and is called *density of states in energy* (compare with Sect. B.5). The constraints now read

$$\int_{E_1}^{E_U} \frac{g(E)\,\mathrm{d}E}{\exp(\alpha+\beta E)+1} = N_S\,, \qquad \int_{E_1}^{E_U} \frac{E\,g(E)\,\mathrm{d}E}{\exp(\alpha+\beta E)+1} = E_S\,, \tag{15.48}$$

with $\beta = 1/(k_B T)$. Consistently, N_r is replaced with $N(E)\,\mathrm{d}E$, where $N(E)$ is the number of particles per unit energy.[9] As a consequence, the relation $N(E) = g(E)\,P(E)$ is fulfilled, where function

$$P(E) = \frac{1}{\exp\left[(E-E_F)\,/\,(k_B\,T)\right]+1} \tag{15.49}$$

is called *Fermi-Dirac statistics*. As $0 < P(E) < 1$, the Fermi-Dirac statistics bears also the meaning of occupation probability of a state at energy E. Its high-energy tail ($E - E_F \gg k_B\,T$) identifies with the Maxwell-Boltzmann distribution law (6.14). The dependence of (15.49) on E is shown in Fig. 15.3 at different temperatures. The states whose energy is substantially lower than the Fermi level ($E - E_F \ll k_B\,T$) are mostly filled with particles, those whose energy is substantially higher than the Fermi level are mostly empty. The energy states in the vicinity of the Fermi

[9] The units are $[g]$, $[N] = \mathrm{J}^{-1}$.

level have a probability around $1/2$ of being filled. For the sake of simplicity the dependence of E_F on T is not considered in the figure; in fact, it is influenced by the form of $g(E)$ and must therefore be determined on a case-by-case basis. A discussion about this issue is found in [70, Sect. 8.6] and is summarized in Prob. 15.2. In the limit $T \to 0$ the function becomes discontinuous, specifically it is $P = 1$ for $E < E_F$ and $P = 0$ for $E > E_F$.

15.8.2 Bose-Einstein Statistics

For a system of Bosons the exclusion principle does not hold, hence it may be $N_r \geq g_r$. Also in this case, to calculate the number of ways of placing N_r particles, not subjected to the exclusion principle, into g_r states, one provisionally assumes that the particles are distinguishable. This yields a number of ways equal to $(g_r + N_r - 1)$ $(g_r + N_r - 2)$ $\ldots$ g_r. Then, to account for indistinguishability one divides the result by $N_r!$ to find[10]

$$W_r = \frac{(g_r + N_r - 1) \ldots g_r}{N_r!} = \frac{(g_r + N_r - 1)!}{N_r!\,(g_r - 1)!} = \binom{g_r + N_r - 1}{N_r} . \tag{15.50}$$

Using the same procedure as in Sect. 6.4 yields

$$\frac{\mathrm{d} \log W_r}{\mathrm{d}N_r} = \frac{\mathrm{d} \log[(g_r + N_r - 1)!]}{\mathrm{d}N_r} - \frac{\mathrm{d} \log(N_r!)}{\mathrm{d}N_r} \simeq \log(g_r + N_r - 1) - \log N_r \tag{15.51}$$

which, combined with (15.42) after neglecting the unity in $g_r + N_r - 1$, yields $N_r = g_r/[\exp(\alpha + \beta E_r) - 1]$. As in the latter it is $N_r, g_r > 0$, it follows $\alpha + \beta E_r > 0$. Still indicating the number of particles with N_S instead of N, the constraints read

$$\sum_{r=1}^{U} \frac{g_r}{\exp(\alpha + \beta\, E_r) - 1} = N_S\,, \qquad \sum_{r=1}^{U} \frac{g_r\, E_r}{\exp(\alpha + \beta\, E_r) - 1} = E_S\,, \tag{15.52}$$

by which the two Lagrange multipliers α and β are determined. The explanation of the upper summation limit U in (15.52) is similar to that given in Sect. 15.8.1. As outlined in Sect. 15.9.5, a comparison with the findings of Thermodynamics shows that in full analogy with the cases of Sects. 6.4 and 15.8.1, β is given by (15.46), whence it is $\alpha = \alpha(T)$. Defining $E_B(T) = -k_B\, T\, \alpha(T)$, the expression of N_r becomes

$$N_r = \frac{g_r}{\exp\left[(E_r - E_B)\,/\,(k_B\, T)\right] - 1}\,. \tag{15.53}$$

[10] Compare with Prob. 15.3.

The inequality $\alpha + \beta E_r > 0$ found above implies $E_r > E_B$. If the constraint on the number of particles does not hold, one lets $\alpha = 0$ whence $E_B = 0$, $E_r > 0$. In general the number of energy levels is large and their separation small so that instead of using the number g_r of states at energy E_r, one disposes of the index and considers the number $g(E)\,\mathrm{d}E$ of states in the infinitesimal interval $\mathrm{d}E$ around E. The constraints now read

$$\int_{E_1}^{E_U} \frac{g(E)\,\mathrm{d}E}{\exp(\alpha + \beta E) - 1} = N_S\,, \qquad \int_{E_1}^{E_U} \frac{E\,g(E)\,\mathrm{d}E}{\exp(\alpha + \beta E) - 1} = E_S\,, \tag{15.54}$$

with $\beta = 1/(k_B T)$. Consistently, N_r is replaced with $N(E)\,\mathrm{d}E$, where $N(E)$ is the number of particles per unit energy. As a consequence, the relation $N(E) = g(E)\,P(E)$ is fulfilled, where function

$$P(E) = \frac{1}{\exp\left[(E - E_B)\,/\,(k_B T)\right] - 1} \tag{15.55}$$

is called *Bose-Einstein statistics*. As $P(E)$ may be larger than unity, the Bose-Einstein statistics is not a probability; rather, it represents the occupation number of a state at energy E. Its high-energy tail ($E - E_B \gg k_B T$) identifies with the Maxwell-Boltzmann distribution law (6.14).

15.9 Complements

15.9.1 Connection with Thermodynamic Functions

The calculation of the equilibrium distribution carried out in Sects. 6.4, 15.8.1, and 15.8.2 respectively for classical particles, fermions, and bosons, entails the maximization of the function $\log W$, subjected to suitable constraints. On the other hand, from the second principle of Thermodynamics one derives that the equilibrium state of a system corresponds to the condition that the system's entropy S has a maximum. For this reason one expects that a functional dependence $W(S)$ exists; to identify its form one notes that if W_1 and W_2 indicate the value of W of two independent systems, the value for the composite system is $W_1\,W_2$ due to the definition of W. On the other hand, entropy is additive, so that the functional dependence sought must be such that $S(W_1\,W_2) = S(W_1) + S(W_2)$, namely, of the logarithmic type. In fact it is

$$S = k_B \log W\,, \tag{15.56}$$

with $k_B = 1.38 \times 10^{-23}\ \mathrm{J\,K^{-1}}$ the Boltzmann constant. The choice of the constant makes (15.56) consistent with the definition $\mathrm{d}S = \mathrm{d}Q/T$ of entropy in Thermodynamics, with $\mathrm{d}Q$ the heat absorbed during an infinitesimal transformation.[11]

Now, consider a system subjected to both constraints (15.38) and (15.39), and assume that the container where the system is placed undergoes an infinitesimal volume change[12] $\mathrm{d}\Omega$ which, in turn, makes all variables N_r and g_r to change; due to (15.41) it is $\partial \log W/\partial N_r = \alpha + \beta\, E_r$ so that [37]

$$\frac{1}{k_B}\,\mathrm{d}S = \mathrm{d}\log W = \sum_{r=1}^{U}\left[(\alpha + \beta\, E_r)\ \mathrm{d}N_r + \frac{\partial \log W}{\partial g_r}\,\mathrm{d}g_r\right]. \tag{15.57}$$

Using the constraints and the relation $\mathrm{d}S = \mathrm{d}Q/T$ transforms (15.57) into

$$\mathrm{d}Q = k_B\, T\, \alpha\, \mathrm{d}N + k_B\, T\, \beta\, \mathrm{d}E_S + k_B\, T \sum_{r=1}^{U}\left(\frac{\partial \log W}{\partial g_r}\,\frac{\mathrm{d}g_r}{\mathrm{d}\Omega}\right)\mathrm{d}\Omega\,. \tag{15.58}$$

Assuming that during the change in volume there is no exchange of matter with the environment, one lets $\mathrm{d}N = 0$ in (15.58); the first principle of Thermodynamics shows that for this type of transformation it is $\mathrm{d}Q = \mathrm{d}E_S + P\,\mathrm{d}\Omega$, with P the pressure at the boundary of Ω. A comparison with (15.58) then yields

$$k_B\, T\, \beta = 1\,, \qquad k_B\, T \sum_{r=1}^{U}\left(\frac{\partial \log W}{\partial g_r}\,\frac{\mathrm{d}g_r}{\mathrm{d}\Omega}\right) = P\,. \tag{15.59}$$

The first of (15.59) coincides with (6.26), and provides the expected relation between one of the Lagrange multiplier of (15.45) and a state function of Thermodynamics.

15.9.2 *Density of States for a Particle in a Three-Dimensional Box*

The density of states $g(E)$ that has been introduced in (15.48) and (15.54) is the number of states per unit energy. Its form depends on the system under consideration. Two examples of the derivation of $g(E)$ are given here (with reference to the problem of an electron in a box) and in Sect. 15.9.4 (with reference to photons).

[11] Compare with the nonequilibrium definition of entropy introduced in Sect. 6.6.3 and the note therein.

[12] The geometrical configuration is kept similar to the original one during the change in volume.

The problem of the particle confined within a one-dimensional box has been tackled in Sect. 8.2.2. The set of eigenvalues is given by the first relation in (8.6), and the corresponding eigenfunctions are (8.8). For a three-dimensional box whose sides have lengths d_1, d_2, d_3, due to the properties of separable operators (Sect. 10.3), the eigenfunctions are products of one-dimensional eigenfunctions,

$$w_{n_1 n_2 n_3}(\mathbf{r}) = \sqrt{8/V}\, \sin(k_{n_1} x_1)\, \sin(k_{n_2} x_2)\, \sin(k_{n_3} x_3)\,, \tag{15.60}$$

where $V = d_1\, d_2\, d_3$ is the volume of the box. In turn, the eigenvalues are

$$E_{n_1 n_2 n_3} = E_{n_1} + E_{n_2} + E_{n_3}\,, \qquad E_{n_i} = \frac{\hbar^2\, k_{n_i}^2}{2\, m}\,, \qquad k_{n_i} = n_i\, \frac{\pi}{d_i}\,, \tag{15.61}$$

namely,

$$E = \frac{\hbar^2\, k^2}{2\, m}\,, \qquad \mathbf{k} = n_1\, \frac{\pi}{d_1}\, \mathbf{i}_1 + n_2\, \frac{\pi}{d_2}\, \mathbf{i}_2 + n_3\, \frac{\pi}{d_3}\, \mathbf{i}_3\,, \qquad n_i = 1, 2, \ldots \tag{15.62}$$

with $\mathbf{i}_1, \ldots$ the unit vectors of the Cartesian axes and $k^2 = \mathbf{k} \cdot \mathbf{k}$. In contrast with the one-dimensional case, the eigenvalues (15.61) are degenerate, because different triads n_1, n_2, n_3 correspond to the same energy.

The density of states could be calculated by the procedure depicted in Sect. (B.5), with the provision that in this case the variables k_{n_i} belong to a discrete set whereas those in Sect. B.5 are continuous. On the other hand, the relations involved here are simple enough to be tackled by a direct calculation. One observes, first, that the distance between two consecutive projections of $\mathbf{k}$ along the ith axis is π/d_i; as a consequence, one may partition the $k_1 k_2 k_3$ space into equal volumes[13] π^3/V, so that each $\mathbf{k}$ vector is associated with one and only one volume: this shows that the density of $\mathbf{k}$ vectors in the $k_1\, k_2\, k_3$ space is $Q_k = V/\pi^3$. Given the electron's energy E, from the geometrical point of view the first relation in (15.62) describes a sphere of radius $(2\, m\, E/\hbar^2)^{1/2}$ in the $k_1\, k_2\, k_3$ space; thus, the total number N_k of π^3/V volumes contained within the sphere is obtained by multiplying the sphere's volume by the density Q_k. Clearly the volumes that are near the boundary produce a ragged surface; the latter is identified with that of the sphere by assuming that the distribution of $\mathbf{k}$ vectors belonging to the spherical surface is very dense. With this provision one finds

$$N_k = Q_k\, \frac{4}{3}\, \pi\, k^3 = \frac{V}{\pi^3}\, \frac{4}{3}\, \pi\, \left(\frac{2\, m\, E}{\hbar^2}\right)^{3/2} = \frac{V}{\pi^2}\, \frac{8\, \sqrt{2}\, m^{3/2}}{3\, \hbar^3}\, E^{3/2}\,. \tag{15.63}$$

As indices n_1, n_2, n_3 take only positive values, the $\mathbf{k}$ vectors belong to 1/8 of the sphere only; it follows that their number is $N_k/8$. Each $\mathbf{k}$ vector is associated with

[13] Here the term "volume" is used in a broader meaning; in fact, the units of π^2/V are m^{-3}.

a triad of quantum numbers n_1, n_2, n_3; however, to completely define the state of the electron a fourth quantum number is necessary, related to spin (Sect. 15.5.1). Remembering that electrons have two possible spin states, one finds that the number of electron states within the sphere is twice the number of **k** vectors, namely, $N_k/4$. Finally, the number of states per unit energy is found by differentiating the latter with respect to energy,[14]

$$\mathrm{d}\frac{N_k}{4} = \frac{\sqrt{2}\,Vm^{3/2}}{\pi^2\hbar^3}\,E^{1/2}\,\mathrm{d}E = g(E)\,\mathrm{d}E\,, \qquad g(E) = \frac{\sqrt{2}\,V\,m^{3/2}}{\pi^2\,\hbar^3}\,E^{1/2}\,. \tag{15.64}$$

Apart from the constants involved, this result is consistent with (B.34). Along with the density of state in energy it is useful to define a combined density of states in energy and coordinate space, that is indicated with γ. In the case of the electron within a box one finds, from (15.64),

$$\gamma(E) = \frac{g(E)}{V} = \frac{\sqrt{2}\,m^{3/2}}{\pi^2\,\hbar^3}\,E^{1/2}. \tag{15.65}$$

15.9.3 Density of States for a Two- or One-Dimensional Box

As observed in Sect. B.5, the functional dependence of g on E is influenced by the number of spatial dimensions: considering by way of example the case of an electron within a two-dimensional box, one must associate with each vector **k** an area of the k_1k_2 space equal to $\pi^2/(d_1\,d_2) = \pi^2/A$, where $A = d_1\,d_2$ is the area of the box. The density of the **k** vectors in the k_1k_2 space is $Q_k = A/\pi^2$, whence the total number of π^2/A areas in a circle of radius $(2\,m\,E/\hbar^2)^{1/2}$ is

$$N_k = Q_k\,\pi\,k^2 = \frac{A}{\pi^2}\,\pi\,\frac{2\,m\,E}{\hbar^2}\,. \tag{15.66}$$

As indices n_1, n_2 take only positive values, the **k** vectors belong to $1/4$ of the circle only; it follows that their number is $N_k/4$. Accounting for spin one finds that the number of states within the circle is twice the number of **k** vectors, namely, $N_k/2$. Finally, the number of states per unit energy is found by differentiating the latter with respect to energy,

$$g(E) = \frac{\mathrm{d}N_k/2}{\mathrm{d}E} = \frac{A\,m}{\pi\,\hbar^2}\,. \tag{15.67}$$

When the energy dependence on the k coordinates is quadratic, the density of states of a two-dimensional case is constant (compare with (B.33)).

[14] As noted above the **k** vectors, hence the values of energy corresponding to them, are distributed very densely. This makes it possible to treat E as a continuous variable.

Finally, considering the case of an electron within a one-dimensional box, one must associate to each vector $\mathbf{k}$ a segment π/d_1 of the k_1 space, to find $Q_k = d_1/\pi$ for the density of the $\mathbf{k}$ vectors. The total number of π/d_1 segments in a domain of length $(2\,m\,E/\hbar^2)^{1/2}$ is

$$N_k = \frac{d_1}{\pi}\left(\frac{2\,m\,E}{\hbar^2}\right)^{1/2}. \tag{15.68}$$

Accounting for spin one finds that the number of states within the domain is twice the number of $\mathbf{k}$ vectors, namely, $2\,N_k$. Finally, the number of states per unit energy is found by differentiating the latter with respect to energy,

$$g(E) = \frac{\mathrm{d}2\,N_k}{\mathrm{d}E} = \frac{d_1\,(2\,m)^{1/2}}{\pi\,\hbar\,E^{1/2}}, \tag{15.69}$$

to be compared with (B.32). Expression (15.67) is useful, e.g., for treating the problem of a two-dimensional charge layer in the channel of a semiconductor device (Sect. 17.6.7); in turn, expression (15.69) is used with reference to nanowires (Sect. 17.6.7);

15.9.4 Density of States for Photons

Consider the case of the electromagnetic field within a box whose sides d_1, d_2, d_3 are aligned with the coordinate axes and start from the origin. It is assumed that no charge is present within the box, so that the calculation of the modes is the same as that illustrated in Sect. 5.5. The wave vectors have the form

$$\mathbf{k} = n_1\,\frac{2\pi}{d_1}\,\mathbf{i}_1 + n_2\,\frac{2\pi}{d_2}\,\mathbf{i}_2 + n_3\,\frac{2\pi}{d_3}\,\mathbf{i}_3\,, \qquad n_i = 0, \pm 1, \pm 2, \ldots\,, \tag{15.70}$$

the angular frequency of the mode corresponding to $\mathbf{k}$ is $\omega = c\,k$, with k the modulus of $\mathbf{k}$, and the energy of each photon of that mode is $E = \hbar\,\omega = \hbar\,c\,k$, with c the speed of light. The calculation of the density of states associated with this case follows the same line as that used in Sect. 15.9.2 for a particle in a three-dimensional box; the differences are that the distance between two consecutive projections of $\mathbf{k}$ along the ith axis is $2\pi/d_i$ instead of π/d_i, the indices n_i are not limited to the positive values, and the $E(\mathbf{k})$ relation is linear instead of being quadratic.

With these premises, and with reference to Fig. 15.4, the volume associated with each $\mathbf{k}$ is $(2\,\pi)^3/(d_1\,d_2\,d_3) = (2\,\pi)^3/V$, where V is the volume of the box, so that the density of the $\mathbf{k}$ vectors in the k space is $Q_k = V/(2\,\pi)^3$; in turn, the total number of $\mathbf{k}$ vectors within a sphere of radius k like that shown in the figure is[15]

[15]The calculation shown here is equivalent to counting the number of elements of volume $(2\,\pi)^3/(d_1\,d_2\,d_3)$ that belong to the spherical shell drawn in Fig. 15.4. The result is then multiplied by 2 to account for spin.

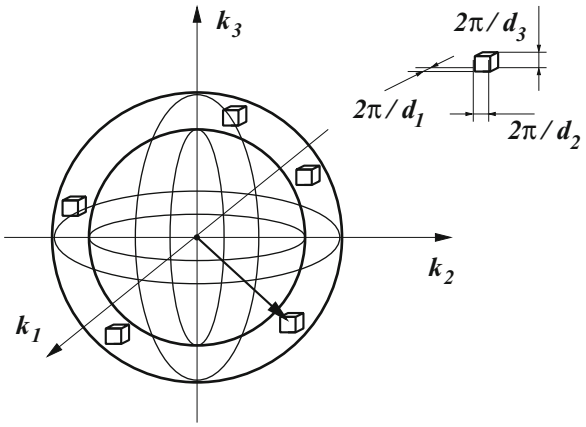


Fig. 15.4 Constant-energy sphere of the **k** space illustrating the procedure for determining the density of states

$$N_k = Q_k \frac{4}{3}\pi k^3 = \frac{V}{(2\pi)^3}\frac{4}{3}\pi\frac{E^3}{\hbar^3 c^3} = \frac{V}{2\pi^2}\frac{E^3}{3\hbar^3 c^3}\,. \tag{15.71}$$

Due to spin, the total number of states is $2N_k$ so that, differentiating with respect to energy,

$$\mathrm{d}2N_k = \frac{V}{\pi^2}\frac{E^2}{\hbar^3 c^3}\mathrm{d}E = g(E)\,\mathrm{d}E\,, \qquad g(E) = \frac{V}{\pi^2\hbar^3 c^3}E^2\,. \tag{15.72}$$

The result expressed by (15.72) can be recast in equivalent forms by using another variable proportional to energy; for instance, letting $G(\omega)$ denote the density of states with respect to angular frequency and $\tilde{g}(\nu)$ denote the density of states with respect to frequency, from the relations

$$\tilde{g}(\nu)\,\mathrm{d}\nu = G(\omega)\,\mathrm{d}\omega = g(E)\,\mathrm{d}E\,, \qquad E = \hbar\,\omega = \hbar\,2\pi\,\nu \tag{15.73}$$

one obtains

$$G(\omega) = \frac{V}{\pi^2 c^3}\omega^2\,, \qquad \tilde{g}(\nu) = 8\pi\frac{V}{c^3}\nu^2\,. \tag{15.74}$$

15.9.5 Derivation of Planck's Law

As illustrated in Sect. 7.4.1, Planck's law provides the black-body's spectral energy density u at equilibrium. From its definition it follows that the product $u\,\mathrm{d}\nu$ is the electromagnetic energy per unit volume in the elementary interval $\mathrm{d}\nu$; using the quantum concepts, the electromagnetic energy in $\mathrm{d}\nu$ may, in turn, be written as

the product of the number $\mathrm{d}N_{\mathrm{ph}}$ of photons belonging to $\mathrm{d}\nu$ by the energy $h\,\nu$ of each photon. Considering that photons are bosons, and that the equilibrium case is considered, the number $\mathrm{d}N_{\mathrm{ph}}$ is given in turn by the product $\tilde{g}(\nu)\,\mathrm{d}\nu\,P(E = h\,\nu)$, with P the Bose-Einstein statistics (15.55). In summary, using the second relation in (15.74), one finds

$$u\,\mathrm{d}\nu = \frac{1}{V}\,h\,\nu\,\mathrm{d}N_{\mathrm{ph}} = \frac{h\,\nu}{V}\,\frac{\tilde{g}(\nu)\,\mathrm{d}\nu}{\exp(\beta\,h\,\nu) - 1} = \frac{8\,\pi\,h\,\nu^3/c^3}{\exp(\beta\,h\,\nu) - 1}\,\mathrm{d}\nu\,. \tag{15.75}$$

The expression of the Bose-Einstein statistic used in (15.75) accounts for the fact that the number of photons is not conserved, so that the only constraint to be considered is that on the total electromagnetic energy within volume V. It follows that the Bose-Einstein statistics has only one Lagrange multiplier, whose value is provisionally left undetermined. From (15.75) one derives Planck's law

$$u(\nu, \beta) = \frac{8\,\pi\,h\,\nu^3/c^3}{\exp(\beta\,h\,\nu) - 1}\,, \tag{15.76}$$

to be compared with (7.18). Like in Sect. 15.9.1, the undetermined parameter β is obtained by comparing the result of the microscopic derivation carried out above with those of Thermodynamics. Letting $W_{\mathrm{em}}^{\mathrm{eq}}$ be the electromagnetic energy within V at equilibrium, the following relations hold (compare with Sect. 7.3):

$$\frac{W_{\mathrm{em}}^{\mathrm{eq}}}{V} = \int_0^\infty u\,\mathrm{d}\nu\,, \qquad \frac{W_{\mathrm{em}}^{\mathrm{eq}}}{V} = \frac{4}{c}\,\sigma\,T^4\,, \tag{15.77}$$

with $\sigma \simeq 5.67\,10^{-12}$ W cm^{-2}K^{-4} the *Stefan-Boltzmann constant*. In fact, the first one derives from the definition of spectral energy density, while the second one (*Stefan law*, 1879) is found experimentally. On the other hand, using (C.137) one finds

$$\int_0^\infty u\,\mathrm{d}\nu = \frac{8\,\pi\,h}{c^3\,(\beta h)^4}\int_0^\infty \frac{(\beta\,h\,\nu)^3}{\exp(\beta\,h\,\nu) - 1}\,\mathrm{d}(\beta\,h\,\nu) = \frac{8\,\pi}{c^3\,h^3\,\beta^4}\,\frac{\pi^4}{15}\,. \tag{15.78}$$

Combining (15.78) with (15.77) yields $1/(\beta T)^4 = 15\,c^2\,h^3\,\sigma/(2\,\pi^5)$; replacing the constants at the right-hand side of the latter shows that the units and numerical value of it are those of k_B^4; it follows that $1/(\beta T)^4 = k_B^4$ and the expected result $\beta = 1/(k_B\,T)$ is found. Conversely, the Stefan-Boltzmann constant can be expressed in terms of the other fundamental constants as

$$\sigma = \frac{\pi^2}{60}\,\frac{k_B^4}{c^2\,\hbar^3}\,. \tag{15.79}$$

Problems

15.1 Estimate the extension of the energy region where the main variation of the Fermi-Dirac statistics (15.49) occurs.

15.2 Using the definition (15.49) of the Fermi-Dirac statistics and the first relation in (15.48), and assuming that the density of states is different from zero around E_F, discuss the temperature dependence of E_F. Consider the case where $E_F > 0$ and $P(E = 0) \simeq 1$.

15.3 Prove (15.50) by combinatorial calculus.

Chapter 16
Separation of Many-Particle Systems

16.1 Introduction

The chapter illustrates a number of steps that are necessary to reduce the many-particle problem to a tractable form. The analysis starts from a system of interacting electrons and nuclei; such a system is not made of identical particles and its Hamiltonian operator is not necessarily separable. Besides that, the number of particles that are present in the typical volume of, e.g., a solid-state device is so huge that the solution of the Schrödinger equation of such a system in the original form is a hopeless task. The first step consists in the application of the adiabatic approximation, by which the system made of the electrons is separated from that of the nuclei. The way in which such a separation is accomplished has the inconvenience that the nuclei are kept fixed in the equilibrium positions; this approximation is too strong, because it prevents the exchange of energy between the two systems from occurring: in fact, it is used provisionally and is removed at a later stage. The next step deals with the electron system which, despite the separation from the nuclei, is still too complicate to be dealt with directly; using the Ritz method, the Schrödinger equation for the electron system is separated into single-particle equations, in which each electron is subjected to the average field of the others. This step yields the Hartree equations and greatly simplifies the problem; in fact, the equations, besides being separated, are also identical to each other, so that the set of eigenvalues and eigenfunction obtained from one of them is applicable to all electrons. The Hartree equations do not comply with the exclusion principle, which must necessarily be fulfilled because the system under consideration is made of identical fermions; a further modification, yielding the Hartree-Fock equations, provides the wave function with the expected antisymmetry property. Finally, the system of nuclei is taken again into consideration for the purpose of eliminating the simplification that the nuclei are fixed in the equilibrium positions: considering the fact that the nuclei are strongly bound together, so that their displacement from the equilibrium position is small, the nuclei are treated as a system of linear harmonic oscillators. In this way, the interaction between

M. Rudan, *Physics of Semiconductor Devices*, DOI 10.1007/978-3-319-63154-7_16

an electron and the nuclei is described (in a later chapter) using the quantum-mechanical, first-order perturbation theory applied to the two-particle collision of an electron with a phonon.

16.2 System of Interacting Electrons and Nuclei

The analysis of a conservative system of identical particles, started in Sect. 15.7, is based on the single-particle Hamiltonian operator (15.28). It must be noted, however, that the typical systems to be dealt with are not made of identical particles but, rather, of a mixture of subsystems; the particles of each subsystem are identical to each other, but different from those of the other subsystems. Moreover, the Hamiltonian operator of each subsystem is not necessarily separable. It follows that the existence of a single-particle Hamiltonian operator like (15.28) is by no means obvious. It is then necessary to tackle the problem from a more general point of view, starting from the consideration of a mixture of systems, and determining the approximations that may simplify the problem to a level that is practically affordable.

To this purpose, consider a conservative system made of K electrons and N nuclei, interacting with each other. The particles are bound together, so that the system's wave function can be assumed to be normalized to unity. The coordinates associated with the electrons are indicated with small letters and grouped into a single vector $\mathbf{r} = (\mathbf{r}_1, \ldots, \mathbf{r}_K)$; those of the nuclei are indicated with capital letters: $\mathbf{R} = (\mathbf{R}_1, \ldots, \mathbf{R}_N)$. The interaction among the particles is assumed to be of the Coulomb type, so that the contributions to the potential energy due to the electron-electron and nucleus-nucleus interactions are, respectively,

$$U_e(\mathbf{r}) = \frac{1}{2} \sum_{i,j=1}^{K} \frac{q^2}{4\,\pi\,\varepsilon_0\,|\mathbf{r}_i - \mathbf{r}_j|}\,, \qquad U_a(\mathbf{R}) = \frac{1}{2} \sum_{i,j=1}^{N} \frac{Z_i\,Z_j\,q^2}{4\,\pi\,\varepsilon_0\,|\mathbf{R}_i - \mathbf{R}_j|}\,, \tag{16.1}$$

with $j \neq i$, where $q > 0$ is the electron charge, ε_0 the vacuum permittivity, and $Z_j\,q$ the charge of the jth nucleus; factor $1/2$ is introduced to avoid a double summation. In addition one must consider the electron–nucleus interaction, whose contribution to the potential energy is

$$U_{ea}(\mathbf{r}, \mathbf{R}) = \sum_{i=1}^{K} \sum_{j=1}^{N} \frac{-Z_j\,q^2}{4\,\pi\,\varepsilon_0\,|\mathbf{r}_i - \mathbf{R}_j|}\,. \tag{16.2}$$

Remembering (15.3), letting $\mathbf{r}_i = (x_{i1}, x_{i2}, x_{i3})$, $\mathbf{R}_i = (X_{j1}, X_{j2}, X_{j3})$, and defining

$$\nabla_i^2 = \frac{\partial^2}{\partial x_{i1}^2} + \frac{\partial^2}{\partial x_{i2}^2} + \frac{\partial^2}{\partial x_{i3}^2}\,, \qquad \nabla_j^2 = \frac{\partial^2}{\partial X_{j1}^2} + \frac{\partial^2}{\partial X_{j2}^2} + \frac{\partial^2}{\partial X_{j3}^2}\,, \tag{16.3}$$

the kinetic parts of the system's Hamiltonian operator are

$$\mathscr{T}_e = -\sum_{i=1}^{K} \frac{\hbar^2}{2\,m} \nabla_i^2 \,, \qquad \mathscr{T}_a = -\sum_{j=1}^{N} \frac{\hbar^2}{2\,m_j} \nabla_j^2 \,. \tag{16.4}$$

The time-independent Schrödinger equation of the system then reads

$$\mathscr{H} w = E w \,, \qquad \mathscr{H} = \mathscr{T}_e + \mathscr{T}_a + U_e + U_a + U_{ea} + U_{\text{ext}} \,, \tag{16.5}$$

where the eigenfunctions depend on all variables, $w = w(\mathbf{r}, \mathbf{R})$. In (16.5), $U_{\text{ext}} = U_{\text{ext}}(\mathbf{r}, \mathbf{R})$ is the potential energy due to an external, conservative field acting on the system. If it were $U_{ea} + U_{\text{ext}} = 0$, the Hamiltonian would be separable with respect to the two subsystems, $\mathscr{H} = (\mathscr{T}_e + U_e) + (\mathscr{T}_a + U_a)$. The time-independent equations resulting from the separation would be

$$(\mathscr{T}_e + U_e)\, u = E_e\, u \,, \qquad (\mathscr{T}_a + U_a)\, v = E_a\, v \,, \tag{16.6}$$

with $u = u(\mathbf{r})$, $v = v(\mathbf{R})$; due to the general properties of separable operators (compare with Sect. 10.3), the eigenvalues and eigenfunctions of (16.5) would have the form $w = u\,v$, $E = E_e + E_a$. As in general it is $U_{ea} + U_{\text{ext}} \neq 0$, the Schrödinger equation (16.5) is not actually separable, and its solution may often present a formidable challenge: considering by way of example that the concentration of atoms in solid matter is of the order of 5×10^{22} cm^{-3}, and that to a nucleus there correspond Z electrons, with Z the atomic number, the number of scalar coordinates necessary to describe a cubic centimeter of solid matter is of the order of $15\,(Z+1) \times 10^{22}$. It is then necessary to make a number of approximations; they are introduced step by step and, as illustrated in Sects. 16.3, 16.4, and 16.5, they are capable to bring the problem back to the solution of single-particle equations.

16.3 Adiabatic Approximation

The problem of solving the Schrödinger equation (16.5) for a system of interacting electrons and nuclei is simplified by observing that the mass of a nucleus is much larger than that of the electrons. As a first consequence, one expects that the interaction with an individual electron influences little the motion of a nucleus; rather, the latter is expected to be influenced by the average interaction with many electrons. A second consequence is that if the particles' dynamics is provisionally considered in classical terms, the motion of the nuclei is much slower than that of the electrons; as a consequence, the classical positions $\mathbf{R}$ of the nuclei can be considered as slowly varying parameters when the electron dynamics is investigated. For this reason, the procedure shown below, based on the latter observation, is called *adiabatic approximation* or also *Born-Oppenheimer approximation*

[78, Sect. 8]. In quantum terms, the approximation leads to the splitting of the original equation (16.5) into a set of two coupled equations. To this purpose, let

$$\mathscr{H}_e = \mathscr{T}_e + U_e + U_{ea} + U_{\text{ext}}\,, \tag{16.7}$$

where the dependence on $\mathbf{R}$ is algebraic only. Considering such a dependence, the Hamiltonian operator $\mathscr{H}_e$ provides an eigenvalue equation whose eigenvalues and eigenfunctions depend on $\mathbf{R}$; they read

$$\mathscr{H}_e\, u = E_e\, u\,, \qquad E_e = E_e(\mathbf{R})\,, \qquad u = u(\mathbf{r}, \mathbf{R})\,. \tag{16.8}$$

Also, from the Schrödinger equation (16.8) one finds that, for any function v that depends only on $\mathbf{R}$, the following holds: $\mathscr{H}_e u\, v = E_e\, u\, v$. As v is undetermined, one may seek a form of v such that $w = u\, v$ is also an eigenfunction of the full Hamiltonian operator $\mathscr{H}$ of (16.5). From the definition of $\mathscr{H}_e$ it follows $\mathscr{H} = \mathscr{T}_a + U_a + \mathscr{H}_e$, whence the original Schrödinger equation (16.5) is recast as

$$(\mathscr{T}_a + U_a + \mathscr{H}_e)\; uv = E\, uv\,. \tag{16.9}$$

The second relation in (16.4) shows that each Laplacian operator in $\mathscr{T}_a$ acts on both u and v, specifically, $\nabla_j^2\, u\, v = u\, \nabla_j^2\, v + v\, \nabla_j^2\, u + 2\, \nabla_j u \cdot \nabla_j v$. Multiplying the latter by $-\hbar^2/(2\, m_j)$ and adding over j transforms (16.9) into

$$u\; (\mathscr{T}_a + U_a + E_e)\; v + v\; \mathscr{T}_a u + \mathscr{G}_a u\, v = E\, u\, v\,, \tag{16.10}$$

where the short-hand notation $\mathscr{G}_a u\, v = -\sum_{j=1}^{N} (\hbar^2/m_j)\, \nabla_j u \cdot \nabla_j v$ has been used. To proceed one notes that the coefficients in the Schrödinger equation (16.8) are real, so that u can be taken real as well. Remembering that the particles described by u are bound, one finds that u is normalizable so that, for any $\mathbf{R}$,

$$\int u^2(\mathbf{r}, \mathbf{R})\, \mathrm{d}r = 1\,, \qquad \mathrm{d}r = \mathrm{d}^3 r_1\, \ldots\, \mathrm{d}^3 r_K\,, \tag{16.11}$$

with the integral extended to the whole K-dimensional space $\mathbf{r}$. Remembering that ∇_j acts on $\mathbf{R}_j$, from (16.11) one derives

$$\int u\, \nabla_j u\, \mathrm{d}r = \frac{1}{2} \int \nabla_j u^2\, \mathrm{d}r = \frac{1}{2}\, \nabla_j \int u^2\, \mathrm{d}r = 0\,, \tag{16.12}$$

whence

$$\int u\, \mathscr{G}_a u\, v\, \mathrm{d}r = -\sum_{j=1}^{N} \frac{\hbar^2}{m_j}\, \nabla_j v \cdot \int u\, \nabla_j u\, \mathrm{d}r = 0\,. \tag{16.13}$$

Left multiplying (16.10) by u, integrating over $\mathbf{r}$, and using (16.11,16.13) yield

$$(\mathscr{T}_a + U_a + E_e + U_u)\, v = E\, v\,, \qquad U_u(\mathbf{R}) = \int u \mathscr{T}_a u \, \mathrm{d}r\,. \tag{16.14}$$

In this way, the original Schrödinger equation (16.5) is eventually split into two coupled equations, (16.8) and (16.14); the former is written in extended form as

$$(\mathscr{T}_e + U_e + U_{ea} + U_{\text{ext}})\, u = E_e\, u\,. \tag{16.15}$$

The first equation, (16.14), is the equation for the nuclei, in fact its kinetic operator $\mathscr{T}_a$ acts on $\mathbf{R}$ only, and its potential-energy term $U_a + E_e + U_u$ depends on $\mathbf{R}$ only; this equation is coupled with the second one, (16.15), because E_e is an eigenvalue of (16.15) and U_u is given by the integral in (16.14) that involves the eigenfunctions of (16.15). In turn, (16.15) is the equation for the electrons, in fact its kinetic operator $\mathscr{T}_e$ acts on $\mathbf{r}$ only; the part $U_{ea} + U_{\text{ext}}$ of its potential-energy term couples (16.15) with (16.14) due to the dependence on $\mathbf{R}$.

In principle, one may solve (16.15) after fixing $\mathbf{R}$ and, from the solution, calculate E_e and U_u to be used in (16.14). From v one then determines the expectation value of $\mathbf{R}$ that updates the starting point for the solution of (16.15). Considering the case of solid matter, and still reasoning in classical terms, a zero-order solution for the iterative procedure depicted above is found by observing that the nuclei, being massive and tightly bound together, are expected to depart little from their equilibrium positions $\mathbf{R}_0$. One then fixes $\mathbf{R} = \mathbf{R}_0$ in (16.15) to find, for the electrons' Schrödinger equation, the approximate form

$$(\mathscr{T}_e + V_e)\, u = E_e\, u\,, \qquad V_e(\mathbf{r}) = U_e(\mathbf{r}) + U_{ea}(\mathbf{r}, \mathbf{R}_0) + U_{\text{ext}}(\mathbf{r}, \mathbf{R}_0)\,. \tag{16.16}$$

This separates completely the equation for the electrons, that may be thought of as forming a separate system of total energy E_e. Clearly, keeping the positions of the nuclei fixed is rather crude an approximation; in this way, in fact, the nuclei cannot exchange energy with the electrons any more. On the other hand, it can be shown that the solution of (16.16) provides an acceptable approximation for the energy of the electrons. Another advantage of keeping the nuclei in the equilibrium positions occurs in the investigation of materials where the nuclei are arranged in periodic structures; in fact, one exploits in this case the properties of differential equations having periodic coefficients (compare with Sects. 13.4 and 17.5.1). In conclusion, the present analysis will continue by provisionally considering only the system of electrons as described by (16.16); obviously the problem of the exchange of energy with the system of nuclei cannot be dispensed with and will be resumed at a later stage (Sect. 16.6).

16.4 Hartree Equations

Equation (16.16), describing the separate system of K electrons after fixing the nuclei to the equilibrium positions $\mathbf{R}_0$, lends itself to the application of the Ritz method outlined in Sect. 16.7.1. For the sake of generality, one starts by assuming that the K particles are distinguishable; for this reason, the symbols $\mathscr{T}_e$, V_e, u, and E_e of (16.16) are not used. Also, the application of the Ritz method will be carried out in the case where the external forces are absent, $U_{\text{ext}} = 0$; such forces will be introduced again into the problem in Sect. 19.2.1, where the single-particle dynamics is described. Given these premises, the operator $\mathscr{A}$, its minimum eigenvalue A_1, and the corresponding eigenfunction v_1 used in Sect. 16.7.1 for describing the Ritz method are indicated here with $\mathscr{A} = \mathscr{H}$, $A_1 = E_1$, $v_1 = w_1$, where

$$\mathscr{H} = \sum_{i=1}^{K} \mathscr{T}_i + \sum_{i=1}^{K} \sum_{j<i} V_{ij} \,, \qquad \mathscr{T}_i = -\frac{\hbar^2 \nabla_i^2}{2\, m_i} \,, \qquad V_{ij} = \frac{\sigma_{ij}\, q^2\, Z_i\, Z_j}{4\, \pi\, \varepsilon_0\, |\mathbf{r}_i - \mathbf{r}_j|} \,, \tag{16.17}$$

with $\sigma_{ij} = \pm 1$. The above describes a system of different particles interacting through Coulomb potentials. Introducing the auxiliary function f as in (16.37), one minimizes the functional $\langle f_1 \ldots f_K | \mathscr{H} | f_1 \ldots f_K \rangle$ subjected to the K constraints $\langle f_i | f_i \rangle = 1$. Using the method of Lagrange multipliers (Sect. B.6), this is equivalent to finding the absolute minimum of

$$F_{\mathscr{H}}[f_1, \ldots, f_K] = \langle f_1 \ldots f_K | \mathscr{H} | f_1 \ldots f_K \rangle - \sum_{i=1}^{K} \epsilon_i \langle f_i | f_i \rangle \,, \tag{16.18}$$

with ϵ_i the multipliers. The terms related to $\mathscr{T}_i$ are separable, while those related to V_{ij} are separable in pairs. From the orthonormalization condition it follows

$$F_{\mathscr{H}} = \sum_{i=1}^{K} \left(\langle f_i | \mathscr{T}_i | f_i \rangle - \epsilon_i \langle f_i | f_i \rangle + \sum_{j=1}^{i-1} \langle f_i f_j | V_{ij} | f_i f_j \rangle \right) . \tag{16.19}$$

The minimum of $F_{\mathscr{H}}$ is found by letting $\delta F_{\mathscr{H}} = F_{\mathscr{H}}[f_1 + \delta f_1, \ldots, f_K + \delta f_K] - F_{\mathscr{H}}[f_1, \ldots, f_K] = 0$. Neglecting the second-order terms and observing that $\mathscr{T}_i$ is Hermitean yield

$$\delta \langle f_i | \mathscr{T}_i | f_i \rangle = \langle \delta f_i | \mathscr{T}_i | f_i \rangle + \langle f_i | \mathscr{T}_i | \delta f_i \rangle = 2\, \Re \langle \delta f_i | \mathscr{T}_i | f_i \rangle \,. \tag{16.20}$$

By the same token one finds $\delta \epsilon_i \langle f_i | f_i \rangle = \epsilon_i\, 2\, \Re \langle \delta f_i | f_i \rangle$ and $\delta \langle f_i f_j | V_{ij} | f_i f_j \rangle = 2\, \Re \langle f_i\, \delta f_j | V_{ij} | f_i f_j \rangle + 2\, \Re \langle f_j\, \delta f_i | V_{ij} | f_i f_j \rangle$. The symmetry of the Coulomb terms yields $V_{ji} = V_{ij}$, whence

$$\sum_{j=1}^{i-1} \left(\langle f_i\, \delta f_j | V_{ij} | f_i f_j \rangle + \langle f_j\, \delta f_i | V_{ij} | f_i f_j \rangle \right) = \sum_{j=1}^{K} \langle f_j\, \delta f_i | V_{ij} | f_i f_j \rangle \,, \tag{16.21}$$

with $j \neq i$ at the right-hand side. The minimization condition of (16.19) then reads

$$2\,\Re \sum_{i=1}^{K} \left(\langle \delta f_i | \mathscr{T}_i | f_i \rangle - \epsilon_i \langle \delta f_i | f_i \rangle + \sum_{j=1}^{K} \langle f_j\, \delta f_i | V_{ij} | f_i f_j \rangle \right) = 0 \,, \tag{16.22}$$

with $j \neq i$ in the inner sum. As the variations δf_i are independent of each other, the term within parentheses in (16.22) must vanish for each i, thus yielding a system of K coupled equations. The inner sum has a double integral in it, and is recast as

$$\sum_{j=1}^{K} \langle f_j\, \delta f_i | V_{ij} | f_i f_j \rangle = \langle \delta f_i | U_i | f_i \rangle \,, \qquad U_i(\mathbf{r}_i) = \sum_{j=1}^{K} \langle f_j | V_{ij} | f_j \rangle \,, \tag{16.23}$$

still with $j \neq i$. The ith equation of the system then reads

$$\langle \delta f_i | \mathscr{T}_i | f_i \rangle - \epsilon_i \langle \delta f_i | f_i \rangle + \langle \delta f_i | U_i | f_i \rangle = \langle \delta f_i | \left(\mathscr{T}_i + U_i - \epsilon_i \right) f_i \rangle = 0 \,. \tag{16.24}$$

As the above holds for any δf_i, the terms on the right-hand side of the scalar products of (16.24) must cancel each other. In conclusion, the minimization condition provides a set of K single-particle equations coupled through the terms U_i:

$$\left(\mathscr{T}_i + U_i \right) f_i = \epsilon_i f_i \,, \qquad i = 1, \ldots, K \,. \tag{16.25}$$

The above are called *Hartree equations* and constitute a set of Schrödinger equations whose potential energy is given by the second relation in (16.23):

$$U_i(\mathbf{r}_i) = \sum_{j=1}^{K} \int \frac{\sigma_{ij}\, q^2\, Z_i\, Z_j}{4\,\pi\,\varepsilon_0\, |\mathbf{r}_i - \mathbf{r}_j|} \, |f_j|^2 \, \mathrm{d}r_j \,, \qquad j \neq i \,. \tag{16.26}$$

The potential energy of the ith particle is the sum of two-particle potential energies averaged with the localization probabilities of the particles different from the ith. The eigenvalue ϵ_i is then the energy of the ith particle in the field of the other $K-1$ particles (for this reason, the sum $\epsilon_1 + \ldots + \epsilon_N$ is not the total energy of the system).

The solution of (16.25) is found by iteration; one starts with $i = 1$ and provides an initial guess for $f_2, \ldots, f_K$, so that the initial guess for U_1 is calculated from (16.26) and used in (16.25). The solution of the latter yields the first iteration for f_1 and, remembering the Ritz method outlined in Sect. 16.7.1, the parameters embedded in f_1 are exploited at this stage to lower the eigenvalue; then, one proceeds with $i = 2$, using the first iteration for f_1 and the initial guess for $f_3, \ldots, f_K$, and so on. It must be noted that the initial guess is used within an integral, whose

effect is that of averaging the difference between the initial guess and the actual solution; therefore, it may happen that the accuracy of the first iteration is sufficient, to the extent that the iterative process may be brought to an end. In this case, the K equations (16.25) become independent of each other and the solution effort is greatly reduced.

16.5 Hartree-Fock Equations

It is now necessary to investigate the problem originally introduced in Sect. 16.3, namely, that of equation (16.16), which describes the separate system of electrons after fixing the nuclei to the equilibrium positions $\mathbf{R}_0$. The Hartree equations (16.25) cannot be applied as they stand, because they have been deduced for a system made of distinguishable particles. To treat a system of electrons, instead, it is necessary to account for the particles' spin that was not considered in Sects. 16.3 and 16.4, and ensure the antisymmetry of the wave function (Sect. 15.7). The procedure is similar to that depicted in Sect. 16.4, the difference being that the auxiliary function f is not expressed as a product like in (16.37), but is given by a Slater determinant (15.37), whose entries depend on the position and spin coordinates of the corresponding particle [78, Sect. 8], [124, Sect. 16.3].

The calculations are rather involved and are not reported here. It is important to mention that like in Sect. 16.4, the derivation is carried out by assuming that the external forces are absent ($U_{\text{ext}} = 0$). The procedure yields a set of K equations, coupled with each other, that generalize the Hartree equations (16.25) and are called *Hartree-Fock equations*. If the accuracy of the first iteration is sufficient, the original Hamiltonian operator of (16.16), describing the system of electrons as a whole, is separated into K single-particle, identical operators:

$$\mathscr{T}_e + V_e = \sum_{i=1}^{K} \left[-\frac{\hbar^2}{2\,m} \nabla_i^2 + V_{ei}(\mathbf{r}_i) \right] . \tag{16.27}$$

In this way the form (15.28) of the Schrödinger equation for the electrons, sought at the beginning of Sect. 16.2, is recovered.

16.6 Schrödinger Equation for the Nuclei

In the process of separating the Schrödinger equation for the electrons, (16.15), from that of the nuclei, (16.14), that is carried out in Sect. 16.3, the positions of the nuclei are fixed to the equilibrium values $\mathbf{R} = \mathbf{R}_0$. This is done to the purpose of calculating the coefficients E_e and U_u (compare with (16.6) and (16.14), respectively); such coefficients depend on $\mathbf{R}$ and can be obtained only by solving the

equation for the electrons. After this step is accomplished, one turns to the equation for the nuclei, in which E_e and U_u are fixed to constants from the previous iteration; the potential energy of (16.14) then becomes

$$V_a(\mathbf{r}) = U_a(\mathbf{r}) + E_e(\mathbf{r}_0) + U_u(\mathbf{r}_0)\,. \tag{16.28}$$

If the positions of the nuclei are kept fixed, that is, the iterative procedure outlined in Sect. 16.3 is brought to an end without solving (16.14), the exchange of energy between the system of electrons and that of nuclei cannot take place. It is then necessary to proceed to the solution of (16.14); such a solution is obtained by means of an approximation shown below, and the iterative procedure for the solution of (16.14) and (16.15) is stopped right after. In this way, one keeps for the energy of the electrons the eigenvalues obtained with the nuclei fixed at $\mathbf{R}_0$, which is a convenient choice due to the advantages illustrated in Sect. 16.3.

As for the nuclei, it has already been observed in Sect. 16.3 with respect to the case of solid matter that the nuclei, being massive and tightly bound together, are expected to depart little from their equilibrium positions $\mathbf{R}_0$. To improve with respect to the zero-order approximation $\mathbf{R} = \mathbf{R}_0$, and provisionally reasoning in classical terms, one assumes that the instantaneous displacement $\mathbf{R}-\mathbf{R}_0$ with respect to the equilibrium point is small, so that V_a in (16.28) can be approximated with a second-order Taylor expansion around $\mathbf{R}_0$. The classical form of the problem is thus brought to the case already solved in Sects. 3.9 and 3.10: indicating with $T_a + V_a$ the classical equivalent of the operator $\mathscr{T}_a + V_a$, the vibrational state of the nuclei is described in terms of the normal coordinates b_σ, whose conjugate momenta are $\dot{b}_\sigma$, and the total energy of the nuclei reads (compare with (3.50)):

$$T_a + V_a = \sum_{\sigma=1}^{3N} H_\sigma + V_{a0}\,, \qquad H_\sigma = \frac{1}{2}\,\dot{b}_\sigma^2 + \frac{1}{2}\,\omega_\sigma^2\, b_\sigma^2\,, \tag{16.29}$$

with $\omega_\sigma > 0$ the angular frequency of the mode. As a consequence, the quantum operator takes the form

$$\mathscr{T}_a + V_a = \sum_{\sigma=1}^{3N} \mathscr{H}_\sigma + V_{a0}\,, \qquad \mathscr{H}_\sigma = -\frac{\hbar^2}{2}\,\frac{\partial^2}{\partial b_\sigma^2} + \frac{1}{2}\,\omega_\sigma^2\, b_\sigma^2 \tag{16.30}$$

which, introduced into (16.14), yields

$$\sum_{\sigma=1}^{3N} \mathscr{H}_\sigma\, v = E' v\,, \qquad E' = E - V_{a0}\,. \tag{16.31}$$

This procedure provides the separation of the degrees of freedom of the nuclei and completes the separation process whose usefulness was anticipated in Sect. 16.2. The solution of (16.31) is illustrated in Sect. 12.5 and shows that the mode's energy

is ascribed to the set of phonons belonging to the mode itself. As the phonons are bosons, the equilibrium distribution of their occupation numbers is given by the Bose-Einstein statistics. The description of the interaction between an electron and the nuclei is obtained from the quantum-mechanical, first-order perturbation theory applied to the two-particle collision of an electron with a phonon (Sect. 17.8).

16.7 Complements

16.7.1 Ritz Method

Let $\mathscr{A}$ be a Hermitean operator with a discrete spectrum and a complete, orthonormal set of eigenfunctions. For the sake of simplicity, the eigenvalues are assumed nondegenerate:

$$\mathscr{A} v_n = A_n v_n \,, \qquad \langle v_m | v_n \rangle = \delta_{mn} \,. \tag{16.32}$$

Consider the expansion of a function f in terms of the eigenfunctions of $\mathscr{A}$, $f = \sum_{n=1}^{\infty} c_n \, v_n$, with $c_n = \langle v_n | f \rangle$. From Parseval theorem (8.41) one finds

$$\langle f | \mathscr{A} | f \rangle = \sum_{n=1}^{\infty} A_n \, |c_n|^2 \,. \tag{16.33}$$

As the eigenvalues are real, one orders them in a nondecreasing sequence: $A_1 \leq A_2 \leq A_3 \leq \ldots$, whence

$$\langle f | \mathscr{A} | f \rangle \geq A_1 \sum_{n=1}^{\infty} |c_n|^2 = A_1 \, \langle f | f \rangle \,. \tag{16.34}$$

More generally, if f is orthogonal to the first $s-1$ eigenfunctions, then $c_1 = c_2 = \ldots = c_{s-1} = 0$, and

$$f = \sum_{n=s}^{\infty} c_n \, v_n \,, \qquad \langle f | \mathscr{A} | f \rangle \geq A_s \, \langle f | f \rangle \,, \tag{16.35}$$

where the equality holds if and only if $f = \text{const} \times v_s$. In conclusion, the functional

$$G_{\mathscr{A}}[f] = \frac{\langle f | \mathscr{A} | f \rangle}{\langle f | f \rangle} \tag{16.36}$$

has a minimum for $f = v_s$, whose value is A_s. The findings above are the basis of the *Ritz method* that provides approximations for the eigenvalues and eigenfunctions of

$\mathscr{A}$. For instance, the minimum eigenvalue A_1 and the corresponding eigenfunction $v_1(\mathbf{r})$ are approximated by letting $v_1 \simeq f(\mathbf{r}, \alpha_1, \alpha_2, \ldots)$, where the form of f is prescribed and $\alpha_1, \alpha_2, \ldots$ are parameters. The latter are then used to minimize $G_{\mathscr{A}}[f]$. The constraint $\langle f|f\rangle = 1$ is imposed along the calculation, yielding $A_1 \simeq \min_\alpha G_{\mathscr{A}}[f] = \langle f|\mathscr{A}|f\rangle$. When A_1, v_1 have been found, the next pair A_2, v_2 is determined by using an approximating function orthonormal to v_1, and so on. For a system made of N particles, the eigenfunctions depend on $3N$ coordinates: $v_n = v_n(\mathbf{r}_1, \ldots, \mathbf{r}_N)$. It is convenient to approximate them by separating the variables:

$$f = f_1(\mathbf{r}_1) \ldots f_N(\mathbf{r}_N)\,, \tag{16.37}$$

where each f_i may also depend on parameters. The normalization constraint then yields

$$\langle f_i|f_i\rangle = 1\,, \qquad i = 1, \ldots, N\,. \tag{16.38}$$

Part V
Applications to Semiconducting Crystals

Chapter 17
Periodic Structures

17.1 Introduction

This chapter outlines a number of concepts that are useful in the description of periodic structures. The first sections describe the geometrical entities (characteristic vectors, direct and reciprocal lattices, translation vectors, cells, and Brillouin zones) used for describing a lattice. The analysis focuses on the lattices of cubic type, because silicon and other semiconductor materials used in the fabrication of integrated circuits have this type of structure. The next sections introduce the mathematical apparatus necessary for solving the Schrödinger equation within a periodic structure, specifically, the translation operators, Bloch theorem, and periodic boundary conditions. Basing on such tools, the Schrödinger equation is solved, leading to the dispersion relation of the electrons, whose properties are worked out and discussed. The analogy between a wave packet in a periodic potential and in free space is also outlined. Then, the parabolic-band approximation is introduced, leading to the concept of effective mass and to the explicit calculation of the density of states. Examples of the structure of the conduction and valence bands of silicon, germanium, and gallium arsenide are provided. Considering the importance of two-dimensional and one-dimensional structures in modern technological applications, a detailed derivation of the subbands and the corresponding density of states is also given. Then, the same mathematical concepts used for solving the Schrödinger equation (Bloch theorem and periodic boundary conditions) are applied to the calculation of the lattice's vibrational spectrum in the harmonic approximation. The properties of the eigenvalues and eigenfunctions of the problem are worked out, leading to the expression of the vibrational modes. Basing on the above, the description of the electron-lattice interaction is worked out. The complements provide a number of details about the description of crystal planes and directions, and about the connection between the symmetries of the Hamiltonian operator and the properties of its eigenvalues. A number of examples of application of the concepts outlined in the chapter are given in the last part of the complements;

M. Rudan, *Physics of Semiconductor Devices*, DOI 10.1007/978-3-319-63154-7_17

specifically, the Kronig-Penney model, showing a one-dimensional calculation of the electrons' dispersion relation, and the derivation of the dispersion relation of the one-dimensional monatomic and diatomic chains. The complements are concluded by a discussion about some analogies between the energy of the electromagnetic field and that of the vibrating lattice, and between the dispersion relation of the electrons and that of the lattice.

17.2 Bravais Lattice

The concepts illustrated in the previous chapters will now be applied to study the properties of a specific class of materials, the *crystals*. In fact, in the majority of cases the solid-state devices are manufactured using a crystal as basic material.[1] A crystal is made of a periodic arrangement of an atom, or a group of atoms, called *basis*. As a periodic structure is unlimited in all spatial directions, the general properties of crystals are derived using the provisional hypothesis that the material extends to infinity. The more realistic case of a finite crystal is considered at a later stage.

To describe the properties of a crystal it is convenient to superimpose to it a geometrical structure, called *Bravais lattice*, made of an infinite array of discrete points generated by translation operations of the form [16]

$$\mathbf{l} = m_1\,\mathbf{a}_1 + m_2\,\mathbf{a}_2 + m_3\,\mathbf{a}_3\,, \tag{17.1}$$

with m_1, m_2, m_3 any integers. In (17.1), $\mathbf{l}$ is called *translation vector* while $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ are the *characteristic vectors*; the discrete points generated by (17.1) are called *nodes*. The set of vectors $\mathbf{l}$ is closed under vector addition. Although the characteristic vectors are not necessarily of equal length, nor are they orthogonal to each other, they form a complete set; it follows that any vector $\mathbf{r}$ of the three-dimensional space is expressible as

$$\mathbf{r} = \mu_1\,\mathbf{a}_1 + \mu_2\,\mathbf{a}_2 + \mu_3\,\mathbf{a}_3\,, \tag{17.2}$$

with μ_1, μ_2, μ_3 real numbers. In a zero-dimensional or one-dimensional space only one Bravais lattice is possible. In a two-dimensional space there are five Bravais lattices, such as *oblique*, *rectangular*, *centered rectangular (rhombic)*, *hexagonal*, and *square* [79]. An example of oblique lattice is shown in Fig. 17.1. Another

[1] Some important exceptions exist. *Thin-Film Transistors* (TFT), commonly used in flat-screen or liquid-crystal displays, are obtained by depositing a semiconductor layer (typically, silicon) over a nonconducting substrate; due to the deposition process, the structure of the semiconductor layer is amorphous or polycrystalline. *Phase-Change Memories* (PCM) exploit the property of specific materials like chalcogenides (for example, $Ge_2Sb_2Te_5$), that switch from the crystalline to the amorphous state, and vice versa, in a controlled way when subjected to a suitable electric pulse.

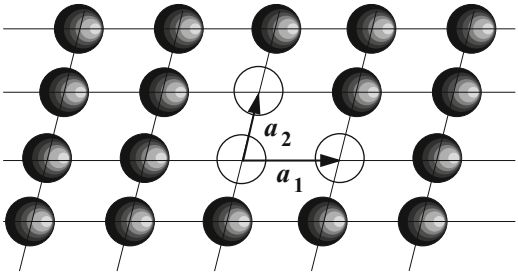


Fig. 17.1 Schematic description of a two-dimensional Bravais lattice of the oblique type. Three atoms have been removed to better show the characteristic vectors. The latter are not orthogonal to each other, and their lengths are different

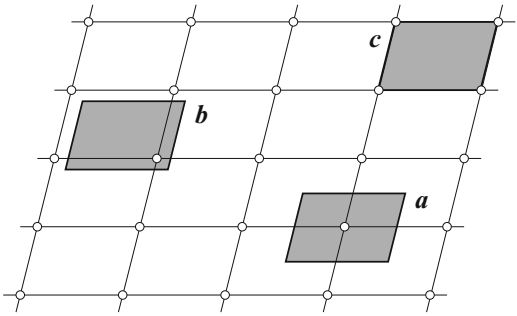


Fig. 17.2 Examples of cells in a two-dimensional Bravais lattice of the oblique type

important concept is that of *cell*. Still considering a two-dimensional lattice of the oblique type, a cell is a two-dimensional surface that is associated with each node and has the following properties: (*i*) the association is one-to-one, (*ii*) the cells are equal to each other, and (*iii*) the union of the cells covers the lattice exactly. The cell so defined is not unique; by way of examples, all the shaded surfaces shown in Fig. 17.2 fulfill the properties listed above. It may seem that the cell of case *c)* is not correct, because it touches four nodes; however, each node is shared by four cells, so that the one-to-one correspondence is maintained. In fact, the type of cell shown in case *c)* is most useful to extend the definitions given so far to the more realistic three-dimensional case. One notes in passing that the common value of the area of the cells depicted in Fig. 17.2 is $A = |\mathbf{a}_1 \wedge \mathbf{a}_2|$, with $\mathbf{a}_1, \mathbf{a}_2$ the characteristic vectors indicated in Fig. 17.1.

In a three-dimensional space the simplest cells are three-dimensional volumes that fulfill the same properties as in the two-dimensional case and have the atoms placed at the corners. Such an arrangement is called *primitive*. It can be shown that seven primitive arrangements are possible; each of them may further be enriched in five possible ways, by adding *a)* one atom at the center of the cell, or *b)* one atom at the center of each face of the cell, or *c)* one atom at the center of each pair of cell faces (this can be done in three different ways). The addition of extra atoms to the primitive arrangement is called *centering*. The total number of three-dimensional arrangements, including the primitive ones, is thus $7 \times 6 = 42$; however, it is found that not all of them are distinct, so that the actual number of distinct three-dimensional cells reduces to 14 [79].

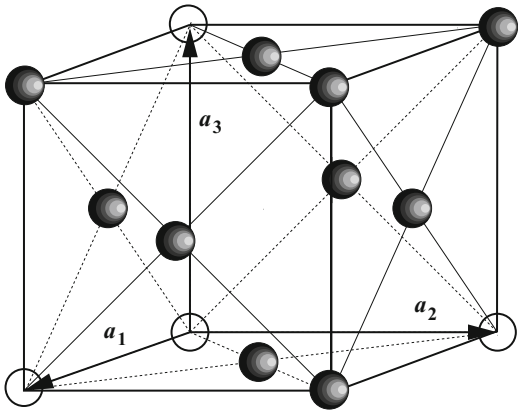


Fig. 17.3 Schematic description of a three-dimensional Bravais lattice of the FCC type. Four atoms have been removed to better show the characteristic vectors. The latter are orthogonal to each other and of equal length

Table 17.1 Crystal constants of silicon and germanium

Material	Lattice constant (nm)	Interatomic distance (nm)
Si	0.543	0.233
Ge	0.566	0.244

A three-dimensional lattice whose characteristic vectors are mutually orthogonal and of equal length is called *cubic*. Besides the primitive one, other two arrangements of cubic type are possible: the first one has one additional atom at the cell's center and is called *body-centered cubic* (BCC), the second one has one additional atom at the center of each cell's face and is called *face-centered cubic* (FCC). A portion of a lattice of the FCC type is shown in Fig. 17.3. Examples of chemical species whose crystalline phase has a cubic cell are carbon (C) in the diamond phase, silicon (Si), and germanium (Ge); in fact, this type of crystallization state is also collectively indicated with *diamond structure*. Elements like C, Si, and Ge belong to the fourth column of the periodic table of elements; Si and Ge are semiconductors, while C in the diamond phase is an insulator.[2] The FCC cell is also exhibited by some compound materials, an example of which is gallium arsenide (GaAs), a semiconductor of the so-called *III-V type*.[3] Some properties of Si, Ge, and GaAs, along with those of other III-V semiconductors, are listed in Tables 17.1 and 17.2. The type of crystallization state of GaAs and the other materials of Table 17.2 is called *zincblende structure*.[4]

[2]The meaning of "insulator" or "semiconductor," as well as that of "conductor," is specified in Sect. 18.2.

[3]The term "III-V" derives from the fact the Ga and As belong, respectively, to the third and fifth column of the periodic table of elements.

[4]*Zincblende* is another name for *Sphalerite*, an economically important mineral whence zinc is extracted. It consists of zinc sulfide in crystalline form with some contents of iron, (Zn,Fe)S.

Table 17.2 Crystal constants of some III-V semiconductors

Material	Lattice constant (nm)	Interatomic distance (nm)
GaAs	0.563	0.244
GaP	0.545	0.236
GaSb	0.609	0.264
InAs	0.605	0.262
InP	0.586	0.254
InSb	0.647	0.280
AlSb	0.613	0.265

17.3 Reciprocal Lattice

As indicated in Sect. 17.2, the type of cell having the nodes at the corners is the most useful one for introducing a number of definitions. For instance, observing that the characteristic vectors coincide with the cell's edges (this property is evident for the FCC cell as shown in Fig. 17.3; however, it applies also to the other types of cells), one obtains for the cell's volume

$$\tau_l = \mathbf{a}_1 \cdot \mathbf{a}_2 \wedge \mathbf{a}_3 = \mathbf{a}_2 \cdot \mathbf{a}_3 \wedge \mathbf{a}_1 = \mathbf{a}_3 \cdot \mathbf{a}_1 \wedge \mathbf{a}_2 \,, \tag{17.3}$$

where the orientation of the characteristic vectors is chosen in such a way as to make τ_l positive.

The major advantage of dealing with a periodic structure is the possibility of using the Fourier series. In fact, the functions describing the physical properties of the structure are expected to be periodic, so that one can take advantage of the Fourier expansion. The latter, in turn, entails a transformation from the coordinate space to another space reciprocal to it. Considering, for instance, the one-dimensional case of the Fourier transform given in (C.17), there the k space is the reciprocal of the x space; similarly, in an n-dimensional space like that considered in (C.20), the $\mathbf{k}$ space is the reciprocal of the $\mathbf{x}$ space: both vectors $\mathbf{k}$, $\mathbf{x}$ are linear combinations of the same set of mutually orthogonal, unit vectors $\mathbf{i}_1, \ldots, \mathbf{i}_n$, so that the scalar product $\mathbf{k} \cdot \mathbf{x}$ yields $k_1 x_1 + \ldots + k_n x_n$. In the case considered here, instead, one must account for the fact that the reference $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ of the space under consideration is made of vectors that, in general, are neither mutually orthogonal nor of equal length. For this reason it is necessary to introduce a *reciprocal lattice* by defining its characteristic vectors as

$$\mathbf{b}_1 = \frac{\mathbf{a}_2 \wedge \mathbf{a}_3}{\tau_l} \,, \qquad \mathbf{b}_2 = \frac{\mathbf{a}_3 \wedge \mathbf{a}_1}{\tau_l} \,, \qquad \mathbf{b}_3 = \frac{\mathbf{a}_1 \wedge \mathbf{a}_2}{\tau_l} \,. \tag{17.4}$$

From the definition (17.3) of τ_l and the mixed-product property (A.32) one finds that the characteristic vectors fulfill the orthogonality and normalization relation

$$\mathbf{a}_i \cdot \mathbf{b}_j = \delta_{ij} \,, \tag{17.5}$$

with δ_{ij} the Kronecker symbol (A.18). The translation vectors and the general vectors of the reciprocal lattice are linear combinations of $\mathbf{b}_1$, $\mathbf{b}_2$, $\mathbf{b}_3$ whose coefficients are, respectively, integer numbers or real numbers. To distinguish the lattice based on $\mathbf{b}_1$, $\mathbf{b}_2$, $\mathbf{b}_3$ from the one based on $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, the latter is also called *direct lattice*. The common value of the cells' volume in the reciprocal lattice is

$$\tau_G = \mathbf{b}_1 \cdot \mathbf{b}_2 \wedge \mathbf{b}_3 = \mathbf{b}_2 \cdot \mathbf{b}_3 \wedge \mathbf{b}_1 = \mathbf{b}_3 \cdot \mathbf{b}_1 \wedge \mathbf{b}_2 \,. \tag{17.6}$$

Observing that $\mathbf{a}_i$ is a length, from (17.4) it follows that $\mathbf{b}_j$ is the inverse of a length; as a consequence, the units of τ_G are m^{-3}, and the product $\gamma = \tau_l \, \tau_G$ is a dimensionless constant. The value of γ is found by combining the first relation in (17.4), which yields $\tau_l \, \mathbf{b}_1 = \mathbf{a}_2 \wedge \mathbf{a}_3$, with (17.6), so that

$$\gamma = \tau_l \, \tau_G = \tau_l \, \mathbf{b}_1 \cdot \mathbf{b}_2 \wedge \mathbf{b}_3 = \mathbf{a}_2 \wedge \mathbf{a}_3 \cdot (\mathbf{b}_2 \wedge \mathbf{b}_3) = \mathbf{a}_2 \cdot \mathbf{a}_3 \wedge (\mathbf{b}_2 \wedge \mathbf{b}_3) \,, \tag{17.7}$$

where the last equality is due to the invariance of the mixed product upon interchange of the "wedge" and "dot" symbols (Sect. A.7). Then, using (A.33) to resolve the double vector product,

$$\tau_l \, \tau_G = \mathbf{a}_2 \cdot (\mathbf{a}_3 \cdot \mathbf{b}_3 \, \mathbf{b}_2 - \mathbf{a}_3 \cdot \mathbf{b}_2 \, \mathbf{b}_3) = \mathbf{a}_2 \cdot (\delta_{33} \, \mathbf{b}_2 - \delta_{32} \, \mathbf{b}_3) = \mathbf{a}_2 \cdot \mathbf{b}_2 = 1 \,. \tag{17.8}$$

From the properties that provide the definition of cell (Sect. 17.2) it follows that for a given lattice, the volume of the cell does not depend on its form. As a consequence it is $\tau_l \, \tau_G = 1$ no matter what the cell's form is. Also, after defining $\mathbf{e}_1 = \mathbf{b}_2 \wedge \mathbf{b}_3 / \tau_G$ one finds, by a similar calculation,

$$\mathbf{e}_1 = \frac{\mathbf{b}_2 \wedge \mathbf{b}_3}{\tau_G} = \frac{\mathbf{b}_2 \wedge (\mathbf{a}_1 \wedge \mathbf{a}_2)}{\tau_l \, \tau_G} = \mathbf{b}_2 \cdot \mathbf{a}_2 \, \mathbf{a}_1 - \mathbf{b}_2 \cdot \mathbf{a}_1 \, \mathbf{a}_2 = \mathbf{a}_1 \,, \tag{17.9}$$

and the like for $\mathbf{e}_2$ and $\mathbf{e}_3$; thus, the direct lattice is the reciprocal of the reciprocal lattice.

Given a direct lattice of characteristic vectors $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, it is convenient to introduce, besides the reciprocal lattice of characteristic vectors $\mathbf{b}_1$, $\mathbf{b}_2$, $\mathbf{b}_3$ defined by (17.4), another lattice called *scaled reciprocal lattice*, whose characteristic vectors are $2\,\pi\,\mathbf{b}_1, 2\,\pi\,\mathbf{b}_2, 2\,\pi\,\mathbf{b}_3$. A translation vector of the scaled reciprocal lattice has the form

$$\mathbf{g} = n_1 \, 2\,\pi \, \mathbf{b}_1 + n_2 \, 2\,\pi \, \mathbf{b}_2 + n_3 \, 2\,\pi \, \mathbf{b}_3 \,, \tag{17.10}$$

with n_1, n_2, n_3 any integers, whereas a general vector of the scaled reciprocal lattice has the form

$$\mathbf{k} = \nu_1 \, 2\,\pi \, \mathbf{b}_1 + \nu_2 \, 2\,\pi \, \mathbf{b}_2 + \nu_3 \, 2\,\pi \, \mathbf{b}_3 \,, \tag{17.11}$$

with ν_1, ν_2, ν_3 any real numbers. From the definitions (17.1,17.10) of $\mathbf{l}$ and $\mathbf{g}$ one finds

$$\mathbf{l} \cdot \mathbf{g} = \sum_{is=1}^{3} m_i \, \mathbf{a}_i \cdot n_s \, 2\,\pi \, \mathbf{b}_s = 2\,\pi \sum_{is=1}^{3} m_i \, n_s \, \delta_{is} = 2\,\pi \sum_{i=1}^{3} m_i \, n_i \,, \tag{17.12}$$

namely, $\mathbf{l} \cdot \mathbf{g} = 2\,\pi\, M$ with M an integer. It follows that

$$\exp\left[\mathrm{i}\,\mathbf{g} \cdot (\mathbf{r} + \mathbf{l})\right] = \exp\left(\mathrm{i}\,\mathbf{g} \cdot \mathbf{r}\right) \exp\left(\mathrm{i}\,2\,\pi\, M\right) = \exp\left(\mathrm{i}\,\mathbf{g} \cdot \mathbf{r}\right) \,, \tag{17.13}$$

that is, $\exp\left(\mathrm{i}\,\mathbf{g} \cdot \mathbf{r}\right)$ is periodic in the $\mathbf{r}$ space. This shows the usefulness of the scaled reciprocal lattice for treating problems related to periodic structures. In fact, given a periodic function in the $\mathbf{r}$ space, $F(\mathbf{r} + \mathbf{l}) = F(\mathbf{r})$, the Fourier expansion is generalized to the non-orthogonal case as

$$F(\mathbf{r}) = \sum_{\mathbf{g}} F_{\mathbf{g}} \, \exp\left(\mathrm{i}\,\mathbf{g} \cdot \mathbf{r}\right) \,, \qquad F_{\mathbf{g}} = \frac{1}{\tau_l} \int_{\tau_l} F(\mathbf{r}) \, \exp\left(-\mathrm{i}\,\mathbf{g} \cdot \mathbf{r}\right) \, \mathrm{d}^3 r \,, \tag{17.14}$$

with $\sum_{\mathbf{g}} = \sum_{n_1} \sum_{n_2} \sum_{n_3}$. The property holds also in reverse, namely,

$$\exp[\mathrm{i}\,\mathbf{l} \cdot (\mathbf{k} + \mathbf{g})] = \exp(\mathrm{i}\,\mathbf{l} \cdot \mathbf{k}) \,, \tag{17.15}$$

so that, given a periodic function in the $\mathbf{k}$ space, $\Phi(\mathbf{k} + \mathbf{g}) = \Phi(\mathbf{k})$, the following expansion holds:

$$\Phi(\mathbf{k}) = \sum_{\mathbf{l}} \Phi_{\mathbf{l}} \, \exp\left(\mathrm{i}\,\mathbf{l} \cdot \mathbf{k}\right) \,, \qquad \Phi_{\mathbf{l}} = \frac{1}{\tau_g} \int_{\tau_g} \Phi(\mathbf{k}) \, \exp\left(-\mathrm{i}\,\mathbf{l} \cdot \mathbf{k}\right) \, \mathrm{d}^3 k \,, \tag{17.16}$$

with $\sum_{\mathbf{l}} = \sum_{m_1} \sum_{m_2} \sum_{m_3}$. From (17.8) one also finds that in the scaled reciprocal lattice the volume of the cell is given by

$$\tau_g = \frac{(2\,\pi)^3}{\tau_G} = \frac{(2\pi)^3}{\tau_l} \,. \tag{17.17}$$

The origin of the reference of the direct or reciprocal space has not been identified so far. After selecting the origin, consider the cell of the $\mathbf{r}$ space whose sides emanate from it; these sides are made to coincide with the characteristic vectors (compare, e.g., with Fig. 17.3), so that, to any point $\mathbf{r} = \mu_1 \, \mathbf{a}_1 + \mu_2 \, \mathbf{a}_2 + \mu_3 \, \mathbf{a}_3$ that belongs to the interior or the boundary of the cell, the following restriction apply: $0 \leq \mu_i \leq 1$. Similarly, if one considers the cell of the $\mathbf{k}$ space whose sides emanate from the origin, for any point $\mathbf{k} = \nu_1 \, 2\,\pi \, \mathbf{b}_1 + \nu_2 \, 2\,\pi \, \mathbf{b}_2 + \nu_3 \, 2\,\pi \, \mathbf{b}_3$ that belongs to the interior or the boundary of the cell it is $0 \leq \nu_i \leq 1$.

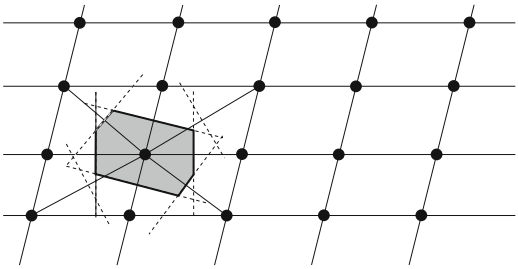


Fig. 17.4 A Wigner-Seitz cell in a two-dimensional, oblique lattice

17.4 Wigner-Seitz Cell—Brillouin Zone

It has been mentioned in Sect. 17.2 that the cell properties do not identify the cell uniquely. Cells of a form different from that shown, e.g., in the two-dimensional example of Fig. 17.2 can be constructed. Among them, of particular interest is the *Wigner-Seitz* cell, whose construction is shown in Fig. 17.4 still considering a two-dimensional lattice of the oblique type. First, the node to be associated with the cell is connected to its nearest neighbors as shown in the figure (the connecting segments are the continuous lines); then, the axis of each segment is drawn (dashed lines). As the axis is the locus of points having the same distance from the extrema of the connecting segment, one can select a contour, made of portions of the axes (thick line), such that the following property holds: any point inside the contour is closer to the node associated with the cell than to any other lattice node. From the construction described above it also follows that the shaded surface so obtained fulfills the requisites listed in Sect. 17.2, so it is a cell proper; such a cell, named after Wigner and Seitz, fulfills the additional property of the closeness of the internal points. Its construction in the three-dimensional case is similar, the only difference being that the axis is replaced with the plane normal to the connecting segment at the midpoint.

The concept of Wigner-Seitz cell, that has been introduced above using the direct lattice by way of example, is applicable also to the reciprocal and scaled-reciprocal lattices. In the scaled reciprocal lattice, the Wigner-Seitz cell associated with the origin is called *first Brillouin zone*. The set of Wigner-Seitz cells adjacent to the first Brillouin zone is called *second Brillouin zone*, and so on. The first Brillouin zone of the FCC lattice is shown in Fig. 17.5; its boundary is made of six square faces and eight hexagonal faces. The center of the zone is called *Γ point*; the k_1 axis belongs to the [100] crystal direction[5] and intercepts the boundary of the Brillouin zone at two opposite positions called *X points*, that coincide with the center of square faces; the k_2, k_3 axes belong to the [010] and [001] directions, respectively, and intercept the boundary at *X* points as well. In turn, the {111} directions intercept the boundary at positions called *L points* that coincide with the center of the hexagonal faces. There

[5]The symbols indicating the crystal directions are illustrated in Sect. 17.9.1.

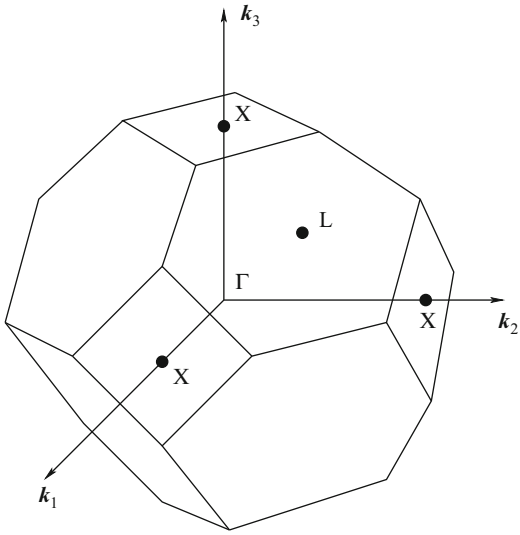


Fig. 17.5 The first Brillouin zone of the FCC lattice

is a total of eight L points, because the set $\{111\}$ is made of the $[111]$, $[\bar{1}11]$, $[\bar{1}\bar{1}1]$, $[1\bar{1}1]$ directions along with those of complementary signs.

17.5 Translation Operators

The typical procedure by which the physical properties of periodic structures, like crystals, are investigated entails the solution of eigenvalue equations generated by quantum-mechanical operators. In this respect it is useful to introduce a class of operators, called *translation operators*, that are associated with the direct-lattice vectors $\mathbf{l}$ defined by (17.1). To ease the notation, the translation operator associated with the ith lattice vector $\mathbf{l}_i$ is indicated with $\mathscr{T}_i = \mathscr{T}(\mathbf{l}_i)$. A translation operator is defined by the property

$$\mathscr{T}_i f(\mathbf{r}) = f(\mathbf{r} + \mathbf{l}_i) \tag{17.18}$$

for all functions f defined over the direct lattice. It is easily found that translation operators are linear and non-Hermitean. Also, they commute with each other; in fact,

$$\mathscr{T}_i \mathscr{T}_s f(\mathbf{r}) = f(\mathbf{r} + \mathbf{l}_i + \mathbf{l}_s) = \mathscr{T}_s \mathscr{T}_i f(\mathbf{r}) \tag{17.19}$$

for all functions f and all indices i, s. Remembering the property derived in Sect. 10.4, it follows that all translation operators have a common set of eigenfunctions v. Combining this result with definition (17.18) provides the form of the

eigenvalues α of the translation operators. For this, consider the operators associated with three arbitrary vectors $\mathbf{l}_i$, $\mathbf{l}_s$, and $\mathbf{l}_u$, and generate the eigenvalue equations

$$\mathscr{T}_i\, v = \alpha(\mathbf{l}_i)\, v\,, \qquad \mathscr{T}_s\, v = \alpha(\mathbf{l}_s)\, v\,, \qquad \mathscr{T}_u\, v = \alpha(\mathbf{l}_u)\, v\,. \tag{17.20}$$

If one now lets $\mathbf{l}_u = \mathbf{l}_i + \mathbf{l}_s$, it follows $\mathscr{T}_i \mathscr{T}_s\, v(\mathbf{r}) = v(\mathbf{r} + \mathbf{l}_i + \mathbf{l}_s) = \mathscr{T}_u\, v(\mathbf{r})$, that is, $\mathscr{T}_i \mathscr{T}_s = \mathscr{T}_u$. On the other hand, the first two eigenvalue equations in (17.20) provide $\mathscr{T}_i \mathscr{T}_s\, v = \mathscr{T}_i\, \alpha(\mathbf{l}_s)\, v = \alpha(\mathbf{l}_i)\, \alpha(\mathbf{l}_s)\, v$ which, combined with the third one, yields

$$\alpha(\mathbf{l}_i)\, \alpha(\mathbf{l}_s) = \alpha(\mathbf{l}_u) = \alpha(\mathbf{l}_i + \mathbf{l}_s)\,. \tag{17.21}$$

The result shows that the functional dependence of α on the translation vector must be of the exponential type:

$$\alpha(\mathbf{l}) = \exp(\mathbf{c} \cdot \mathbf{l})\,, \qquad \mathbf{c} = \sum_{s=1}^{3} (\Re\chi_s + \mathrm{i}\, \Im\chi_s)\, 2\,\pi\, \mathbf{b}_s\,, \tag{17.22}$$

where $\mathbf{c}$ is a complex vector whose units are the inverse of a length. For this reason $\mathbf{c}$ is given the general form shown in the second relation of (17.22), with $\mathbf{b}_s$ a characteristic vector of the reciprocal lattice, $s = 1, 2, 3$, and χ_s a dimensionless complex number that is momentarily left unspecified.

17.5.1 Bloch Theorem

The expression (17.22) of the eigenvalues of the translation operators makes it possible to specify some property of the eigenfunctions; in fact, combining the definition of translation operator, $\mathscr{T} v(\mathbf{r}) = v(\mathbf{r} + \mathbf{l})$, with the eigenvalue equation $\mathscr{T} v(\mathbf{r}) = \exp(\mathbf{c} \cdot \mathbf{l})\, v(\mathbf{r})$ yields

$$v_{\mathbf{c}}(\mathbf{r} + \mathbf{l}) = \exp(\mathbf{c} \cdot \mathbf{l})\, v_{\mathbf{c}}(\mathbf{r})\,, \tag{17.23}$$

called *Bloch theorem (first form)*. The importance of this result can be appreciated by observing that if v is known within a lattice cell, and $\mathbf{c}$ is given, then the eigenfunction can be reconstructed everywhere else. The index in (17.23) reminds one that the eigenfunction depends on the choice of $\mathbf{c}$. The theorem can be recast differently by defining an auxiliary function $u_{\mathbf{c}}(\mathbf{r}) = v_{\mathbf{c}}(\mathbf{r})\, \exp(-\mathbf{c} \cdot \mathbf{r})$, so that

$$v_{\mathbf{c}}(\mathbf{r} + \mathbf{l}) = \exp(\mathbf{c} \cdot \mathbf{l})\, v_{\mathbf{c}}(\mathbf{r}) = \exp(\mathbf{c} \cdot \mathbf{l})\, u_{\mathbf{c}}(\mathbf{r})\, \exp(\mathbf{c} \cdot \mathbf{r})\,. \tag{17.24}$$

In turn, from the definition of $u_{\mathbf{c}}$ one draws $v_{\mathbf{c}}(\mathbf{r} + \mathbf{l}) = u_{\mathbf{c}}(\mathbf{r} + \mathbf{l})\, \exp[\mathbf{c} \cdot (\mathbf{r} + \mathbf{l})]$ which, combined with (17.24), yields the *Bloch theorem (second form)*:

$$v_{\mathbf{c}}(\mathbf{r}) = u_{\mathbf{c}}(\mathbf{r})\, \exp(\mathbf{c}\cdot\mathbf{r})\,, \qquad u_{\mathbf{c}}(\mathbf{r}+\mathbf{l}) = u_{\mathbf{c}}(\mathbf{r})\,. \tag{17.25}$$

The second form of the Bloch theorem shows that the auxiliary function $u_{\mathbf{c}}$ is periodic in the direct lattice, so that the eigenfunctions of the translation operators are the product of an exponential function times a function having the lattice periodicity. One notes the similarity of this result with that expressed by (13.39); in fact, the Bloch theorem is a form of the Floquet theorem (Sect. 13.4).

The eigenfunctions of the translation operators play an important role in the description of the physical properties of periodic structures. For this reason, vectors $\mathbf{c}$ of the general form are not acceptable because their real part would make $v_{\mathbf{c}}(\mathbf{r}) = u_{\mathbf{c}}(\mathbf{r})\, \exp(\mathbf{c}\cdot\mathbf{r})$ to diverge as $\mathbf{r}$ departs more and more from the origin.[6] It is then necessary to impose the restriction $\mathbf{c} = \mathrm{i}\,\mathbf{k}$, with $\mathbf{k}$ real. This is achieved by letting $\Re\chi_s = 0$ and $\Im\chi_s = \nu_s$ in the second relation of (17.22), so that the eigenvalues of the translation operators become

$$\alpha(\mathbf{l}) = \exp(\mathrm{i}\,\mathbf{k}\cdot\mathbf{l})\,, \qquad \mathbf{k} = \sum_{s=1}^{3} \nu_s\, 2\,\pi\, \mathbf{b}_s\,. \tag{17.26}$$

Remembering (17.15), such eigenvalues are periodic in the scaled reciprocal lattice. In turn, the first and second forms of the Bloch theorem become, respectively,

$$v_{\mathbf{k}}(\mathbf{r}+\mathbf{l}) = \exp(\mathrm{i}\,\mathbf{k}\cdot\mathbf{l})\, v_{\mathbf{k}}(\mathbf{r})\,, \tag{17.27}$$

$$v_{\mathbf{k}}(\mathbf{r}) = u_{\mathbf{k}}(\mathbf{r})\, \exp(\mathrm{i}\,\mathbf{k}\cdot\mathbf{r})\,, \qquad u_{\mathbf{k}}(\mathbf{r}+\mathbf{l}) = u_{\mathbf{k}}(\mathbf{r})\,. \tag{17.28}$$

Eigenfunctions of the form (17.27,17.28) are also called *Bloch functions*. They fulfill the eigenvalue equation $\mathscr{T} v_{\mathbf{k}}(\mathbf{r}) = \exp(\mathrm{i}\,\mathbf{k}\cdot\mathbf{l})\, v_{\mathbf{k}}(\mathbf{r})$ so that, observing that $\mathscr{T}$ is real and taking the conjugate of the eigenvalue equation yield

$$\mathscr{T} v^*_{\mathbf{k}}(\mathbf{r}) = \exp(-\mathrm{i}\,\mathbf{k}\cdot\mathbf{l})\, v^*_{\mathbf{k}}(\mathbf{r})\,. \tag{17.29}$$

If, instead, one replaces $\mathbf{k}$ with $-\mathbf{k}$ in the original equation, the following is found:

$$\mathscr{T} v_{-\mathbf{k}}(\mathbf{r}) = \exp(-\mathrm{i}\,\mathbf{k}\cdot\mathbf{l})\, v_{-\mathbf{k}}(\mathbf{r})\,. \tag{17.30}$$

Comparing (17.30) with (17.29) shows that $v^*_{\mathbf{k}}$ and $v_{-\mathbf{k}}$ belong to the same eigenvalue. Moreover, comparing the second expression in (17.26) with (17.11) shows that $\mathbf{k}$ is a vector of the scaled reciprocal lattice.

A further reasoning demonstrates that the variability of $\mathbf{k}$ in (17.27,17.28) can be limited to a single cell of the scaled reciprocal lattice; for instance, to the first Brillouin zone or, alternatively, to the cell of the $\mathbf{k}$ space whose sides emanate from the origin, so that the coefficients of (17.11) fulfill the relation $0 \leq \nu_i \leq 1$ as shown

[6] This aspect is further elaborated in Sect. 17.5.3.

in Sect. 17.3. The property derives from the periodicity of the eigenvalues in the scaled reciprocal lattice, $\exp[\mathrm{i}\,(\mathbf{k}+\mathbf{g})\cdot\mathbf{l}] = \exp(\mathrm{i}\,\mathbf{k}\cdot\mathbf{l})$, due to which the values of $\exp(\mathrm{i}\,\mathbf{k}\cdot\mathbf{l})$, with $\mathbf{k}$ ranging over a single cell, provide the whole set of the operator's eigenvalues. As $\exp[\mathrm{i}\,(\mathbf{k}+\mathbf{g})\cdot\mathbf{l}]$ is the same eigenvalue as $\exp(\mathrm{i}\,\mathbf{k}\cdot\mathbf{l})$, the Bloch function $v_{\mathbf{k}+\mathbf{g}}(\mathbf{r})$ is the same as $v_{\mathbf{k}}(\mathbf{r})$. Note that the reasoning does not prevent the eigenvalue from being degenerate: if this is the case, one finds

$$v^{(1)}_{\mathbf{k}+\mathbf{g}} = v^{(1)}_{\mathbf{k}}\,, \qquad v^{(2)}_{\mathbf{k}+\mathbf{g}} = v^{(2)}_{\mathbf{k}}\,, \qquad \ldots \tag{17.31}$$

17.5.2 Periodic Operators

An operator $\mathscr{A}$ is *periodic* in the direct lattice if $\mathscr{A}(\mathbf{r}+\mathbf{l}) = \mathscr{A}(\mathbf{r})$ for all vectors $\mathbf{r}$ and all translation vectors $\mathbf{l}$ of the direct lattice. Periodic operators commute with translation operators: this is shown by letting $v' = \mathscr{A}v$, so that

$$\begin{aligned} &\mathscr{T}(\mathbf{l})\mathscr{A}(\mathbf{r})v(\mathbf{r}) = \mathscr{T}(\mathbf{l})v'(\mathbf{r}) = v'(\mathbf{r}+\mathbf{l}) = \\ &= \mathscr{A}(\mathbf{r}+\mathbf{l})v(\mathbf{r}+\mathbf{l}) = \mathscr{A}(\mathbf{r})v(\mathbf{r}+\mathbf{l}) = \mathscr{A}(\mathbf{r})\mathscr{T}(\mathbf{l})v(\mathbf{r})\,. \end{aligned} \tag{17.32}$$

From the commutativity property $\mathscr{T}\mathscr{A} = \mathscr{A}\mathscr{T}$ it follows that $\mathscr{T}$ and $\mathscr{A}$ have a common set of eigenfunctions, so that the eigenfunctions of a periodic operator are Bloch functions; letting $A_{\mathbf{k}}$ be the eigenvalue, one has

$$\mathscr{A}v_{\mathbf{k}} = A_{\mathbf{k}}\,v_{\mathbf{k}}\,, \tag{17.33}$$

with $\mathbf{k}$ belonging to a single cell of the scaled reciprocal lattice, and $v_{\mathbf{k}+\mathbf{g}} = v_{\mathbf{k}}$. Since an eigenfunction belongs to one eigenvalue only, it follows

$$A_{\mathbf{k}+\mathbf{g}} = A_{\mathbf{k}}\,, \tag{17.34}$$

namely, if an operator is periodic in the direct lattice, its eigenvalues are periodic in the scaled reciprocal lattice.

17.5.3 Periodic Boundary Conditions

In the derivation of the Bloch theorem, carried out in Sect. 17.5.1, it has been observed that complex vectors $\mathbf{c}$ in the expression $v_{\mathbf{c}}(\mathbf{r}) = u_{\mathbf{c}}(\mathbf{r})\,\exp(\mathbf{c}\cdot\mathbf{r})$ of the eigenfunctions are not acceptable because their real part would make the function to diverge. In fact, as noted in Sect. 17.5.2, such eigenfunctions belong also to the

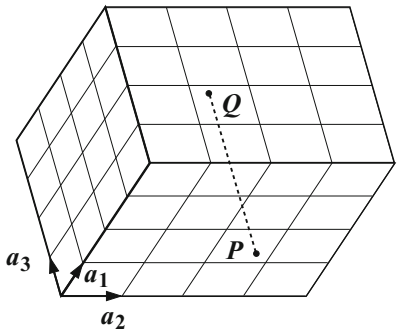


Fig. 17.6 A finite block of material obtained by sectioning a crystal by means of three pairs of parallel crystal planes

operators that commute with the translation operators and may describe physical properties of the crystal, so that diverging solutions must be discarded. On the other hand, such a divergence is due to the assumption that the crystal is unlimited; in the more realistic case of a finite block of material, diverging solutions would not appear. Unfortunately, a finite block of material is not periodic, hence the useful concepts and properties worked out so far in this chapter are not applicable to it.

To further investigate on the subject one notes that a finite block may be thought of as derived from the original crystal by sectioning the latter using three pairs of parallel crystal planes, as shown in Fig. 17.6. One of the vertices of the block coincides with the origin of the reference of the direct lattice, and the block's sides are aligned with the characteristic vectors. Also, the type of cell chosen here is the one whose sides coincide with the characteristic vectors themselves. The relation between the total number N_c of the block's cells, the block's volume Ω, and that of the cell is easily found to be

$$N_c = N_1 N_2 N_3 = \frac{\Omega}{\tau_l} , \tag{17.35}$$

with N_s the number of cell sides that match the side of the block in the sth direction. In a finite volume of material, the number of cells that belong to the interior is typically much larger than the number of cells that are adjacent to the boundaries; when solving a single-electron Schrödinger equation within such a structure, it is found that in the interior of it the eigenfunctions practically coincide with Bloch functions, whereas the effect of the real part of vector $\mathbf{c}$ in $v_{\mathbf{c}}(\mathbf{r}) = u_{\mathbf{c}}(\mathbf{r}) \exp(\mathbf{c} \cdot \mathbf{r})$ becomes relevant only when the position under investigation is close to a boundary. In fact, the real part of $\mathbf{c}$ is such that the eigenfunctions become vanishingly small far away from the volume considered [24, Sects. F-XI, O-III].

The considerations above show that for practical purposes one may keep the analysis based on the original periodicity of the unlimited structure by replacing the vanishing boundary conditions with a different type of conditions, able to formally restore periodicity in a finite structure. This is accomplished by imposing the identity of the Bloch functions corresponding to two boundary points facing each other along the direction of a characteristic vector. This prescription, called *periodic*

boundary condition or *Born-Von Karman boundary condition*, is illustrated with the aid of Fig. 17.6. Consider for instance point $\mathbf{r} = \mu_1 \mathbf{a}_1 + \mu_2 \mathbf{a}_2$ (labeled P in the figure), that belongs to the boundary plane defined by $\mathbf{a}_1, \mathbf{a}_2$. Point Q facing P on the opposite boundary plane is such that $Q - P = \mathbf{l} = N_3 \mathbf{a}_3$ whence, applying the first form (17.23) of Bloch's theorem, one obtains $v_c(\mathbf{r} + N_3 \mathbf{a}_3) = \exp(N_3 \mathbf{c} \cdot \mathbf{a}_3)\, v_c(\mathbf{r})$. Imposing $v_c(\mathbf{r} + N_3 \mathbf{a}_3) = v_c(\mathbf{r})$ yields $N_3 \mathbf{c} \cdot \mathbf{a}_3 = \mathrm{i}\, n_3\, 2\pi$, with n_3 any integer, so that, using expression (17.22) for $\mathbf{c}$,

$$N_3 \mathbf{c} \cdot \mathbf{a}_3 = 2\pi N_3 (\Re\chi_3 + \mathrm{i}\,\Im\chi_3) = \mathrm{i}\, n_3\, 2\pi \,. \tag{17.36}$$

In conclusion, $\Re\chi_3 = 0$ and $\Im\chi_3 = 2\pi n_3/N_3$. The same reasoning is repeated along the other directions to finally yield

$$\mathbf{c} = \mathrm{i}\,\mathbf{k}\,, \qquad \mathbf{k} = \sum_{s=1}^{3} \frac{n_s}{N_s} 2\pi \mathbf{b}_s \,. \tag{17.37}$$

In summary, the application of the periodic boundary conditions gives $\mathbf{c}$ the same imaginary form $\mathbf{c} = \mathrm{i}\,\mathbf{k}$ that was found in an unlimited structure, the difference being that in a finite structure the components of $\mathbf{k}$ are discrete instead of being continuous: given the size of the structure, which is prescribed by N_1, N_2, N_3, each $\mathbf{k}$ vector of the scaled reciprocal lattice is associated with a triad of integers n_1, n_2, n_3.

Note that the reasoning carried out at the end of Sect. 17.5.1 about the variability of $\mathbf{k}$ still holds; as a consequence, $\mathbf{k}$ can be restricted to a single cell of the scaled reciprocal lattice, so that its coefficients $\nu_s = n_s/N_s$ fulfill the relation $0 \le \nu_s \le 1$ as shown in Sect. 17.3. In fact, as $\nu_s = 0$ and $\nu_s = 1$ are redundant, the above relation must more appropriately be recast as $0 \le n_s/N_s < 1$, corresponding to $n_s = 0, 1, \ldots, N_s - 1$ or, alternatively, $0 < n_s/N_s \le 1$, corresponding to $n_s = 1, 2, \ldots, N_s$. In both cases, n_s can take N_s distinct values, so that the total number of distinct $\mathbf{k}$ vectors in a cell of the scaled reciprocal lattice is $N_1 N_2 N_3 = N_c$. From (17.35) one finds that such a number equals the number of the structure's cells in the direct lattice. Also, as the $\mathbf{k}$ vectors are equally spaced in each direction, their density in the reciprocal scaled lattice is uniform; it is given by the ratio N_c/τ_g, with τ_g the cell's volume. Remembering that the latter is invariant (Sect. 17.3), one may think of $\mathbf{k}$ as restricted to the first Brillouin zone. Combining (17.8) with (17.17) and (17.35) yields for the density

$$\frac{N_c}{\tau_g} = \frac{\Omega/\tau_l}{\tau_g} = \frac{\Omega}{(2\pi)^3} \,. \tag{17.38}$$

One can also define a combined density of the $\mathbf{k}$ vectors in the $\mathbf{r}, \mathbf{k}$ space, which is obtained by dividing (17.38) by the volume Ω. This yields the dimensionless combined density

$$\frac{1}{\Omega} \frac{N_c}{\tau_g} = \frac{1}{(2\pi)^3} \,. \tag{17.39}$$

17.6 Schrödinger Equation in a Periodic Lattice

The concepts introduced in the previous sections of this chapter are applied here to the solution of the Schrödinger equation in a periodic lattice. It is assumed provisionally that the lattice is unlimited; as a consequence, the components of the **k** vector are continuous. The Schrödinger equation to be solved is the single-electron equation (16.27) obtained from the separation procedure outlined in Sects. 16.2 through 16.5. This means that the nuclei are kept fixed and the force acting on the electron derives from a potential energy[7] having the periodicity of the direct lattice: $V(\mathbf{r}+\mathbf{l}) = V(\mathbf{r})$, with **l** given by (17.1). As mentioned in Sects. 16.4 and 16.5, the external forces are absent ($U_{\text{ext}} = 0$). The equation then reads

$$\mathscr{H} w = E\, w\,, \qquad \mathscr{H} = -\frac{\hbar^2}{2\,m}\,\nabla^2 + V\,. \tag{17.40}$$

Replacing **r** with $\mathbf{r}+\mathbf{l}$ is equivalent to add a constant to each component of **r**, say, $x_i \leftarrow x_i + l_i$, hence the partial derivatives in (17.40) are unaffected. As a consequence, the Hamiltonian operator as a whole has the lattice periodicity, so that its eigenfunctions are Bloch functions. Remembering (17.28), they read

$$w_{\mathbf{k}}(\mathbf{r}) = u_{\mathbf{k}}(\mathbf{r})\, \exp(\mathrm{i}\,\mathbf{k}\cdot\mathbf{r})\,, \qquad u_{\mathbf{k}}(\mathbf{r}+\mathbf{l}) = u_{\mathbf{k}}(\mathbf{r})\,, \tag{17.41}$$

with **k** belonging to the first Brillouin zone (Sect. 17.5.3). Letting $k^2 = |\mathbf{k}|^2$, (17.41) yields $\nabla^2 w_{\mathbf{k}} = \exp(\mathrm{i}\,\mathbf{k}\cdot\mathbf{r})\,(-k^2 + \nabla^2 + 2\,\mathrm{i}\,\mathbf{k}\cdot\mathrm{grad})\,u_{\mathbf{k}}$ whence, if $E_{\mathbf{k}}$ is the eigenvalue corresponding to $w_{\mathbf{k}}$, the Schrödinger equation (17.40) becomes

$$V\,u_{\mathbf{k}} = \left[E_{\mathbf{k}} + \frac{\hbar^2}{2\,m}\,\left(-k^2 + \nabla^2 + 2\,\mathrm{i}\,\mathbf{k}\cdot\mathrm{grad}\right)\right] u_{\mathbf{k}}\,. \tag{17.42}$$

As both V and $u_{\mathbf{k}}$ have the periodicity of the lattice, they can be expanded in terms of the translation vectors of the scaled reciprocal lattice $\mathbf{g} = n_1\,2\,\pi\,\mathbf{b}_1 + n_2\,2\,\pi\,\mathbf{b}_2 + n_3\,2\,\pi\,\mathbf{b}_3$:

$$V(\mathbf{r}) = \sum_{\mathbf{g}} V_{\mathbf{g}}\, \exp(\mathrm{i}\,\mathbf{g}\cdot\mathbf{r})\,, \qquad u_{\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{g}} s_{\mathbf{k}\mathbf{g}}\, \exp(\mathrm{i}\,\mathbf{g}\cdot\mathbf{r})\,, \tag{17.43}$$

where $\sum_{\mathbf{g}} = \sum_{n_1}\sum_{n_2}\sum_{n_3}$ and

$$V_{\mathbf{g}} = \frac{1}{\tau_l}\int_{\tau_l} V(\mathbf{r})\, \exp(-\mathrm{i}\,\mathbf{g}\cdot\mathbf{r})\,\mathrm{d}^3 r\,, \qquad s_{\mathbf{k}\mathbf{g}} = \frac{1}{\tau_l}\int_{\tau_l} u_{\mathbf{k}}(\mathbf{r})\, \exp(-\mathrm{i}\,\mathbf{g}\cdot\mathbf{r})\,\mathrm{d}^3 r\,. \tag{17.44}$$

[7] The indices of (16.27) are dropped for simplicity.

Letting $g^2 = |\mathbf{g}|^2$, from the expansion of $u_\mathbf{k}$ it follows $(\nabla^2 + 2\,\mathrm{i}\,\mathbf{k}\cdot\mathrm{grad})u_\mathbf{k} = -\sum_\mathbf{g}(g^2 + 2\,\mathbf{k}\cdot\mathbf{g})\,s_\mathbf{kg}\,\exp(\mathrm{i}\,\mathbf{g}\cdot\mathbf{r})$ whence, using $g^2 + 2\,\mathbf{k}\cdot\mathbf{g} + k^2 = |\mathbf{g}+\mathbf{k}|^2$,

$$V\,u_\mathbf{k} = \sum_\mathbf{g}\left[E_\mathbf{k} - \frac{\hbar^2}{2\,m}\,|\mathbf{g}+\mathbf{k}|^2\right] s_\mathbf{kg}\,\exp(\mathrm{i}\,\mathbf{g}\cdot\mathbf{r})\,. \tag{17.45}$$

In turn, the left-hand side of (17.45) reads

$$V\,u_\mathbf{k} = \sum_{\mathbf{g}'}\sum_{\mathbf{g}''} V_{\mathbf{g}'}\,s_{\mathbf{kg}''}\,\exp[\mathrm{i}\,(\mathbf{g}'+\mathbf{g}'')\cdot\mathbf{r}] = \sum_{\mathbf{g}'}\sum_{\mathbf{g}-\mathbf{g}'} V_{\mathbf{g}'}\,s_{\mathbf{k},\mathbf{g}-\mathbf{g}'}\,\exp(\mathrm{i}\,\mathbf{g}\cdot\mathbf{r})\,, \tag{17.46}$$

with $\mathbf{g} = \mathbf{g}' + \mathbf{g}''$. Note that the last expression on the right of (17.46) is left unchanged if $\sum_{\mathbf{g}-\mathbf{g}'}$ is replaced with $\sum_\mathbf{g}$. In fact, as for each vector $\mathbf{g}'$ the indices n_i of $\mathbf{g}$ span from $-\infty$ to $+\infty$, all ∞^6 combinations of indices of $\mathbf{g}$ and $\mathbf{g}'$ are present in either form of the expansion; using $\sum_\mathbf{g}$ instead of $\sum_{\mathbf{g}-\mathbf{g}'}$ merely changes the order of summands. Combining (17.45) with (17.46) then yields

$$\sum_\mathbf{g}\exp(\mathrm{i}\,\mathbf{g}\cdot\mathbf{r})\left\{\sum_{\mathbf{g}'} V_{\mathbf{g}'}\,s_{\mathbf{k},\mathbf{g}-\mathbf{g}'} - \left[E_\mathbf{k} - \frac{\hbar^2}{2\,m}\,|\mathbf{g}+\mathbf{k}|^2\right] s_\mathbf{kg}\right\} = 0\,. \tag{17.47}$$

As the factors $\exp(\mathrm{i}\,\mathbf{g}\cdot\mathbf{r})$ are linearly independent from each other for all $\mathbf{r}$, to fulfill (17.47) it is necessary that the term in braces vanishes. To proceed it is useful to associate[8] a single index b with the triad (n_1, n_2, n_3) defining $\mathbf{g}$, and another single index b' to the triad (n_1', n_2', n_3') defining $\mathbf{g}'$. Remembering that at the beginning of this section the assumption of an unlimited lattice has been made, $\mathbf{k}$ must be considered a continuous variable, so that s and E become functions of $\mathbf{k}$ proper. In conclusion, (17.47) transforms into

$$\sum_{b'} s_{b-b'}(\mathbf{k})\,V_{b'} = [E(\mathbf{k}) - T_b(\mathbf{k})]\;s_b(\mathbf{k})\,, \qquad b = 0, \pm1, \pm2, \ldots\,, \tag{17.48}$$

with $T_b(\mathbf{k})$ the result of the association $b \leftrightarrow (n_1, n_2, n_3)$ in $\hbar^2\,|\mathbf{g}+\mathbf{k}|^2/(2\,m)$. For each $\mathbf{k}$, (17.48) is a linear, homogeneous algebraic system in the infinite unknowns s_b and coefficients $V_{b'}$, $E(\mathbf{k}) - T_b(\mathbf{k})$, with $E(\mathbf{k})$ yet undetermined.

[8] The association $b \leftrightarrow (n_1, n_2, n_3)$ can be accomplished in a one-to-one fashion by, first, distributing the triads into groups having a common value of $d = |n_1| + |n_2| + |n_3|$, then ordering the groups in ascending order of d: for example, $d = 0$ corresponds to $(0, 0, 0)$, $d = 1$ to $[(0, 0, 1), (0, 1, 0), (1, 0, 0), (0, 0, -1), (0, -1, 0), (-1, 0, 0)]$, and so on. As each group is made by construction of finite number of triads, the latter are numbered within each group using a finite set of values of b; in order to have b ranging from $-\infty$ to $+\infty$, one associates a positive (negative) value of b with the triads in which the number of negative indices is even (odd).

The solution of (17.48) provides an infinite set of eigenvalues $E_1(\mathbf{k})$, $E_2(\mathbf{k}), \ldots,$ $E_i(\mathbf{k}), \ldots$ associated with the given $\mathbf{k}$. As the latter ranges over the first Brillouin zone, the functions $E_i(\mathbf{k})$ are thought of as branches[9] of a many-valued function. For each branch-index i, the function $E_i(\mathbf{k})$ is called *dispersion relation*, and the set of values spanned by $E_i(\mathbf{k})$ as $\mathbf{k}$ runs over the Brillouin zone is called *energy band* of index i. Being an eigenvalue of a periodic operator, $E_i(\mathbf{k})$ is periodic within the reciprocal, scaled lattice (compare with (17.34)); also, it can be shown that $E_i(\mathbf{k})$ is even with respect to $\mathbf{k}$ (Sect. 17.9.3):

$$E_i(\mathbf{k}+\mathbf{g}) = E_i(\mathbf{k})\,, \qquad E_i(-\mathbf{k}) = E_i(\mathbf{k})\,. \tag{17.49}$$

When a finite structure is considered, supplemented with the periodic boundary condition discussed in Sect. 17.5.3, vector $\mathbf{k}$ is discrete. On the other hand, for the derivation of (17.47) it is irrelevant whether $\mathbf{k}$ is continuous or discrete; hence, the analysis carried out in this section still holds for a discrete $\mathbf{k}$, provided the additional relations derived in Sect. 17.5.3, that describe the form of $\mathbf{k}$ and the corresponding densities, are accounted for. It must be remarked that the number N_s of cells along each direction in the direct lattice is typically very large.[10] As a consequence, a change by one unity of n_s in (17.37) is much smaller than the corresponding denominator N_s, so that for all practical purposes $E_i(\mathbf{k})$ is treated as a function of continuous variables when the derivatives with respect to the components of $\mathbf{k}$ enter the calculations.

The analysis carried out in this section clarifies the role of $\mathbf{k}$. In fact, for a given band index i, $\mathbf{k}$ labels the energy eigenvalue; for this reason, remembering the discussion about spin carried out in Sect. 15.5, $\mathbf{k}$ and the quantum number associated with spin determine the state of the particle. For fermions, the quantum number associated with spin has two possible values, so that two states with opposite spins are associated with each $\mathbf{k}$ vector. When the periodic boundary conditions are considered, the density of $\mathbf{k}$ vectors in the $\mathbf{k}$ space is given by (17.38), and the combined density of $\mathbf{k}$ vectors in the $\mathbf{r}, \mathbf{k}$ space is given by (17.39). As a consequence, the *density of states* in the $\mathbf{k}$ space and in the $\mathbf{r}, \mathbf{k}$ space are given, respectively, by

$$Q_k = 2\,\frac{N_c}{\tau_g} = \frac{\Omega}{4\,\pi^3}\,, \qquad Q = 2\,\frac{1}{\Omega}\,\frac{N_c}{\tau_g} = \frac{1}{4\,\pi^3}\,. \tag{17.50}$$

[9] Typically, a graphic representation of $E_i(\mathbf{k})$ is achieved by choosing a crystal direction and drawing the one-dimensional restriction of E_i along such a direction. Examples are given in Sect. 17.6.5.

[10] For instance, in a cube of material with an atomic density of 6.4×10^{27} m^{-3}, the number of atoms per unit length in each direction is 4,000 μm^{-1}.

17.6.1 Wave Packet in a Periodic Potential

From the solution of the Schrödinger equation worked out from (17.48) one reconstructs the periodic part of the Bloch function using the second relation in (17.43). Such a function inherits the band index i, so that the Bloch functions[11] read $w_{i\mathbf{k}} = \zeta_{i\mathbf{k}} \exp(\mathrm{i}\,\mathbf{k}\cdot\mathbf{r})$; they form a complete set so that, letting $\omega_{i\mathbf{k}} = \omega_i(\mathbf{k}) = E_i(\mathbf{k})/\hbar$, the expansion of the wave function $\psi(\mathbf{r},t)$ in terms of the eigenfunctions of the periodic Hamiltonian operator (17.40) reads

$$\psi(\mathbf{r},t) = \sum_{i\mathbf{k}} c_{i\mathbf{k}}\, w_{i\mathbf{k}}(\mathbf{r})\, \exp(-\mathrm{i}\,\omega_{i\mathbf{k}}\, t) = \sum_{i\mathbf{k}} c_{i\mathbf{k}}\, \zeta_{i\mathbf{k}}(\mathbf{r})\, \exp[\mathrm{i}\,(\mathbf{k}\cdot\mathbf{r} - \omega_{i\mathbf{k}}\, t)]\,, \tag{17.51}$$

with $c_{i\mathbf{k}} = \langle w_{i\mathbf{k}}|\psi\rangle_{t=0}$ a set of constants. The expansion (17.51) bears a strong similarity with that of the wave packet describing a free particle (compare with (9.1)), the only difference between the two expansions being the periodic factor $\zeta_{i\mathbf{k}}(\mathbf{r})$. The similarity suggests that an approximate expression of the wave packet is achieved by following the same reasoning as in Sect. 9.6, namely, by expanding $\omega_i(\mathbf{k})$ around the average value $\mathbf{k}_0$ of the wave vector and retaining the first-order term of the expansion[12]:

$$\omega_i(\mathbf{k}) \simeq \omega_i(\mathbf{k}_0) + \mathbf{u}_i(\mathbf{k}_0)\cdot(\mathbf{k}-\mathbf{k}_0)\,, \qquad \mathbf{u}_i(\mathbf{k}_0) = (\mathrm{grad}_{\mathbf{k}}\omega_i)_0\,, \tag{17.52}$$

with $\mathbf{u}_i$ the group velocity of the ith band. The approximation holds as long as $|R_i|\,t \ll 2\,\pi$, where R_i is the rest of the expansion. Letting $\omega_{i0} = \omega_i(\mathbf{k}_0)$, $\zeta_{i0} = \zeta_i(\mathbf{k}_0)$, and $\Phi_{i0} = \mathbf{k}_0\cdot\mathbf{r} - \omega_{i0}\,t$, the approximate expression of ψ reads

$$\psi(\mathbf{r},t) \simeq \sum_{i\mathbf{k}} c_{i\mathbf{k}}\, \zeta_{i0}\, \exp(\mathrm{i}\,\Phi_{i0})\, \exp\left[\mathrm{i}\,(\mathbf{r}-\mathbf{u}_i\,t)\cdot(\mathbf{k}-\mathbf{k}_0)\right]\,. \tag{17.53}$$

The envelope function is now defined as in (9.26), the difference being that a sum is used here instead of an integral:

$$A\,(\mathbf{r}-\mathbf{u}_i\,t\,;\mathbf{k}_0) = \sum_{\mathbf{k}} c_{i\mathbf{k}} \exp\left[\mathrm{i}\,\,(\mathbf{r}-\mathbf{u}_i\,t)\cdot(\mathbf{k}-\mathbf{k}_0)\right]\,, \tag{17.54}$$

so that

$$\psi(\mathbf{r},t) \simeq \sum_{i} \zeta_{i0}\, \exp(\mathrm{i}\,\Phi_{i0})\, A\,(\mathbf{r}-\mathbf{u}_i\,t\,;\mathbf{k}_0)\,. \tag{17.55}$$

[11] Here the periodic part of $w_{i\mathbf{k}}$ is indicated with $\zeta_{i\mathbf{k}}$ to avoid confusion with the group velocity.

[12] In the case of a free particle (Sect. 9.6) the approximation neglects only the second order because $\omega(\mathbf{k})$ has a quadratic dependence on the components of $\mathbf{k}$. Here, instead, the expansion has in general all terms due to the more complicate form of $\omega_i(\mathbf{k})$, so the neglected rest R_i contains infinite terms.

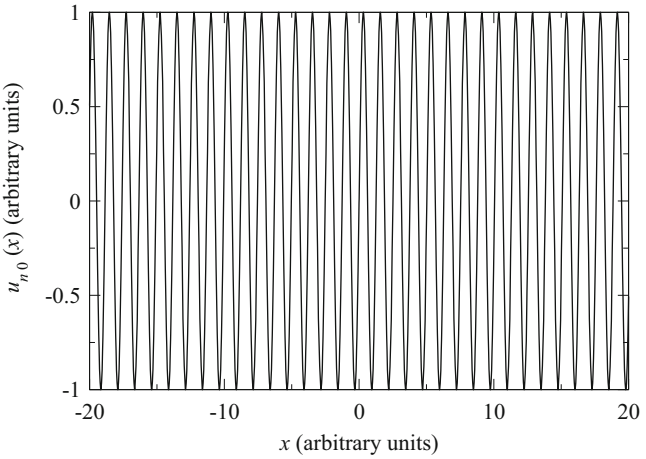


Fig. 17.7 A one-dimensional example of the periodic factor ζ_{n0} of (17.56)

As a further approximation one considers the fact that the number of $\mathbf{k}$ vectors of the first Brillouin zone is in general very large, because it equals the number N_c of direct-lattice cells. It follows that although the set of eigenfunctions belonging to a single branch is not complete, such a set is still able to provide an acceptable description of the wave packet. In this case one fixes the branch index, say, $i = n$, so that $\psi(\mathbf{r}, t) \simeq \zeta_{n0} \exp(\mathrm{i}\, \Phi_{n0})\, A(\mathbf{r} - \mathbf{u}_n\, t\,; \mathbf{k}_0)$. It follows

$$|\psi(\mathbf{r}, t)|^2 \simeq |\zeta_{n0}(\mathbf{r})|^2\, |A\,(\mathbf{r} - \mathbf{u}_n\, t\,; \mathbf{k}_0)\,|^2\,. \tag{17.56}$$

In (17.56), the periodic factor $|u_{n0}(\mathbf{r})|^2$ is a rapidly oscillating term whose period is of the order of the lattice constant; such a term does not provide any information about the particle's localization. This information, in fact, is carried by the envelope function, like in the case of a free particle outlined in Sect. 9.6. A one-dimensional example about how (17.56) is built up is given in Figs. 17.7, 17.8, 17.9, and 17.10.

17.6.2 Parabolic-Band Approximation

The dispersion relation $E_n(\mathbf{k})$ obtained from the solution of the Schrödinger equation (17.40) in a periodic lattice fulfills the periodicity condition given by the first expression in (17.49). As a consequence, $E_n(\mathbf{k})$ has necessarily a number of extremum points within the first Brillouin zone or at the boundary of it.[13]

[13] As mentioned in Sect. 17.6, $E_n(\mathbf{k})$ is considered as a function of a continuous vector variable $\mathbf{k}$ even when the periodic boundary conditions are assumed.

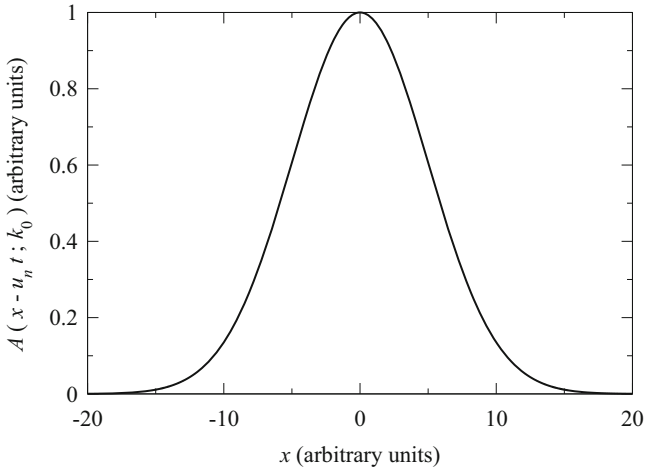


Fig. 17.8 A one-dimensional example of the envelope function $A(\mathbf{r} - \mathbf{u}_n\, t\,; \mathbf{k}_0)$ of (17.56)

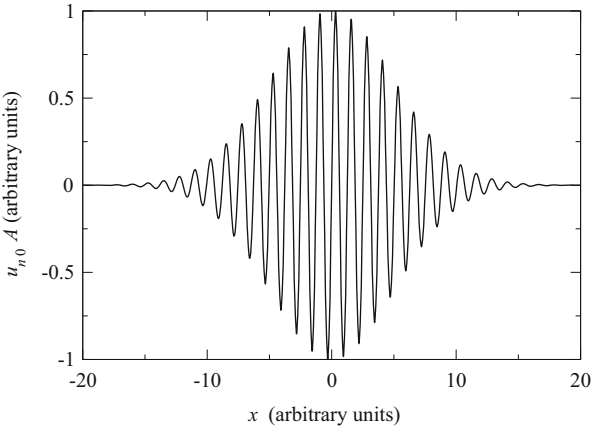


Fig. 17.9 Product of the two functions shown in Figs. 17.7 and 17.8

In view of further developments of the theory it is useful to investigate the form of $E_n(\mathbf{k})$ in the vicinity of such extremum points. To this purpose, the absolute minima are considered first; for a given branch index n assume that the number of such minima is M_C, and let $\mathbf{k}_a$ be the value of $\mathbf{k}$ at the ath minimum, $a = 1, \ldots, M_C$, with $E_C = E_n(\mathbf{k}_a)$. At $\mathbf{k} = \mathbf{k}_a$ the Hessian matrix of $E_n(\mathbf{k})$ is symmetric and positive definite, hence it can be diagonalized with positive real eigenvalues. In other terms, the reference in the $\mathbf{k}$ space can be chosen in such a way as to make the Hessian

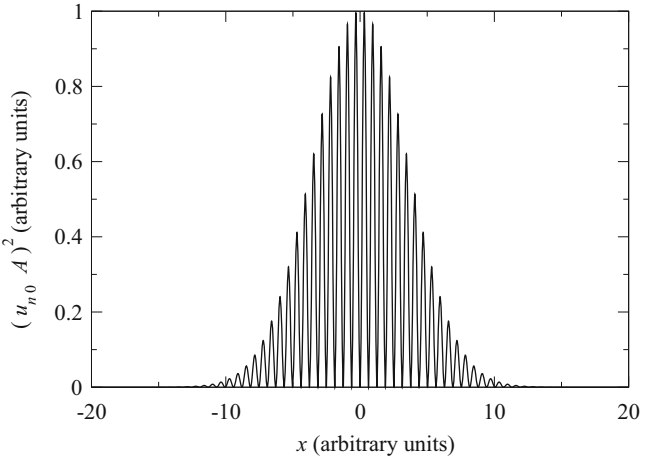


Fig. 17.10 The function of Fig. 17.9 squared

matrix of $E_n(\mathbf{k})$ diagonal; using such a reference, the second-order expansion of $E_n(\mathbf{k})$ around $\mathbf{k}_a$ reads

$$E_n(\mathbf{k}) \simeq E_C + \frac{1}{2}\sum_{i=1}^{3}\left(\frac{\partial^2 E_n}{\partial k_i^2}\right)_a (k_i - k_{ia})^2 \geq E_a\,, \quad a = 1, \ldots, M_C\,. \tag{17.57}$$

The first derivatives are missing from (17.57) because the expansion is carried out at an extremum. The coefficients $(\partial^2 E_n/\partial k_i^2)_a$ are in general different from each other, so that the sum in (17.57) may be thought of as a positive-definite quadratic form generated by a 3×3 diagonal matrix. Noting the units of the matrix entries one defines the *inverse, effective-mass tensor* of the ath minimum as

$$(\hat{m}_a)^{-1} = \begin{bmatrix} 1/m_{1\,a} & 0 & 0 \\ 0 & 1/m_{2\,a} & 0 \\ 0 & 0 & 1/m_{3\,a} \end{bmatrix}, \qquad \frac{1}{m_{ia}} = \frac{1}{\hbar^2}\left(\frac{\partial^2 E_n}{\partial k_i^2}\right)_a > 0 \tag{17.58}$$

so that, using the notation $E_{ne}(\mathbf{k}) = E_n(\mathbf{k}) - E_C \geq 0$, $\delta k_i = k_i - k_{ia}$, (17.57) takes the form

$$E_{ne} = \sum_{i=1}^{3}\frac{\hbar^2}{2\,m_{ia}}\,(k_i - k_{ia})^2 = \frac{1}{2}\,\hbar\,\delta\mathbf{k}\cdot(\hat{m}_a)^{-1}\,\hbar\,\delta\mathbf{k}\,. \tag{17.59}$$

Being the inverse, effective-mass tensor diagonal, the *effective-mass tensor* $\hat{m}_a$ is given by

$$\hat{m}_a = \begin{bmatrix} m_{1\,a} & 0 & 0 \\ 0 & m_{2\,a} & 0 \\ 0 & 0 & m_{3\,a} \end{bmatrix} . \tag{17.60}$$

The approximation shown above, that consists in replacing the dispersion relation with its second-order expansion near an extremum, is called *parabolic-band approximation*. The group velocity to be associated with a $\mathbf{k}$ vector in the vicinity of a minimum is found by applying to (17.59) the second relation in (17.52):

$$\mathbf{u}_n(\mathbf{k}) = \frac{1}{\hbar} \operatorname{grad}_{\mathbf{k}} E_n(\mathbf{k}) = \frac{1}{\hbar} \operatorname{grad}_{\mathbf{k}} E_{ne}(\mathbf{k}) = (\hat{m}_a)^{-1} \hbar\, \delta\mathbf{k} . \tag{17.61}$$

The calculation in the vicinity of an absolute maximum is similar.[14] Assume that the number of maxima in the nth branch of the dispersion relation is M_V, and let $\mathbf{k}_a$ be the value of $\mathbf{k}$ at the ath maximum, $a = 1, \ldots, M_V$, with $E_V = E_n(\mathbf{k}_a)$. The second-order expansion of $E_n(\mathbf{k})$ around $\mathbf{k}_a$ reads

$$E_n(\mathbf{k}) \simeq E_V + \frac{1}{2} \sum_{i=1}^{3} \left(\frac{\partial^2 E_n}{\partial k_i^2} \right)_a (k_i - k_{ia})^2 \leq E_a , \quad a = 1, \ldots, M_V , \tag{17.62}$$

where the Hessian matrix is negative definite. For this reason, the inverse, effective-mass tensor at the ath maximum is defined as

$$(\hat{m}_a)^{-1} = \begin{bmatrix} 1/m_{1\,a} & 0 & 0 \\ 0 & 1/m_{2\,a} & 0 \\ 0 & 0 & 1/m_{3\,a} \end{bmatrix} , \qquad \frac{1}{m_{ia}} = -\frac{1}{\hbar^2} \left(\frac{\partial^2 E_n}{\partial k_i^2} \right)_a > 0 \tag{17.63}$$

so that, using the notation $E_{nh}(\mathbf{k}) = E_V - E_n(\mathbf{k}) \geq 0$, (17.57) takes the form

$$E_{nh} = \sum_{i=1}^{3} \frac{\hbar^2}{2\, m_{ia}} (k_i - k_{ia})^2 = \frac{1}{2} \hbar\, \delta\mathbf{k} \cdot (\hat{m}_a)^{-1} \hbar\, \delta\mathbf{k} . \tag{17.64}$$

The group velocity to be associated with a $\mathbf{k}$ vector in the vicinity of a maximum reads

$$\mathbf{u}_n(\mathbf{k}) = \frac{1}{\hbar} \operatorname{grad}_{\mathbf{k}} E_n(\mathbf{k}) = -\frac{1}{\hbar} \operatorname{grad}_{\mathbf{k}} E_{nh}(\mathbf{k}) = -(\hat{m}_a)^{-1} \hbar\, \delta\mathbf{k} . \tag{17.65}$$

[14] The parabolic-band approximation is not necessarily limited to absolute minima or absolute maxima; here it is worked out with reference to such cases because they are the most interesting ones. However, it applies as well to relative minima and relative maxima. The different values of the inverse, effective-mass tensor's entries between an absolute and a relative minimum of a branch in GaAs give rise to interesting physical effects (Sect. 17.6.6).

It is important to note that the expressions of the parabolic-band approximation given in this section have been worked out in a specific reference of the **k** space, namely, the reference where the Hessian matrix is diagonal. In so doing, the reference of the direct space **r** has been fixed as well, because the two references are reciprocal to each other (Sect. 17.3). In other terms, when diagonal expressions like (17.59) or (17.64) are used in a calculation of dynamic properties, the reference in the **r** space cannot be chosen arbitrarily.

17.6.3 Density of States in the Parabolic-Band Approximation

Calculations related to many-particle systems often involve the density of states in energy (e.g., Sects. 15.8.1,15.8.2). The calculation of this quantity is relatively simple for the dispersion relation $E_n(\mathbf{k})$ in the parabolic-band approximation, because the dispersion relation is quadratic in the components of **k** and, in turn, the density of the **k** vectors is constant. In fact, it is found from (B.34) that in the three-dimensional case a quadratic expression $A = u^2 + v^2 + w^2$ yields a density of states equal to $A^{1/2}$. It is then sufficient to reduce (17.59) and (17.64) to the quadratic expression above. Taking (17.59) by way of example, one disposes of the multiplicative factors and the shift in the origin by applying the *Herring-Vogt transformation*

$$\eta_i = \frac{\hbar}{\sqrt{2\, m_{ia}}} (k_i - k_{ia}) \,, \tag{17.66}$$

to find

$$E_{ne} = \sum_{i=1}^{3} \eta_i^2 = \eta^2 \,, \qquad \mathrm{d}k_i = \frac{\sqrt{2\, m_{ia}}}{\hbar} \,\mathrm{d}\eta_i \,, \tag{17.67}$$

with $\eta > 0$, and

$$\mathrm{d}^3 k = \mathrm{d}k_1 \,\mathrm{d}k_2 \,\mathrm{d}k_3 = 2 \frac{\sqrt{2}}{\hbar^3} \, m_{ea}^{3/2} \,\mathrm{d}^3 \eta \,, \qquad m_{ea} = (m_{1a}\, m_{2a}\, m_{3a})^{1/3} \,. \tag{17.68}$$

Turning to spherical coordinates $\eta_1 = \eta \sin\vartheta \cos\varphi$, $\eta_2 = \eta \sin\vartheta \sin\varphi$, $\eta_3 = \eta \cos\vartheta$ yields $\mathrm{d}^3\eta = \eta^2 \mathrm{d}\eta \sin\vartheta \,\mathrm{d}\vartheta \,\mathrm{d}\varphi$ (Sect. B.1), where the product $\eta^2 \,\mathrm{d}\eta$ is found by combining the relations $2\,\eta \,\mathrm{d}\eta = \mathrm{d}E_{ne}$ and $\mathrm{d}\eta = \mathrm{d}E_{ne}/(2\,\sqrt{E_{ea}})$:

$$\eta^2 \,\mathrm{d}\eta = \frac{1}{2} \sqrt{E_{ne}} \,\mathrm{d}E_{ne} \,. \tag{17.69}$$

The number of states belonging to the elementary volume $\mathrm{d}^3 k$ is $\mathrm{d}N = Q_k \,\mathrm{d}^3 k$, with Q_k the density of states in the **k** space given by the first expression in (17.50). If the

elementary volume is centered on a $\mathbf{k}$ vector close to the ath minimum of $E_n(\mathbf{k})$, so that the parabolic-band approximation holds, one has

$$\mathrm{d}N_a = Q_k \,\mathrm{d}^3k = \frac{\Omega}{4\,\pi^3}\, 2\, \frac{\sqrt{2}}{\hbar^3}\, m_{ea}^{3/2}\, \frac{1}{2}\, \sqrt{E_{ne}}\, \mathrm{d}E_{ne}\, \sin\vartheta\, \mathrm{d}\vartheta\, \mathrm{d}\varphi. \tag{17.70}$$

The integral over the angles yields 4π, whence

$$\int_{\varphi=0}^{2\pi} \int_{\vartheta=0}^{\pi} \mathrm{d}N_a = \Omega\, \frac{\sqrt{2}}{\pi^2\,\hbar^3}\, m_{ea}^{3/2}\, \sqrt{E_{ne}}\, \mathrm{d}E_{ea} = g_a(E_{ne})\, \mathrm{d}E_{ne} \tag{17.71}$$

where, by construction, $g_a(E_{ne})$ is the density of states in energy around the ath minimum. Adding g_a over the M_C absolute minima yields the total density of states in energy,

$$g(E_{ne}) = \sum_{a=1}^{M_C} g_a = \Omega\, \frac{\sqrt{2}}{\pi^2\,\hbar^3}\, M_C\, m_e^{3/2}\, \sqrt{E_{ne}}\,, \qquad m_e = \left(\frac{1}{M_C} \sum_{a=1}^{M_C} m_{ea}^{3/2} \right)^{2/3}, \tag{17.72}$$

with m_e the *average effective mass* of the absolute minima. The *combined density of states* in the energy and $\mathbf{r}$ spaces then reads

$$\gamma(E_{ne}) = \frac{g(E_{ne})}{\Omega} = \frac{\sqrt{2}}{\pi^2\,\hbar^3}\, M_C\, m_e^{3/2}\, \sqrt{E_{ne}}\,. \tag{17.73}$$

Note that apart from the different symbol used to indicate the volume in the $\mathbf{r}$ space, and the replacement of m with $M_C\, m_e^{3/2}$, the relations (17.72) and (17.73) are identical, respectively, to (15.64) and (15.65), expressing the density of states and combined density of states in a box.

The calculation of the density of states in energy in the vicinity of the M_V absolute maxima is identical to the above and yields

$$g(E_{nh}) = \sum_{a=1}^{M_V} g_a = \Omega\, \frac{\sqrt{2}}{\pi^2\hbar^3}\, M_V\, m_h^{3/2}\, \sqrt{E_{nh}}\,, \quad m_h = \left(\frac{1}{M_V} \sum_{a=1}^{M_V} m_{ha}^{3/2} \right)^{2/3}, \tag{17.74}$$

where m_h is the average effective mass of the absolute maxima. In turn it is $m_{ha} = (m_{1a}\, m_{2a}\, m_{3a})^{1/3}$, with m_{ia} given by the second relation in (17.63).

17.6.4 Crystals of Si, Ge, and GaAs

Among semiconductors, silicon (Si), germanium (Ge), and gallium arsenide (GaAs) are very important for the electronic industry. This section is devoted to illustrating some properties of their crystal and energy-band structures. The crystals of silicon and germanium are of the face-centered, cubic type; the reciprocal lattices have the

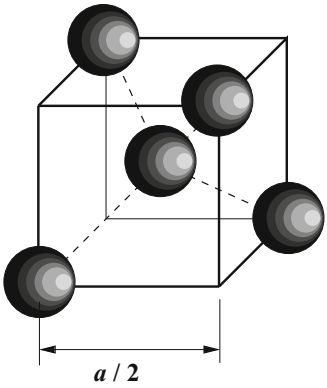


Fig. 17.11 Tetrahedral organization of the elementary, body-centered cubic block of silicon or germanium. The side of the cube is one half the lattice constant a

body-centered, cubic structure. The *lattice constants*, that is, the physical sizes of the unit cell, are the same in the [100], [010], and [001] directions (Sect. 17.9.1). Their values at $T = 300$ K are given in Table 17.1 [99]. The crystals of the materials under consideration are formed by elementary blocks like that shown in Fig. 17.11. Each atom has four electrons in the external shell, so that it can form four chemical bonds with other identical atoms; the latter place themselves symmetrically in space, to build up the tetrahedral structure shown in the figure. In this structure, which is of the body-centered cubic type with a side equal to one half the lattice constant a, the chemical bonds of the central atom are saturated, whereas the atoms placed at the vertices still have three bonds unsaturated; as a consequence, they may behave as centers of new tetrahedral structures identical to the original one.

An example of this is given in Fig. 17.12: the top half of the figure shows two replicas of the elementary block of Fig. 17.11 sharing an atom belonging to an upper corner, while the bottom half of the figure shows again two replicas, this time sharing an atom belonging to a lower corner. The atoms drawn in white do not belong to any of the elementary blocks considered in the figure, and serve the purpose of demonstrating how the rest of the crystal is connected to them. Note that the structure in the bottom half of Fig. 17.12 is identical to that of the top half, the difference being simply that one structure is rotated by 90 degrees with respect to the other on a vertical axis. The construction is now completed by bringing the two halves together, as shown in Fig. 17.13; this provides the diamond structure mentioned in Sect. 17.2. Such a structure is of the face-centered, cubic type, with an additional atom at the center of each tetrahedral block.

The minimum distance d among the atoms (*interatomic distance*) is the distance from the atom in the center of the tetrahedral elementary block to any of the atoms at its vertices; its relation with the lattice constant is easily found to be

$$d = \frac{\sqrt{3}}{4}\, a\,. \tag{17.75}$$

The description is similar for gallium arsenide [99], and for a number of semiconductors of the III-V type, whose crystal constants are listed in Table 17.2.

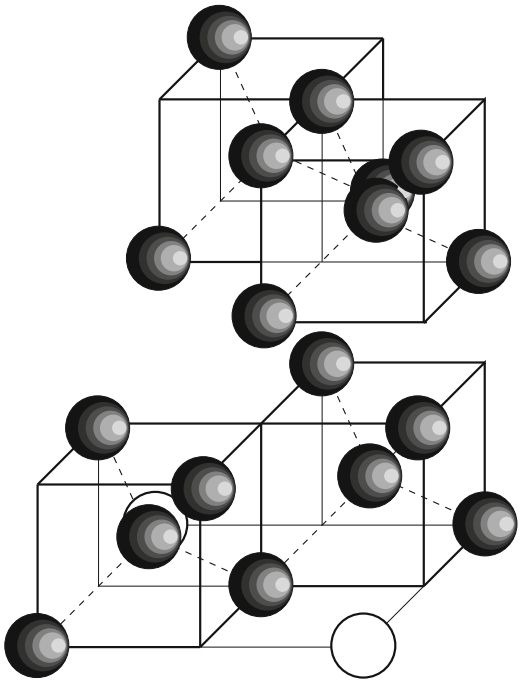


Fig. 17.12 Diamond structure. The top and bottom halves are shown separately

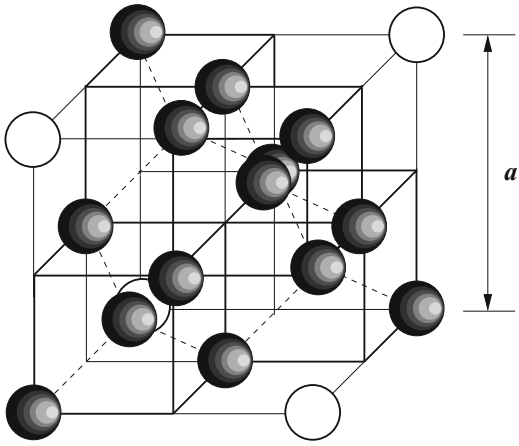


Fig. 17.13 Diamond structure obtained by joining together the top and bottom halves shown separately in Fig. 17.12

17.6.5 Band Structure of Si, Ge, and GaAs

Coming now to the description of the band structure, it is important to focus on the bands that are able to contribute to the electric conduction of the material. In fact, considering the aim of manufacturing electronic devices out of these materials, the bands that do not contribute to the electric current are not relevant. It is intuitive

that a band with no electrons, that is, whose states have a zero probability of being occupied, is not able to provide any conduction; it is less intuitive (in fact, this is demonstrated in Sect. 19.3) that a band whose states are fully occupied does not provide any conduction either. It follows that the only bands of interest are those where only a fraction of the electronic states are occupied. Although a discussion involving the electric current must necessarily refer to a nonequilibrium condition, it is easier to base the reasoning upon the equilibrium condition at some temperature T; in fact, in this case the occupation probability of the electronic states is given by the Fermi-Dirac statistics (15.49). As a consequence, the number of electrons belonging to a band whose energy values range, say, from E_a to E_b, is given by the first relation in (15.48) with $\alpha + \beta\, E = (E - E_F)/(k_B\, T)$, namely,

$$N_{ab} = \int_{E_a}^{E_b} \frac{g(E)}{\exp[(E - E_F)/(k_B\, T)] + 1}\, \mathrm{d}E\,. \tag{17.76}$$

As mentioned in Sect. 17.6 each branch of the dispersion relation $E_i(\mathbf{k})$ spans an energy band. In many cases the bands are disjoint from each other, namely, energy intervals exist that contain no eigenvalue of the Schrödinger equation (17.40). Such intervals are called *forbidden bands* or *gaps*. In the equilibrium condition the energy of an electron can never belong to a gap, no matter what the value of the occupation probability is, because the density of states is zero there. Also, at a given temperature the position of the Fermi level E_F is either within a band (edges included), or within a gap; the latter case, typical of semiconductors, is illustrated with the aid of Fig. 17.14, where it is assumed (using the units of $(E - E_F)/(k_B\, T)$) that a gap exists between the energies E_V, E_C such that $(E_V - E_F)/(k_B\, T) = -0.15$ and $(E_C - E_F)/(k_B\, T) = +0.15$. In other terms, E_V is the upper energy edge of a band, and E_C the lower energy edge of the next band. These assumptions also imply that the Fermi level coincides with the gap's midpoint. As will become apparent below, the two bands that are separated by the Fermi level are especially important; for this reason they are given specific names: the band whose absolute maximum is E_V is called *valence band*, that whose absolute minimum is E_C is called *conduction band*.

As shown in the figure, the case is considered (typical of Si, Ge, and GaAs) where the gap's width contains the main variation of the Fermi-Dirac statistics[15]; as a consequence, the occupation probability becomes vanishingly small as the difference $E - E_C$ becomes larger, so that only the energy states near the absolute minimum E_C have a nonvanishing probability of being occupied. Thank to this reasoning, to the purpose of calculating (17.76) one can replace the density of states $g(E)$ with the simplified expression (17.72) deduced from the parabolic-band approximation; such an expression, $g(E) \propto \sqrt{E - E_C}$, is shown in Fig. 17.14 in arbitrary units, along with the $g(E)\, P(E)$ product (thick line), that represents

[15]The extension of the energy region where the main variation of the Fermi-Dirac statistics occurs is estimated in Prob. 15.1.

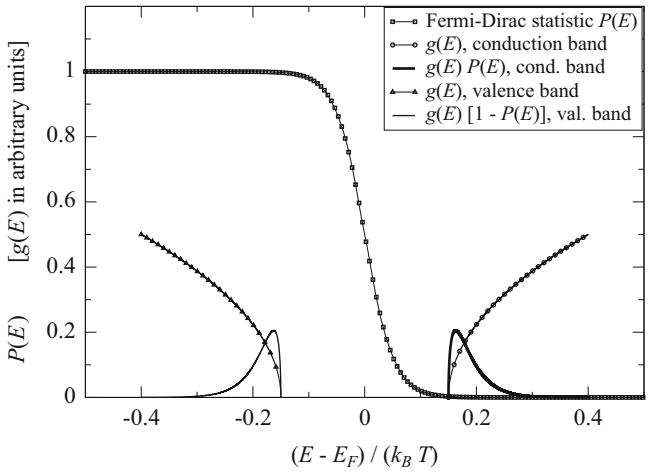


Fig. 17.14 Calculation of the particles' population in the conduction and valence bands of a semiconductor. To make them more visible, the products $g(E)\,P(E)$ and $g(E)\,[1 - P(E)]$ have been amplified with respect to $g(E)$ alone. The gap's extension is arbitrary and does not refer to any specific material

the integrand of (17.76) with reference to the conduction band. To make it more visible, the $g(E)\,P(E)$ product is drawn in a scale amplified by 10^3 with respect to that of $g(E)$ alone. The number of electrons belonging to the conduction band is proportional to the area subtended by the $g(E)\,P(E)$ curve.

Coming now to the valence band, the probability $1 - P(E)$ that a state at energy E is empty becomes vanishingly small as the difference $E_V - E$ becomes larger, so that only the energy states near the absolute maximum E_V have a nonvanishing probability of being empty. Empty states are also called *holes*. The number of holes is given by an integral similar to (17.76), where $P(E)$ is replaced with $1 - P(E)$. This calculation is made easier by observing that, due to the form of the Fermi-Dirac statistics, it is

$$1 - \frac{1}{\exp[(E - E_F)/(k_B\,T)] + 1} = \frac{1}{\exp[(E_F - E)/(k_B\,T)] + 1}\,. \tag{17.77}$$

Also in this case one can use for the density of states the parabolic-band approximation; such an expression, $g(E) \propto \sqrt{E_V - E}$, is shown in Fig. 17.14 in arbitrary units, along with the $g(E)\,[1 - P(E)]$ product (thin line). As before, the product is drawn in a scale amplified by 10^3 with respect to that of $g(E)$ alone. The number of holes belonging to the valence band is proportional to the area subtended by the $g(E)\,[1 - P(E)]$ curve.

Thanks to the spatial uniformity of the crystal, the concentration of the electrons in the conduction band is obtained by dividing their number by the crystal volume Ω or, equivalently, by replacing the density of states in energy g with the combined

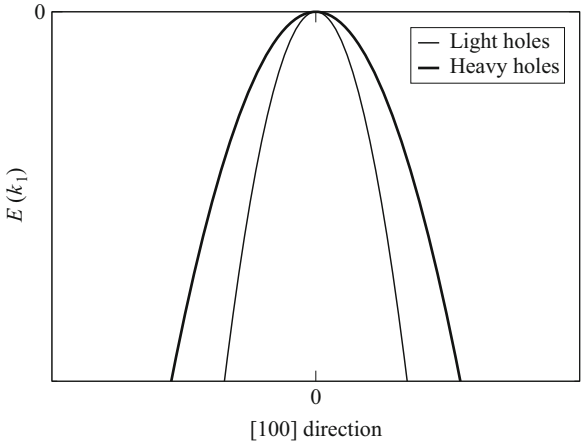


Fig. 17.15 Schematic view of the two branches of the valence band of Si, Ge, or GaAs in the [100] direction

density of states in energy and volume γ given by (17.73). A similar reasoning holds for holes. The explicit expressions of the concentrations are given in Sect. 18.3. Here it is important to remark that the perfect symmetry of the curves $g(E)\,P(E)$ and $g(E)\,[1-P(E)]$ in Fig. 17.14 is due to the simplifying assumptions that E_F coincides with the gap's midpoint and that $M_V\,m_h^{3/2} = M_C\,m_e^{3/2}$ (compare with (17.73) and (17.74)). Neither hypothesis is actually true, so that in real cases the two curves are not symmetric; however, as shown in Sect. 18.3, the areas subtended by them are nevertheless equal to each other.

Valence Band

The valence band of Si, Ge, and GaAs is made of two branches of $E(\mathbf{k})$, having the same absolute maximum E_V at $\mathbf{k} = 0$ (so that $M_V = 2$), but different curvatures. They are shown in Fig. 17.15, where the horizontal axis coincides with the [100] direction in the $\mathbf{k}$ space, corresponding to the scalar variable k_1. As a consequence, the origin of the horizontal axis coincides with the Γ point (Sect. 17.4); the axis intersects the boundary of the first Brillouin zone at the X points (not shown in the figure). The origin of the vertical axis coincides with E_V. The two branches are not spherically symmetric; in fact, letting $E_V = 0$, the dependence of each of them on the spherical coordinates k, ϑ, φ has the form [70, Sect. 8.7]

$$-\frac{\alpha}{2}\,k^2\,[1 \pm j(\vartheta,\varphi)]\,, \qquad \alpha > 0\,, \tag{17.78}$$

called *warped*. In the parabolic-band approximation the angular part j is neglected with respect to unity, and the two branches become spherically symmetric around $\mathbf{k} = 0$; still with $E_V = 0$, the dependence on k_1 of each branch has the form $E = -\alpha\,k_1^2/2$, where the constant α is smaller in the upper branch (indicated by the thick line in Fig. 17.15), and larger in the lower one. As a consequence, the corresponding component of the effective-mass tensor (17.63), that reads in this case $m_1 = \hbar^2/\alpha$,

Table 17.3 Normalized effective masses of the valence band of Si, Ge, and GasAs

Material	$m_{hh}(T_a)/m_0$	$m_{hl}(T_a)/m_0$
Si	0.5	0.16
Ge	0.3	0.04
GaAs	0.5	0.12

is larger in the upper branch and smaller in the lower one. For this reason, the holes associated with the energy states of the upper branch are called *heavy holes*, those associated with the lower branch are called *light holes*.

The analysis is identical in the other two directions [010] and [001] so that, for each branch of the valence band, the diagonal entries of the effective-mass tensor are equal to each other. Such tensors then read $m_{hh}\,\mathscr{I}$, $m_{hl}\,\mathscr{I}$, with $\mathscr{I}$ the identity tensor; the first index of the scalar effective mass stands for "hole," while the second one stands for "heavy" or "light." The second-order expansions around E_V take respectively the form[16]

$$E_V - E_h(\mathbf{k}) = \frac{\hbar^2}{2\,m_{hh}} \sum_{i=1}^{3} k_i^2\,, \qquad E_V - E_l(\mathbf{k}) = \frac{\hbar^2}{2\,m_{hl}} \sum_{i=1}^{3} k_i^2\,. \tag{17.79}$$

Due to (17.79), the constant-energy surfaces $E_V - E_h(\mathbf{k}) = \text{const}$ and $E_V - E_l(\mathbf{k}) = \text{const}$ are spheres, whose radius squared is $2\,m_{hh}\,(E_V - E_h)/\hbar^2$ and $2\,m_{hl}\,(E_V - E_l)/\hbar^2$, respectively. The values of m_{hh} and m_{hl} at room temperature T_a are listed in Table 17.3 [128, Sect. 2-3]; they are normalized to the rest mass of the free electron, $m_0 \simeq 9.11 \times 10^{-31}$ kg. The effective masses depend in general on temperature because a change in the latter modifies the lattice constants: as a consequence, the characteristic vectors of the reciprocal lattice change as well, thus deforming the dispersion relation $E(\mathbf{k})$; on the other hand, the variation of the effective masses with temperature is weak, so it is typically neglected.[17]

Conduction Band

The conduction band of Si, Ge, and GaAs has only one branch. However, the absolute minima (also called *valleys*) are placed differently. In GaAs there is only one absolute minimum at $\mathbf{k} = 0$, with spherical symmetry. In the parabolic-band approximation, the constant-energy surface is given by

[16] From now on the band index n introduced in (17.56) is omitted from the notation.

[17] In contrast, the temperature dependence of the energy gap, due to the deformation of the dispersion relation, cannot be neglected because of its strong effect on the carrier concentration (Sect. 18.3).

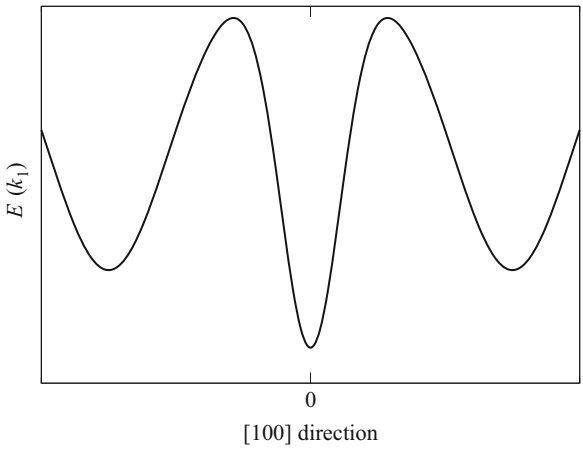


Fig. 17.16 Schematic view of the conduction band of GaAs in the [100] direction

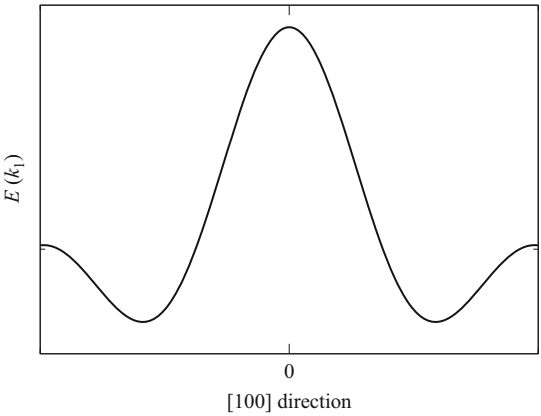


Fig. 17.17 Schematic view of the conduction band of Si in the [100] direction

$$E(\mathbf{k}) - E_C = \frac{\hbar^2}{2\,m_e} \sum_{i=1}^{3} k_i^2 \,, \tag{17.80}$$

namely, a sphere whose radius squared is $2\,m_e\,(E - E_C)/\hbar^2$. The band exhibits also secondary minima at $E_C + \Delta E$, with $\Delta E \simeq 0.36$ eV (Fig. 17.16).

The conduction band of Si has six absolute minima ($M_C = 6$), grouped into three pairs. The latter belong to the [100], [010], and [001] directions, respectively, and are symmetrically placed with respect to the Γ point $\mathbf{k} = 0$. Their coordinates are

$$[100] : (\pm k_m, 0, 0)\,, \qquad [010] : (0, \pm k_m, 0)\,, \qquad [001] : (0, 0, \pm k_m)\,, \tag{17.81}$$

where $k_m \simeq 0.85\,k_B > 0$, with k_B the distance between the Γ and X points (Fig. 17.17). In the parabolic-band approximation, the surfaces at constant energy

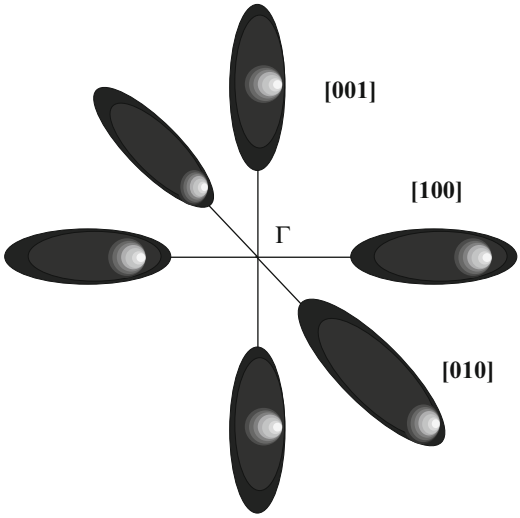


Fig. 17.18 Constant-energy surfaces of the conduction band of silicon

of the conduction band of Si are ellipsoids of revolution about the [100], [010], or [001] axes (Fig. 17.18). Their expressions are

$$[100]: \qquad E_{e1} = E(\mathbf{k}) - E_C = \frac{\hbar^2}{2}\left[\frac{(k_1 - k_m)^2}{m_l} + \frac{k_2^2}{m_t} + \frac{k_3^2}{m_t}\right], \tag{17.82}$$

$$[010]: \qquad E_{e2} = E(\mathbf{k}) - E_C = \frac{\hbar^2}{2}\left[\frac{k_1^2}{m_t} + \frac{(k_2 - k_m)^2}{m_l} + \frac{k_3^2}{m_t}\right], \tag{17.83}$$

$$[001]: \qquad E_{e3} = E(\mathbf{k}) - E_C = \frac{\hbar^2}{2}\left[\frac{k_1^2}{m_t} + \frac{k_2^2}{m_t} + \frac{(k_3 - k_m)^2}{m_l}\right]. \tag{17.84}$$

Similarly, E_{e4}, E_{e5}, E_{e6} are derived from E_{e1}, E_{e2}, E_{e3}, respectively, by letting $k_m \leftarrow -k_m$. The effective masses m_l and m_t are called *longitudinal* and *transverse* mass, respectively.

The conduction band of Ge has eight absolute minima, grouped into four pairs. The pairs belong to the four {111} directions and are placed at the boundary of the first Brillouin zone (Fig. 17.19); thus, only four absolute minima must be accounted for ($M_C = 4$). In the parabolic-band approximation, the surfaces at constant energy of the conduction band of Ge are ellipsoids of revolution about the corresponding axis; like in silicon, for each ellipsoid the longitudinal mass corresponds to the direction of the axis itself, while the transverse masses correspond to the directions normal to it. The values of m_l and m_t at room temperature T_a, normalized to the rest mass of the free electron, are listed in Table 17.4 [128, Sect. 2-3].

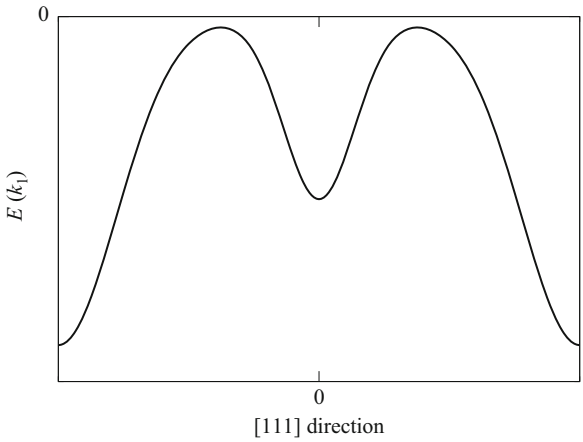


Fig. 17.19 Schematic view of the conduction band of Ge in the [111] direction

Table 17.4 Normalized effective masses of the conduction band of Si, Ge, and GasAs

Material	$m_l(T_a)/m_0$	$m_t(T_a)/m_0$
Si	0.97	0.19
Ge	1.6	0.082
GaAs[a]	0.068	0.068

[a]The effective masses of GaAs are equal to each other due to the band's isotropy

17.6.6 Further Comments About the Band Structure

As better detailed in Sects. 19.5.2 and 19.5.3, among the coefficients of the equations describing the transport phenomena in a semiconductor are the electron and hole *mobilities*, that enter the relation between current density and electric field in a uniform material. For the conduction band of the semiconductors considered here, and in the parabolic-band approximation, the electron mobility μ_n turns out to be proportional to $1/m_n = (2/m_t + 1/m_l)/3$, that is, a weighed average of the entries of the inverse, effective-mass tensor.[18] Table 17.4 shows that GaAs has the largest value of $1/m_n$; thus, it is expected to have the largest mobility, which is indeed the case. As far as holes are concerned, the effective masses of heavy holes of Si, Ge, and GaAs are similar to each other; also the effective masses of light holes have the same order of magnitude. Besides, considering that the valence band has two branches of $E(\mathbf{k})$, the effective masses do not combine in the simple way as for the conduction band.

The secondary minima of GaAs, placed at an energy $E_C + \Delta E$ with $\Delta E \simeq 0.36$ eV (Fig. 17.16), have a larger effective mass than the absolute minimum; due to

[18]If the magnitudes of m_t and m_l are significantly different, the smaller effective mass dictates the magnitude of m_n.

this, the mobility of the electrons in the upper valleys is smaller than that of the electrons populating the absolute minimum. As ΔE is relatively small, the population of the secondary minima is not negligible; in a nonequilibrium condition, the scattering events tend to increase the electron population of the upper valleys at the expense of that of the absolute minimum, with a ratio between the upper and lower population that depends on the applied electric field. This gives rise to a negative differential resistivity in the current-to-voltage curve of the material, i.e., an operating region exists where the current density decreases as the electric field increases. The phenomenon is called *Ridley-Watkins-Hilsum mechanism* [128, Sect. 14-3].

In semiconductors, the absorption of energy from an electromagnetic field may induce the transition of an electron from a state belonging to the valence band to a state belonging to the conduction band. Such a transition increases by one the number of electrons in the conduction band and, at the same time, increases by one the number of holes in the valence band; for this reason it is called *generation of an electron-hole pair*. The opposite phenomenon may also occur, namely, a release of electromagnetic energy due to the transition of an electron from a state belonging to the conduction band to a state belonging to the valence band. Such a transition decreases by one the number of electrons in the conduction band and, at the same time, decreases by one the number of holes in the valence band (*recombination of an electron-hole pair*). It is worth pointing out that generation and recombination events may also occur with an energy absorption from, or release to, an external agent different from the electromagnetic field (e.g., the agent could be a vibrational mode of the lattice); for this reason, the phenomena considered here better specified as *generations-recombinations of the radiative type*. In GaAs, the minimum of the conduction band and the maxima of the two branches of the valence band correspond to the same value of $\mathbf{k}$; semiconductors fulfilling this condition are called *direct-gap semiconductors*. Instead, Si and Ge are *indirect-gap semiconductors*, because the maxima of the valence band correspond to $\mathbf{k} = 0$, whereas the minima of the conduction band correspond to $\mathbf{k} \neq 0$. Direct- and indirect-gap semiconductors behave differently as far as generation-recombination of the radiative type are concerned; in fact, the probability of such events is much higher in direct-gap semiconductors. This explains why some classes of solid-state optical devices like, e.g., lasers, are manufactured using direct-gap semiconductors.[19]

[19]The reasoning seems to contradict the fact that large-area, solid-state optical sensors used in cameras and video cameras, based on the CCD or CMOS architecture, are made of silicon. In fact, the complex structure of these several-megapixel sensors and related signal-management circuitry can be realized only with the much more advanced technology of silicon. The relative ease of fabricating complex structures largely compensates for the poorer optical properties of the material.

17.6.7 Subbands

The calculations of the density of states carried out so far have been based on the assumption that all components of the **k** vector can be treated as continuous variables. In particular, the adoption of the parabolic-band approximation in the case of a periodic lattice (Sect. 17.6) leads to expressions for the density of states $g(E)$ and combined density of states $\gamma(E)$ that are formally identical to those obtained for a particle in a three-dimensional box (Sect. 15.9.2). However, in some situations it happens that not all components of **k** may be treated as continuous. To describe this case it is convenient to use the example of the box first; that of the periodic lattice is worked out later, in the frame of the parabolic-band approximation.

To proceed, consider like in Sect.. 15.9.2 a three-dimensional box whose sides have lengths d_1, d_2, d_3, so that the eigenvalues of the Schrödinger equation are $E_{n_1 n_2 n_3} = \hbar^2 k^2/(2\,m)$, where k^2 is the square of

$$\mathbf{k} = n_1 \frac{\pi}{d_1}\,\mathbf{i}_1 + n_2 \frac{\pi}{d_2}\,\mathbf{i}_2 + n_3 \frac{\pi}{d_3}\,\mathbf{i}_3\,, \qquad n_i = 1, 2, \ldots \tag{17.85}$$

The distance between two consecutive projections of **k** along the ith side is $\Delta k_i = \pi/d_i$, and the volume associated with each **k** is $\Delta k_1\,\Delta k_2\,\Delta k_3 = \pi^3/V$, with $V = d_1\,d_2\,d_3$ the volume of the box in the **r** space. The density of the **k** vectors in the **k** space is $Q_k = V/\pi^3$.

Two-Dimensional Layer

Now, in contrast to what was implicitly assumed in Sect. 15.9.2, let one side of the box be much different from the others, for instance, $d_2 \sim d_1$, $d_3 \ll d_1, d_2$. It follows that $\Delta k_3 \gg \Delta k_1, \Delta k_2$. If the magnitudes involved are such that k_1, k_2 may still be considered continuous variables, while k_3 cannot, one must calculate the density of states by treating k_1, k_2 differently from k_3. Considering $k_1 = n_1\,\pi/d_1, k_2 = n_2\,\pi/d_2$ as continuous, fix E and s in the relations

$$\frac{2\,m}{\hbar^2}\,E = k_1^2 + k_2^2 + n_3^2\,\frac{\pi^2}{d_3^2}\,, \qquad n_3 = s < \frac{d_3}{\pi}\,\frac{\sqrt{2\,m\,E}}{\hbar}\,. \tag{17.86}$$

For each integer $s = 1, 2, \ldots$ the two relations (17.86) determine in the k_1, k_2 plane a circumference of radius $c_s = \sqrt{c_s^2}$, with

$$c_s^2 = k_1^2 + k_2^2\,, \qquad c_s^2 = \frac{2\,m\,E}{\hbar^2} - s^2\,\frac{\pi^2}{d_3^2}\,. \tag{17.87}$$

It is $\min_s c_s > 0$ because $s_{\max} < d_3 \sqrt{2mE}/(\pi\hbar)$, and $\max_s c_s = c_1 > 0$. For a fixed s the states are distributed over the circumference of radius c_s: such a set of states is also called *subband*.

The density of states in energy of the subband thus defined is calculated following the same reasoning as in Sect. 15.9.3: in fact, one observes that the density of **k** vectors in the two-dimensional space k_1, k_2 is $d_1 d_2/\pi^2$, namely, the inverse of the area $\pi^2/(d_1 d_2)$ associated with each **k** belonging to the given circumference. Then, the total number of **k** vectors in a circle of radius c_s is

$$N_{ks} = \frac{d_1\, d_2}{\pi^2}\, \pi\, c_s^2 = \frac{d_1\, d_2}{\pi} \left(\frac{2\,m\,E}{\hbar^2} - s^2 \frac{\pi^2}{d_3^2} \right) . \tag{17.88}$$

Remembering that indices n_1, n_2 are positive, it is necessary to consider only the first quadrant; as a consequence, N_{ks} must be divided by 4. Further, it is necessary to multiply it by 2 to account for electron spin. In conclusion, the density of states of the two-dimensional subbands is

$$g_{2\mathrm{D}}(E) = \frac{\mathrm{d}(2\,N_{ks}/4)}{\mathrm{d}E} = \frac{d_1\, d_2\, m}{\pi\,\hbar^2} = \text{const}\,, \tag{17.89}$$

to be compared with (15.67). Note that (17.89) is independent of index s. This result is useful, e.g., for treating the problem of a two-dimensional charge layer in the channel of a semiconductor device.

Wire

Now, assume that $d_2 \sim d_3$, and $d_2, d_3 \ll d_1$. It follows that $\Delta k_2, \Delta k_3 \gg \Delta k_1$. If the magnitudes involved are such that k_1 may still be considered a continuous variable, while k_2, k_3 cannot, one must calculate the density of states by treating k_1 differently from k_2, k_3. Considering $k_1 = n_1 \pi/d_1$ as continuous, fix E, r, s in the relations

$$\frac{2\,m}{\hbar^2}\,E = k_1^2 + n_2^2 \frac{\pi^2}{d_2^2} + n_3^2 \frac{\pi^2}{d_3^2}\,, \qquad n_2 = r\,, \quad n_3 = s\,, \tag{17.90}$$

with $r^2/d_2^2 + s^2/d_3^2 \le 2\,m\,E/(\pi^2\,\hbar^2)$. For each pair of integers $r, s = 1, 2, \ldots$, (17.90) determine in the **k** space two points given by the relation

$$\kappa_{rs}^2 = \frac{2\,m\,E}{\hbar^2} - \frac{r^2\,\pi^2}{d_2^2} - \frac{s^2\,\pi^2}{d_3^2}\,, \qquad \kappa_{rs} = \sqrt{\kappa_{rs}^2}\,. \tag{17.91}$$

It is $\min_{rs} \kappa_{rs} > 0$ and $\max_{rs} \kappa_{rs} = \kappa_{11}$. For a fixed pair r, s the states are placed at the ends of the segment $[-\kappa_{rs}, +\kappa_{rs}]$ parallel to k_1. The density of states in energy of such a segment is calculated following the same reasoning as in Sect. 15.9.3: in fact, one observes that the density of **k** vectors in the one-dimensional space k_1 is

d_1/π, namely, the inverse of the length π/d_1 associated with each $\mathbf{k}$ belonging to the segment. Then, the total number of $\mathbf{k}$ vectors in the segment of length $2\,\kappa_{rs}$ is

$$N_{krs} = \frac{d_1}{\pi}\,2\kappa_{rs} = \frac{2d_1}{\pi}\left(\frac{2mE}{\hbar^2} - r^2\frac{\pi^2}{d_2^2} - s^2\frac{\pi^2}{d_3^2}\right)^{1/2} . \tag{17.92}$$

Remembering that index n_1 is nonnegative, it is necessary to consider only the positive half of the segment. As a consequence, N_{krs} must be divided by 2. Further, it is necessary to multiply it by 2 to account for electron spin. The density of states of the one-dimensional case then reads

$$g_{1\mathrm{D}}(E) = \frac{\mathrm{d}(2\,N_{krs}/2)}{\mathrm{d}E} = \frac{2d_1\,m}{\pi\,\hbar^2\,\kappa_{rs}} , \tag{17.93}$$

to be compared with (15.67). Note that, in contrast with the two-dimensional case (17.89), here the result depends on both indices r, s.

A device with $d_2, d_3 \ll d_1$ is also called *wire*. When the device size is such that the transport of a particle in it must be studied by means of Quantum Mechanics, it is also called *quantum wire*. The $E(\kappa_{rs})$ relation may be recast as

$$\frac{\hbar^2}{2\,m}\,\kappa_{rs}^2 = E - E_{rs} , \qquad E_{rs} = \frac{\pi^2\,\hbar^2}{2\,m}\left(\frac{r^2}{d_2^2} + \frac{s^2}{d_3^2}\right) . \tag{17.94}$$

As E_{rs} is an increasing function of the indices, its minimum is attained for $r = s = 1$ and represents the ground state in the variables k_2, k_3. It is interesting to note that if the total energy E is prescribed, e.g., by injecting the particle from an external source, such that $E_{11} < E < \min(E_{12}, E_{21})$, then the particle's wave function has the form

$$\psi = \sqrt{\frac{8}{V}}\,\sin(\kappa_{11}\,x_1)\,\sin(\pi\,x_2/d_2)\,\sin(\pi\,x_3/d_3)\,\exp(-\mathrm{i}\,E\,t/\hbar) \tag{17.95}$$

(compare with (15.60)). Remembering the expression (15.64) of the density of states in a box where $d_1 \sim d_2 \sim d_3$, the results obtained so far are summarized as:

$$g_{3\mathrm{D}}(E) = \frac{V\sqrt{2\,m^3\,E}}{\pi^2\,\hbar^3} , \quad g_{2\mathrm{D}}(E) = \frac{d_1\,d_2\,m}{\pi\,\hbar^2} , \quad g_{1\mathrm{D}}(E) = \frac{d_1\sqrt{2\,m}}{\pi\,\hbar\sqrt{E - E_{rs}}} . \tag{17.96}$$

17.6.8 Subbands in a Periodic Lattice

The calculations leading to (17.96) consider the case of a box within which the potential energy is zero (as a consequence, the total energy E is purely kinetic), and prescribe a vanishing wave function at the boundaries. If a periodic lattice is present, with the provisions indicated in Sect. 17.5.3 one can apply the periodic boundary conditions. In this case, the spacing between the components of $\mathbf{k}$ in each direction doubles ($n_i \pi/d_i \leftarrow 2 n_i \pi/d_i$), but the number of components doubles as well ($n_i = 1, 2, \ldots \leftarrow n_i = 0, \pm 1, \pm 2, \ldots$), so the density of states remains the same. In a semiconductor the calculation leading to the density of states is made more complicate by the presence of the lattice. However, the analysis may be brought to a simple generalization of that carried out in a box by means of the following simplifications:

- It is assumed that a band structure exists even if the size of the device is small in one or two spatial directions. In fact, it can be shown that the presence of a number of atomic planes of the order of ten is sufficient to form a band structure.
- The analysis is limited to the case of parabolic bands.

The case of the conduction band of silicon is considered by way of example, with the k_1, k_2, k_3 axes placed along the [100], [010], [001] directions. The parabolic-band approximation yields for the kinetic energies E_{e1}, E_{e2}, $E_{e3} \geq 0$ the expressions given in (17.82), (17.83), and (17.84); the other three kinetic energies E_{e4}, E_{e5}, E_{e6} are derived from E_{e1}, E_{e2}, E_{e3}, respectively, by letting $k_m \leftarrow -k_m$. Apart from the constant E_C, the energies $E_{e1}, E_{e2}, \ldots$ are simplified forms of the eigenvalues of the Schrödinger equation (17.40). Conversely, to the purpose of determining the corresponding eigenfunctions, one may view $E_{e1}, E_{e2}, \ldots$ as the exact eigenvalues of simplified forms of the original Hamiltonian operator (17.40), that hold near the band's minima; such simplified forms are expected to be of the purely kinetic type. They are found by replacing k_i with $-\mathrm{i}\,\mathrm{d}/\mathrm{d}x_i$ in (17.82), (17.83), and (17.84), thus yielding

$$[100]: \quad \mathscr{H}_{e1} = \frac{\hbar^2}{2}\left[+\frac{1}{m_l}\left(\mathrm{i}\frac{\partial}{\partial x_1} + k_m\right)^2 - \frac{1}{m_t}\frac{\partial^2}{\partial x_2^2} - \frac{1}{m_t}\frac{\partial^2}{\partial x_3^2}\right], \tag{17.97}$$

$$[010]: \quad \mathscr{H}_{e2} = \frac{\hbar^2}{2}\left[-\frac{1}{m_t}\frac{\partial^2}{\partial x_1^2} + \frac{1}{m_l}\left(\mathrm{i}\frac{\partial}{\partial x_2} + k_m\right)^2 - \frac{1}{m_t}\frac{\partial^2}{\partial x_3^2}\right], \tag{17.98}$$

$$[001]: \quad \mathscr{H}_{e3} = \frac{\hbar^2}{2}\left[-\frac{1}{m_t}\frac{\partial^2}{\partial x_1^2} - \frac{1}{m_t}\frac{\partial^2}{\partial x_2^2} + \frac{1}{m_l}\left(\mathrm{i}\frac{\partial}{\partial x_3} + k_m\right)^2\right], \tag{17.99}$$

with m_l and m_t the longitudinal and transverse masses. Considering $\mathscr{H}_{e1}$ first, the solution of the time-independent Schrödinger equation generated by it,

$$\mathscr{H}_{e1} w_1(\mathbf{r}, n_1, n_2, n_3) = E(n_1, n_2, n_3)\, w_1(\mathbf{r}, n_1, n_2, n_3)\,, \tag{17.100}$$

is found by separation, specifically, by letting $E = E_\alpha(n_1) + E_\beta(n_2) + E_\gamma(n_3)$, $w_1 = \exp(\mathrm{i}\, k_m\, x_1)\, \alpha(x_1, n_1)\, \beta(x_2, n_2)\, \gamma(x_3, n_3)$. One finds for the $\mathbf{k}$ vector the expression

$$\mathbf{k} = n_1 \frac{\pi}{d_1}\, \mathbf{i}_1 + n_2 \frac{\pi}{d_2}\, \mathbf{i}_2 + n_3 \frac{\pi}{d_3}\, \mathbf{i}_3\,, \tag{17.101}$$

with $n_i = 1, 2, \ldots$, while the eigenfunctions and eigenvalues read

$$w_1 = \sqrt{\frac{8}{V}}\, \exp(\mathrm{i}\, k_m\, x_1)\, \sin\left(\frac{n_1\, \pi}{d_1}\, x_1\right) \sin\left(\frac{n_2\, \pi}{d_2}\, x_2\right) \sin\left(\frac{n_3\, \pi}{d_3}\, x_3\right), \tag{17.102}$$

$$E = \frac{\hbar^2}{2\, m_l}\, n_1^2\, \frac{\pi^2}{d_1^2} + \frac{\hbar^2}{2\, m_t}\, n_2^2\, \frac{\pi^2}{d_2^2} + \frac{\hbar^2}{2\, m_t}\, n_3^2\, \frac{\pi^2}{d_3^2}\,. \tag{17.103}$$

The eigenvalues and eigenfunctions of $\mathscr{H}_{e2}$, $\mathscr{H}_{e3}$ are found by a cyclic permutation of the indices, $1 \leftarrow 2 \leftarrow 3 \leftarrow 1$, while those of $\mathscr{H}_{e4}$, $\mathscr{H}_{e5}$, $\mathscr{H}_{e6}$ are derived from those of $\mathscr{H}_{e1}$, $\mathscr{H}_{e2}$, $\mathscr{H}_{e3}$, respectively, by letting $k_m \leftarrow -k_m$.

One notes that the $\mathbf{k}$ vectors and the eigenfunctions are not influenced by the effective masses, whereas the eigenvalues are. As a consequence, the density of states is affected as well. In the case where $d_1 \sim d_2 \sim d_3$ the density of states associated with the minimum of index 1 is found by the same procedure as that leading to the first relation in (17.96); the result is

$$g_{3\mathrm{D}}^{(1)}(E) = \frac{d_1\, d_2\, d_3\, \sqrt{2\, m_l\, m_t^2}}{\pi^2\, \hbar^3}\, \sqrt{E}\,. \tag{17.104}$$

Such a density of states is not affected by interchanging the effective masses; thus, the total density of states is found by adding over the densities of states of the M_C minima of the conduction band:

$$g_{3\mathrm{D}}(E) = M_C\, \frac{d_1\, d_2\, d_3\, \sqrt{2\, m_l\, m_t^2}}{\pi^2\, \hbar^3}\, \sqrt{E}\,. \tag{17.105}$$

Like in the case of a box, the distance between two consecutive projections of $\mathbf{k}$ along the ith side is $\Delta k_i = \pi/d_i$, and the volume associated with each $\mathbf{k}$ is $\Delta k_1\, \Delta k_2\, \Delta k_3 = \pi^3/V$, with $V = d_1 d_2 d_3$. The density of the $\mathbf{k}$ vectors in the $\mathbf{k}$ space is $Q_k = V/\pi^3$.

Consider now the case of a two-dimensional layer, namely, $d_2 \sim d_1$, while $d_3 \ll d_1, d_2$. Let $k_1 = n_1\pi/d_1$, $k_2 = n_2\pi/d_2$, and fix $n_3 = 1$ whence, for the minima of indices 1 and 4,

$$E = \frac{\hbar^2}{2\, m_l}\, k_1^2 + \frac{\hbar^2}{2\, m_t}\, k_2^2 + \frac{\hbar^2}{2\, m_t}\, \frac{\pi^2}{d_3^2}\,, \qquad E \geq \frac{\hbar^2}{2\, m_t}\, \frac{\pi^2}{d_3^2}\,. \tag{17.106}$$

A calculation similar to that carried out in a box provides, for the minima of indices 1 and 4, an expression similar to that of the second relation in (17.96):

$$g_{2D}^{(1)} = g_{2D}^{(4)} = \frac{d_1\, d_2 \sqrt{m_l\, m_t}}{\pi\, \hbar^2} \,. \tag{17.107}$$

For the other pairs of minima one finds

$$g_{2D}^{(2)} = g_{2D}^{(5)} = \frac{d_1\, d_2\, \sqrt{m_l\, m_t}}{\pi\, \hbar^2} \,, \qquad g_{2D}^{(3)} = g_{2D}^{(6)} = \frac{d_1\, d_2\, m_t}{\pi\, \hbar^2} \,. \tag{17.108}$$

In conclusion, for a two-dimensional layer with $d_3 \ll d_1, d_2$ and $n_3 = 1$, within the parabolic-band approximation, the density of states for the minima of indices 1, 2, 4, and 5 is the same constant for all energies $E \geq \hbar^2\, \pi^2/(2\, m_t\, d_3^2)$. The total density of states for these minima is

$$g_{2D}^{(1,2,4,5)} = 4\, \frac{d_1\, d_2\, \sqrt{m_l\, m_t}}{\pi\, \hbar^2} \,. \qquad E \geq E_t = \frac{\hbar^2\, \pi^2}{2\, m_t\, d_3^2} \,, \qquad n_3 = 1 \,, \tag{17.109}$$

while $g_{2D,n_3=1}^{(1,2,4,5)} = 0$ for $E < E_t$. Similarly, still with $n_3 = 1$, the density of states for the minima of indices 3 and 6 is another constant for all energies $E \geq \hbar^2\, \pi^2/(2\, m_l\, d_3^2)$. The total density of states for these minima is

$$g_{2D}^{(3,6)} = 2\, \frac{d_1\, d_2\, m_t}{\pi\, \hbar^2} \,, \qquad E \geq E_l = \frac{\hbar^2\, \pi^2}{2\, m_l\, d_3^2} \,, \qquad n_3 = 1 \,, \tag{17.110}$$

while $g_{2D,n_3=1}^{(3,6)} = 0$ for $E < E_l$. Now, let $n_3 = 2$; it is easily found that the value of $g_{2D,n_3=2}^{(1,2,4,5)}$ is the same as above, however, it holds for $E \geq 4\, E_t$. It adds up to the value found for $n_3 = 1$, giving rise to a stair-like form of $g_{2D}^{(1,2,4,5)}$ as a function of energy. The same is obtained for $g_{2D,n_3=2}^{(3,6)}$ when $E \geq 4\, E_l$, and so on. An example of such a density of states is sketched in Fig. 17.20, where the ratio $g_{2D}^{(1,2,4,5)}/g_{2D,n_3=1}^{(1,2,4,5)}$ is shown as a function of E/E_t. The total density of states is found by adding up the two stair-like functions. From Table 17.4 one finds that in silicon at room temperature it is $m_l \simeq 0.97\, m_0$, $m_t \simeq 0.19\, m_0$, whence $E_t \simeq 5.1\, E_l$ and $g_{2D,n_3=1}^{(1,2,4,5)} \simeq 4.47\, g_{2D,n_3=1}^{(3,6)}$.

As shown by Fig. 17.20, the derivative of the density of states with respect to energy diverges at some points. Such divergences are called *Van Hove singularities* [3, Chap. 8].

Finally, consider the case of a wire, namely, $d_2 \sim d_3$, while $d_2, d_3 \ll d_1$. Let $k_1 = n_1 \pi/d_1$ and fix $n_2 = n_3 = 1$ whence, for the minima of indices 1 and 4,

$$E = \frac{\hbar^2}{2\, m_l}\, k_1^2 + \frac{\hbar^2}{2\, m_t}\, \frac{\pi^2}{d_2^2} + \frac{\hbar^2}{2\, m_t}\, \frac{\pi^2}{d_3^2} \,, \qquad E \geq \frac{\pi^2\, \hbar^2}{2\, m_t} \left(\frac{1}{d_2^2} + \frac{1}{d_3^2} \right) = E_{11}^{(1,4)} \,. \tag{17.111}$$

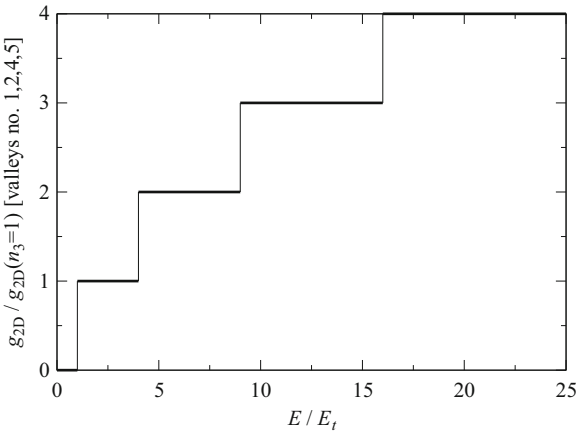


Fig. 17.20 Normalized, two-dimensional density of states (17.109) for the $1, 2, 4, 5$ valleys of silicon, as a function of E/E_t, in the parabolic-band approximation

A calculation similar to that carried out in a box provides, for the minima of indices 1 and 4, an expression similar to that of the third relation in (17.96):

$$g_{1D}^{(1)} = g_{1D}^{(4)} = \frac{2\, d_1\, m_l}{\pi\, \hbar^2\, \kappa_{11}^{(1,4)}} = \frac{d_1\, \sqrt{2\, m_l}}{\pi\, \hbar\, \sqrt{E - E_{11}^{(1,4)}}}\,, \tag{17.112}$$

while $g_{1D}^{(1,4)} = 0$ if $E < E_{11}^{(1,4)}$. For the other minima one finds

$$g_{1D}^{(2,5)} = \frac{d_1\, \sqrt{2\, m_t}}{\pi\, \hbar\, \sqrt{E - E_{11}^{(2,5)}}}\,, \qquad E \geq E_{11}^{(2,5)} = \frac{\pi^2\, \hbar^2}{2} \left(\frac{1}{m_l\, d_2^2} + \frac{1}{m_t\, d_3^2} \right), \tag{17.113}$$

with $g_{1D}^{(2,5)} = 0$ if $E < E_{11}^{(2,5)}$, and

$$g_{1D}^{(3,6)} = \frac{d_1\, \sqrt{2\, m_t}}{\pi\, \hbar\, \sqrt{E - E_{11}^{(3,6)}}}\,, \qquad E \geq E_{11}^{(3,6)} = \frac{\pi^2\, \hbar^2}{2} \left(\frac{1}{m_t\, d_2^2} + \frac{1}{m_l\, d_3^2} \right), \tag{17.114}$$

with $g_{1D}^{(3,6)} = 0$ if $E < E_{11}^{(3,6)}$. In conclusion, for a wire with $d_2, d_3 \ll d_1$ and $n_2 = n_3 = 1$, within the parabolic-band approximation, the density of states of each pair of minima is the sum of expressions of the form (17.112,17.113,17.114); the latter are complicate because all possible pairs of indices r, s combine with the two lengths d_2, d_3, that in general are not commensurable with each other. A somewhat easier description is obtained by considering (17.112) alone and letting $d_2 = d_3$ in it; this yields $E_{21} = E_{12} = 2.5\, E_{11}$, $E_{22} = 4\, E_{11}$, $E_{31} = E_{13} = 5\, E_{11}$, and so on, and

$$g_{1D}^{(1,4)} = \frac{g_{1D}^{(1,4)}(E = 2\, E_{11})}{\sqrt{E/E_{11} - 1}}\,, \qquad E_{11} < E \leq E_{21}\,. \tag{17.115}$$

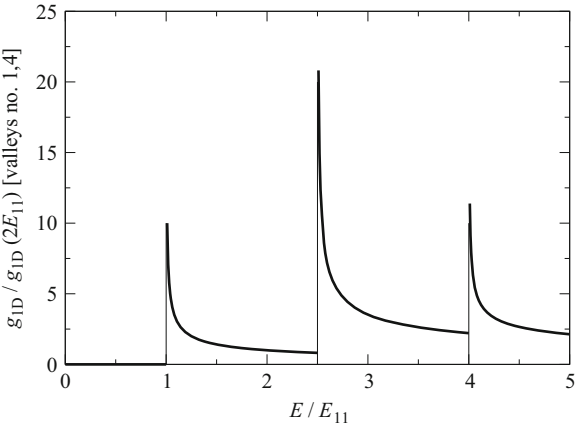


Fig. 17.21 Normalized, one-dimensional density of states for the 1, 4 valleys of silicon, as a function of E/E_{11}, in the parabolic-band approximation and with $d_2 = d_3$

In the next interval $E_{21} < E \leq E_{22}$, the density of states is the sum of (17.115) and $2/\sqrt{E/E_{11} - 2.5}$, where factor 2 accounts for the $(r = 2, s = 1)$, $(r = 1, s = 2)$ degeneracy; in the interval $E_{22} < E \leq E_{31}$ one adds the further summand $1/\sqrt{E/E_{11} - 4}$, and so on. The normalized density of states $g_{1D}(E)/g_{1D}(2E_{11})$ is shown in Fig. 17.21 as a function of E/E_{11}. Also in this case the Van Hove singularities are present; in addition, the density of states itself diverges at such points. However, such divergences are integrable; consider for instance an integral of the form

$$\int_{E_0}^{\infty} \frac{c}{\sqrt{E - E_0}} P(E)\, \mathrm{d}E\,, \tag{17.116}$$

with c a constant and $0 < P < 1$ a distribution function. Splitting the integration domain into two intervals $E_0 \leq E \leq E'$ and $E' \leq E < \infty$, with $E' > E_0$, one finds for the first integral, that contains the singularity,

$$\int_{E_0}^{E'} \frac{c}{\sqrt{E - E_0}} P(E)\, \mathrm{d}E \leq \int_{E_0}^{E'} \frac{c}{\sqrt{E - E_0}}\, \mathrm{d}E < \infty\,. \tag{17.117}$$

17.7 Calculation of Vibrational Spectra

The discussion carried out in Sect. 16.6 has led to the conclusion that in the case of solid matter the nuclei, being massive and tightly bound together, are expected to depart little from their equilibrium positions $\mathbf{R}_0$. The classical description of the nuclear motion is thus brought to the case already solved in Sects. 3.9 and 3.10: the vibrational state of the nuclei is described in terms of the normal coordinates b_σ,

whose conjugate momenta are $\dot{b}_\sigma$, and the total energy of the nuclei reads (compare with (3.50))

$$T_a + V_a = \sum_{\sigma=1}^{3N} H_\sigma + V_{a0}\,, \qquad H_\sigma = \frac{1}{2}\,\dot{b}_\sigma^2 + \frac{1}{2}\,\omega_\sigma^2\, b_\sigma^2\,, \tag{17.118}$$

where each H_σ corresponds to one degree of freedom and $\omega_\sigma > 0$ is the angular frequency of the corresponding mode. The system is completely separable in the normal coordinates, and each normal coordinate evolves in time as a linear harmonic oscillator. The calculation is based on Classical Mechanics; it is carried out in this chapter because it exploits the periodicity properties of the material and, in this respect, presents several analogies with the solution of the Schrödinger equation in a periodic lattice. To determine the vibrational frequencies ω_σ it is necessary to solve the eigenvalue equation (3.43), namely,

$$\mathbf{C}\,\mathbf{g}_\sigma = \omega_\sigma^2\,\mathbf{M}\,\mathbf{g}_\sigma\,, \qquad \sigma = 1, \ldots, 3N\,, \tag{17.119}$$

with $\mathbf{g}_\sigma$ the eigenvectors. The entries of $\mathbf{C}$, $\mathbf{M}$ are given by

$$c_{pn} = [\mathbf{C}]_{pn} = \left(\frac{\partial^2 V_a}{\partial h_p\,\partial h_n}\right)_0\,, \qquad [\mathbf{M}]_{pn} = \mu_n\,\delta_{pn}\,, \tag{17.120}$$

where V_a is the potential energy, h_p the displacement of the pth degree of freedom with respect to the equilibrium position, μ_p the mass associated with the pth degree of freedom, and δ_{pn} the Kronecker symbol (A.18).

The calculation is in principle the same for any system of particles; however, if the system has special properties, they reflect into the form of the eigenvalues and eigenvectors. A particularly important case is that of a periodic structure, such as a crystal. Considering this case, let the crystal be made of N_c elementary cells, with a basis made of N_b nuclei (the definition of basis is in Sect. 17.2). It follows that the total number of nuclei is $N = N_b\,N_c$, and the total number of degrees of freedom is $3\,N$. With respect to a given origin O (Fig. 17.22), the mth cell of the lattice is identified by the corresponding translation vector of the direct lattice, $\mathbf{l}_m$; the latter determines a local origin within the mth cell. In turn, the equilibrium position of the αth nucleus of the mth cell with respect to the local origin is identified by a vector $\mathbf{e}_\alpha$ of the direct lattice.

17.7.1 Labeling the Degrees of Freedom—Dynamic Matrix

To proceed, it is convenient to label the degrees of freedom in such a way as to distinguish the indices of the cells from those of the basis and of the coordinate

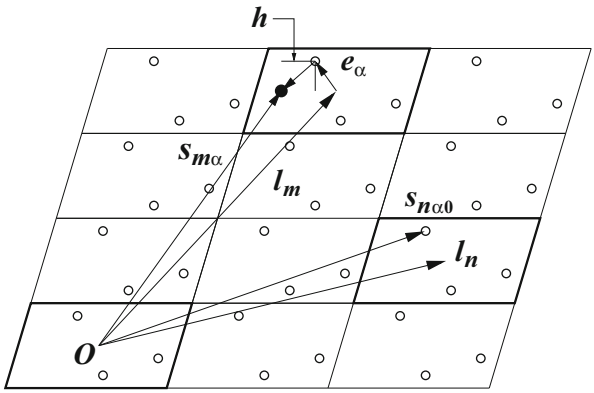


Fig. 17.22 Definition of the labels used to identify the degrees of freedom in a periodic lattice

axes. To this purpose, one observes that the component along the uth coordinate axis of the equilibrium position of the jth nucleus is

$$X_{ju0} = s_{q0}\,, \qquad q = u + 3\,(j-1)\,, \qquad j = \alpha + N_b\,(m-1)\,, \tag{17.121}$$

with $u = 1,2,3$; $\alpha = 1,\ldots,N_b$; $m = 1,\ldots,N_c$. The same applies to the displacements, which are more conveniently expressed in terms of three indices:

$$h_q \longleftarrow h_{m\alpha u}\,, \qquad h_r \longleftarrow h_{n\beta w}\,. \tag{17.122}$$

The entries of $\mathbf{C}$ are identified in the same manner:

$$c_{qr} = \left(\frac{\partial^2 V_a}{\partial h_q \partial h_r}\right)_0 \qquad \longleftarrow \qquad c^{n\beta w}_{m\alpha u} = \left(\frac{\partial^2 V_a}{\partial h_{m\alpha u}\,\partial h_{n\beta w}}\right)_0\,, \tag{17.123}$$

with

$$m,n = 1,\ldots,N_c\,, \qquad \alpha,\beta = 1,\ldots,N_b\,, \qquad u,w = 1,2,3\,. \tag{17.124}$$

The order of derivation is irrelevant, so that $c^{n\beta w}_{m\alpha u} = c^{m\alpha u}_{n\beta w}$. As the number of nuclei is finite, the crystal is not actually periodic; as indicated in Sect. 17.5.3, periodicity is recovered by imposing periodic boundary conditions to the quantities of interest (Sect. 17.5.3).[20] With this provision, the entries of $\mathbf{C}$ are invariant with respect to the lattice translations. The latter are related only to the cell indices m,n and are obtained by the replacements $\mathbf{l}_m \leftarrow \mathbf{l}_m + \mathbf{l}_\nu$, $\mathbf{l}_n \leftarrow \mathbf{l}_n + \mathbf{l}_\nu$, with ν any integer. In particular, taking $\mathbf{l}_\nu = -\mathbf{l}_n$ yields

[20] As mentioned in Sect. 17.5.3, the periodic boundary conditions are actually an approximation; however, the interatomic interactions typically give rise to short-range forces, hence the above reasoning holds for all the cells that are not too close to the boundaries.

$$c^{n\beta w}_{m\alpha u} = c^{\beta w}_{\alpha u}(\mathbf{l}_m, \mathbf{l}_n) = c^{\beta w}_{\alpha u}(\mathbf{l}_m - \mathbf{l}_n, 0) = c^{\beta w}_{\alpha u}(\mathbf{l}_m - \mathbf{l}_n)\,. \tag{17.125}$$

The above shows that the entries of $\mathbf{C}$ depend on the relative positions of the cells. Due to the invariance of $\mathbf{C}$ with respect to the lattice translations one sees that, given α, u and β, w, there are only N_c distinct entries of $\mathbf{C}$ out of N_c^2, namely, the distinct entries are those such that $m - n = 0, m - n = 1, \ldots, m - n = N_c - 1$. In fact, all remaining $N_c^2 - N_c$ entries are derived from the first N_c ones by suitable translations of the indices. In turn, using the new indices (17.124) the entries of $\mathbf{M}$ read

$$\mu_r\,\delta_{qr} \qquad \longleftarrow \qquad \mu_{n\beta w}\,\delta^{n\beta w}_{m\alpha u} = \mu_\beta\,\delta^{n\beta w}_{m\alpha u}, \tag{17.126}$$

where the last equality is due to the fact that the mass of a given nucleus of the cell does not depend on the cell position within the crystal nor on the coordinate axis. In the new indices the eigenvalue equation (17.119) becomes

$$\sum_{n\beta w} c^{n\beta w}_{m\alpha u}\,g_{n\beta w} = \omega^2\,\mu_\alpha\,g_{m\alpha u}\,, \tag{17.127}$$

where the indices' ranges are given in (17.124). The indices of the eigenvalue and eigenvector have been omitted for simplicity. Defining

$$d^{n\beta w}_{m\alpha u} = \frac{c^{n\beta w}_{m\alpha u}}{\sqrt{\mu_\alpha\,\mu_\beta}}\,, \qquad z_{m\alpha u} = \sqrt{\mu_\alpha}\,g_{m\alpha u}\,, \qquad z_{n\beta w} = \sqrt{\mu_\beta}\,g_{n\beta w}\,, \tag{17.128}$$

transforms (17.127) into

$$\sum_{n\beta w} d^{n\beta w}_{m\alpha u}\,z_{n\beta w} = \omega^2\,z_{m\alpha u}\,. \tag{17.129}$$

The latter form of the eigenvalue equation is more convenient because it eliminates the coefficient μ_α from the right-hand side. Matrix $\mathbf{D}$ of entries $d^{n\beta w}_{m\alpha u}$ is called *dynamic matrix* and, due to the properties of $\mathbf{C}$, is symmetric ($d^{n\beta w}_{m\alpha u} = d^{m\alpha u}_{n\beta w}$) and translationally invariant:

$$d^{n\beta w}_{m\alpha u} = d^{\beta w}_{\alpha u}(\mathbf{l}_m, \mathbf{l}_n) = d^{\beta w}_{\alpha u}(\mathbf{l}_m - \mathbf{l}_n, 0) = d^{\beta w}_{\alpha u}(\mathbf{l}_m - \mathbf{l}_n)\,. \tag{17.130}$$

17.7.2 Application of the Bloch Theorem

As a consequence of the translation invariance of **D**, Bloch's theorem (17.23) applies,[21] namely, for any eigenvector of indices $p\gamma e$, and letting $\mathbf{l}_0 = 0$, the following holds:

$$\mathbf{z}_{\gamma e}(\mathbf{l}_p) = \exp(\mathbf{c} \cdot \mathbf{l}_p)\, \mathbf{z}_{\gamma e}(0)\,. \tag{17.131}$$

In (17.131), **c** is any complex vector of the reciprocal lattice, and $p = 0, \ldots, N_c - 1$; $\gamma = 1, \ldots, N_b$; $e = 1, 2, 3$. The complex form of the eigenvectors is adopted for convenience; at the end of the calculation, a set of real eigenvectors is recovered from suitable combinations of the complex ones. Using the periodic boundary conditions, the expression of **c** is found to be

$$\mathbf{c} = \mathrm{i}\,\mathbf{q}\,, \qquad \mathbf{q} = \sum_{s=1}^{3} \frac{\nu_s}{N_s}\, 2\,\pi\, \mathbf{b}_s\,, \tag{17.132}$$

with N_1, N_2, N_3 the number of cells along the directions of the characteristic vectors of the direct lattice, $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ the characteristic vectors of the reciprocal lattice, and ν_1, ν_2, ν_3 integers, with $\nu_s = 0, 1, \ldots, N_s - 1$. The total number of distinct **q** vectors is thus $N_1 N_2 N_3 = N_c$. Comparing (17.132) with (17.37) shows that the structure of the **q** vector is the same as that of the **k** vector found in the solution of the Schrödinger equation (Sect. 17.5.3). Inserting (17.130) into (17.129) yields, for the line of indices $m\alpha u$ of the eigenvalue equation,

$$\sum_{n\beta w} A^{n\beta w}_{m\alpha u}\, z_{\beta w}(0) = \omega^2 z_{\alpha u}(0)\,, \tag{17.133}$$

with

$$A^{n\beta w}_{m\alpha u} = \frac{1}{\sqrt{\mu_\alpha \mu_\beta}}\, c^{\beta w}_{\alpha u}(\mathbf{l}_m - \mathbf{l}_n)\, \exp[\mathrm{i}\,\mathbf{q} \cdot (\mathbf{l}_n - \mathbf{l}_m)]\,. \tag{17.134}$$

As the eigenvalues ω^2 are real, the matrix made of the entries $A^{n\beta w}_{m\alpha u}$ must be Hermitean; in fact, this is easily found by observing that **D** is real and symmetric:

$$A^{m\alpha u}_{n\beta w} = d^{\alpha u}_{\beta w}(\mathbf{l}_n - \mathbf{l}_m)\, \exp[\mathrm{i}\,\mathbf{q} \cdot (\mathbf{l}_m - \mathbf{l}_n)] = \left(A^{n\beta w}_{m\alpha u}\right)^*\,. \tag{17.135}$$

Another property stems from the expression at the left-hand side of (17.133),

[21] The Bloch theorem was derived in Sect. 17.5.1 with reference to eigenfunctions that depend on a continuous parameter like, e.g., vector **c**; the theorem equally holds for a translationally invariant operator in the discrete case, like the dynamic matrix considered here.

$$\sum_{n\beta w} A^{n\beta w}_{m\alpha u}\, z_{\beta w}(0) = \sum_{\beta w} \left(\sum_{n} A^{n\beta w}_{m\alpha u} \right) z_{\beta w}(0)\,, \tag{17.136}$$

where $A^{n\beta w}_{m\alpha u}$ is translationally invariant because it depends on the cell indices only through the difference $\mathbf{l}_m - \mathbf{l}_n$. It follows that $\sum_n A^{n\beta w}_{m\alpha u}$ does not depend on m. This is easily verified by carrying out the sum first with, say, $m = 1$, then with $m = 2$, and observing that the terms of the second sum are the same as in the first one, displaced by one position. In summary, letting $\mathbf{A}$ be the $3N_b \times 3N_b$, Hermitean matrix of entries

$$A^{\beta w}_{\alpha u}(\mathbf{q}) = \sum_{n=1}^{N_c} d^{\beta w}_{\alpha u}(\mathbf{l}_m - \mathbf{l}_n)\, \exp[\mathrm{i}\,\mathbf{q}\cdot(\mathbf{l}_n - \mathbf{l}_m)]\,, \tag{17.137}$$

(17.133) becomes

$$\sum_{\beta w} A^{\beta w}_{\alpha u}(\mathbf{q})\, z_{\beta w}(0) = \omega^2 z_{\alpha u}(0)\,. \tag{17.138}$$

For a given $\mathbf{q}$, (17.138) is an eigenvalue equation of order $3N_b$, whose eigenvalues are found by solving the algebraic equation

$$\det\left(\mathbf{A} - \omega^2\,\mathbf{I}\right) = 0\,, \tag{17.139}$$

with $\mathbf{I}$ the identity matrix. As the entries of $\mathbf{A}$ depend on $\mathbf{q}$, the calculation of the $3N_b$ eigenvalues of (17.138) must be repeated for each distinct value of $\mathbf{q}$, namely, N_c times. The total number of eigenvalues thus found is $3N_b \times N_c = 3N$, as should be. This result shows that while the translation invariance eliminates the dependence on $\mathbf{l}_m$, it introduces that on $\mathbf{q}$. As the number of different determinations of the two vectors $\mathbf{l}_m$ and $\mathbf{q}$ is the same, namely, N_c, the total number of eigenvalues is not affected. Letting the N_c determinations of $\mathbf{q}$ be numbered as $\mathbf{q}_1, \mathbf{q}_2 \ldots \mathbf{q}_p \ldots$, the algebraic system (17.138) is recast as

$$\sum_{\beta w} A^{\beta w}_{\alpha u}(\mathbf{q}_p)\, z_{\beta w}(0, \mathbf{q}_p) = \omega^2(\mathbf{q}_p)\, z_{\alpha u}(0, \mathbf{q}_p)\,, \qquad p = 1, 2, \ldots N_c \tag{17.140}$$

which, for each $\mathbf{q}_p$, yields $3N_b$ eigenvalues ω^2 and $3N_b$ eigenvectors of length $3N_b$; as a consequence, the set of column vectors made of the eigenvectors associated with $\mathbf{q}_p$ forms a $3N_b \times 3N_b$ matrix, indicated here with $\mathbf{Z}_{1p}$. By letting $\mathbf{q}_p$ span over all its N_c determinations, the total number of eigenvalues turns out to be $3N$, namely,

$$\omega^2_{\gamma e}(\mathbf{q}_1)\,, \ldots\,, \omega^2_{\gamma e}(\mathbf{q}_{N_c})\,, \qquad \gamma = 1, \ldots, N_b\,, \qquad e = 1, 2, 3\,. \tag{17.141}$$

Similarly, the total number of eigenvectors (of order $3\,N_b$) turns out to be $3\,N$,

$$\mathbf{z}_{\gamma e}(0,\mathbf{q}_1)\,,\ \ldots\ ,\ \mathbf{z}_{\gamma e}(0,\mathbf{q}_{N_c})\,,\qquad \gamma=1,\ldots,N_b\,,\qquad e=1,2,3\,. \tag{17.142}$$

They provide the set of N_c square matrices of order $3\,N_b$, indicated with $\mathbf{Z}_{11},\mathbf{Z}_{12},\ldots,$ $\mathbf{Z}_{1N_c}$. Finally, each $\mathbf{z}_{\gamma e}(0,\mathbf{q}_p)$ provides an eigenvector of order $3\,N$ whose entries are

$$z^{\alpha u}_{\gamma e}(\mathbf{l}_m,\mathbf{q}_p)=\exp(\mathrm{i}\,\mathbf{q}_p\cdot\mathbf{l}_m)\ z^{\alpha u}_{\gamma e}(0,\mathbf{q}_p)\,, \tag{17.143}$$

where, as usual, $\alpha,\gamma=1,\ldots,N_b$; $u,e=1,2,3$ and, in turn, $m=0,\ldots,N_c-1$; $p=1,\ldots,N_c$. The first index of matrices $\mathbf{Z}_{11},\mathbf{Z}_{12},\ldots$ corresponds to $m=0$. Similarly, index $m=1$ provides a new set of matrices $\mathbf{Z}_{21},\mathbf{Z}_{22},\ldots$, and so on. The whole set of N_c^2 matrices $\mathbf{Z}_{mp}$ is equivalent to the $3\,N\times 3\,N$ matrix $\mathbf{Z}$ of the eigenvectors of the dynamic matrix, according to the following scheme:

$$\mathbf{Z}=\begin{bmatrix}\mathbf{Z}_{11} & \mathbf{Z}_{12} & \ldots & \mathbf{Z}_{1N_c}\\ \mathbf{Z}_{21} & \mathbf{Z}_{22} & \ldots & \mathbf{Z}_{2N_c}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{Z}_{N_c1} & \mathbf{Z}_{N_c2} & \ldots & \mathbf{Z}_{N_cN_c}\end{bmatrix}. \tag{17.144}$$

17.7.3 Properties of the Eigenvalues and Eigenvectors

Remembering that $\mathbf{A}$ (defined in (17.137)) is Hermitean, and ω^2 is real, one finds $(\mathbf{A}-\omega^2\,\mathbf{I})^*=\mathbf{A}^*-\omega^2\,\mathbf{I}=\mathbf{A}^T-\omega^2\,\mathbf{I}=(\mathbf{A}-\omega^2\,\mathbf{I})^T$, whence

$$\det\left[(\mathbf{A}-\omega^2\,\mathbf{I})^*\right]=\det\left[(\mathbf{A}-\omega^2\,\mathbf{I})^T\right]=\det\left(\mathbf{A}-\omega^2\,\mathbf{I}\right). \tag{17.145}$$

This shows that the eigenvalue equation $\mathbf{A}(\mathbf{q}_p)\,\mathbf{z}(0,\mathbf{q}_p)=\omega^2\,(\mathbf{q}_p)\,\mathbf{z}(0,\mathbf{q}_p)$, and its conjugate, $\mathbf{A}^*(\mathbf{q}_p)\,\mathbf{z}^*(0,\mathbf{q}_p)=\omega^2\,(\mathbf{q}_p)\,\mathbf{z}^*(0,\mathbf{q}_p)$ have the same eigenvalues. Moreover, as the entries (17.137) of $\mathbf{A}$ are polynomials in $\exp[\mathrm{i}\,\mathbf{q}_p\cdot(\mathbf{l}_n-\mathbf{l}_m)]$ with real coefficients, the following hold:

$$A^{\beta w}_{\alpha u}(-\mathbf{q}_p)=\left[A^{\beta w}_{\alpha u}(\mathbf{q}_p)\right]^*\,,\qquad \mathbf{A}(-\mathbf{q}_p)=\mathbf{A}^*(\mathbf{q}_p)\,. \tag{17.146}$$

The above properties give rise to other important consequences for the eigenvalues and eigenvectors. In fact, from the property $\mathbf{A}(-\mathbf{q}_p)=\mathbf{A}^*(\mathbf{q}_p)$ and the hermiticity of $\mathbf{A}$ one finds

$$\det[\mathbf{A}(-\mathbf{q}_p)-\omega^2\,\mathbf{I}]=\det\{[\mathbf{A}(\mathbf{q}_p)-\omega^2\,\mathbf{I}]^T\}=\det[\mathbf{A}(\mathbf{q}_p)-\omega^2\,\mathbf{I}]\,, \tag{17.147}$$

showing that the eigenvalues calculated from $\mathbf{A}(-\mathbf{q}_p)$ are the same as those calculated from $\mathbf{A}(\mathbf{q}_p)$. It follows that ω is an even function of $\mathbf{q}_p$:

$$\omega(-\mathbf{q}_p) = \omega(\mathbf{q}_p)\,, \qquad \mathbf{A}(-\mathbf{q}_p)\,\mathbf{z}(0,-\mathbf{q}_p) = \omega^2(\mathbf{q}_p)\,\mathbf{z}(0,-\mathbf{q}_p)\,. \tag{17.148}$$

Taking the conjugate of the second equation in (17.148) and using again the relation $\mathbf{A}(-\mathbf{q}_p) = \mathbf{A}^*(\mathbf{q}_p)$ yield $\mathbf{A}(\mathbf{q}_p)\,\mathbf{z}^*(0,-\mathbf{q}_p) = \omega^2(\mathbf{q}_p)\,\mathbf{z}^*(0,-\mathbf{q}_p)$. Comparing the above with the original eigenvalue equation $\mathbf{A}(\mathbf{q}_p)\,\mathbf{z}(0,\mathbf{q}_p) = \omega^2\,(\mathbf{q}_p)\,\mathbf{z}(0,\mathbf{q}_p)$ provides a relation between the eigenvectors:

$$\mathbf{z}(0,-\mathbf{q}_p) = \mathbf{z}^*(0,\mathbf{q}_p)\,. \tag{17.149}$$

From Bloch's theorem (17.131) it follows $z^{\alpha u}_{\gamma e}(\mathbf{l}_m,\mathbf{q}_p) = \exp(\mathrm{i}\,\mathbf{q}_p\cdot\mathbf{l}_m)\; z^{\alpha u}_{\gamma e}(0,\mathbf{q}_p)$ which, combined with (17.149), allows one to recover a set of real eigenvectors of the dynamic matrix:

$$z^{\alpha u}_{\gamma e}(\mathbf{l}_m,\mathbf{q}_p) + z^{\alpha u}_{\gamma e}(\mathbf{l}_m,-\mathbf{q}_p) = z^{\alpha u}_{\gamma e}(0,\mathbf{q}_p)\;\exp(\mathrm{i}\,\mathbf{q}_p\cdot\mathbf{l}_m) + z^{*\,\alpha u}_{\gamma e}(0,\mathbf{q}_p)\;\exp(-\mathrm{i}\,\mathbf{q}_p\cdot\mathbf{l}_m)$$

where, as usual, indices $p\gamma e$ count the eigenvectors and indices $m\alpha u$ count the entries. Using the results of Sect. 3.10, the displacements of the particles from the equilibrium position are given by $\mathbf{h} = \mathbf{G}\mathbf{b}$, where $\mathbf{G}$ is the matrix of the eigenvalues of (17.119) and the entries of $\mathbf{b}$ have the form (3.49), namely,

$$b_{p\gamma e}(t) = \frac{1}{2}\,\left\{\tilde{b}_{p\gamma e0}\;\exp[-\mathrm{i}\,\omega_{\gamma e}(\mathbf{q}_p)\,t] + \tilde{b}^*_{p\gamma e0}\;\exp[\mathrm{i}\,\omega_{\gamma e}(\mathbf{q}_p)\,t]\right\}\,, \tag{17.150}$$

with $\tilde{b}_{p\gamma e0}$ depending on the initial conditions $b_{p\gamma e0}(0)$, $\dot{b}_{p\gamma e0}(0)$. In turn, the entries of matrix $\mathbf{g}$ are $g^{m\alpha u}_{p\gamma e}$, where the lower indices refer to the columns and count the eigenvectors, the upper ones refer to the rows and count the entries of each eigenvector. Due to (17.128), such entries equal the corresponding terms of the real eigenvector of the dynamic matrix, divided by $\sqrt{\mu_\alpha}$. In conclusion, from $\mathbf{h} = \mathbf{G}\mathbf{b}$, the displacements are given by

$$h_{m\alpha u} = \sum_{p\gamma e} g^{m\alpha u}_{p\gamma e}\,b_{p\gamma e} = \sum_{p\gamma e}\frac{1}{\sqrt{\mu_\alpha}}\,z^{m\alpha u}_{p\gamma e}\,b_{p\gamma e}\,. \tag{17.151}$$

Using (17.150) yields

$$h_{m\alpha u} = \frac{1}{\sqrt{\mu_\alpha}}\,\Re\sum_{p\gamma e} z^{\alpha u}_{\gamma e}\,(0,\mathbf{q}_p)\,\left[\tilde{b}_{p\gamma e0}\;\exp(\mathrm{i}\,\Phi^m_{p\gamma e}) + \tilde{b}^*_{p\gamma e0}\;\exp(\mathrm{i}\,\Psi^m_{p\gamma e})\right]\,, \tag{17.152}$$

where the phases are defined by

$$\Phi^m_{p\gamma e} = \mathbf{q}_p\cdot\mathbf{l}_m - \omega_{\gamma e}(\mathbf{q}_p)\,t\,, \qquad \Psi^m_{p\gamma e} = \mathbf{q}_p\cdot\mathbf{l}_m + \omega_{\gamma e}(\mathbf{q}_p)\,t\,. \tag{17.153}$$

The above result shows that in the harmonic approximation, the displacements have the form of a superposition of plane and monochromatic waves, whose wave vector and angular frequency are $\mathbf{q}_p$, $\omega_{\gamma e}(\mathbf{q}_p)$. The wave corresponding to a given $\mathbf{q}_p$ is called *vibrational mode*. Typically, the number of $\mathbf{q}_p$ vectors is very large; in such cases, the same reasoning made in Sect. 17.6 with reference to the $\mathbf{k}$ vectors holds, and $\mathbf{q}$ is considered a continuous variable ranging over the first Brillouin zone.

Function $\omega_{\gamma e}(\mathbf{q})$ is also called *dispersion relation*, and is viewed as a multi-valued function of $\mathbf{q}$ having $3N_b$ branches. For each branch, letting $q = |\mathbf{q}|$, the *wavelength* is defined by $\lambda = 2\pi/q$ and the *phase velocity* by $u_f = \omega/q = \lambda\,\nu$, with $\nu = \omega/(2\pi)$ the frequency. The *group velocity* is defined by $\mathbf{u} = \mathrm{grad}_{\mathbf{q}}\omega$. As shown in Sect. 3.10, the total energy of the system is the sum of the mode energies, and in the classical description is expressed in terms of the initial conditions as

$$T_a + V_a = V_{a0} + \sum_{\sigma=1}^{3N} E_\sigma\,, \qquad E_\sigma = \frac{1}{2}\,\dot{b}_\sigma^2(0) + \frac{1}{2}\,\omega_\sigma^2\,b_\sigma^2(0)\,. \tag{17.154}$$

As remarked in Sect. 12.5, the classical expression of the energy associated with each mode has the same form as that of a mode of the electromagnetic field. In turn, the energy quantization shows that each mode energy is made of terms of the form $\hbar\omega_{\gamma e}(\mathbf{q})$, thus leading to the concept of phonon (Eqs. (12.35,12.36)).

17.8 Interaction of an Electron with the Lattice

The calculation of the vibrational spectra, carried out in Sect. (17.7) using the harmonic approximation, yields expression (17.152) for the displacements of the nuclei with respect to the equilibrium positions. More precisely, (17.152) provides, in classical terms, the uth component of the instantaneous displacement of the αth atom of the mth cell, where μ_α is the mass of the αth atom of the basis and $\mathbf{q}_p$ a vector of the scaled reciprocal lattice belonging to the first Brillouin zone; the form of this vector is given by the second relation in (17.132), namely, $\mathbf{q}_p = \sum_{s=1}^{3}(p_s/N_s)\,2\,\pi\,\mathbf{b}_s$, with $p_s = 0, 1, \ldots, N_s - 1$. The quantity $z_{\gamma e}^{\alpha u}(0, \mathbf{q}_p)$ in (17.152) is the component of indices $\alpha\,u$ of the eigenvector of indices $\gamma\,e$ of the $3\,N_b \times 3\,N_b$ dynamic matrix associated with $\mathbf{q}_p$. In turn, $\tilde{b}_{p\gamma e0}$ is a constant that depends on the initial conditions of the normal coordinate $b_{p\gamma e}(t)$. Finally, the phases in (17.152) are defined by (17.153), where the angular frequency $\omega_{\gamma e}(\mathbf{q}_p) > 0$ is the positive square root of the eigenvalue of indices $\gamma\,e$ of the $3\,N_b \times 3\,N_b$ dynamic matrix associated with $\mathbf{q}_p$. In summary, in the harmonic approximation the displacements have the form of a superposition of plane and monochromatic waves, with wave vector $\mathbf{q}_p$ and angular frequency $\omega_{\gamma e}(\mathbf{q}_p)$.

It must be remembered that the solution of the single-electron Schrödinger equation in a periodic lattice, carried out in Sect. 17.6, is based on the hypothesis that the nuclei are kept fixed in the equilibrium positions, so that the force acting on the

electron derives from a potential energy V having the periodicity of the lattice. As a consequence, the set of eigenvalues and eigenfunctions obtained from the solution of (17.40) do not account for the displacements of the nuclei produced by the lattice vibrations. As remarked in Sect. 16.6, if the positions of the nuclei are kept fixed, the exchange of energy between the system of electrons and that of nuclei cannot take place; in order to let such an exchange occur, it is necessary to incorporate the lattice vibrations into the problem. This task is carried out below using the perturbative approach, namely, starting from the set of eigenvalues and eigenfunctions of the unperturbed problem and accounting for the effect of the lattice vibrations using the first-order perturbative approach of Chap. 14.

Let the equilibrium positions of the nuclei be $\mathbf{R}_{0m\alpha} = \mathbf{l}_m + \mathbf{e}_\alpha$, with $m = 1, \ldots, N_c$ and $\alpha = 1, \ldots, N_b$. In the above, $\mathbf{e}_\alpha$ is the vector connecting the center of the mth cell with the equilibrium position of the αth atom of that cell; the potential energy of the electron in the unperturbed case is $V = V(\mathbf{r}, \mathbf{R}_{011}, \mathbf{R}_{012}, \ldots)$. Similarly, the displaced positions of the atoms are denoted with $\mathbf{R}_{m\alpha} = \mathbf{R}_{0m\alpha} + \mathbf{h}_{m\alpha}$, where $\mathbf{h}_{m\alpha}$ is the displacement vector of components $h_{m\alpha u}$, $u = 1, 2, 3$; the corresponding potential energy is $V_{\text{dis}} = V_{\text{dis}}(\mathbf{r}, \mathbf{R}_{11}, \mathbf{R}_{12}, \ldots)$ and the force acting on the electron is $-\operatorname{grad} V_{\text{dis}}$, where the gradient is calculated with respect to the components of $\mathbf{r}$. At this point it is necessary to better analyze the form of the potential energy: if the latter were simply due to the combined effects of hydrogen-like, Coulomb centers placed at the equilibrium positions of the atoms, its form would be

$$V^{\text{C}} = -\frac{q^2}{4\,\pi\,\varepsilon_0} \sum_{\alpha=1}^{N_b} C_\alpha\,, \qquad C_\alpha = \frac{1}{|\mathbf{r} - \mathbf{e}_\alpha|} + \frac{1}{|\mathbf{r} - \mathbf{l}_1 - \mathbf{e}_\alpha|} + \frac{1}{|\mathbf{r} - \mathbf{l}_2 - \mathbf{e}_\alpha|} + \ldots\,. \tag{17.155}$$

The sum C_α in (17.155) is made of N_c terms, each one containing a different translation vector from $\mathbf{l}_0 = 0$ to $\mathbf{l}_{N_c-1}$. The perturbed form of the potential energy, still in the hypothesis of hydrogen-like, Coulomb centers, is obtained by introducing the displacements $\mathbf{h}_{m\alpha}$ into the denominators of (17.155). Actually, due to the screening effect of the electrons the potential energy is more complicate than a superposition of interactions of the Coulomb type. However, it is still reasonable to assume that its form is of the type

$$V_{\text{dis}} \approx \sum_{m=1}^{N_c} \sum_{\alpha=1}^{N_b} U\left(\mathbf{r} - \mathbf{l}_m - \mathbf{e}_\alpha - \mathbf{h}_{m\alpha}\right)\,. \tag{17.156}$$

Another observation about the potential energy is that in a form like (17.155), the largest contribution to the sum is given by the summand having the minimum denominator. For example, assume that $\mathbf{r}$ is close to the equilibrium position of, say, an atom belonging to the mth cell. In this case the denominator $|\mathbf{r} - \mathbf{l}_m - \mathbf{e}_\alpha - \mathbf{h}_{m\alpha}|$ is small with respect to the other denominators, thus making the contribution of the mth cell to the sum dominant. This property becomes even more pronounced when the screening effect is accounted for, and may conveniently be exploited when evaluating integrals over Ω like, e.g., (17.157) below.

Note that the potential energies V or, similarly, V_{dis}, are a different concept from the potential energy V_a used for the harmonic approximation of the lattice vibrations: the latter is in fact the potential energy of the system of interacting nuclei and, as a consequence, the first-order term of its Taylor expansion with respect to the displacements is zero when the series is started from the equilibrium condition, namely, from the minimum of V_a. Instead, the single-electron potential energy has no special property even when the nuclei are in the equilibrium positions.

If V_{dis} as a whole is considered as the perturbation, the eigenfunctions of the unperturbed problem, namely, those corresponding to $V_{\text{dis}} = 0$, are the spatial part of planar waves; this yields for the perturbation-matrix entry[22]

$$H_{\mathbf{k}\,\mathbf{k}'} = \sum_{m\alpha} \int_{\Omega} U\left(\mathbf{r} - \mathbf{l}_m - \mathbf{e}_\alpha - \mathbf{h}_{m\alpha}\right) \exp\left[-\mathrm{i}\,\left(\mathbf{k}' - \mathbf{k}\right) \cdot \mathbf{r}\right] \mathrm{d}^3 r\,, \tag{17.157}$$

where each summand is proportional to the spatial Fourier transform of U. Noting that U is real one finds, as expected, $H_{\mathbf{k}'\,\mathbf{k}} = H^*_{\mathbf{k}\,\mathbf{k}'}$. To calculate the integral in (17.157) one lets $\mathbf{r}' = \mathbf{r} - \mathbf{l}_m - \mathbf{e}_\alpha - \mathbf{h}_{m\alpha}$. The integration domain in (17.157) should be shifted accordingly; however, as observed above, only the small portion of the integration domain around the atom of indices $m\,\alpha$ contributes to the integral, so the original domain Ω may be kept. Remembering that the displacements are small, one lets

$$\exp\left[-\mathrm{i}\,\left(\mathbf{k}' - \mathbf{k}\right) \cdot \mathbf{h}_{m\alpha}\right] \simeq 1 - \mathrm{i}\,\left(\mathbf{k}' - \mathbf{k}\right) \cdot \mathbf{h}_{m\alpha}\,, \tag{17.158}$$

to find

$$H_{\mathbf{k}\,\mathbf{k}'} \simeq \sum_{m\alpha} \left[1 - \mathrm{i}\,\left(\mathbf{k}' - \mathbf{k}\right) \cdot \mathbf{h}_{m\alpha}\right] \exp\left[-\mathrm{i}\,\left(\mathbf{k}' - \mathbf{k}\right) \cdot \mathbf{l}_m\right] J_\alpha(\mathbf{k}, \mathbf{k}')\,, \tag{17.159}$$

with

$$J_\alpha(\mathbf{k}, \mathbf{k}') = \int_{\Omega} U(\mathbf{r}') \exp\left[-\mathrm{i}\,\left(\mathbf{k}' - \mathbf{k}\right) \cdot \left(\mathbf{r}' + \mathbf{e}_\alpha\right)\right] \mathrm{d}^3 r'\,. \tag{17.160}$$

As before, one finds that interchanging $\mathbf{k}$ with $\mathbf{k}'$ transforms the integral in (17.160) into its conjugate; thus, this property is retained also by the simplified form (17.159) of $H_{\mathbf{k}\,\mathbf{k}'}$. From (17.159) one also finds that the perturbation-matrix entry is made of different contributions; they are treated separately below.[23]

[22] The perturbation-matrix entry is indicated here with a capital letter to avoid confusion with the displacements.

[23] Treating the contributions separately from each other may seem inconsistent due to the fact that eventually the transition probability is found by squaring the modulus of the Fourier transform of $H_{\mathbf{k}\,\mathbf{k}'}$ (Chap. 14). However, as shown below, the terms of $H_{\mathbf{k}\,\mathbf{k}'}$ are mutually incompatible due to

17.8.1 Rigid Lattice

Consider the contribution of the unity in the first factor of (17.159), which corresponds to the case of vanishing displacements. One carries out the sum over m first, after letting $\mathbf{k} - \mathbf{k}' = 2\,\pi\,\gamma_1\,\mathbf{b}_1 + 2\,\pi\,\gamma_2\,\mathbf{b}_2 + 2\,\pi\,\gamma_3\,\mathbf{b}_3$; the sum is treated in the same manner as in Sect. 14.8.6 and problems therein, to yield

$$\sum_{m=0}^{N_c-1} \exp[\mathrm{i}\,(\mathbf{k} - \mathbf{k}') \cdot \mathbf{l}_m] = N_c\,\delta[\mathbf{k} - \mathbf{k}' - \mathbf{g}_\gamma]\,, \tag{17.161}$$

where $\mathbf{g}_\gamma$ is an arbitrary translation vector of the scaled reciprocal lattice. As illustrated in Chap. 14, the transition probability from state $\mathbf{k}$ to state $\mathbf{k}'$ is obtained from the Fourier transform with respect to time of the perturbation-matrix entry; in particular, due to the absence of displacements the case considered here corresponds to a perturbation constant in time (Sect. 14.8.1), so that the transition probability is proportional to $\delta(E_\mathbf{k} - E_{\mathbf{k}'})$. In summary, the electron's energy is conserved and the following relations hold:

$$\mathbf{k} = \mathbf{k}' + \mathbf{g}_\gamma\,, \qquad E_\mathbf{k} = E_{\mathbf{k}'}\,. \tag{17.162}$$

This result describes a situation where the lattice acts as a rigid body and no exchange of energy with the electron takes place. If $E_\mathbf{k}$ has a spherical symmetry, from the second relation in (17.162) the equality of the moduli follows, $k' = k$; letting $2\,\vartheta$ be the angle between $\mathbf{k}$ and $\mathbf{k}'$ (Fig. 17.23), one obtains $g_\gamma = 2\,k\,\sin\vartheta$. The latter relation is recast in an interesting form by choosing an auxiliary set of characteristic vectors $\mathbf{a}'_1$, $\mathbf{a}'_2$, $\mathbf{a}'_3$ of the Bravais lattice such that the characteristic

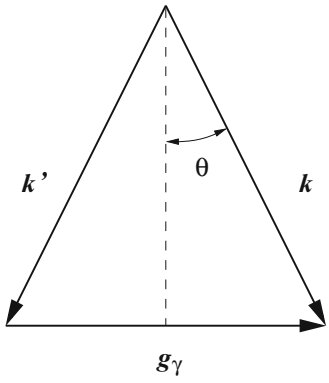


Fig. 17.23 Vectors used for determining the Bragg relation

the selection rules, namely, if one them contributes to the final result, the others vanish; thus, no inconsistency occurs.

vector $\mathbf{b}'_3$ of the reciprocal lattice is parallel to $\mathbf{g}_\gamma$. It follows $\mathbf{g}_\gamma = 2\,\pi\,i\,\mathbf{b}'_3$, with i an integer. Applying (17.4) to the auxiliary set, one finds that $\mathbf{b}'_3 = \mathbf{a}'_1 \wedge \mathbf{a}'_2/\tau'_l$ is normal to the lattice plane Π defined by $\mathbf{a}'_1$ and $\mathbf{a}'_2$; the closest plane parallel to Π is reached by the characteristic vector $\mathbf{a}'_3$ which, in general, is not normal to Π: thus, the distance d between two consecutive parallel planes of the Bravais lattice, in the chosen reference, is obtained by projecting $\mathbf{a}'_3$ along the direction of $\mathbf{b}'_3$. Using (17.5), one finds $d = \mathbf{a}'_3 \cdot \mathbf{b}'_3/b'_3 = 1/b'_3$, so that the modulus of $\mathbf{g}_\gamma$ turns out to be $g_\gamma = 2\,\pi\,i/d$. Combining the latter with $g_\gamma = 2\,k\,\sin\vartheta$ and letting $k = 2\,\pi/\lambda$ yields the *Bragg law*

$$2\,d\,\sin\vartheta = i\,\lambda\,, \tag{17.163}$$

of fundamental importance in the investigation of diffraction phenomena in crystals. Using (17.162) within (17.160) yields

$$J_\alpha(\mathbf{k}, \mathbf{g}_\gamma) = \int_\Omega U(\mathbf{r}')\,\exp\left[\mathrm{i}\,\mathbf{g}_\gamma \cdot (\mathbf{r}' + \mathbf{e}_\alpha)\right]\,\mathrm{d}^3 r'\,, \tag{17.164}$$

whence the part of (17.159) corresponding to vanishing displacements becomes

$$H_{\mathbf{k}\,\mathbf{g}_\gamma} = N_c \sum_\alpha J_\alpha(\mathbf{k}, \mathbf{g}_\gamma)\,. \tag{17.165}$$

17.8.2 Energy Exchange Between Electron and Lattice

If there is an exchange of energy between electron and lattice, $E_\mathbf{k} \neq E_{\mathbf{k}'}$, then the contribution of the unity in the first factor of (17.159) vanishes due to $\delta(E_\mathbf{k} - E_{\mathbf{k}'})$, so that one must consider the contribution deriving from $\mathrm{i}\,\left(\mathbf{k} - \mathbf{k}'\right) \cdot \mathbf{h}_{m\alpha}$. To proceed, consider the first summand in the expression (17.152) of the displacement,

$$\mathbf{h}'_{m\alpha}(t) = \frac{1}{\sqrt{\mu_\alpha}}\,\Re \sum_{p\gamma e} \mathbf{z}^\alpha_{\gamma e}\,(0, \mathbf{q}_p)\,\tilde{b}_{p\gamma e 0}\,\exp\left\{\mathrm{i}\,\left[\mathbf{q}_p \cdot \mathbf{l}_m - \omega_{\gamma e}(\mathbf{q}_p)\,t\right]\right\}\,, \tag{17.166}$$

where $\mathbf{z}^\alpha_{\gamma e}$ is a vector of components $z^{\alpha u}_{\gamma e}$, $u = 1, 2, 3$. Remembering that $\Re\lambda = (\lambda + \lambda^*)/2$, it follows that (17.166) provides in fact two contributions: from the first one, the sum over m acquires the extra factor $\exp(\mathrm{i}\,\mathbf{q}_p \cdot \mathbf{l}_m)$, whence (17.161) must be replaced with

$$\sum_{m=0}^{N_c - 1} \exp[\mathrm{i}\,(\mathbf{k} - \mathbf{k}' + \mathbf{q}_p) \cdot \mathbf{l}_m] = N_c\,\delta[\mathbf{k} - \mathbf{k}' + \mathbf{q}_p - \mathbf{g}_\gamma]\,. \tag{17.167}$$

In parallel, the time part acquires the extra factor $\exp[-\mathrm{i}\,\omega_{\gamma e}(\mathbf{q}_p)\,t]$, whence $\delta(E_{\mathbf{k}} - E_{\mathbf{k}'})$ must be replaced with $\delta[E_{\mathbf{k}} - E_{\mathbf{k}'} - \hbar\,\omega_{\gamma e}(\mathbf{q}_p)]$ (compare with the analysis of the harmonic perturbation in Sect. 14.8.2). Thus, (17.162) is replaced with

$$\mathbf{k} + \mathbf{q}_p = \mathbf{k}' + \mathbf{g}_\gamma\,, \qquad E_{\mathbf{k}} = E_{\mathbf{k}'} + \hbar\,\omega_{\gamma e}(\mathbf{q}_p)\,. \tag{17.168}$$

Conversely, the second contribution from $\mathbf{h}'_{m\alpha}$ provides the extra factors $\exp(-\mathrm{i}\,\mathbf{q}_p \cdot \mathbf{l}_m)$ and $\exp[\mathrm{i}\,\omega_{\gamma e}(\mathbf{q}_p)\,t]$, whence

$$\mathbf{k} - \mathbf{q}_p = \mathbf{k}' + \mathbf{g}_\gamma\,, \qquad E_{\mathbf{k}} = E_{\mathbf{k}'} - \hbar\,\omega_{\gamma e}(\mathbf{q}_p)\,. \tag{17.169}$$

As $\omega_{\gamma e}(\mathbf{q}_p) > 0$, the relations involving energy in (17.168) and (17.169) are mutually incompatible; they are also incompatible with (17.162). When (17.168) applies, the electron emits a phonon[24] of energy $\hbar\,\omega_{\gamma e}(\mathbf{q}_p)$, which is absorbed by the lattice; correspondingly, the electron's momentum changes by $\hbar\,\mathbf{k}' - \hbar\,\mathbf{k} = \hbar\,\mathbf{q}_p - \hbar\,\mathbf{g}_\gamma$. Here one lets $\mathbf{k}' - \mathbf{k} = \mathbf{q}_p - \mathbf{g}_\gamma$ in the definition (17.160); the result is further simplified by imposing that the three vectors $\mathbf{k}$, $\mathbf{k}'$, and $\mathbf{q}_p$ belong to the first Brillouin zone. In this case, following the same reasoning as in Sect. 14.8.6 one finds that if two of such vectors are prescribed, then the third one is uniquely defined, and the possible combinations of components of $\mathbf{g}_\gamma$ are those listed in Table 14.1; also, as $\mathbf{q}_p$ is fixed, the sum over p must eventually be dropped from (17.166).

When (17.169) applies, the electron absorbs a phonon of energy $\hbar\,\omega_{\gamma e}(\mathbf{q}_p)$, which is emitted by the lattice; correspondingly, the electron's momentum changes by $\hbar\,\mathbf{k}' - \hbar\,\mathbf{k} = -\hbar\,\mathbf{q}_p - \hbar\,\mathbf{g}_\gamma$. The further prescription that $\mathbf{k}$, $\mathbf{k}'$, and $\mathbf{q}_p$ belong to the first Brillouin zone yields for the components of $\mathbf{g}_\gamma$ the combinations listed in Table 14.2.

On must now consider the second summand in the expression (17.152) of the displacement, namely,

$$\mathbf{h}''_{m\alpha}(t) = \frac{1}{\sqrt{\mu_\alpha}}\,\Re \sum_{p\gamma e} \mathbf{z}^\alpha_{\gamma e}\,(0, \mathbf{q}_p)\,\tilde{b}^*_{p\gamma e 0}\,\exp\left\{\mathrm{i}\,\left[\mathbf{q}_p \cdot \mathbf{l}_m + \omega_{\gamma e}(\mathbf{q}_p)\,t\right]\right\}\,. \tag{17.170}$$

The analysis is identical to that of $\mathbf{h}'_{m\alpha}(t)$ and leads to the following result: when the energy balance of (17.168) applies, then the relation between the wave vectors is $\mathbf{k} - \mathbf{q}_p = \mathbf{k}' + \mathbf{g}_\gamma$; observing that $\mathbf{q}_p \neq 0$, the above is incompatible with the first relation in (17.168), hence only one of the two cases of phonon emission may occur. Similarly, when the energy balance of (17.169) applies, then the relation between the wave vectors is $\mathbf{k} + \mathbf{q}_p = \mathbf{k}' + \mathbf{g}_\gamma$, incompatible with the first relation in (17.169), hence only one of the two cases of phonon absorption may occur.

It is worth observing that the approach of this section and that of Sect. 14.8.6 are different; however, the final result is the same. In Sect. 14.8.6 it is assumed that the unperturbed problem derives from the solution of the Schrödinger equation

[24] The concept of phonon is introduced in Sect. 12.5 and elaborated further in Sect. 16.6.

with a periodic potential, whence the eigenfunctions of the unperturbed problem are Bloch functions; then, it is assumed that the spatial part of the perturbation has the form of a plane wave, $W(\mathbf{r}) = W_0 \exp(\mathrm{i}\,\mathbf{k}_d \cdot \mathbf{r})$. Here, the whole lattice potential V_{dis} is considered as the perturbation, whence the eigenfunctions of the unperturbed problem are planar waves.[25] The key point is (17.158), which brings in the displacement $\mathbf{h}_{m\alpha}$; the latter, in turn, brings in factor $\exp(\mathrm{i}\,\mathbf{q}_p \cdot \mathbf{l}_m)$ through (17.166), or factor $\exp(-\mathrm{i}\,\mathbf{q}_p \cdot \mathbf{l}_m)$ through (17.170). Comparing the expressions shows in fact that $\mathbf{k}_d$ in (14.62) is the analogue of $\mathbf{q}_p$ in (17.167). As a consequence, the selection rules are the same in the two approaches.

17.9 Complements

17.9.1 Crystal Planes and Directions in Cubic Crystals

From the general definition (17.1) of the translation vector, which provides the positions of all nodes of the crystal, it follows that a *crystal plane* is defined by the set of all triads of integers m_1, m_2, m_3 such that, for a given vector $\mathbf{g}_0$ of the scaled reciprocal lattice, the quantity $(m_1\,\mathbf{a}_1 + m_2\,\mathbf{a}_2 + m_3\,\mathbf{a}_3) \cdot \mathbf{g}_0/(2\,\pi)$ equals a fixed integer. Such a plane is normal to $\mathbf{g}_0$. In turn, given two crystal planes defined as above using, respectively, two nonparallel vectors $\mathbf{g}_1$ and $\mathbf{g}_2$, a *crystal direction* is defined by the set of all triads of integers m_1, m_2, m_3 that belong to the two crystal planes so prescribed. In cubic crystals, the typical method by which the crystal planes are identified is outlined below [128, Sect. 2-2].

Let the plane be indicated with Π. After labeling the nodes by the respective triads of integers m_1, m_2, m_3, as shown in Fig. 17.24, one starts by finding the intercepts of Π with the directions of the characteristic vectors. Letting such intercepts be (m_1^*, m_2^*, m_3^*), the triad $(r\,m_1^*, r\,m_2^*, r\,m_3^*)$ with $r \neq 0$ an integer, spans a set of planes parallel to Π. If M is the largest divisor of m_1^*, m_2^*, m_3^*, then the new triad $m_i' = m_i^*/M$ identifies the plane Π' parallel to Π and closest to the origin. Then, the inverse of the triad's elements is taken: $1/m_1'$, $1/m_2'$, $1/m_3'$. This avoids the occurrence of infinities; in fact, if Π was parallel to one of the characteristic vectors, say, $\mathbf{a}_i$, then m_i^* and m_i' would become infinite. One the other hand, using the inverse indices may bring to fractional numbers, a circumstance that must be avoided as well; so, as the last step, the new elements $1/m_i'$ are multiplied by the least multiple N of the m_i' that are not infinite:

$$(m_1'', m_2'', m_3'') = \left(\frac{N}{m_1'}, \frac{N}{m_2'}, \frac{N}{m_3'}\right) . \tag{17.171}$$

[25] Both approaches provide the same factor $\exp\left[-\mathrm{i}\,\left(\mathbf{k}' - \mathbf{k}\right) \cdot \mathbf{r}\right]$ that enters the calculation of the perturbation-matrix entry, (14.55) and (17.157) respectively.

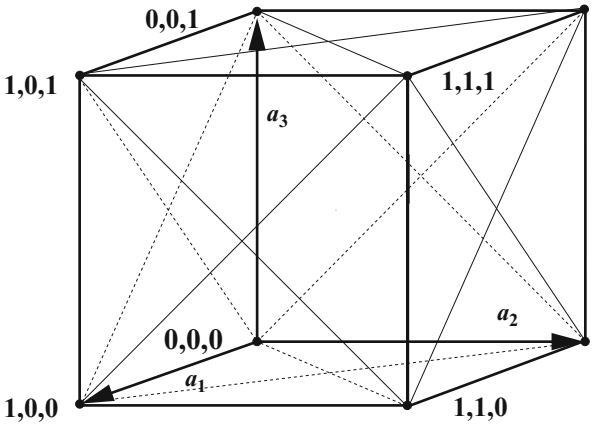


Fig. 17.24 Example of node labeling in the cubic lattice

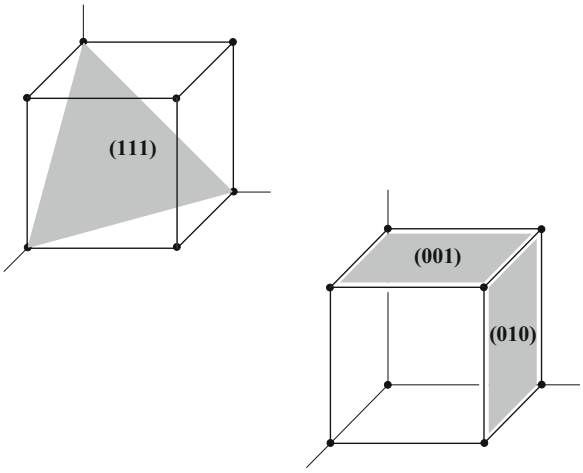


Fig. 17.25 Schematic representation of the (111) plane (top left) and of the (001) and (010) planes (bottom right) in a cubic crystal

The elements m_i'' thus found are the *Miller indices* of Π. They are enclosed in parentheses as in (17.171). By way of example, if $m_1^* = \infty$, $m_2^* = 2$, $m_3^* = 4$, then $M = 2$ so that $m_1' = \infty$, $m_2' = 1$, $m_3' = 2$. Calculating the inverse indices yields $1/m_1' = 0$, $1/m_2' = 1$, $1/m_3' = 1/2$; the least multiple is $N = 2$, so that the Miller indices are found to be $(0, 2, 1)$.

The indices that turn out to be negative are marked with a bar; for instance, in $(h\bar{k}l)$ the second index is negative. Some planes have the same symmetry; in cubic crystals this happens, for instance, to planes (100), (010), (001), $(\bar{1}00)$, $(0\bar{1}0)$, and $(00\bar{1})$. A set of planes with the same symmetry is indicated with braces, e.g., {100}. Examples of the (111), (001), and (010) planes are given in Fig. 17.25. As remarked in Sect. 24.4 about the silicon-oxidation process, the (111) plane has the highest concentration of atoms, followed by the {100} planes.

Symbols using three integers are also used to identify the crystal directions. To distinguish them from the symbols introduced so far, such triads of integers are

enclosed in brackets. Consider, for instance, the line connecting nodes P and Q, oriented from P to Q. Letting m_{1P}, m_{2P}, m_{3P} be the coordinates of P, and the like for those of Q, one forms the new triad $m'_1 = m_{1Q} - m_{1P}$, $m'_2 = m_{2Q} - m_{2P}$, and $m'_3 = m_{2Q} - m_{2P}$. Then, the indices of the crystal direction are obtained as

$$[m''_1, m''_2, m''_3] = \left[\frac{m'_1}{M}, \frac{m'_2}{M}, \frac{m'_3}{M}\right], \tag{17.172}$$

with M the largest divisor of m'_1, m'_2, m'_3. Also in this case, negative indices are marked with a bar. By way of examples, the characteristic vectors $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ in Fig. 17.24 are aligned, respectively, with the [100], [010], and [001] directions.

17.9.2 Examples of Translation Operators

A one-dimensional example of translation operator is easily found by considering the Taylor expansion of a function f around some position x:

$$f(x+l) = \sum_{n=0}^{\infty} \frac{l^n}{n!} \left(\frac{\mathrm{d}^n f}{\mathrm{d}x^n}\right)_{l=0} = \sum_{n=0}^{\infty} \frac{l^n}{n!} \frac{\mathrm{d}^n}{\mathrm{d}x^n} f(x) = \exp\left(l \frac{\mathrm{d}}{\mathrm{d}x}\right) f(x)\,, \tag{17.173}$$

where the expression on the right stems from a formal application of the Taylor expansion of the exponential function, in which a numerical factor within the exponent is replaced with the operator $\mathrm{d}/\mathrm{d}x$. Extending the above reasoning to three dimensions yields

$$\mathscr{T}(\mathbf{l}) = \exp(\mathbf{l} \cdot \mathrm{grad})\,. \tag{17.174}$$

17.9.3 Symmetries of the Hamiltonian Operator

Given an operator $\mathscr{R}$, a second operator $\mathscr{A}$ is associated with $\mathscr{R}$ in the following manner [74, Sect. 1.5]:

$$\mathscr{A} f(\mathbf{r}) = f(\mathscr{R}^{\dagger} \mathbf{r}) \tag{17.175}$$

for all functions f. Thus, the action of $\mathscr{A}$ on f at $\mathbf{r}$ is the same as calculating the original function at $\mathbf{r}' = \mathscr{R}^{\dagger}\mathbf{r}$. Let $\mathscr{R}$ be unitary (Sect. 8.6.2), whence $\mathbf{r} = \mathscr{R}\mathbf{r}'$. A unitary operator acting on $\mathbf{r}$ leaves the norm $r = |\mathbf{r}|$ unchanged; as a consequence, the unitary operations possible on the coordinates are only those that perform a

rotation or a reflexion of the coordinate axes, or both. It follows that the unit volume $\mathrm{d}\tau = \mathrm{d}^3 r$ is also invariant: $\mathrm{d}^3 \mathscr{R} r = \mathrm{d}^3 r$, thus showing that $\mathscr{A}$ is unitary as well:

$$\int_\tau |\mathscr{A} f(\mathbf{r})|^2 \, \mathrm{d}^3 r = \int_{\tau'} |f(\mathbf{r}')|^2 \, \mathrm{d}^3 \mathscr{R} r' = \int_{\tau'} |f(\mathbf{r}')|^2 \, \mathrm{d}^3 r' \,, \tag{17.176}$$

where τ' is the transformed domain. This reasoning does not apply to the translation operators $\mathscr{T}$. In fact, the operation $\mathscr{T}\mathbf{r} = \mathbf{r} + \mathbf{l}$ does not leave the norm of $\mathbf{r}$ unchanged. This shows in passing that $\mathscr{T}$ is not unitary. Other consequences of the above definitions and of the proof that $\mathscr{A}$ is unitary are

$$\mathscr{A} f(\mathscr{R}\mathbf{r}') = f(\mathbf{r}') \,, \qquad \mathscr{A}^\dagger f(\mathbf{r}') = f(\mathscr{R}\mathbf{r}') \,. \tag{17.177}$$

Also, $\mathscr{A}$ commutes with the operators that are invariant under the transformation $\mathbf{r} \leftarrow \mathscr{R}^\dagger \mathbf{r}$. In fact, if $\mathscr{B}$ is such an operator,

$$\mathscr{A}\mathscr{B}(\mathbf{r}) f(\mathbf{r}) = \mathscr{B}(\mathscr{R}^\dagger \mathbf{r}) f(\mathscr{R}^\dagger \mathbf{r}) = \mathscr{B}(\mathbf{r}) \mathscr{A} f(\mathbf{r}) \tag{17.178}$$

for all functions f. As $\mathscr{R}^\dagger$ is the inverse of $\mathscr{R}$, then $\mathscr{B}$ is also invariant under the transformation $\mathbf{r} \leftarrow \mathscr{R}\mathbf{r}$. As a consequence, $\mathscr{B}$ commutes also with $\mathscr{A}^\dagger$.

Let $\mathscr{B} v_n = b_n v_n$ be the eigenvalue equation for $\mathscr{B}$ (a discrete spectrum is assumed for the sake of simplicity). If b_n is s-fold degenerate, and $v_n^{(1)}$, $v_n^{(2)}$, …, $v_n^{(s)}$ are s linearly independent eigenfunctions corresponding to b_n, then

$$\mathscr{B} \sum_{i=1}^{s} c_i \, v_n^{(i)} = \sum_{i=1}^{s} c_i \, \mathscr{B} v_n^{(i)} = \sum_{i=1}^{s} c_i \, b_n \, v_n^{(i)} = b_n \sum_{i=1}^{s} c_i \, v_n^{(i)} \,, \tag{17.179}$$

namely, any nonvanishing linear combination of the form $\varphi_n = \sum_{i=1}^{s} c_i \, v_n^{(i)}$ is also an eigenfunction of $\mathscr{B}$ belonging to b_n. Let M be the space of all linear combinations of the form of φ_n; from (17.179) it follows that all members of M are eigenfunctions of $\mathscr{B}$ belonging to b_n. Conversely, all eigenfunctions of $\mathscr{B}$ belonging to b_n are members of M: letting $q_n \neq 0$ be one such eigenfunction, if q_n was not a member of M it would be $q_n - \sum_{i=1}^{s} c_i \, v_n^{(i)} \neq 0$ for all choices of the coefficients c_i. But this would imply that q_n, $v_n^{(1)}$, $v_n^{(2)}$, …, $v_n^{(s)}$ are $s+1$ linearly independent eigenfunctions of b_n, thus contradicting the hypothesis that the latter's degeneracy is of order s. Finally, if $\mathscr{A}$ commutes with $\mathscr{B}$ it is

$$\mathscr{B}\mathscr{A}\varphi_n = \mathscr{A}\mathscr{B}\varphi_n = \mathscr{A} b_n \varphi_n = b_n \mathscr{A}\varphi_n \,, \tag{17.180}$$

namely, $\mathscr{A}\varphi_n$ belongs to M.

In crystals, the unitary coordinate transformations $\mathbf{r}' = \mathscr{R}\mathbf{r}$ that leave the Hamiltonian operator $\mathscr{H}$ invariant are of particular interest. In fact, such coordinate transformations provide a method to study the degenerate eigenvalues of $\mathscr{H}$.

Let $\mathscr{B} = \mathscr{H}$, and let $\mathscr{H}$ be invariant under a coordinate transformation $\mathscr{R}\mathbf{r}$. If, in addition, $\mathscr{H}$ is translationally invariant and the periodic boundary conditions apply (Sect. 17.5.3), then the eigenfunctions w of $\mathscr{H}$ are Bloch functions, namely, they fulfill the Bloch theorem

$$w_i(\mathbf{r}+\mathbf{l},\mathbf{k}) = \exp(\mathrm{i}\,\mathbf{k}\cdot\mathbf{l})\, w_i(\mathbf{r},\mathbf{k})\,, \tag{17.181}$$

with $\mathbf{l}$ a translation vector and i the band index. Let $\mathscr{A}$ be the operator associated with $\mathscr{R}$. Then, from $\mathscr{H}\mathscr{A}^\dagger = \mathscr{A}^\dagger\mathscr{H}$,

$$\mathscr{H}\mathscr{A}^\dagger w_i(\mathbf{r},\mathbf{k}) = E_i(\mathbf{k})\, \mathscr{A}^\dagger w_i(\mathbf{r},\mathbf{k})\,, \tag{17.182}$$

with $E_i(\mathbf{k})$ the eigenvalue. One infers from (17.182) that if $w_i(\mathbf{r},\mathbf{k})$ and $\mathscr{A}^\dagger w_i(\mathbf{r},\mathbf{k})$ are linearly independent, then the eigenvalue is degenerate. Such a degeneracy does not depend on the detailed form of the Hamiltonian operator, but only on its symmetry properties. For this reason, the degeneracy is called *essential*. If further degeneracies exist, that depend on the detailed form of $\mathscr{H}$, they are called *accidental*.

Let $M(\mathbf{k})$ be the space made of the linearly independent eigenfunctions of $E(\mathbf{k})$, and of any nonvanishing linear combination of them, and define

$$v_i(\mathbf{r},\mathbf{k}') = \mathscr{A}^\dagger w_i(\mathbf{r},\mathbf{k}) = w_i(\mathscr{R}\mathbf{r},\mathbf{k})\,, \tag{17.183}$$

where symbol $\mathbf{k}'$ accounts for a possible influence on $\mathbf{k}$ of the coordinate transformation $\mathscr{R}\mathbf{r}$. Being an eigenfunction of $\mathscr{H}$, $v_i(\mathbf{r},\mathbf{k}')$ is a Bloch function,

$$v_i(\mathbf{r}+\mathbf{l},\mathbf{k}') = \exp(\mathrm{i}\,\mathbf{k}'\cdot\mathbf{l})\, v_i(\mathbf{r},\mathbf{k}')\,, \tag{17.184}$$

where $v_i(\mathbf{r}+\mathbf{l},\mathbf{k}') = w_i(\mathscr{R}\mathbf{r}+\mathscr{R}\mathbf{l},\mathbf{k})$. On the other hand, Bloch's theorem applied to $w_i(\mathscr{R}\mathbf{r}+\mathscr{R}\mathbf{l},\mathbf{k})$ yields

$$w_i(\mathscr{R}\mathbf{r}+\mathscr{R}\mathbf{l},\mathbf{k}) = \exp(j\mathbf{k}\cdot\mathscr{R}\mathbf{l})\, w_i(\mathscr{R}\mathbf{r},\mathbf{k})\,, \tag{17.185}$$

where the equality $\mathbf{k}\cdot\mathscr{R}\mathbf{l} = \mathscr{R}^\dagger\mathbf{k}\cdot\mathbf{l}$ holds due to the definition of adjoint operator. Comparison with the expression of the Bloch theorem applied to $v_i(\mathbf{r}+\mathbf{l},\mathbf{k}')$ provides $\mathbf{k}' = \mathscr{R}^\dagger\mathbf{k}$, whence

$$w_i(\mathscr{R}\mathbf{r},\mathbf{k}) = v_i(\mathbf{r},\mathscr{R}^\dagger\mathbf{k})\,. \tag{17.186}$$

In conclusion, if $w_i(\mathbf{r},\mathbf{k})$ is a Bloch function belonging to $M(\mathbf{k})$ and $\mathscr{R}\mathbf{r}$ a coordinate transformation that leaves the Hamiltonian operator invariant, then the eigenfunction obtained by such a transformation also belongs to $M(\mathbf{k})$ and is labeled by $\mathscr{R}^\dagger\mathbf{k}$. The following also holds true,

$$\mathscr{H}\, v_i(\mathbf{r},\mathscr{R}^\dagger\mathbf{k}) = E_i(\mathscr{R}^\dagger\mathbf{k})\, v_i(\mathbf{r},\mathscr{R}^\dagger\mathbf{k}) \tag{17.187}$$

which, compared with $\mathscr{H} w_i(\mathbf{r}, \mathbf{k}) = E_i(\mathbf{k})\, w_i(\mathbf{r}, \mathbf{k})$, shows that

$$E_i(\mathscr{R}^\dagger \mathbf{k}) = E_i(\mathbf{k})\,. \tag{17.188}$$

The theory of this section is applied by way of example to the Hamiltonian operator of a system of K electrons and N nuclei, interacting through electrostatic forces, that was introduced in Sect. 16.2. The potential energy is (compare with (16.5))

$$U_e(\mathbf{r}) + U_a(\mathbf{r}) + U_{ea}(\mathbf{r}, \mathbf{R}) + U_{\text{ext}}(\mathbf{r}, \mathbf{R})\,, \tag{17.189}$$

with

$$U_e(\mathbf{r}) = \sum_{i,j=1}^{K} \frac{q^2}{4\,\pi\,\varepsilon_0\,|\mathbf{r}_i - \mathbf{r}_j|}\,, \qquad j \neq i\,. \tag{17.190}$$

Similar expressions hold for U_a and U_{ea} (the second relation in (16.1) and (16.2), respectively). If $U_{\text{ext}} = 0$, the potential energy is invariant upon the reflexion transformation $\mathscr{R}\mathbf{r} = -\mathbf{r}$, $\mathscr{R}\mathbf{R} = -\mathbf{R}$. Clearly, the kinetic part of the Hamiltonian operator is also invariant. In the adiabatic approximation (Sect. 16.3), the coordinates of the nuclei are fixed to the equilibrium positions $\mathbf{R}_0$, which preserves the reflexion invariance. Finally, the reflexion invariance is still preserved in the Hartree and Hartree-Fock approximations (Sects. 16.4 and 16.5, respectively), which also provide single-electron Hamiltonian operators that are translationally invariant. Due to lattice periodicity, the eigenfunctions of the Hamiltonian operator are Bloch functions. Denoting now with $\mathbf{r}$ the coordinates associated with a single electron, the transformation $\mathscr{R}\mathbf{r} = -\mathbf{r}$ corresponds to $\mathscr{R}^\dagger \mathbf{k} = -\mathbf{k}$ whence, from (17.188),

$$E_i(-\mathbf{k}) = E_i(\mathbf{k})\,. \tag{17.191}$$

This type of degeneracy is accidental because it depends on the detailed form of the Hamiltonian operator. If the crystal has also a reflection symmetry, then the reflexion invariance of the single-electron Hamiltonian operators occurs irrespective of the form of the interactions. In this case, the degeneracy is essential.

17.9.4 Kronig-Penney Model

The general method for solving the Schrödinger equation in a periodic lattice, shown in Sect. 17.6, is applied here to a one-dimensional case, where the potential energy is described as the series of equal barriers shown in Fig. 17.26. The approach is called *Kronig-Penney model*; it is amenable to an analytical solution and, despite its simplicity, is able to capture the main properties of the dispersion relation $E(\mathbf{k})$.

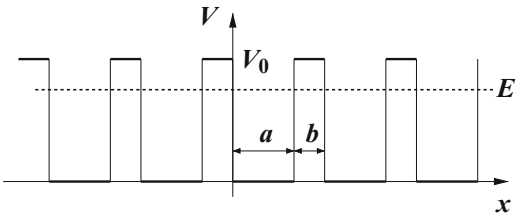


Fig. 17.26 Potential energy in the Kronig-Penney model

As shown in the figure, the potential energy is prescribed as $V = 0$ for $n\,(a + b) < x < n\,(a + b) + a$, and $V = V_0 > 0$ for $n\,(a + b) - b < x < n\,(a + b)$, with $n = 0, \pm 1, \pm 2 \ldots$ There is only one characteristic vector in the direct lattice, $\mathbf{a}_1 = (a + b)\,\mathbf{i}_1$; the corresponding characteristic vector of the reciprocal lattice is

$$\mathbf{b}_1 = \frac{\mathbf{i}_1}{a + b}\,. \tag{17.192}$$

As a consequence, the first Brillouin zone extends from $-\pi/(a+b)$ to $+\pi/(a+b)$ in the $\mathbf{i}_1$ direction. From the general properties of the time-independent Schrödinger equation (Sect. 8.2.3) it follows $E \geq 0$. As shown in Fig. 17.26, the case $0 < E < V_0$ is considered. A non-localized wave function w is expected even in the $E < V_0$ case due to the tunnel effect. From the Bloch theorem, the wave function has the form

$$w_k = u_k\, \exp(\mathrm{i}\, k\, x)\,, \qquad u_k(x + a + b) = u_k(x)\,, \tag{17.193}$$

where k belongs to the first Brillouin zone. In the intervals where $V = 0$ the Schrödinger equation reads

$$-\,w'' = \alpha^2\, w\,, \qquad \alpha = \sqrt{2\,m\,E}/\hbar > 0\,. \tag{17.194}$$

Replacing (17.193) into (17.194) yields

$$u_k'' + 2\,\mathrm{i}\,k\,u_k' - (k^2 - \alpha^2)\,u_k = 0\,, \tag{17.195}$$

whose associate algebraic equation has the roots

$$s = -\mathrm{i}\,k \pm \sqrt{-k^2 + (k^2 - \alpha^2)} = -\mathrm{i}\,k \pm \mathrm{i}\,\alpha\,. \tag{17.196}$$

The solution of (17.194) then reads

$$u_k^+ = c_1\, \exp[\mathrm{i}\,(\alpha - k)\,x] + c_2\, \exp[-\mathrm{i}\,(\alpha + k)\,x]\,, \tag{17.197}$$

with c_1, c_2 undetermined coefficients. The procedure is similar in the intervals where $V = V_0$ and yields

$$w'' = \beta^2 w\,, \quad \beta = \sqrt{2\,m\,(V_0 - E)}/\hbar\,, \quad u_k'' + 2\,\mathrm{i}\,k\,u_k' - (k^2 + \beta^2)\,u_k = 0\,, \tag{17.198}$$

$$s = -\mathrm{i}\,k \pm \sqrt{-k^2 + (k^2 + \beta^2)} = -\mathrm{i}\,k \pm \beta\,, \tag{17.199}$$

whence

$$u_k^- = c_3\,\exp[(\beta - \mathrm{i}\,k)\,x] + c_4\,\exp[-(\beta + \mathrm{i}\,k)\,x]\,, \tag{17.200}$$

with c_3, c_4 undetermined coefficients. The regional solutions u_k^+, u_k^- must fulfill the continuity conditions imposed by the general properties of the Schrödinger equation; in addition, they must fulfill the periodicity condition prescribed by the Bloch theorem (second relation in (17.193)). To proceed, one focuses on the period $-b \le x \le a$, so that the continuity conditions at $x = 0$ for the function, $u_k^+(0) = u_k^-(0)$, and first derivative, $(u_k^+)'(0) = (u_k^-)'(0)$, provide

$$c_1 + c_2 = c_3 + c_4\,, \qquad \mathrm{i}\,\alpha\,(c_1 - c_2) = \beta\,(c_3 - c_4)\,. \tag{17.201}$$

Combining (17.201),

$$c_1 = \sigma c_3 + \sigma^*\,c_4\,, \qquad c_2 = \sigma^*\,c_3 + \sigma c_4\,, \qquad 2\,\sigma = 1 - \mathrm{i}\,\beta/\alpha\,. \tag{17.202}$$

In turn, from the periodicity of u, namely, $u_k^+(a) = u_k^-(-b)$, and of u', namely, $(u_k^+)'(a) = (u_k^-)'(-b)$, one finds

$$c_1\,A + \frac{c_2}{A} = K\,L\,\left(\frac{c_3}{B} + c_4\,B\right)\,, \qquad c_1\,A - \frac{c_2}{A} = -K\,L\,\left(\frac{c_3}{B} - c_4\,B\right)\,\mathrm{i}\,\frac{\beta}{\alpha}\,, \tag{17.203}$$

with

$$A = \exp(\mathrm{i}\,\alpha\,a)\,, \quad B = \exp(\beta\,b)\,, \quad K = \exp(\mathrm{i}\,k\,a)\,, \quad L = \exp(\mathrm{i}\,k\,b)\,. \tag{17.204}$$

Combining (17.203),

$$c_1 = \frac{K\,L}{A}\,\left(\frac{\sigma}{B}\,c_3 + \sigma^*\,B\,c_4\right)\,, \qquad c_2 = A\,K\,L\,\left(\frac{\sigma^*}{B}\,c_3 + \sigma\,B\,c_4\right)\,. \tag{17.205}$$

Eliminating c_1, c_2 between (17.202) and (17.205) finally provides an algebraic system in the two unknowns $c3, c4$:

$$\sigma\,\left(1 - \frac{K\,L}{A\,B}\right)\,c_3 + \sigma^*\,\left(1 - \frac{B\,K\,L}{A}\right)\,c_4 = 0\,, \tag{17.206}$$

$$\sigma^*\,\left(1 - \frac{A\,K\,L}{B}\right)\,c_3 + \sigma\,\left(1 - A\,B\,K\,L\right)\,c_4 = 0\,. \tag{17.207}$$

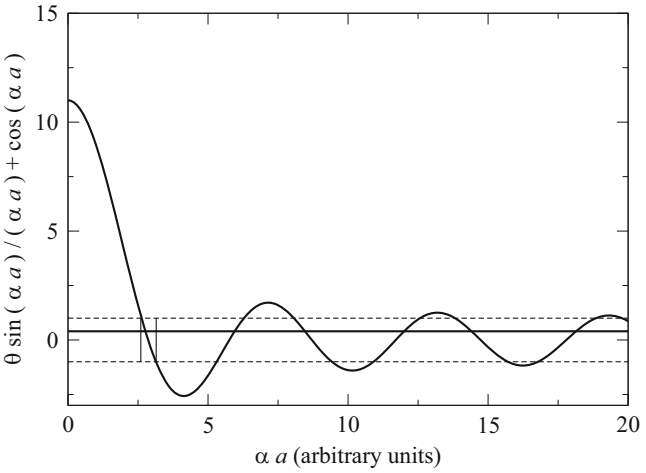


Fig. 17.27 Graphic solution of 17.210, with $\vartheta = 10$. The two vertical lines mark the values of $\alpha\, a$ delimiting the lowest band

As expected, the system is homogeneous, so a solution is possible only if the determinant vanishes. This in turn determines a relation between $\alpha(E)$, $\beta(E)$, and k, that eventually provides the dispersion relation $E(k)$. The determinant vanishes if

$$(\sigma^{*2} - \sigma^2)\left(K\,L + \frac{1}{K\,L}\right) = \sigma^{*2}\left(\frac{A}{B} + \frac{B}{A}\right) - \sigma^2\left(A\,B + \frac{1}{A\,B}\right). \qquad (17.208)$$

Introducing the expressions (17.202,17.204) of σ, A, B, K, L transforms (17.208) into

$$\frac{\beta^2 - \alpha^2}{2\,\alpha\,\beta}\,\sin(\alpha\,a)\,\sinh(\beta\,b) + \cos(\alpha\,a)\,\cosh(\beta\,b) = \cos[k\,(a+b)]\,, \qquad (17.209)$$

which has the form $F(E) = G(k)$. From this, the relation $E = E(k)$ can be determined. Note that $G(-k) = G(k)$ and $G[k + 2\,\pi/(a+b)] = G(k)$. As a consequence, the function $E(k)$ is even and has the periodicity of the reciprocal scaled lattice, as should be.

To the purpose of completing the analysis one may simplify (17.209) by considering a limiting case, namely, $V_0 \gg E$ so that, from (17.194,17.198), the limit $\beta^2 \gg \alpha^2$ would result. This, however, would eliminate the tunnel effect and reduce the problem to that of a series of boxes. To avoid this outcome, the proper limiting case is $b \to 0$ and $V_0 \to \infty$, in such a way as to leave the area

$b\,V_0$ of each barrier unchanged.[26] In other terms, one lets $b = \text{const}/V_0$, so that $\beta^2\,b \to \text{const} \neq 0$ while $\beta\,b \to 0$. It follows $\sinh(\beta\,b) \to \beta\,b$, $\cosh(\beta\,b) \to 1$ so that, letting $\vartheta = \lim(a\,b\,\beta^2/2)$, the $F(E) = G(k)$ relation (17.209) simplifies to

$$\vartheta\,\frac{\sin(\alpha\,a)}{\alpha\,a} + \cos(\alpha\,a) = \cos(k\,a)\,, \qquad \vartheta > 0\,, \qquad \alpha = \frac{\sqrt{2\,m\,E}}{\hbar}\,. \tag{17.210}$$

The function $E = E(k)$ can be determined by inverting (17.210); alternatively, it may be obtained in graphic form as shown in Fig. 17.27, where ϑ has been fixed to 10: given k, the right-hand side of (17.210) is fixed at some value $-1 \leq \cos(k\,a) \leq 1$. The energy E is then found by seeking $\alpha\,a$ such that the two sides become equal. The horizontal, dashed lines in the figure correspond to $\cos(ka) = 1$ and $\cos(k\,a) = -1$; they limit the interval where (17.210) has real solutions. The horizontal, continuous line corresponds to $\cos(k\,a) = 0.4$, while the oscillating curve represents the left-hand side of (17.210). The latter intercepts the $\cos(k\,a) = 0.4$ line at infinite points $\alpha_1\,a$, $\alpha_2\,a$, ...; from each α_i thus found, one determines the energy corresponding to the given k from the relation $\alpha_i = \sqrt{2\,m\,E_i}/\hbar$. Each branch of the multi-valued function $E(k)$ is then found by repeating the procedure for all values of k within the first Brillouin zone, thus making $\cos(k\,a)$ to range from -1 to 1. In the figure, the two vertical lines mark the values of $\alpha\,a$ delimiting the lowest band. The following are also worth noting:

- Letting λ indicate the left-hand side of (17.210), there are no real solutions for $\lambda > 1$ or $\lambda < -1$; the intervals with no real solutions are the forbidden bands. In fact, the k solutions in the forbidden bands are complex: it is $k\,a = \pm\mathrm{i}\,\log(\lambda + \sqrt{\lambda^2 - 1})$ when $\lambda > 1$, and $k\,a = \pi \pm \mathrm{i}\,\log(|\lambda| + \sqrt{\lambda^2 - 1})$ when $\lambda < -1$.
- At large energies the (17.210) relation tends to $\cos(\alpha a) = \cos(ka)$, namely, to the free-particle one: $k = \alpha = \sqrt{2\,m\,E}/\hbar$.
- Like in the general case, for a finite structure where the periodic boundary conditions are applied, the above calculation still holds, with k a discrete variable.

17.9.5 Linear, Monatomic Chain

The calculation of vibrational spectra has been carried out in general form in Sect. 17.7. Simple examples of application, with reference to a one-dimensional lattice, are given in this section and in the next one. Like the Kronig-Penney model used in Sect. 17.9.4 for determining the dispersion relation of electrons, the one-dimensional models of the lattice vibrations are amenable to analytical solutions; the latter, as shown below, are able to provide the explicit expression of the dispersion relation.

[26] The same type of limit is applicable to the single-barrier case, whose transmission coefficient is given in (11.22).

To begin, consider a one-dimensional monatomic lattice made of N_c cells (*linear, monatomic chain*). Let the lattice be aligned with the x axis, and the corresponding characteristic vector be $\mathbf{a} = a\,\mathbf{i}$, $a > 0$, with $\mathbf{i}$ the unit vector of the x axis. Finally, let the positions of the $N_c + 1$ nodes be $0, a, 2\,a, \ldots, n\,a, \ldots$. The translation vector associated with the nth node is $\mathbf{l}_n = n\,a\,\mathbf{i}$. Finally, it is assumed that the motion of each atom is constrained to the x axis, and the periodic boundary conditions are applied.

Due to the periodic boundary conditions the nodes of indices $n = 0$ and $n = N_c$ are actually the same node. As a one-dimensional case is considered, with $N_b = 1$, the total number of atoms is $N = N_c$. The number of the lattice's degrees of freedom is N_c, and the correspondence with the indices used in the general theory (compare with (17.124)) is

$$m\,, n = 1, \ldots, N_c\,, \qquad \alpha\,, \beta = 1\,, \qquad u\,, w = 1\,. \tag{17.211}$$

As only one atom per cell is present, one may assume that the equilibrium position of each nucleus coincides with that of a node. In the harmonic approximation the force acting on the rth nucleus is a linear function of the displacements:

$$F_r = -\sum_{k=1}^{N_c} c_{rk}\, h_k\,, \tag{17.212}$$

where all coefficients c_{rk} in general differ from 0. In real crystals, however, the interaction between nuclei becomes rapidly negligible as the distance increases. As a consequence, the dynamics of a nucleus may be tackled in a simplified manner by considering only the interaction with the neighboring nuclei to be effective. This is equivalent to letting $c_{rk} = 0$ when $|r - k| > 1$, whence

$$F_r = -c_r^{r-1}\, h_{r-1} - c_r^r\, h_r - c_r^{r+1}\, h_{r+1} = F_r(h_{r-1}, h_r, h_{r+1})\,. \tag{17.213}$$

In the coefficients of (17.213), the lower index refers to the node being acted upon by the force at the left-hand side, the upper index refers to the node whose displacement contributes to such a force. When the nuclei of indices $r - 1$, r, $r + 1$ are in the equilibrium positions it is $F_r(0, 0, 0) = 0$ for all r. On the other hand, it is also $F_r(\delta, \delta, \delta) = 0$, with $\delta \neq 0$ an arbitrary displacement. In fact, when all displacements are equal, the interatomic distance remains the same as in the equilibrium condition. From $F_r(\delta, \delta, \delta) = 0$ it follows that the coefficients are connected by the relation

$$c_r^{r-1} + c_r^r + c_r^{r+1} = 0\,. \tag{17.214}$$

Moreover, on account of the fact that all atoms are identical and all equilibrium distances are also identical, it is $F_r(-\delta, 0, \delta) = 0$, $\delta \neq 0$, whence

$$c_r^{r-1} = c_r^{r+1} \,. \tag{17.215}$$

From (17.214,17.215) it follows

$$c_r^r = -c_r^{r-1} - c_r^{r+1} = -2\, c_r^{r-1} = -2\, c_r^{r+1} \,. \tag{17.216}$$

Finally, the relation $F_r(0, \delta, 0) = -c_r^r\, \delta$, on account of the fact that $(0, 0, 0)$ is an equilibrium condition, shows that $c_r^r > 0$. As shown in Sect. 17.7 for the general case, due to the translation invariance the elastic coefficients do not depend on the cell index, but on the difference between cell indices (compare with (17.125)); in conclusion, letting

$$\chi = -c_r^{r-1} = -c_r^{r+1} > 0 \,, \qquad c_r^r = 2\,\chi \,, \tag{17.217}$$

and letting μ be the common mass of the nuclei, the dynamics of the rth nucleus is described by the equation

$$\mu\, \ddot{h}_r = -\chi\, (2\, h_r - h_{r+1} - h_{r-1}) \,. \tag{17.218}$$

The general theory shows that the displacement has the form

$$h_r = h_0\, \exp\left(\mathrm{i}\, q\, r\, a - \mathrm{i}\, \omega\, t\right) \,, \tag{17.219}$$

(compare with (17.152)), where h_0 is a complex constant, $q\, r\, a = \mathbf{q} \cdot \mathbf{l}_r = q\, \mathbf{i} \cdot r\, a\, \mathbf{i}$, and $\omega = \omega(q)$. Replacing (17.219) in (17.218) and dividing by h_r yield

$$\mu\, \omega^2 = \chi\, [2 - \exp(\mathrm{i}\, q\, a) - \exp(-\mathrm{i}\, q\, a)] = 4\, \chi\, \sin^2(q\, a/2) \,. \tag{17.220}$$

Defining $\tilde{\omega} = \sqrt{\chi/\mu}$ and remembering that ω is nonnegative, one finds the dispersion relation

$$\omega(q) = 2\tilde{\omega}\, |\sin(q\, a/2)| \,. \tag{17.221}$$

From the periodic boundary condition $h(r = N_c) = h(r = 0)$ one finds, with ν an integer,

$$\exp(\mathrm{i}\, q\, N_c\, a) = 1 \,, \qquad q\, N_c\, a = 2\, \pi\, \nu \,, \qquad q = \frac{\nu}{N_c}\, \frac{2\, \pi}{a} \,. \tag{17.222}$$

Replacing this form of q within (17.219) and (17.221) shows that using $\nu + N_c$ instead of ν leaves h_r and ω unchanged. As a consequence, it is sufficient to consider only N_c consecutive values of ν, say, $\nu = 0, 1, \ldots, N_c - 1$, which in turn limit the possible values of q to an interval of length $2\, \pi/a$. This was expected, because the values of the indices in (17.211) are such that the number of eigenvalues of the problem is N_c. Thus, the dispersion relation has only one branch, given by (17.221).

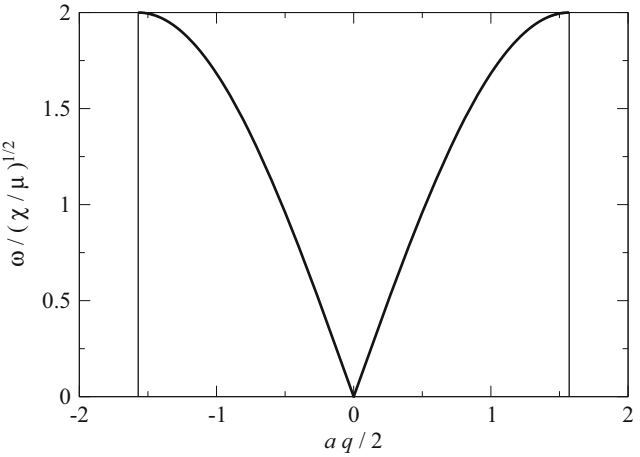


Fig. 17.28 Normalized dispersion relation of a linear, monatomic chain. The vertical lines, placed at $a\,q/2 = \pm\pi/2$, are the limits of the first Brillouin zone

One also notes that $2\,\pi/a$ is the size of the first Brillouin zone in the one-dimensional case. Typically, the interval of q is made to coincide with the first Brillouin zone, namely, $-\pi/a \le q < +\pi/a$. Also, as mentioned in Sect. 17.7, in most cases q is treated as a continuous variable. The phase and group velocities are

$$u_f = \frac{\omega}{q} = \pm a\,\tilde{\omega}\,\frac{\sin(q\,a/2)}{q\,a/2}\,, \qquad u = \frac{\mathrm{d}\omega}{\mathrm{d}q} = \pm a\,\tilde{\omega}\,\cos(q\,a/2)\,, \tag{17.223}$$

respectively, where the positive (negative) sign holds when q is positive (negative). At the boundary of the Brillouin zone it is $q\,a/2 = \pi/2$, whence $\omega = 2\,\tilde{\omega}$, $u_f = \pm a\tilde{\omega}/\pi$, $u = 0$. Near the center of the Brillouin zone it is $\omega \simeq a\tilde{\omega}\,|q|$, $u_f \simeq u \simeq \pm a\tilde{\omega}$. At the center it is $\omega = 0$.

The dispersion relation (17.221) normalized to $\tilde{\omega} = \sqrt{\chi/\mu}$ is shown in Fig. 17.28 as a function of $q\,a/2$. The range of the first Brillouin zone is $-\pi/2 \le q\,a/2 \le +\pi/2$. Remembering that the wavelength corresponding to q is $\lambda = 2\,\pi/q$, the interval near the origin, where the phase and group velocities are equal to each other and independent of q, corresponds to the largest values of the vibrations' wavelength. As some of these wavelengths fall in the audible range, the branch is called *acoustic branch*.

17.9.6 Linear, Diatomic Chain

As a second example, consider a one-dimensional lattice made of N_c cells, with a two-atom basis. Let the lattice be aligned with the x axis, and the corresponding

characteristic vector be $\mathbf{a} = a\,\mathbf{i}$, $a > 0$, with $\mathbf{i}$ the unit vector of the x axis. Finally, let the positions of the $N_c + 1$ nodes be $0, a, 2\,a, \ldots, n\,a, \ldots$. The translation vector associated with the nth node is $\mathbf{l}_n = n\,a\,\mathbf{i}$. Finally, it is assumed that the motion of each atom is constrained to the x axis, and the periodic boundary conditions are applied.

Due to the periodic boundary conditions the nodes of indices $n = 0$ and $n = N_c$ are actually the same node. As a one-dimensional case is considered, with $N_b = 2$, the total number of atoms is $N = 2\,N_c$. The number of degrees of freedom of the lattice is $2\,N_c$, and the correspondence with the indices (17.124) used in the general theory of Sect. 17.7 is

$$m\,, n = 1, \ldots, N_c\,, \qquad \alpha\,, \beta = 1, 2\,, \qquad u\,, w = 1\,. \tag{17.224}$$

As two atoms per cell are present, one may assume that the equilibrium position of one type of nucleus coincides with that of a node. Such nuclei will be given the index $\alpha, \beta = 1$, while the other nuclei will be given the index $\alpha, \beta = 2$. In the harmonic approximation the force acting on a nucleus is a linear function of the displacements:

$$F_r = -\sum_{k=1}^{2\,N_c} c_{rk}\,h_k\,. \tag{17.225}$$

In real crystals, the interaction between nuclei becomes rapidly negligible as the distance increases. Following the same reasoning as in Sect. 17.9.5, the dynamics of a nucleus is tackled in a simplified manner by considering only the interaction with the neighboring nuclei to be effective. This is equivalent to letting $c_{rk} = 0$ when $|r - k| > 1$. For a nucleus of type 1 the neighboring nuclei are of type 2. It is assumed that the node numbering is such, that the neighbors of interest belong to the cells of indices $r - 1$ and r. It follows

$$F_{r,1} = -c_{r,1}^{r-1,2}\,h_{r-1,2} - c_{r,1}^{r,1}\,h_{r,1} - c_{r,1}^{r,2}\,h_{r,2} = F_{r,1}(h_{r-1,2}\,, h_{r,1}\,, h_{r,2})\,. \tag{17.226}$$

In the coefficients of (17.226), the left-lower index refers to the node being acted upon by the force at the left-hand side, the left-upper index refers to the node whose displacement contributes to such a force, the right-lower and right-upper indices refer to the nucleus type. When the nuclei involved are in the equilibrium positions it is $F_{r,1}(0, 0, 0) = 0$. On the other hand, it is also $F_{r,1}(\delta, \delta, \delta) = 0$, with $\delta \neq 0$ an arbitrary displacement. In fact, when all displacements are equal the interatomic distance remains the same as in the equilibrium condition. From $F_{r,1}(\delta, \delta, \delta) = 0$ it follows

$$c_{r,1}^{r-1,2} + c_{r,1}^{r,1} + c_{r,1}^{r,2} = 0\,. \tag{17.227}$$

As, on the other hand, there is no special symmetry in the interaction of the nucleus of indices $r, 1$ with the neighboring ones, it is in general (in contrast to the case of a monatomic linear chain) $c_{r,1}^{r-1,2} \neq c_{r,1}^{r,2}$. Finally, the relation $F_{r,1}(0, \delta, 0) = -c_{r,1}^{r,1}\,\delta$, on account of the fact that $(0, 0, 0)$ is an equilibrium condition, shows that $c_{r,1}^{r,1} > 0$.

The calculation is then repeated for a nucleus of type 2, whose neighboring nuclei are of type 1. Due to the node numbering chosen here, the neighbors of interest belong to the cells of indices r and $r + 1$. As a consequence,

$$F_{r,2} = -c_{r,2}^{r,1}\,h_{r,1} - c_{r,2}^{r,2}\,h_{r,2} - c_{r,2}^{r+1,1}\,h_{r+1,1} = F_{r,2}(h_{r,1}\,, h_{r,2}\,, h_{r+1,1})\,. \tag{17.228}$$

By the same reasoning leading to (17.227) one finds

$$c_{r,2}^{r,1} + c_{r,2}^{r,2} + c_{r,2}^{r+1,1} = 0\,, \tag{17.229}$$

where, like in the case of (17.227), it is in general $c_{r,2}^{r,1} \neq c_{r,2}^{r+1,1}$. Finally, the relation $F_{r,2}(0, \delta, 0) = -c_{r,2}^{r,2}\,\delta$, on account of the fact that $(0, 0, 0)$ is an equilibrium condition, shows that $c_{r,2}^{r,2} > 0$. Due to the lattice periodicity the coefficients do not depend on the cell index, whence one lets

$$-\chi_1 = c_{r-1,1}^{r-1,2} = c_{r,1}^{r,2} = c_{r+1,1}^{r+1,2} = \ldots \quad -\chi_2 = c_{r,1}^{r-1,2} = c_{r+1,1}^{r,2} = c_{r+2,1}^{r+1,2} = \ldots \tag{17.230}$$

Remembering the invariance relation $c_{m\alpha u}^{n\beta w} = c_{n\beta w}^{m\alpha u}$ of the general theory one also finds

$$c_{r,2}^{r,1} = c_{r,1}^{r,2} = -\chi_1\,, \qquad c_{r,2}^{r+1,1} = c_{r+1,1}^{r,2} = c_{r,1}^{r-1,2} = -\chi_2\,, \tag{17.231}$$

whence

$$c_{r,1}^{r,1} = c_{r,2}^{r,2} = \chi_1 + \chi_2 > 0\,. \tag{17.232}$$

Finally, from the relations $F_{r,1}(0, 0, \delta) = \chi_1\,\delta$, $F_{r,1}(\delta, 0, 0) = \chi_2\,\delta$, it follows that $\chi_1 > 0$, $\chi_2 > 0$, on account of the fact that $(0, 0, 0)$ is an equilibrium condition. Letting μ_1, μ_2 be the masses of the two types of nuclei, the dynamics of the rth nuclei is described by the equations

$$\mu_1\,\ddot{h}_{r,1} = -\chi_1\,(h_{r,1} - h_{r,2}) - \chi_2\,(h_{r,1} - h_{r-1,2})\,, \tag{17.233}$$

$$\mu_2\,\ddot{h}_{r,2} = -\chi_1\,(h_{r,2} - h_{r,1}) - \chi_2\,(h_{r,2} - h_{r+1,1})\,, \tag{17.234}$$

where the displacements have the form

$$h_{r,1(2)} = h_{0,1(2)}\,\exp\left(\mathrm{i}\,qra - \mathrm{i}\,\omega t\right)\,, \tag{17.235}$$

thanks to the general theory. In (17.235), $h_{0,1}$, $h_{0,2}$ are complex constants, $q\,r\,a = \mathbf{q}\cdot\mathbf{l}_r = q\,\mathbf{i}\cdot r\,a\,\mathbf{i}$, and $\omega = \omega(q)$. Replacing (17.235) in (17.233,17.234) and dividing by $h_{r,1}$, $h_{r,2}$, respectively, yield

$$\mu_1\,\omega^2\,h_{0,1} = \chi_1\,(h_{0,1} - h_{0,2}) + \chi_2\,[h_{0,1} - h_{0,2}\,\exp(-\mathrm{i}\,q\,a)]\,, \tag{17.236}$$

$$\mu_2\,\omega^2\,h_{0,2} = \chi_1\,(h_{0,2} - h_{0,1}) + \chi_2\,[h_{0,2} - h_{0,1}\,\exp(+\mathrm{i}\,q\,a)]\,. \tag{17.237}$$

Defining $A_{11}, A_{12}, A_{21}, A_{22}$ such that

$$\mu_1\,A_{11} = \chi_1 + \chi_2\,, \qquad -\mu_1\,A_{12} = \chi_1 + \chi_2\,\exp(-\mathrm{i}\,q\,a)\,, \tag{17.238}$$

$$\mu_2\,A_{22} = \chi_1 + \chi_2\,, \qquad -\mu_2\,A_{21} = \chi_1 + \chi_2\,\exp(+\mathrm{i}\,q\,a)\,, \tag{17.239}$$

the homogeneous algebraic system (17.236,17.237) transforms into

$$\left(A_{11} - \omega^2\right)\,h_{0,1} + A_{12}\,h_{0,2} = 0\,, \qquad A_{21}\,h_{0,1} + \left(A_{22} - \omega^2\right)\,h_{0,2} = 0\,. \tag{17.240}$$

The trace $T = A_{11} + A_{22}$ and determinant $D = A_{11}\,A_{22} - A_{12}\,A_{21}$ of the matrix formed by $A_{11}, A_{12}, A_{21}, A_{22}$ read

$$T = \frac{\mu_1 + \mu_2}{\mu_1\mu_2}\,(\chi_1 + \chi_2)\,, \qquad D = 2\,\frac{\chi_1\,\chi_2}{\mu_1\,\mu_2}\,[1 - \cos(q\,a)]\,. \tag{17.241}$$

The eigenvalues ω^2 are found by solving the algebraic equation $(\omega^2)^2 - T\,\omega^2 + D = 0$, whose discriminant is

$$\Delta(q) = T^2 - 4\,D = \left[\frac{(\chi_1 + \chi_2)\,(\mu_1 + \mu_2)}{\mu_1\,\mu_2}\right]^2 + 8\,\frac{\chi_1\,\chi_2}{\mu_1\,\mu_2}\,[\cos(q\,a) - 1]\,. \tag{17.242}$$

Remembering that $\chi_1, \chi_2 > 0$, the minimum Δ_m of (17.242) occurs for $q = \pm\pi/2$. Letting $K_\chi = (\chi_1 - \chi_2)^2/(4\,\chi_1\,\chi_2) \geq 0$ and $K_\mu = (\mu_1 - \mu_2)^2/(4\,\mu_1\,\mu_2) \geq 0$, one finds the relation

$$\Delta(q) \geq \Delta_m = 16\,\frac{\chi_1\,\chi_2}{\mu_1\,\mu_2}\,\left[\left(1 + K_\chi\right)\left(1 + K_\mu\right) - 1\right] \geq 0\,, \tag{17.243}$$

showing that the discriminant is nonnegative. It follows that the eigenvalues ω^2 are real, as should be. The solution of the algebraic equation provides two branches of the dispersion relation to be found by taking the square root of

$$\omega^2 = \frac{T}{2} \pm \frac{1}{2}\,\sqrt{\Delta(q)}\,. \tag{17.244}$$

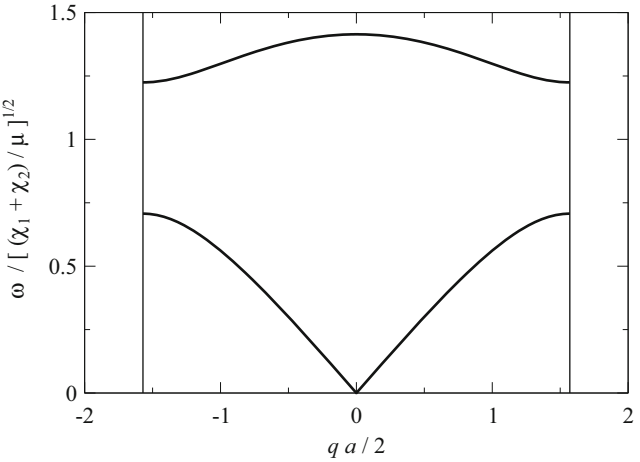


Fig. 17.29 Normalized dispersion relation of a linear, diatomic chain with $\mu_1 = \mu_2 = \mu$ and $\chi_1 = 3\,\chi_2$. The vertical lines, placed at $a\,q/2 = \pm\pi/2$, are the limits of the first Brillouin zone

Observing that $\Delta(0) = T^2$, one finds that selecting the minus sign in (17.244) provides the branch that contains $\omega = 0$. As in the case of the monatomic chain, this branch is called *acoustic branch*. In the other branch it is always $\omega > 0$; in ionic crystals like, e.g., sodium chloride, the frequencies typical of this branch are excited by infrared radiation. For this reason, the branch is called *optical branch*.

The acoustic and optical branch of a linear, diatomic chain are shown in Fig. 17.29, where $\mu_1 = \mu_2 = \mu$ is assumed for simplicity. Letting $\tilde{\omega} = \sqrt{(\chi_1 + \chi_2)/\mu}$, one finds for the acoustic branch

$$\frac{\omega_{\rm ac}^2}{\tilde{\omega}^2} = 1 - \left[1 - 4\,\frac{\chi_1\,\chi_2}{(\chi_1 + \chi_2)^2}\,\sin^2\left(\frac{q\,a}{2}\right)\right]^{1/2}. \tag{17.245}$$

At the center of the first Brillouin zone it is $\omega_{\rm ac} = 0$, while the maximum of $\omega_{\rm ac}$ is reached at the boundary $qa/2 = \pm\pi/2$ of the zone. For the optical branch one finds

$$\frac{\omega_{\rm op}^2}{\tilde{\omega}^2} = 1 + \left[1 - 4\,\frac{\chi_1\,\chi_2}{(\chi_1 + \chi_2)^2}\,\sin^2\left(\frac{q\,a}{2}\right)\right]^{1/2}. \tag{17.246}$$

At the center of the first Brillouin zone $\omega_{\rm op}$ reaches its maximum. The minimum of $\omega_{\rm op}$ is reached at the boundary $qa/2 = \pm\pi/2$ of the zone. The discretization of q due to the periodic boundary conditions, and the definitions of the phase and group velocity, are the same as those already given for the monatomic lattice.

From (17.245,17.246), the distance $G = \omega_{\rm op}(q = \pm\pi/a) - \omega_{\rm ac}(q = \pm\pi/a)$ between the minimum of the optical branch and the maximum of the acoustic branch fulfills the relation

$$\frac{G}{\tilde{\omega}} = \left(1 + \frac{|\chi_1 - \chi_2|}{\chi_1 + \chi_2}\right)^{1/2} - \left(1 - \frac{|\chi_1 - \chi_2|}{\chi_1 + \chi_2}\right)^{1/2} . \tag{17.247}$$

It is $G > 0$ if $\chi_2 \neq \chi_1$, whereas $G = 0$ if $\chi_2 = \chi_1$. The latter case is called *degenerate*.

17.9.7 Analogies

It is interesting to note the analogy between the expression of the energy of the electromagnetic field *in vacuo*, described as a superposition of modes (Eqs. (5.38, 5.40)), and that of a system of vibrating nuclei (Eqs. (3.48) and (17.118)). In essence, the two expressions derive from the fact that in both cases the energy is a positive-definite, symmetric quadratic form. In the case of the electromagnetic field the form is exact because of the linearity of the Maxwell equations; for the vibrating nuclei the form is approximate because of the neglect of the anharmonic terms (Sect. 3.13.1).

Other analogies exist between the dispersion relation $E(\mathbf{k})$ of the electrons subjected to a periodic potential energy, worked out in Sect. 17.6, and the dispersion relation $\omega(\mathbf{q})$, worked out in Sect. (17.7). Both relations are even and periodic in the reciprocal, scaled lattice; both have a branch structure, the difference being that the number of branches of $\omega(\mathbf{q})$ is finite because the number of degrees of freedom of the vibrating lattice is finite, whereas that of $E(\mathbf{k})$ is infinite.

Chapter 18
Electrons and Holes in Semiconductors at Equilibrium

18.1 Introduction

The purpose of this chapter is to provide the equilibrium expressions of the electron and hole concentrations in a semiconductor. For comparison, the cases of insulators and conductors are discussed qualitatively, and the concepts of conduction band, valence band, and generation of an electron-hole pair are introduced. The important issue of the temperature dependence of the concentrations is also discussed. Then, the general expressions of the concentrations in an intrinsic semiconductor are worked out, followed by an estimate of the Fermi level's position. Next, the equilibrium expressions are worked out again, this time in the case where substitutional impurities of the donor or acceptor type are present within the semiconductor. The mechanism by which donor-type dopants provide electrons to the conduction band and acceptor-type dopants provide holes to the valence band is explained. An important outcome of the analysis is that the introduction of suitable dopants makes the concentration of majority carriers practically independent of temperature, at least in a range of temperatures of practical interest for the functioning of integrated circuits. The simplifications due to the complete-ionization and nondegeneracy conditions are illustrated, along with the compensation effect. Finally, the theory is extended to the case of a nonuniform doping distribution, where the concentrations must be calculated self-consistently with the electric potential by solving the Poisson equation. The last section illustrates the band-gap narrowing phenomenon. In the complements, after a brief description of the relative importance of germanium, silicon, and gallium arsenide in the semiconductor industry, a qualitative analysis of the impurity levels is carried out by an extension of the Kronig-Penney model, and the calculation of the position of the impurity levels with respect to the band edges is carried out.

M. Rudan, *Physics of Semiconductor Devices*, DOI 10.1007/978-3-319-63154-7_18

18.2 Equilibrium Concentration of Electrons and Holes

The expressions worked out in this chapter are obtained by combining the information about the band structure of the crystal under investigation, given in Chap. 17, with that of the equilibrium distribution of fermions (Chap. 15). A qualitative description is given first, starting from the simplified case where the structures of the conduction and valence band are symmetric (Fig. 17.14).

Considering again a case where the Fermi level E_F coincides with the gap's midpoint and $M_V\, m_h^{3/2} = M_C\, m_e^{3/2}$, let the only difference with respect to the material of Fig. 17.14 be that the energy gap is larger (Fig. 18.1). Despite the fact that the situations illustrated in the two figures may appear similar to each other, it must be realized that a 2×10^{31} amplification factor, with respect to the scale of $g(E)$ alone, is necessary to make the products $g(E)\,P(E)$, $g(E)\,[1 - P(E)]$ visible in Fig. 18.1, in contrast to the 10^3 factor used in Fig. 17.14. In practice, for the material of Fig. 18.1 the states in the conduction band are empty and those of the valence band are full.[1]

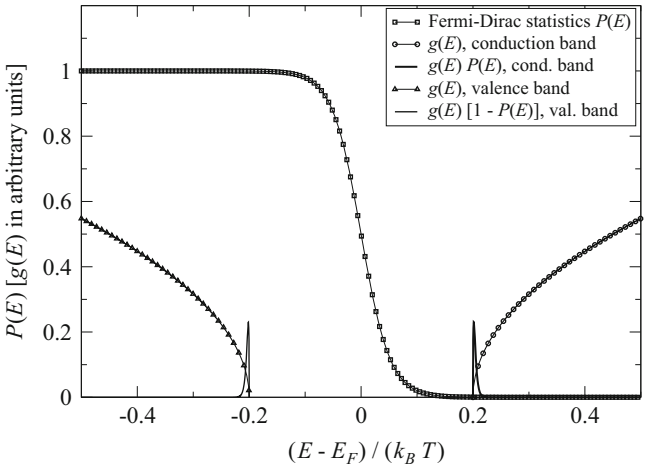


Fig. 18.1 Description of the particles' population in the conduction and valence bands of an insulator. To make them more visible, the products $g(E)\,P(E)$ and $g(E)\,[1 - P(E)]$ have been amplified, with respect to $g(E)$ alone, by a factor 2×10^{31} (compare with Figs. 17.14 and 18.2). The gap's extension is arbitrary and does not refer to any specific material

[1] The curves of Figs. 17.14, 18.1 are drawn in arbitrary units. To better appreciate the difference in a practical situation, one may use the equilibrium concentration of electrons in silicon at $T = 300$ K, which is about 10^{16} m^{-3}. Thus, if Fig. 17.14 is thought of as representing silicon, the ratio of the amplification factors used in Figs. 17.14 and 18.1 produces, in the latter, a concentration of about one electron in 1000 km^3.

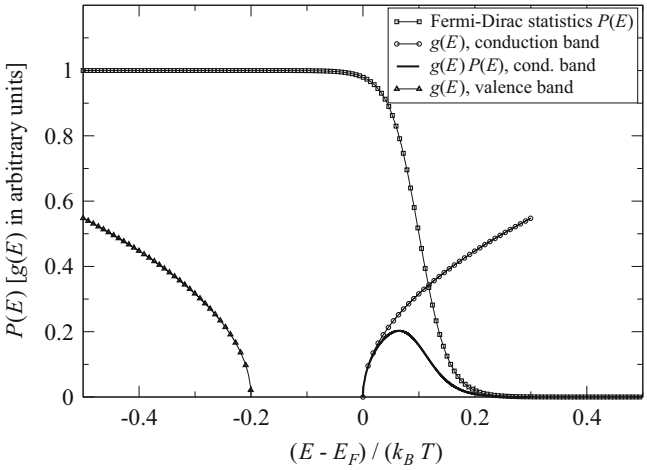


Fig. 18.2 Description of the electron population in the conduction band of a conductor. The product $g(E)\,P(E)$ is drawn in the same scale as $g(E)$ alone (compare with Figs. 17.14 and 18.1). The gap's extension is arbitrary and does not refer to any specific material

Remembering that a band whose states are empty and, similarly, a band whose states are fully occupied, do not provide any conduction, it turns out that the electrical conductivity of a material like that of Fig. 18.1 is expected to vanish: the material is an *insulator*.

A different case is found when the Fermi level is inside a band, like in the crystal illustrated in Fig. 18.2. The name *conduction band* is given in this case to the band where the Fermi level belongs; the band beneath is called *valence band* also in this case. Due to the position of the Fermi level, the parabolic-band approximation is grossly mistaken, and is used for a qualitative discussion only. In the figure, the $g(E)\,P(E)$ product is drawn without using any amplification factor, thus showing that the electron concentration in the conduction band is much larger than for a semiconductor; the latter band, in turn, is the only one that contributes to conduction, because the concentration of holes in the valence band is negligible. Due to the much larger concentration of charges one expects that the conductivity of the crystal under investigation be large; in fact, this is found to be the case, and the crystal is a *conductor*.

In summary, the combination of a few factors: position of the Fermi level with respect to the bands, gap's width, and structure of the density of state, dictates the presence of partially filled bands and the concentration of charges in them. Other situations may occur besides those depicted in Figs. 17.14, 18.1, and 18.2, e.g., the Fermi level may be positioned within the gap, but closer to one of the band edges than to the other. They are illustrated in Sects. 18.4.1 and 18.4.2.

Another important issue is the dependence of the electron and hole populations on temperature. Considering the case of a semiconductor first, the discussion is carried out with reference to Figs. 17.14 and 15.3; the latter illustrates the tempera-

ture dependence of the Fermi-Dirac statistics, using the simplifying hypothesis that the position of the Fermi level does not change with temperature.[2] Following the reasoning carried out in Sect. 17.6.5, the temperature dependence of the effective masses is neglected as well. When temperature increases, the occupation probability of the conduction-band states increases, thus making the concentration of electrons in the conduction band to increase; at the same time, the occupation probability of the valence-band states decreases, thus making the concentration of holes in the valence band to increase as well. This outcome is easily understood if one thinks that the increase in temperature is achieved by transferring energy from an external reservoir to the semiconductor; part of this energy is absorbed by the nuclei and produces a change in the equilibrium distribution of phonons (Sect. 16.6), while the other part is absorbed by the electrons and produces a redistribution of the latter within the available energy states. The absorption of energy by electrons that, prior to the temperature increase, belonged to valence-band states makes some of them to transit to the conduction band. As each electron that makes a transition leaves a hole in the valence band, this is an example of *generation of an electron-hole pair* (compare with Sect. 17.6.6).

As shown in Sect. 19.5.5, the conductivity of a semiconductor is proportional to $\mu_n n + \mu_p p$, where n is the concentration of conduction-band electrons, p the concentration of valence-band holes, and μ_n, μ_p the electron and hole mobilities, respectively. Mobilities account for the scattering events undergone by electrons and holes during their motion within the material, and are found to decrease when temperature increases. It follows that the decrease in mobility competes with the increase of the concentrations in determining the temperature dependence of conductivity in a semiconductor. In practice, the increase in the concentrations is much stronger due to the exponential form of the Fermi-Dirac statistics, so that the conductivity of a semiconductor strongly increases with temperature.[3]

The qualitative analysis of conductivity is the same for a conductor where, as holes are absent, the conductivity is proportional to $\mu_n n$. However, the outcome is different; in fact, while μ_n, like that of a semiconductor, decreases when temperature increases, the electron concentration n is unaffected by temperature. In fact, from Fig. 18.2 one finds that the deformation of the Fermi-Dirac statistics due to a temperature variation produces a rearrangement of the electron distribution within the conduction band itself, and no hole-pair generations; as a consequence, the energies of some electrons of the conduction band change, while the electron number (hence the concentration n) does not. As mobility depends weakly on temperature, the conductivity of a conductor turns out to slightly decrease as temperature increases.

[2]The temperature dependence of the Fermi level is influenced by the shape of the density of states ([70] and Prob. 15.2). Such a dependence is neglected in the qualitative discussion carried out here.

[3]This is a negative aspect because the electrical properties of the material are strongly influenced by the ambient temperature. The drawback is absent in doped semiconductors, at least in the range of temperatures that are typical of the operating conditions of semiconductor devices (Sect. 18.4.1).

The calculation of the equilibrium electron concentration in the conduction band of a semiconductor is based on (17.76), where the density of states in energy g is replaced with the combined density of states in energy and volume γ given by (17.73). The concentration of electrons and the Fermi level are indicated here with n_i, E_{Fi} instead of n, E_F to remark the fact that the semiconductor is free from impurities (for this reason, the semiconductor is called *intrinsic*). From the discussion of Sect. 17.6.5, the parabolic-band approximation is acceptable, so that the combined density of states is given by (17.73).[4] Remembering that the lower edge of the conduction band is indicated with E_C, and letting E_{CU} be the upper edge of the same band, one finds

$$n_i = \int_{E_C}^{E_{CU}} \gamma(E)\, P(E)\, \mathrm{d}E \simeq \int_{E_C}^{+\infty} \frac{\sqrt{2}\, M_C\, m_e^{3/2}\, \sqrt{E - E_C}/(\pi^2\, \hbar^3)}{\exp[(E - E_{Fi})/(k_B\, T)] + 1}\, \mathrm{d}E\,, \tag{18.1}$$

where the upper integration limit has been replaced with $+\infty$ on account of the fact that the integrand vanishes exponentially at high energies. Using the auxiliary variables

$$\zeta_e = E_C - E_{Fi}\,, \qquad x = \frac{E - E_C}{k_B\, T} \geq 0\,, \qquad \xi_e = -\frac{\zeta_e}{k_B\, T}\,, \tag{18.2}$$

with x, ξ_e dimensionless, transforms (18.1) into

$$n_i = \frac{\sqrt{2}}{\pi^2 \hbar^3}\, M_C m_e^{3/2}\, (k_B\, T)^{3/2} \int_0^{+\infty} \frac{\sqrt{x}}{\exp(x - \xi_e) + 1}\, \mathrm{d}x\,. \tag{18.3}$$

From the definition (C.116) of the Fermi integral of order $1/2$ it then follows

$$n_i = N_C\, \Phi_{1/2}(\xi_e)\,, \qquad N_C = 2\, M_C \left(\frac{m_e}{2\,\pi\,\hbar^2}\, k_B\, T\right)^{3/2}\,, \tag{18.4}$$

with N_C the *effective density of states* of the conduction band. Observing that $\Phi_{1/2}$ is dimensionless, the units of N_C are m^{-3}.

The concentration of holes in the valence band is determined in a similar manner, namely, starting from an integral of the form (17.76), where $P(E)$ is replaced with $1 - P(E)$ and the density of states in energy g is replaced with the combined density of states in energy and volume γ. The concentration of holes is indicated here with p_i and, as for the electrons, the parabolic-band approximation is adopted, so that the combined density of states is obtained from (17.74). Finally, $1 - P(E)$ is expressed through (17.77). Remembering that the upper edge of the valence band is indicated with E_V, and letting E_{VL} be the lower edge of the same band, one finds

[4] As before, the band index is dropped from (17.73).

$$p_i = \int_{E_{VL}}^{E_v} \gamma(E)\,[1 - P(E)]\,\mathrm{d}E \simeq \int_{-\infty}^{E_V} \frac{\sqrt{2}\,M_V\,m_h^{3/2}\,\sqrt{E_V - E}/(\pi^2\,\hbar^3)}{\exp[(E_{Fi} - E)/(k_B\,T)] + 1}\,\mathrm{d}E\,, \tag{18.5}$$

where the lower integration limit has been replaced with $-\infty$ on account of the fact that the integrand vanishes exponentially at low energies. Using the auxiliary variables

$$\zeta_h = E_{Fi} - E_V\,, \qquad x = \frac{E_V - E}{k_B\,T} \geq 0\,, \qquad \xi_h = -\frac{\zeta_h}{k_B\,T}\,, \tag{18.6}$$

transforms (18.5) into

$$p_i = \frac{\sqrt{2}}{\pi^2\hbar^3}\,M_V m_h^{3/2}\,(k_B\,T)^{3/2} \int_0^{+\infty} \frac{\sqrt{x}}{\exp(x - \xi_h) + 1}\,\mathrm{d}x\,, \tag{18.7}$$

whence, introducing the *effective density of states* of the valence band,

$$p_i = N_V\,\Phi_{1/2}(\xi_h)\,, \qquad N_V = 2\,M_V \left(\frac{m_h}{2\,\pi\,\hbar^2}\,k_B\,T\right)^{3/2}. \tag{18.8}$$

18.3 Intrinsic Concentration

Equations (18.2,18.4) and (18.6,18.8) express the equilibrium concentrations of electron and holes in a semiconductor in terms of temperature and of the distance of the Fermi level from the edge E_C of the conduction band or, respectively, from the edge E_V of the valence band. Obviously the two distances are not independent from each other; from (18.2,18.6) one finds in fact

$$-\,(\xi_e + \xi_h) = \frac{\zeta_e + \zeta_h}{k_B\,T} = \frac{E_C - E_{Fi} + E_{Fi} - E_V}{k_B\,T} = \frac{E_G}{k_B\,T}\,, \tag{18.9}$$

where $E_G = E_C - E_V$ is the extension of the semiconductor's gap, known from the calculation of the band structure, or also from electrical or optical measurements. As the band structure is influenced by temperature, the gap depends on temperature as well (Sect. 17.6.5); such a dependence is important because it strongly influences the concentration of electrons and holes. The results of gap's calculations or measurements, that show that E_G decreases when temperature increases, are usually rendered in compact form by means of interpolating expressions, an example of which is

$$E_G(T) \approx E_{G0} - \alpha\,\frac{T^2}{T + \beta}\,, \tag{18.10}$$

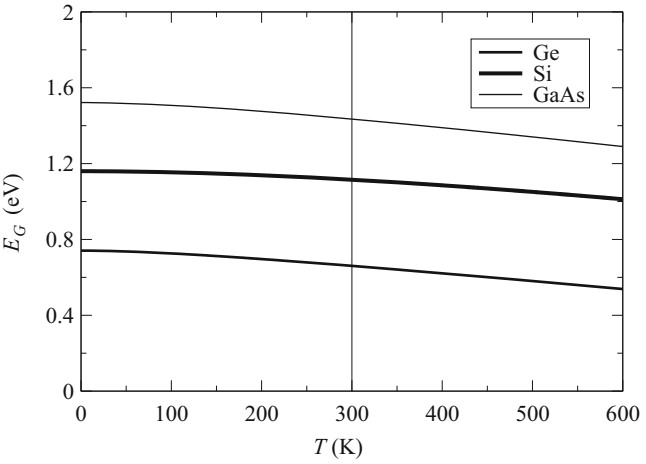


Fig. 18.3 Plot of the gap as a function of temperature for Ge, Si, and GaAs. The vertical line marks $T = 300$ K

Table 18.1 Gap and average effective masses of silicon, germanium, and gallium arsenide

Material	E_{G0} (eV)	α (eV/K)	β (K)	$E_G(T_a)$	m_e/m_0	m_h/m_0
Si	1.160	7.02×10^{-4}	1,108	1.12	0.33	0.56
Ge	0.741	4.56×10^{-4}	210	0.66	0.22	0.31
GaAs	1.522	5.80×10^{-4}	300	1.43	0.68	0.50

Note: Symbol T_a indicates the room temperature

where E_{G0} is the gap's width extrapolated at $T = 0$ and $\alpha > 0$, $\beta > 0$ are material's parameters. Table 18.1 reports the parameters related to the gap's width for Si, Ge, and GaAs, along with the values of the average effective masses normalized to the rest mass of the free electron [128, Chap. 2-3]. The plot of $E_G(T)$ is shown in Fig. 18.3 for the three semiconductors of Table 18.1.

Note that expressions (18.4,18.8) can be used only if the position of the Fermi level is known; in fact, the latter (which is unknown as yet) enters the definitions (18.2,18.6) of parameters ξ_e, ξ_h. To proceed one remembers that in the $T \to 0$ limit the Fermi-Dirac statistics becomes discontinuous, specifically it is $P = 1$ for $E < E_{Fi}$ and $P = 0$ for $E > E_{Fi}$ (Sect. 15.8.1); on the other hand, the experimental evidence shows that in the $T \to 0$ limit the conductivity of a semiconductor vanishes: this corresponds to a situation where all states of the conduction band are empty while those of the valence bands are filled. In conclusion, when $T \to 0$ the Fermi level is still positioned in the gap and it is $n_i = p_i = 0$. If, starting from this situation, the temperature is brought again to some finite value $T > 0$, such that some of the valence-band electrons transit to the conduction band,

the total number of holes thus formed equals that of the transited electrons. Due to the spatial uniformity of the material, the concentrations are equal to each other as well; in conclusion it is[5]

$$n_i = p_i\,, \qquad M_C\, m_e^{3/2}\, \Phi_{1/2}(\xi_e) = M_V\, m_h^{3/2}\, \Phi_{1/2}(\xi_h)\,, \tag{18.11}$$

the second of which has been obtained by deleting the common factors from (18.4) and (18.8). Relations (18.9) and (18.11) are a set of two equations in the unknowns ξ_e, ξ_h, whose solution allows one to determine the position of the Fermi level through (18.2) or (18.6).

The second relation in (18.11) can also be exploited for carrying out an estimate of ξ_e, ξ_h, basing on the values of the masses given in Table 18.1. To this purpose one observes[6] that the ratio $M_C\, m_e^{3/2}/(M_V\, m_h^{3/2})$ is about 1.4, 1.2, 0.8 for Si, Ge, and GaAs, respectively, so that a crude estimate is obtained by letting $\Phi_{1/2}(\xi_e) \simeq \Phi_{1/2}(\xi_h)$. As the Fermi integral is a monotonic function of the argument (Sect. C.13), it follows $\xi_e \simeq \xi_h$. Replacing from (18.2,18.6) yields $E_{Fi} \simeq (E_C + E_V)/2$, and $\zeta_e \simeq \zeta_h \simeq E_G/2 > 0$.

The usefulness of this estimate actually lies in that it simplifies the Fermi integrals. Taking for instance the case of room temperature, it is $k_B\, T_a \simeq 26$ meV; using the values of $E_G(T_a)$ from Table 18.1 shows that $E_G \gg k_B T_a$, whence $-\xi_e \gg 1$ and $-\xi_h \gg 1$, so that the approximate expression (C.112) applies: $\Phi_{1/2}(\xi) \simeq \exp(\xi)$. In conclusion, (18.4,18.8) simplify to $n_i \simeq N_C \exp(\xi_e)$, $p_i \simeq N_V \exp(\xi_h)$, namely,

$$n_i \simeq N_C \exp\left(-\frac{E_C - E_{Fi}}{k_B\, T}\right), \qquad p_i \simeq N_V \exp\left(-\frac{E_{Fi} - E_V}{k_B\, T}\right). \tag{18.12}$$

As the two concentrations are equal to each other it is customary to use the same symbol n_i for both; the product of the two expressions (18.12) combined with (18.9) yields

$$n_i\, p_i = n_i^2 \simeq N_C\, N_V \exp(\xi_e + \xi_h) = N_C\, N_V \exp[-E_G/(k_B\, T)]\,. \tag{18.13}$$

The expression of the *intrinsic concentration* thus reads

$$n_i \simeq \sqrt{N_C\, N_V}\, \exp\left(-\frac{E_G}{2\, k_B\, T}\right), \qquad N_C\, N_V = \frac{M_C\, M_V}{2\, \pi^3\, \hbar^6}\, (m_e\, m_h)^{3/2}\, (k_B\, T)^3\,. \tag{18.14}$$

[5]Note that the reasoning leading to the first relation in 18.11 is not limited to the case of the parabolic-band approximation, but holds for a general form of the densities of states.

[6]Remembering the discussion carried out in Sect. 17.6.5 it is $M_C(\text{Si}) = 6$, $M_C(\text{Ge}) = 4$, $M_C(\text{GaAs}) = 1$, $M_V = 2$.

Table 18.2 Intrinsic concentrations of silicon, germanium, and gallium arsenide

Material	$N_C(T_a)$ (cm^{-3})	$N_V(T_a)$ (cm^{-3})	$n_i(T_a)$ (cm^{-3})
Si	2.82×10^{19}	1.05×10^{19}	7.61×10^{9}
Ge	1.04×10^{19}	0.43×10^{19}	1.39×10^{12}
GaAs	4.45×10^{19}	0.99×10^{19}	2.40×10^{6}

Note: In this field it is customary to express the concentrations in cm^{-3} instead of m^{-3}

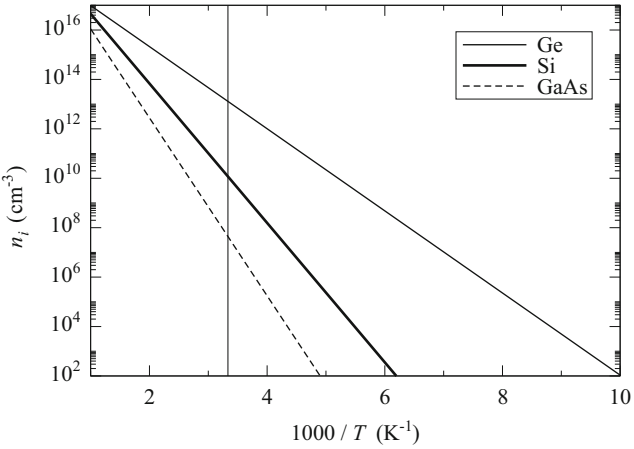


Fig. 18.4 Arrhenius plot of the intrinsic concentration in Ge, Si, and GaAs. The vertical line marks $T = 300$ K

Its values at room temperature are listed in Table 18.2 along with those of the effective densities of states, for Si, Ge, and GaAs. As for the temperature dependence of n_i one notes that, besides appearing in the exponent's denominator in the first relation of (18.14), the lattice temperature also influences the numerator E_G and the $N_C N_V$ factor. Among these dependencies, that of the exponent's denominator is by far the strongest; it follows that a first-hand description of $n_i(T)$ can be given by considering only the latter. This yields to the Arrhenius plots shown in Fig. 18.4, where the relations $n_i(T)$ reduce to straight lines. It is important to note that despite the similarity of the effective densities of states for the three semiconductors considered here (Table 18.2), the intrinsic concentrations differ by orders of magnitude. This is due to the exponential dependence of n_i on E_G, which amplifies the differences in E_G (visible in Table 18.1) of the three semiconductors.

The estimate of the position of the Fermi level in an intrinsic semiconductor carried out above has led to the conclusion that the Fermi integrals used for calculating the intrinsic concentrations can be replaced with exponentials.[7] Such a

[7] This is not necessarily true for an *extrinsic* semiconductor, where suitable impurity atoms are introduced into the semiconductor lattice (Sects. 18.4.1, 18.4.2).

conclusion can now be exploited for a more precise calculation of the Fermi level's position, where the coefficients of (18.11) are kept. Using the exponentials in (18.11) one finds

$$M_C\, m_e^{3/2}\, \exp(\xi_e) = M_V\, m_h^{3/2}\, \exp(\xi_h)\,. \tag{18.15}$$

Taking the logarithm of both sides and using (18.9) yield $\xi_h - \xi_e = \xi_h + \xi_e - 2\,\xi_e = \log[M_C\, m_e^{3/2}/(M_V\, m_h^{3/2})]$, whence

$$E_C - E_{Fi} = \frac{E_G}{2} + \frac{k_B\, T}{2}\, \log\left(\frac{M_C\, m_e^{3/2}}{M_V\, m_h^{3/2}}\right)\,. \tag{18.16}$$

The second term at the right-hand side of (18.16) is the correction with respect to the estimate carried out earlier. In the $T \to 0$ limit, the Fermi level in the intrinsic semiconductors under consideration coincides with the gap's midpoint. When the temperature increases, if $M_C\, m_e^{3/2} > M_V\, m_h^{3/2}$ the distance $E_C - E_{Fi}$ becomes larger, that is, the Fermi level moves towards the valence band; the opposite happens if $M_C\, m_e^{3/2} < M_V\, m_h^{3/2}$. For all practical purposes, considering that the argument of the logarithm in (18.16) is close to unity and the coefficient $k_B\, T$ is always small with respect to E_G, the position of the Fermi level can be thought of as coinciding with the gap's midpoint.

18.4 Uniform Distribution of Impurities

As described in Chap. 23, the fabrication of integrated circuits (IC) requires the introduction into the semiconductor material of atoms (called *impurities* or *dopants*) belonging to specifically selected chemical species. Dopants are divided into two classes, termed *n-type* and *p-type*. With reference to silicon (Si), the typical *n*-type dopants are phosphorus (P), arsenic (As), and antimony (Sb), while the typical *p*-type dopants are boron (B), aluminum (Al), gallium (Ga), and Indium (In). For a qualitative introduction to the effects of dopants it is instructive to start with the case of an intrinsic semiconductor; the analysis is based on the simplified picture in which the original arrangement of the atoms in space, like that shown in Fig. 17.11, is deformed in such a way as to become two dimensional. This representation, shown in Fig. 18.5 with reference to silicon, is convenient for the description carried out below. Each silicon atom has four electrons in the external shell, so that it can form four covalent bonds[8] with other identical atoms; each pair of lines connecting two atoms in Fig. 18.5 stands for a pair of shared electrons. In the $T \to 0$ limit, the electrons are permanently bound to the atoms because the energy necessary to

[8] In a *covalent* bond atoms share their outermost electrons.

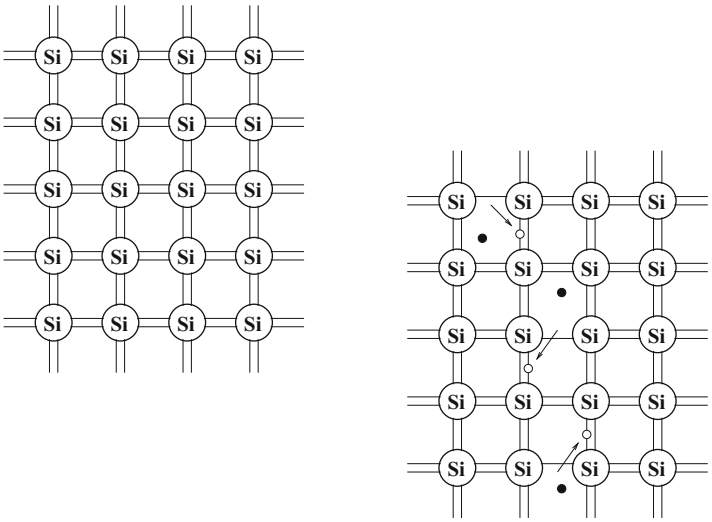


Fig. 18.5 Two-dimensional representation of the intrinsic silicon lattice. The upper-left part of the figure shows the $T \to 0$ limit

ionize is not available; this situation is drawn in the upper-left part of the figure. If, instead, the temperature is brought to some finite value $T > 0$ by transferring energy from an external reservoir to the semiconductor, part of this energy is absorbed by some of the electrons; the latter break the bond and become free to move within the material.[9] This situation is depicted in the lower-right part of Fig. 18.5, where the free electrons are represented with black dots.

When an electron becomes free and departs from an atom, the unbalanced positive charge left behind in the nucleus deforms the shape of the potential energy in the vicinity of the nucleus itself. The deformation is such that the potential-energy barrier that separates the nucleus from the neighboring ones becomes lower and thinner; this, in turn, enhances the probability of tunneling across the barrier by an electron belonging to a shared pair. Such a tunneling event restores the original pair, but leaves an unbalanced positive charge behind; instead of considering it as the motion of an electron from a complete to an incomplete pair, the tunneling event is more conveniently described as the opposite motion of an empty state, that is, a hole. This is indicated in Fig. 18.5 by the combinations of arrows and white dots.

The above description, based on a spatial picture of the material, is able to provide a qualitative explanation of the existence of electrons and holes in a semiconductor,[10] and completes the description given in Sect. 18.3, that focuses on

[9] Remembering that an equilibrium situation is considered here, the contributions to the electric current of these electrons cancel each other.

[10] A similar description could as well apply to a conductor. However, in such a case the barrier deformation is small and tunneling does not occur.

an energy picture. It also constitutes the basis for analyzing the case of a doped semiconductor, as shown in the following sections.

18.4.1 Donor-Type Impurities

As explained in Chap. 23, when a dopant atom is introduced into the semiconductor lattice, in order to properly act as a dopant it must replace an atom of the semiconductor, namely, it must occupy a lattice position (*substitutional impurity*). The concentration of the dopant atoms that are introduced into a semiconductor is smaller by orders of magnitude than the concentration of the semiconductor atoms themselves. As a consequence, the average distance between dopant atoms within the lattice is much larger than that between the semiconductor atoms. Due to this, when the band structure is calculated, the modification in the potential energy introduced by the dopant atoms can be considered as a perturbation with respect to the periodic potential energy of the lattice; the resulting band structure is therefore the superposition of that of the intrinsic semiconductor and of a set of additional states, whose characteristics will be described later. For the moment being, the spatial description is considered, still using the two-dimensional picture, where it is assumed that phosphorus is used as dopant material (Fig. 18.6).

As the dopant concentration is small with respect to the semiconductor's, each phosphorus atom is surrounded by silicon atoms. Phosphorus has five electrons in the external shell: thus, it forms four covalent bonds with silicon, while the

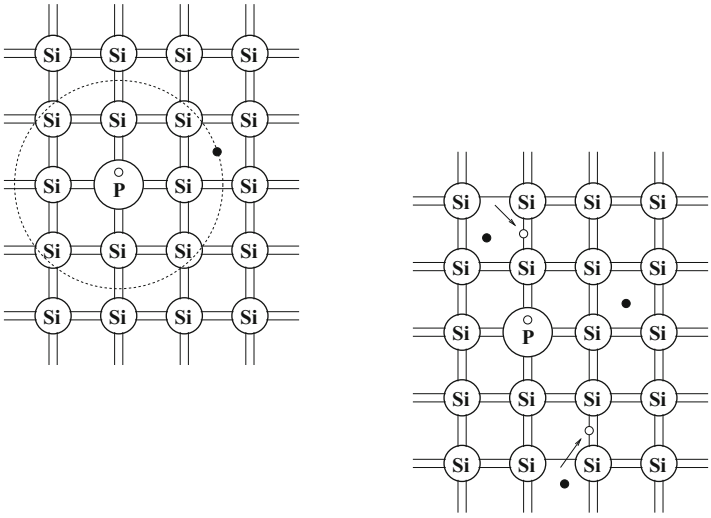


Fig. 18.6 Two-dimensional representation of the *n*-doped silicon lattice. The upper-left part of the figure shows the $T \rightarrow 0$ limit

remaining electron does not participate in any bond. In the $T \to 0$ limit, the electrons are permanently bound to the atoms as in the intrinsic case; this situation is drawn in the upper-left part of the figure, where the electron that does not form any bond is represented, in a particle-like picture, as orbiting around the phosphorus atom. The orbit's radius is relatively large because the binding force is weak.[11] The white dot inside the phosphorus atom indicates the positive nuclear charge that balances the orbiting electron. If, instead, the temperature is brought to some finite value $T > 0$ by transferring energy from an external reservoir to the semiconductor, a fraction of the orbiting electrons breaks the bond and become free to move within the material. At the same time, electrons belonging to shared pairs may also break their bonds, like in intrinsic silicon. This situation is depicted in the lower-right part of Fig. 18.6, where the free electrons are represented with black dots.

The deformation in the shape of the potential energy in the vicinity of a phosphorus nucleus is different from that near a silicon nucleus. In the latter case, the phenomenon is the same as in the intrinsic material: a series of tunneling events takes place, leading to the motion of holes from one site to a neighboring one. This is still represented by the combinations of arrows and white dots in Fig. 18.6. In the case of phosphorus, instead, the barrier deformation is small and tunneling does not occur: thus, the phosphorus atoms provide free electrons to the lattice, but no holes.

It is intuitive that the insertion of a prescribed amount of n-type dopants into the semiconductor lattice provides a method for controlling the number of free electrons and, through them, the material's conductivity. Moreover, if the ionization of the dopant atoms is not influenced by temperature, at least in a range of temperatures of practical interest, and the number of electrons made available by the dopant is dominant with respect to that provided by the intrinsic semiconductor, the drawback mentioned in Sect. 18.2 is eliminated: conductivity becomes temperature independent and its value is controlled by the fabrication process. This analysis is better specified below, starting from the consideration of the density of states in an n-type semiconductor, shown in Fig. 18.7.

Concentration of Ionized Impurities (Donor Type)

Atoms like that of phosphorus contribute electrons to the lattice; for this reason they are also called *donors*. The concentration of donors is indicated with N_D and, in this section, is assumed to be uniform in space. When donor atoms are present, the equilibrium concentrations of electrons and holes are different from those of the intrinsic case; they are indicated with n and p, respectively, and are termed *extrinsic concentrations*. Their derivation is identical to that of the intrinsic case and yields

$$n = N_C \, \Phi_{1/2}(\xi_e) \,, \qquad p = N_V \, \Phi_{1/2}(\xi_h) \,, \tag{18.17}$$

[11] In a more precise, quantum-mechanical description the electron is described as a stationary wave function extending over several lattice cells.

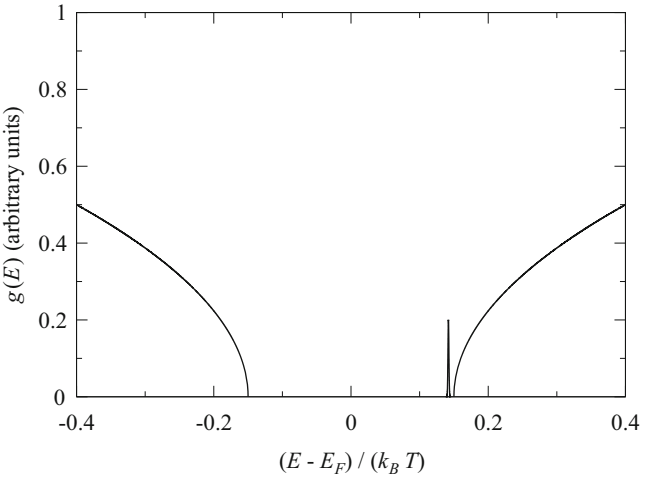


Fig. 18.7 Density of states in an *n*-doped semiconductor. The gap's extension is arbitrary and does not refer to any specific material

with N_C, N_V given by (18.4,18.8), respectively, and

$$\xi_e = -\frac{\zeta_e}{k_B\,T} = -\frac{E_C - E_F}{k_B\,T}\,, \qquad \xi_h = -\frac{\zeta_h}{k_B\,T} = -\frac{E_F - E_V}{k_B\,T}\,. \tag{18.18}$$

The above expression of ξ_e, ξ_h is similar to (18.2,18.6), the only difference being that the intrinsic Fermi level E_{Fi} is replaced here with the *extrinsic Fermi level* E_F.

In addition to the electrons of the conduction band and holes of the valence band, a third population of particles must be accounted for, namely, that associated with the dopants. Let $N_D^+ \leq N_D$ be the spatially constant concentration of the ionized donors.[12] The difference $n_D = N_D - N_D^+$ is the concentration of the donor atoms that have not released the orbiting electron to the lattice. Equivalently, n_D may be thought of as the concentration of orbiting electrons that have not been released; such a concentration is given by

$$n_D = \int_{\Delta E_D} \gamma_D(E)\, P_D(E)\, \mathrm{d}E\,, \tag{18.19}$$

where $\gamma_D(E)$ is the combined density of states produced by the donor atoms, ΔE_D the energy range where $\gamma_D(E) \neq 0$, and $P_D(E)$ the occupation probability of such

[12] A phosphorous atom that has not released the orbiting electron is electrically neutral; when the orbiting electron breaks the bond and becomes free to move within the lattice, the atom becomes positively ionized. The symbol N_D^+ reminds one of that. A second ionization does not occur because the energy necessary for it is too high.

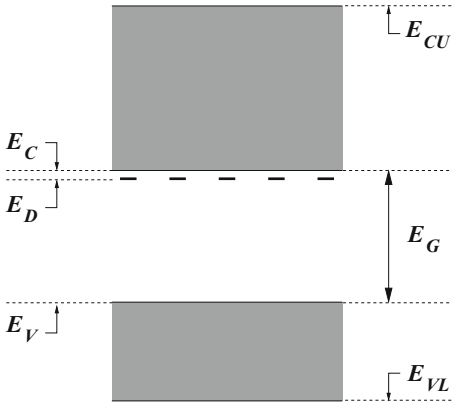


Fig. 18.8 Schematic representation of the donor states

states. The form of γ_D depends on the concentration N_D; for low or moderately high values of the donor concentration, γ_D has the form shown in Fig. 18.7. Such a density of states can be approximated by a Dirac delta centered at an energy E_D, called *donor level*, positioned within the gap at a small distance from the edge E_C of the conduction band:

$$\gamma_D(E) \simeq N_D \, \delta(E - E_D) \, . \tag{18.20}$$

The coefficient in (18.20) is such that (18.19) yields $n_D = N_D$ when $P_D = 1$. The distance of the donor level from the minimum of the conduction band is calculated in Sect. 18.7.3; for the typical n-type dopants used with silicon it is $E_C - E_D \simeq 0.033$ eV. Another important feature of the donor levels (still in the case of low or moderately high values of the impurity concentration) is that they are localized in space, as schematically illustrated in Fig. 18.8. For this reason, the Fermi-Dirac statistics describing the equilibrium distribution of the electrons within the donor states is slightly different from that used for the band electrons and reads

$$P_D(E) = \frac{1}{(1/d_D) \, \exp[(E - E_F)/(k_B \, T)] + 1} \, , \tag{18.21}$$

with d_D the *donors' degeneracy coefficient* ([33, Sect. 3.4] and Prob. 18.1). In silicon, germanium, and gallium arsenide it is $d_D = 2$ [128, Sect. 2-4]. Combining (18.19) with (18.20) and (18.21) yields $n_D = N_D \, P_D(E_D)$, namely,

$$N_D^+ = N_D \, [1 - P_D(E_D)] = \frac{N_D}{d_D \, \exp[(E_F - E_D)/(k_B \, T)] + 1} \, . \tag{18.22}$$

Like in the intrinsic case discussed in Sect. 18.3, expressions (18.17,18.22) can be used only if the position of the Fermi level is known. To proceed one considers again the $T \to 0$ limit, where the Fermi-Dirac statistics becomes discontinuous and the experimental evidence shows that the conductivity of a doped semiconductor

vanishes: this corresponds to a situation where all states of the conduction band are empty, those of the valence bands are filled, and the donor states are filled as well (in other terms, no dopant atoms are ionized). In conclusion, when $T \to 0$ the Fermi level of an n-doped semiconductor is positioned between E_D and E_C, so that $n = p = N_D^+ = 0$. If, starting from this situation, the temperature is brought again to some finite value $T > 0$, such that some of the valence-band electrons and some of the electrons belonging to the dopant atoms transit to the conduction band, the total number of holes and ionized donors thus formed is equal to that of the transited electrons. Due to the spatial uniformity of the material, the same relation holds among the concentrations, so that

$$n = p + N_D^+ \,, \tag{18.23}$$

whose limit for $N_D \to 0$ yields (18.11) as should be. Now, inserting (18.17,18.22) into (18.23) after letting $\xi_D = (E_D - E_C)/(k_B\,T) < 0$ provides

$$N_C\,\Phi_{1/2}(\xi_e) = N_V\,\Phi_{1/2}(\xi_h) + \frac{N_D}{d_D\,\exp(\xi_e - \xi_D) + 1} \,. \tag{18.24}$$

The latter, along with the relation $\xi_e + \xi_h = -E_G/(k_B\,T)$, forms a system in the two unknowns ξ_e, ξ_h, whose solution determines the position of E_F with respect to the band edges E_C and E_V at a given temperature $T > 0$ and donor concentration N_D. It is easily found that for a fixed temperature, the argument $\xi_e = (E_F - E_C)/(k_B\,T)$ increases when N_D increases. In fact, using the short-hand notation $f(\xi_e) = d_D\,\exp(\xi_e - \xi_D) + 1 > 1$ one finds from (18.24)

$$\frac{\mathrm{d}N_D}{\mathrm{d}\xi_e} = N_D\,\frac{f-1}{f} + f\left[N_C\,\frac{\mathrm{d}\Phi_{1/2}(\xi_e)}{\mathrm{d}\xi_e} - N_V\,\frac{\mathrm{d}\Phi_{1/2}(\xi_h)}{\mathrm{d}\xi_h}\,\frac{\mathrm{d}\xi_h}{\mathrm{d}\xi_e}\right] , \tag{18.25}$$

where the right-hand side is positive because $\mathrm{d}\xi_h/\mathrm{d}\xi_e = -1$ and $\Phi_{1/2}$ is a monotonically increasing function of the argument (Sect. C.13). In the intrinsic case $N_D = 0$ it is $E_F = E_{Fi}$ and $\xi_e < 0$; as N_D increases, the Fermi level moves towards E_C, thus making ξ_e less and less negative. At extremely high concentrations of the donor atoms, the Fermi level may reach E_C and even enter the conduction band, thus making ξ_e positive. As n, p, and N_D^+ are nonnegative, it is intuitive that an increasing concentration of donor atoms makes n larger than p; for this reason, in an n-doped semiconductor electrons are called *majority carriers* while holes are called *minority carriers*.[13]

The behavior of the Fermi level with respect to variations in temperature, at a fixed dopant concentration, cannot be discussed as easily as that with respect to the dopant variations, because (18.24) is not expressible in the form $T = T(E_F)$

[13] The electrons of the conduction band and the holes of the valence band are collectively indicated as *carriers*.

or its inverse. The analytical approach is made easier when some approximations are introduced into (18.24), as shown below. From a qualitative standpoint, one may observe that at very high temperatures the concentration of electron-hole pairs generated by the semiconductor prevails over the concentration of electrons provided by the dopant atoms, so that position of the Fermi level must be close to the intrinsic one; in contrast, when $T \to 0$ the Fermi level is positioned between E_D and E_C as remarked above. In conclusion, for an n-doped semiconductor one expects that $\mathrm{d}E_F/\mathrm{d}T < 0$.

In view of further elaborations it is convenient to associate with the Fermi level an electric potential φ_F, called *Fermi potential*, defined by

$$q\,\varphi_F = E_{Fi} - E_F\,, \tag{18.26}$$

with $q > 0$ the elementary charge. In an n-type semiconductor it is $E_F > E_{Fi}$; thus, the Fermi potential is negative.

Complete Ionization and Nondegenerate Condition (Donor Type)

When the Fermi level's position and temperature are such that

$$\xi_e - \xi_D < -1\,, \qquad E_F < E_D - k_B\,T\,, \tag{18.27}$$

the exponential term at the right-hand side of (18.22) is negligible with respect to unity, so that $N_D^+ \simeq N_D$. This condition is indicated with *complete ionization*. Note that, since $\xi_D < 0$, the complete-ionization condition in an n-doped semiconductor also implies $\xi_e < -1$; as a consequence, the approximation $\Phi_{1/2}(\xi_e) \simeq \exp(\xi_e)$ for the Fermi integral holds (Sect. C.13). Moreover, it is $\xi_h < \xi_e < -1$ because the Fermi level belongs to the upper half of the gap; thus, the same approximation applies to $\Phi_{1/2}(\xi_h)$ as well. When both equilibrium concentrations n and p are expressible through the approximation $\Phi_{1/2}(\xi) \simeq \exp(\xi)$, the semiconductor is called *nondegenerate* and the balance relation (18.24) reduces to

$$N_C\,\exp\left(-\frac{E_C - E_F}{k_B\,T}\right) \simeq N_V\,\exp\left(-\frac{E_F - E_V}{k_B\,T}\right) + N_D\,. \tag{18.28}$$

From the exponential expressions of n and p it follows, using (18.13),

$$n\,p \simeq N_C\,N_V\,\exp\left(-\frac{E_C - E_V}{k_B\,T}\right) = N_C\,N_V\,\exp\left(-\frac{E_G}{k_B\,T}\right) = n_i^2\,. \tag{18.29}$$

Letting $N_D^+ = N_D$ in (18.23) and multiplying both sides of it by n provides, thanks to (18.29), an easy algebraic derivation of the electron and hole concentrations in the nondegenerate and spatially uniform case:

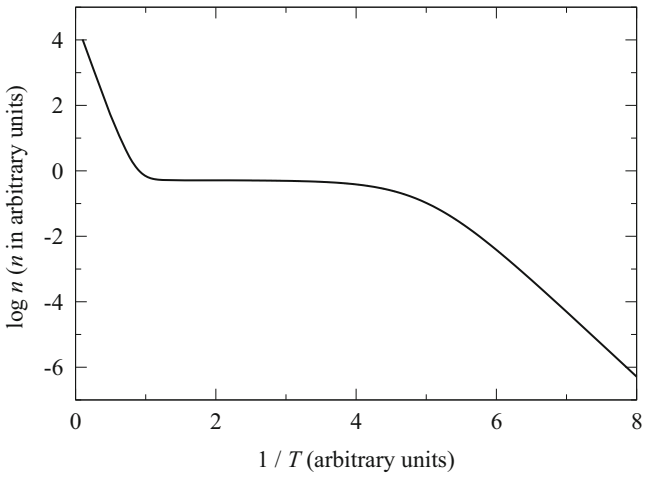


Fig. 18.9 Arrhenius plot of $n(T)$ for an n-type semiconductor, in arbitrary units

$$n^2 - N_D\, n - n_i^2 = 0\,, \qquad n = \frac{N_D}{2} + \sqrt{\frac{N_D^2}{4} + n_i^2}\,, \qquad p = \frac{n_i^2}{n}\,. \tag{18.30}$$

The range of validity of (18.30) is quite vast; in silicon at room temperature the approximation (18.28) holds up to $N_D \simeq 10^{17}$ cm^{-3}; also, even for the lowest doping concentrations (about 10^{13} cm^{-3}) it is $N_D \gg n_i$, so that (18.30) can further be approximated as $n \simeq N_D$, $p \simeq n_i^2/N_D$. Thus, the concentration of majority carriers turns out to be independent of temperature, while that of minority carriers still depends on temperature through n_i. Assuming by way of example $N_D = 10^{15}$ cm^{-3} and taking a temperature such that $n_i = 10^{10}$ cm^{-3} (compare with Table 18.2), one finds $n \simeq 10^{15}$ cm^{-3}, $p \simeq 10^5$ cm^{-3}. This result is very important because it demonstrates that the concentration of minority carriers is negligible; as anticipated above, the inclusion of a suitable concentration of dopants makes the material's conductivity independent of temperature. The Arrhenius plot of $n(T)$ for an n-type semiconductor is shown in Fig. 18.9 in arbitrary units. The left part of the curve, called *intrinsic range*, corresponds to the situation where the intrinsic concentration prevails due to the high temperature; the plateau, called *saturation range*, corresponds to the situation where $n \simeq N_D$; finally, the right part of the curve, called *freeze-out range*, corresponds to the case where $n \simeq N_D^+ < N_D$, with N_D^+ decreasing as temperature decreases. From the practical standpoint, the important outcome is that the saturation region covers the range of temperatures within which the integrated circuits operate.[14]

[14]For instance, for silicon with $N_D = 10^{15}$ cm^{-3} the saturation region ranges from $T_{\min} \simeq 125$ K to $T_{\max} \simeq 370$ K [128, Sect. 2–4].

Coupling $n = N_D$ with the expression of n given by the left-hand side of (18.28) and considering the values of N_C given in Table 18.2 yield

$$E_F = E_C - k_B\,T\,\log\left(\frac{N_C}{N_D}\right) < E_C\,, \tag{18.31}$$

whence, as anticipated,

$$\frac{\mathrm{d}E_F}{\mathrm{d}N_D} = \frac{k_B\,T}{N_D} > 0\,, \qquad \frac{\mathrm{d}E_F}{\mathrm{d}T} = k_B\,\log\left(\frac{N_D}{N_C}\right) = -\frac{E_C - E_F}{T} < 0\,. \tag{18.32}$$

From definition (18.26) of the Fermi potential it follows $E_C - E_F = E_C - E_{Fi} + q\,\varphi_F$; replacing the latter in $n = N_C\,\exp[-(E_C - E_F)/(k_B\,T)]$ yields an alternative expression of the equilibrium concentration and of the Fermi potential itself,

$$n = n_i\,\exp\left(-\frac{q\,\varphi_F}{k_B\,T}\right)\,, \qquad \varphi_F = -\frac{k_B\,T}{q}\,\log\left(\frac{N_D}{n_i}\right) < 0\,. \tag{18.33}$$

18.4.2 Acceptor-Type Impurities

The analysis carried out in Sect. 18.4.1 is repeated here in a shorter form with reference to a substitutional impurity of the acceptor type like, e.g., boron (Fig. 18.10).

As the dopant concentration is small with respect to the semiconductor's, each boron atom is surrounded by silicon atoms. Boron has three electrons in the external shell: while these electrons form covalent bonds with silicon, the remaining unsaturated bond deforms the shape of the potential energy in the vicinity of the boron atom. This attracts an electron from a shared pair of a neighboring silicon. In other terms, to form four covalent bonds with silicon, boron generates an electron-hole pair as shown in the figure. In the $T \to 0$ limit, the holes are permanently bound to the atoms as in the intrinsic case; this situation is drawn in the upper-left part of the figure, where the hole is represented, in a particle-like picture, as orbiting around the boron atom. The orbit's radius is relatively large because the binding force is weak. The black dot inside the boron atom indicates the negative charge that balances the orbiting hole. If, instead, the temperature is brought to some finite value $T > 0$ by transferring energy from an external reservoir to the semiconductor, a fraction of the orbiting holes break the bond and become free to move within the material. At the same time, electrons belonging to shared pairs may also break their bonds, like in intrinsic silicon. This situation is depicted in the lower-right part of Fig. 18.10, where the free holes are represented with white dots. The negative charge within the boron atom remains trapped within the atom itself: thus, the boron atoms provide free holes to the lattice, but no electrons.

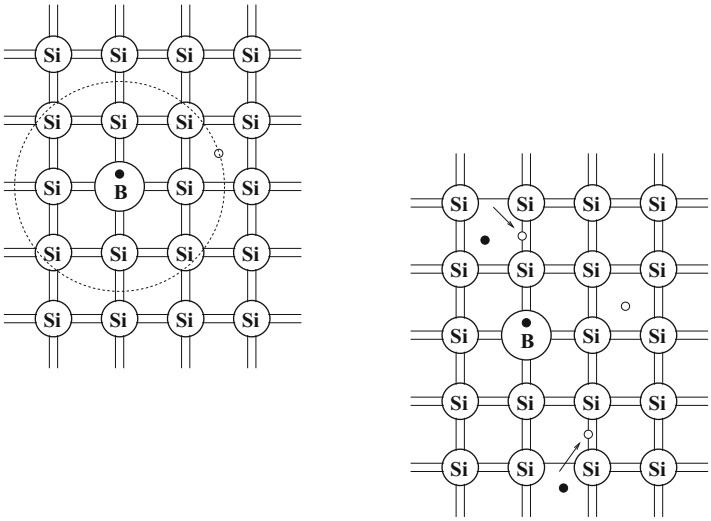


Fig. 18.10 Two-dimensional representation of the p-doped silicon lattice. The upper-left part of the figure shows the $T \to 0$ limit

The analysis is better specified below, starting from the consideration of the density of states in a p-type semiconductor, shown in Fig. 18.11.

Concentration of Ionized Impurities (Acceptor Type)

Atoms like that of boron trap electrons from the lattice; for this reason they are also called *acceptors*. The concentration of acceptors is indicated with N_A and, in this section, is assumed to be uniform in space. When acceptor atoms are present, the equilibrium concentrations of electrons and holes are different from those of the intrinsic semiconductor; their derivation is identical to that of the intrinsic or n-type case and yields again (18.17), with N_C, N_V given by (18.4,18.8), respectively, and ξ_e, ξ_h given by (18.18).

In addition to the electron of the conduction band and holes of the valence band, a third population of particles must be accounted for, namely, that associated with the dopants. Let $N_A^- \leq N_A$ be the spatially constant concentration of the acceptors that have released the orbiting hole to the lattice.[15] Equivalently, N_A^- may be thought of as the concentration of electrons that have been captured by the acceptor atoms;

[15] A boron atom that has not released the orbiting hole is electrically neutral; when the orbiting hole breaks the bond and becomes free to move within the lattice, the atom becomes negatively ionized, because the release of a hole is actually the capture of a valence-band electron by the boron atom. The symbol N_A^- reminds one of that. The release of a second hole does not occur.

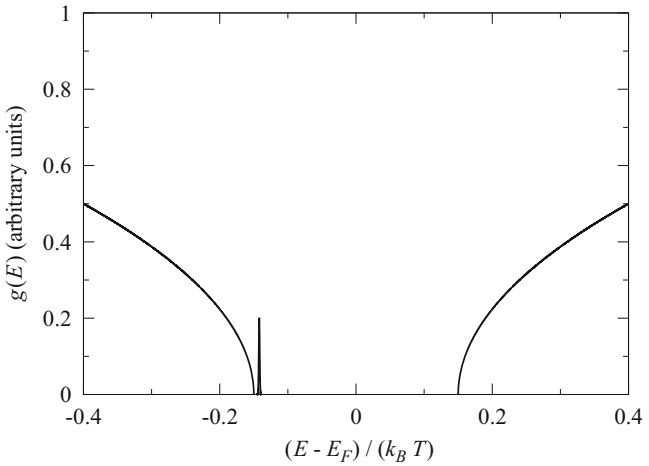


Fig. 18.11 Density of states in a *p*-doped semiconductor. The gap's extension is arbitrary and does not refer to any specific material

$$N_A^- = \int_{\Delta E_A} \gamma_A(E)\, P_A(E)\, \mathrm{d}E\,, \tag{18.34}$$

where $\gamma_A(E)$ is the combined density of states produced by the acceptor atoms, ΔE_A the energy range where $\gamma_A(E) \neq 0$, and $P_A(E)$ the occupation probability of such states. The form of γ_A depends on the concentration N_A; for low or moderately high values of the acceptor concentration, γ_A has the form shown in Fig. 18.11. Such a density of states can be approximated by a Dirac delta centered at an energy E_A, called *acceptor level*, positioned within the gap at a small distance from the edge E_V of the conduction band:

$$\gamma_A(E) \simeq N_A\, \delta(E - E_A)\,. \tag{18.35}$$

The coefficient in (18.35) is such that (18.34) yields $N_A^- = N_A$ when $P_A = 1$. The distance of the acceptor level from the maximum of the valence band is calculated in Sect. 18.7.3; for the typical *p*-type dopants used with silicon it is $E_A - E_V \simeq 0.05$ eV for the heavy holes and $E_A - E_V \simeq 0.016$ eV for the light holes. Another important feature of the acceptor levels (still in the case of low or moderately high values of the impurity concentration) is that they are localized in space, as schematically illustrated in Fig. 18.12. For this reason, the Fermi-Dirac statistics describing the equilibrium distribution of the electrons within the acceptor states is slightly different from that used for the band electrons and reads

$$P_A(E) = \frac{1}{(1/d_A)\, \exp[(E - E_F)/(k_B\, T)] + 1}\,, \tag{18.36}$$

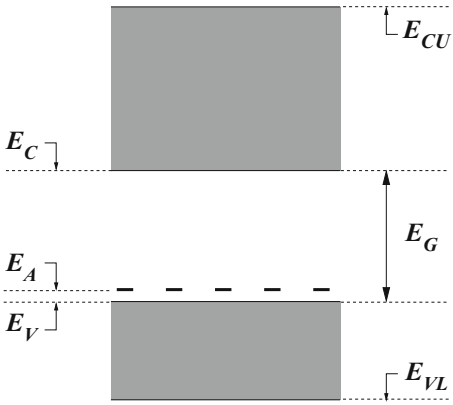


Fig. 18.12 Schematic representation of the acceptor states

with d_A the *acceptors' degeneracy coefficient*. In silicon, germanium, and gallium arsenide it is $d_A = 4$ [128, Sect. 2–4]. Combining (18.34) with (18.35) and (18.36) yields

$$N_A^- = \frac{N_A}{(1/d_A)\,\exp[(E_D - E_F)/(k_B\,T)] + 1}\,. \tag{18.37}$$

The position of the Fermi level is calculated by the same token as for the n-type dopant and is based on the balance relation

$$n + N_A^- = p\,, \tag{18.38}$$

whose limit for $N_A \to 0$ yields (18.11) as should be. Now, inserting (18.17,18.37) into (18.38) after letting $\xi_A = (E_V - E_A)/(k_B\,T) < 0$, provides

$$N_C\,\Phi_{1/2}(\xi_e) + \frac{N_A}{(1/d_A)\,\exp(\xi_h - \xi_A) + 1} = N_V\,\Phi_{1/2}(\xi_h)\,. \tag{18.39}$$

The latter, along with the relation $\xi_e + \xi_h = -E_G/(k_B\,T)$, forms a system in the two unknowns ξ_e, ξ_h, whose solution determines the position of E_F with respect to the band edges E_C and E_V at a given temperature $T > 0$ and acceptor concentration N_A. It is easily found that for a fixed temperature, the argument $\xi_h = (E_V - E_F)/(k_B\,T)$ increases when N_A increases. The demonstration is identical to that leading to (18.25). In the intrinsic case $N_A = 0$ it is $E_F = E_{Fi}$ and $\xi_h < 0$; as N_A increases, the Fermi level moves towards E_V, thus making ξ_h less and less negative. At extremely high concentrations of the acceptor atoms, the Fermi level may reach E_V and even enter the valence band, thus making ξ_h positive. As n, p, and N_A^- are nonnegative, it is intuitive that an increasing concentration of acceptor atoms makes p larger than n; for this reason, in a p-doped semiconductor holes are called *majority carriers* while electrons are called *minority carriers*.

To discuss from a qualitative standpoint the dependence of the Fermi level with respect to variations in temperature one may observe that, like in an n-doped material, at very high temperatures the concentration of electron-hole pairs generated by the semiconductor prevails over the concentration of holes provided by the dopant atoms, so that position of the Fermi level must be close to the intrinsic one; in contrast, when $T \to 0$ the Fermi level is positioned between E_V and E_A. In conclusion, for a p-doped semiconductor one expects that $\mathrm{d}E_F/\mathrm{d}T > 0$. Also, in a p-type semiconductor it is $E_F < E_{Fi}$; thus, the Fermi potential (18.26) is positive.

Complete Ionization and Nondegenerate Condition (Acceptor Type)

When the Fermi level's position and temperature are such that

$$\xi_h - \xi_A < -1\,, \qquad E_A < E_F - k_B\,T\,, \tag{18.40}$$

the exponential term at the right-hand side of (18.37) is negligible with respect to unity, so that $N_A^- \simeq N_A$. This condition is indicated with *complete ionization*. Note that, since $\xi_A < 0$, the complete-ionization condition in a p-doped semiconductor also implies $\xi_h < -1$; as a consequence, the approximation $\Phi_{1/2}(\xi_h) \simeq \exp(\xi_h)$ for the Fermi integral holds (Sect. C.13). Moreover, it is $\xi_e < \xi_h < -1$ because the Fermi level belongs to the lower half of the gap; thus, the same approximation applies to $\Phi_{1/2}(\xi_e)$ as well. In conclusion, the semiconductor is nondegenerate and the balance relation (18.39) reduces to

$$N_C\,\exp\left(-\frac{E_C - E_F}{k_B\,T}\right) + N_A \simeq N_V\,\exp\left(-\frac{E_F - E_V}{k_B\,T}\right)\,. \tag{18.41}$$

From the exponential expressions of n and p it follows, like in the n-type case, that the product of the equilibrium concentrations fulfills (18.29). The algebraic derivation of the electron and hole concentrations in the nondegenerate and spatially uniform case is also similar:

$$p^2 - N_A\,p - n_i^2 = 0\,, \qquad p = \frac{N_A}{2} + \sqrt{\frac{N_A^2}{4} + n_i^2}\,, \qquad n = \frac{n_i^2}{p}\,, \tag{18.42}$$

with the further approximation $p \simeq N_A$, $n \simeq n_i^2/N_A$ holding within the low and moderately high range of dopant concentrations, and around room temperature.

Coupling $p = N_A$ with the expression of p given by the left-hand side of (18.41) and considering the values of N_V given in Table 18.2 yield

$$E_F = E_V + k_B\,T\,\log\left(\frac{N_V}{N_A}\right) > E_V\,, \tag{18.43}$$

whence, as anticipated,

$$\frac{\mathrm{d}E_F}{\mathrm{d}N_A} = -\frac{k_B\,T}{N_A} < 0\,, \qquad \frac{\mathrm{d}E_F}{\mathrm{d}T} = k_B\,\log\left(\frac{N_V}{N_A}\right) = \frac{E_F - E_V}{T} > 0\,. \tag{18.44}$$

From definition (18.26) of the Fermi potential it follows $E_F - E_V = E_{Fi} - E_V - q\,\varphi_F$; replacing the latter in $p = N_V\,\exp[-(E_F - E_V)/(k_B\,T)]$ yields an alternative expression of the equilibrium concentration and of the Fermi potential itself,

$$p = n_i\,\exp\left(\frac{q\,\varphi_F}{k_B\,T}\right)\,, \qquad \varphi_F = \frac{k_B\,T}{q}\,\log\left(\frac{N_A}{n_i}\right) > 0\,. \tag{18.45}$$

18.4.3 Compensation Effect

The architecture of semiconductor devices is such that donor and acceptor dopants are present in the same region (Fig. 18.13). Letting N_D, N_A be the corresponding concentrations, still assumed uniform in space, the equilibrium concentrations of electrons and holes are expressed by (18.17) as in the other cases discussed above, with N_C, N_V and ξ_e, ξ_h given by (18.4,18.8) and (18.18), respectively. In turn, the concentrations of ionized donors and acceptors are given by (18.22) and (18.37), respectively. The balance equation reads

$$n + N_A^- = p + N_D^+\,, \tag{18.46}$$

namely, $N_C\,\Phi_{1/2}(\xi_e) + N_A\,P_A(E_A) = N_V\,\Phi_{1/2}(\xi_h) + N_D\,[1 - P_D(E_D)]$. One observes that if $N_D^+ = N_A^-$, the balance equation (18.46) coincides with that of an intrinsic semiconductor (compare with (18.11)). As a consequence, the position of the Fermi

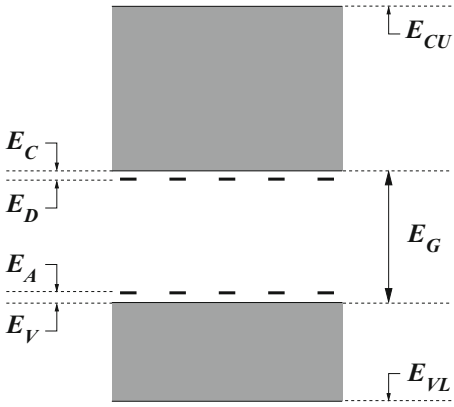


Fig. 18.13 Schematic representation of a semiconductor with both donor and acceptor states

level is given by (18.16), and the electrons released by the donor atoms are trapped by the acceptor ones. In this case the semiconductor is *fully compensated*. If, instead, it is $N_D^+ \neq N_A^-$, the semiconductor is *partially compensated*, and one must distinguish between two cases of the balance relation (18.46), depending on the sign of the *net ionized impurity concentration* $N = N_D^+ - N_A^-$. In the first case,

$$N > 0\,, \qquad n = p + N\,, \tag{18.47}$$

the balance relation is identical to that of an n-doped semiconductor (compare with (18.23)), with an effective donor-dopant concentration equal to N. In the second case,

$$N < 0\,, \qquad n + |N| = p\,, \tag{18.48}$$

the balance relation is identical to that of a p-doped semiconductor (compare with (18.38)), with an effective acceptor-dopant concentration equal to $|N|$. In the nondegenerate case, (18.29) still holds. If complete ionization also occurs, then $N = N_D - N_A$; when the donor-type dopant prevails, electrons are the majority carriers, and the same calculation as that leading to (18.30) yields

$$n = \frac{N}{2} + \sqrt{\frac{N^2}{4} + n_i^2}\,. \tag{18.49}$$

If $N \gg n_i$, then

$$n \simeq N\,, \qquad p \simeq \frac{n_i^2}{N}\,, \qquad \varphi_F = -\frac{k_B\,T}{q}\,\log\left(\frac{N}{n_i}\right) < 0\,. \tag{18.50}$$

If, on the contrary, the acceptor-type dopant prevails, holes are the majority carriers, and the same calculation as that leading to (18.42) yields

$$p = -\frac{N}{2} + \sqrt{\frac{N^2}{4} + n_i^2}\,. \tag{18.51}$$

If $-N \gg n_i$, then

$$p \simeq -N = |N|\,, \qquad n \simeq \frac{n_i^2}{|N|}\,, \qquad \varphi_F = \frac{k_B\,T}{q}\,\log\left(\frac{|N|}{n_i}\right) > 0\,. \tag{18.52}$$

18.5 Nonuniform Distribution of Dopants

This section deals with the more realistic case where the donor concentration N_D, the acceptor concentration N_A, or both, depend on position. A qualitative reasoning shows that the balance equations (18.23), (18.38), or (18.46) do not hold anymore. To show this, consider the simple case where only the n-type dopant is present, $N_D = N_D(\mathbf{r})$, $N_A = 0$, and select two nearby positions, say, $\mathbf{r}_1$ and $\mathbf{r}_2$, in the limiting case $T \to 0$; then, let the temperature increase such that a fraction of the donor atoms ionizes. If $N_D(\mathbf{r}_1) \neq N_D(\mathbf{r}_2)$, the numbers of electrons transiting to the conduction band at $\mathbf{r}_1$ and $\mathbf{r}_2$ is different, so that the concentration in one position, say, $\mathbf{r}_1$ is larger than that in the other; thus, some electrons diffuse[16] from the former position to the latter. On the other hand, as the position of the positive charges within the ionized donors is fixed, the spatial rearrangement of the conduction-band electrons unbalances the negative charges with respect to the positive ones, so that the local charge density differs from zero. The same reasoning applies when both donor and acceptor dopants are present, so that in a general nonuniform case it is

$$\varrho(\mathbf{r}) = q\left[p(\mathbf{r}) - n(\mathbf{r}) + N_D^+(\mathbf{r}) - N_A^-(\mathbf{r})\right] \neq 0\,. \tag{18.53}$$

A nonvanishing charge density produces in turn an electric field $\mathbf{E} = \mathbf{E}(\mathbf{r})$ whose action balances the diffusion. In conclusion, the equilibrium condition is kept by an exact balance of the two transport mechanisms. Considering that the equilibrium condition is time-independent, the Maxwell equations in the semiconductor reduce to $\varepsilon_{sc}\ \mathrm{div}\,\mathbf{E} = \varrho$, with $\mathbf{E} = -\,\mathrm{grad}\,\varphi$ (Sect. 4.4), φ and ε_{sc} being the electric potential and semiconductor permittivity, respectively.[17] In other terms, the electric field due to the nonuniformity of ϱ is found by solving the Poisson equation.

The effect onto the total energy of the electrons due to the presence of an electric field also influences the statistical distribution of the electrons in the energy states. It will be demonstrated in Sect. 19.2.2 that the statistical distribution to be used here is a modified form of the Fermi-Dirac statistics where E is replaced with $E - q\,\varphi(\mathbf{r})$:

$$P(E, \mathbf{r}) = \frac{1}{\exp\left[(E - q\,\varphi(\mathbf{r}) - E_F)\,/\,(k_B\,T)\right] + 1} \tag{18.54}$$

Similarly, (18.21) and (18.36) become

$$P_{D(A)}(E, \mathbf{r}) = \frac{1}{(1/d_{D(A)})\,\exp[(E - q\,\varphi(\mathbf{r}) - E_F)/(k_B\,T)] + 1}\,, \tag{18.55}$$

[16] Diffusive transport is introduced in Sect. 19.3.3.

[17] Note that the material's permittivity must be used here instead of vacuum's. This is coherent with the use of charge density and, in a nonequilibrium situation, of current density, which entail averages over volumes of space.

Note that the calculations leading to the concentrations are carried out in the same manner as in Sects. 18.4.1,18.4.2, because they involve integrals over energy only. As a consequence, (18.17) generalizes to

$$n(\mathbf{r}) = N_C\, \Phi_{1/2}\left(\xi_e(\mathbf{r})\right)\,, \qquad \xi_e(\mathbf{r}) = -\frac{\zeta_e(\mathbf{r})}{k_B\,T} = -\frac{E_C - E_F - q\,\varphi(\mathbf{r})}{k_B\,T}\,, \tag{18.56}$$

and

$$p(\mathbf{r}) = N_V\, \Phi_{1/2}\left(\xi_h(\mathbf{r})\right)\,, \qquad \xi_h(\mathbf{r}) = -\frac{\zeta_h(\mathbf{r})}{k_B\,T} = -\frac{E_F + q\,\varphi(\mathbf{r}) - E_V}{k_B\,T}\,. \tag{18.57}$$

Similarly, (18.22,18.34) become

$$N_D^+(\mathbf{r}) = N_D(\mathbf{r})\,\left[1 - P_D(E_D, \mathbf{r})\right]\,, \qquad N_A^-(\mathbf{r}) = N_A(\mathbf{r})\,P_A(E_A, \mathbf{r})\,. \tag{18.58}$$

The summands at the right-hand side of (18.53) depend on position through the electric potential φ and also through the explicit dependence of N_D and N_A; as a consequence, inserting the expression of the charge density into the right-hand side of the Poisson equation yields

$$-\,\varepsilon_{\rm sc}\,\nabla^2\varphi = \varrho(\varphi, \mathbf{r})\,, \tag{18.59}$$

namely, a second-order, partial differential equation in the unknown φ, with position-dependent coefficients. The equation must be supplemented with suitable boundary conditions.[18] Equation (18.59) is the generalization of the balance equation (18.46) to the nonuniform case: in the case of (18.46), the problem is algebraic and yields E_F, whereas in the case of (18.59) it is differential and yields $E_F + q\,\varphi$. After solving (18.59) one reconstructs n, p, N_D^+, and N_A^- at each point through (18.56), (18.57), and (18.58). In the nondegenerate case, (18.56) becomes

$$n = n^{(0)}\,\exp\left(\frac{q\,\varphi}{k_B\,T}\right)\,, \quad n^{(0)} = N_C\,\exp\left(\frac{E_F - E_C}{k_B\,T}\right) = n_i\,\exp\left(\frac{-q\varphi_F}{k_B\,T}\right)\,, \tag{18.60}$$

with $n^{(0)}$ the value of the electron concentration in the position(s) where $\varphi = 0$. The last expression of $n^{(0)}$ is obtained by combining definition (18.26) with the first relation in (18.12). Similarly,

$$p = p^{(0)}\,\exp\left(\frac{-q\,\varphi}{k_B\,T}\right)\,, \quad p^{(0)} = N_V\,\exp\left(\frac{E_V - E_F}{k_B\,T}\right) = n_i\,\exp\left(\frac{q\varphi_F}{k_B\,T}\right)\,. \tag{18.61}$$

[18] Such boundary conditions must be coherent with the choice of the zero point of the total energy. An example is given in Sect. 21.2.

From (18.60,18.61) one finally obtains

$$n = n_i \exp\left[\frac{q\ (\varphi - \varphi_F)}{k_B\, T}\right], \qquad p = n_i \exp\left[\frac{q\ (\varphi_F - \varphi)}{k_B\, T}\right]. \tag{18.62}$$

One observes that (18.29) still holds, namely, $np = n^{(0)} p^{(0)} = n_i^2$: in the nondegenerate case, the equilibrium product does not depend on position. If complete ionization also occurs, then $N_D^+(\mathbf{r}) = N_D(\mathbf{r})$, $N_A^-(\mathbf{r}) = N_A(\mathbf{r})$: the ionized-dopant concentrations do not depend on the electric potential.

18.6 Band-Gap Narrowing

When the dopant concentration is large, the density of states associated with the dopant atoms can no longer be described as in Figs. 18.7, 18.11. Rather, considering by way of example an *n*-type dopant, the form of the density of states is similar to that shown in Fig. 18.14, namely, it overlaps the lower portion of the conduction band forming the so-called *impurity band*. In addition, the dopant atoms are close to each other, to the extent that the probability of tunneling of an electron, from a neutral to a nearby, ionized donor atom, is not negligible (Sect. 18.7.2). In a nonequilibrium condition, the tunneling electrons give rise to a current density; the phenomenon is referred to as *impurity-band conduction*. From the practical standpoint, the union of the conduction and impurity bands is viewed as a broader

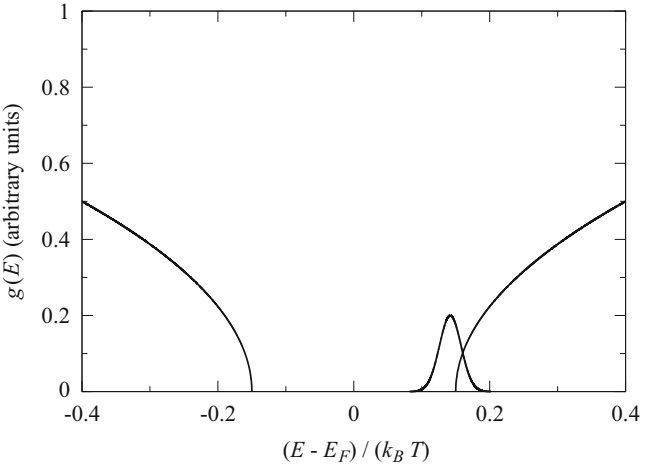


Fig. 18.14 Density of states in an *n*-doped semiconductor, where the high concentration of the dopant produces the band-gap narrowing. The gap's extension is arbitrary and does not refer to any specific material

conduction band whose lower edge is shifted with respect to the undoped, or moderately doped case; the effect is also called *band-gap narrowing*. The analysis is similar when a large concentration of acceptor atoms is present. In conclusion, band-gap narrowing is in general produced by the lowering of the conduction-band edge, $\Delta E_C(\mathbf{r}) > 0$, combined with the lifting of the valence-band edge, $\Delta E_V(\mathbf{r}) > 0$. Both quantities are position dependent because in general the dopant concentrations are such. Indicating with E_{Ci}, E_{Vi} the lower edge of the conduction band and, respectively, the upper edge of the valence band in the undoped or moderately doped case, and observing that the variations due to heavy doping are positive, one has for the actual positions of the band edges:

$$E_C(\mathbf{r}) = E_{Ci} - \Delta E_C(\mathbf{r})\,, \qquad E_V(\mathbf{r}) = E_{Vi} + \Delta E_V(\mathbf{r})\,, \tag{18.63}$$

whence

$$E_G(\mathbf{r}) = E_{Gi} - \Delta E_G(\mathbf{r})\,, \quad E_{Gi} = E_{Ci} - E_{Vi}\,, \quad \Delta E_G = \Delta E_C + \Delta E_V > 0\,. \tag{18.64}$$

To calculate the carrier concentrations when band-gap narrowing is present, one must replace E_C with $E_{Ci} - \Delta E_C(\mathbf{r})$ in the second relation of (18.56), and E_V with $E_{Vi} + \Delta E_V(\mathbf{r})$ in the second relation of (18.57), to find

$$\xi_e(\mathbf{r}) = -\frac{E_{Ci} - E_F - q\,\varphi(\mathbf{r})}{k_B\,T} + \frac{\Delta E_C(\mathbf{r})}{k_B\,T}\,, \tag{18.65}$$

$$\xi_h(\mathbf{r}) = -\frac{E_F - E_{Vi} + q\,\varphi(\mathbf{r})}{k_B\,T} + \frac{\Delta E_V(\mathbf{r})}{k_B\,T}\,. \tag{18.66}$$

Band-gap narrowing makes ξ_e, ξ_h to increase with respect to the moderately doped case; as a consequence, the equilibrium carrier concentrations n and p are larger as well. As mentioned in Sect. 18.3, band gap is measured by either electrical or optical methods. The results of gap's measurements, that show that E_G decreases when the dopant concentration exceeds some limiting value N_r, are usually rendered in compact form by means of interpolating expressions, an example of which is, for silicon [125], [126],

$$\Delta E_G = E_r \left(F + \sqrt{F^2 + 0.5}\right)\,, \qquad F = \log\left(\frac{N_D + N_A}{N_r}\right)\,, \tag{18.67}$$

with $E_r = 9$ meV, $N_r = 10^{17}$ cm^{-3}. The function described by (18.67) is shown in normalized form in Fig. 18.15. Expressions like (18.67) describe the cumulative effect of the total doping concentration, without distinguishing between the donor or acceptor contribution to band-gap narrowing. For this reason, when band-gap narrowing is accounted for in numerical calculations, ΔE_G is equally distributed between the two bands, namely, $\Delta E_C = \Delta E_V = \Delta E_G/2$.

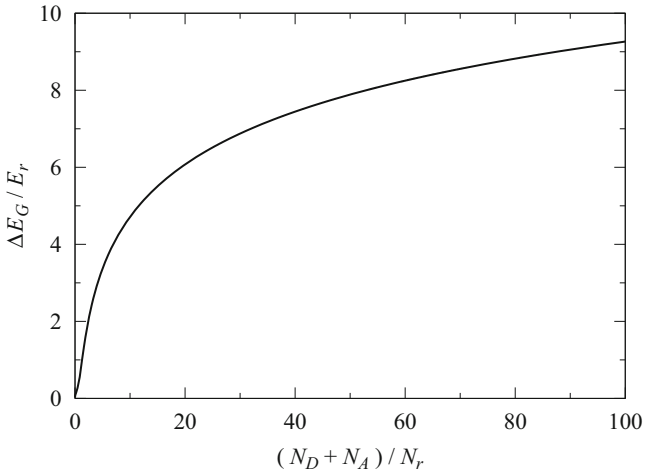


Fig. 18.15 Band-gap narrowing as a function of the total doping concentration, in normalized form, using the experimental expression (18.67)

It is interesting to note that the onset of band-gap narrowing corresponds to a total dopant concentration of about 10^{17} cm^{-3}, where the nondegeneracy condition still holds. As a consequence, a range of dopant concentrations exists where the exponential approximation can be used for the equilibrium carrier concentrations, whereas the band-gap narrowing effect must be accounted for.[19] The nondegeneracy condition reads in this case

$$E_{Ci} - E_F - q\,\varphi - \Delta E_C > k_B\,T\,, \qquad E_F - E_{Vi} + q\,\varphi - \Delta E_V > k_B\,T\,. \tag{18.68}$$

Remembering (18.62), the equilibrium concentrations become

$$n = n_e \exp\left[\frac{q\,(\varphi - \varphi_F)}{k_B\,T}\right], \qquad n_e = n_i \exp\left(\frac{\Delta E_C}{k_B\,T}\right) > n_i\,, \tag{18.69}$$

$$p = p_e \exp\left[\frac{q\,(\varphi_F - \varphi)}{k_B\,T}\right], \qquad p_e = n_i \exp\left(\frac{\Delta E_V}{k_B\,T}\right) > n_i\,, \tag{18.70}$$

where $p_e = n_e$ on account of $\Delta E_V = \Delta E_C$. The common value n_e is called *effective intrinsic concentration*. The equilibrium product then reads

$$n\,p = n_e^2\,, \qquad n_e^2 = n_i^2 \exp\left(\frac{\Delta E_G}{k_B\,T}\right). \tag{18.71}$$

[19] Such a range may be quite large if one considers the compensation effect (Sect. 18.4.3).

18.7 Complements

18.7.1 Si, Ge, GaAs in the Manufacturing of Integrated Circuits

As noted in Sect. 18.3, silicon, germanium, and gallium arsenide have similar effective densities of states, but different gap extensions (Table 18.1); in this respect, silicon is considered as a reference material, so that germanium is indicated as a *narrow-gap material* while gallium arsenide is indicated as a *wide-gap material*. The differences in the gap extension produce huge differences in the intrinsic concentration n_i (Table 18.2); the latter, in turn, has a strong influence on the functioning of the integrated circuits. In fact, the saturation current of a *p-n* junction is proportional to n_i^2 (Sect. 21.3.1); as a consequence, this parameter determines the current of the junction when a reverse bias is applied to it. When many junctions are present, as is typically the case in integrated circuits, the inverse currents may build up and give rise to a substantial parasitic current. From this standpoint, gallium arsenide is preferable with respect to silicon, which in turn is preferable with respect to germanium. Gallium arsenide is also preferable because of the smaller effective mass of the electrons, which makes the electron mobility larger (Sect. 17.6.6). On the other hand, silicon is much less expensive; in fact it is the second most abundant element in Earth's crust (the first one is oxygen); gallium, germanium, and arsenic are much rarer.

The historical development of the semiconductor-device manufacture has followed, instead, a different path. Until the mid-sixties of the last century, germanium was preferred; the reason for this was that a technological process, able to purify the material to the level required by the electronic industry, was available for germanium first. As soon as the purification method became available for silicon as well, the latter replaced germanium in the fabrication of semiconductor devices and, soon after, of integrated circuits. The silicon technology developed with a steady pace, giving rise to decades of exponential miniaturization and development in the integrated-circuit manufacture. The miniaturization of gallium-arsenide-based circuits did not proceed with the same pace because the technology of compound materials is more complicate.

In 1980, practically 100% of the worldwide market share of integrated circuit was silicon based, almost equally distributed between the bipolar and MOSFET technologies [21]. In the following years the MOSFET technology became dominant, reaching a market share of 88% in the year 2000; of the remaining 12%, the bipolar, silicon-based technology kept an 8% share, while the remaining 4% was taken by integrated circuits using III-V compounds.

In 1989, germanium was introduced again in the silicon integrated-circuit technology to form silicon-germanium alloys ($Si_{1-x}Ge_x$). The alloy makes a flexible band-gap tuning possible; it is used for manufacturing heterojunction bipolar transistors, yielding higher forward gain and lower reverse gain than traditional bipolar transistors. Another application of the alloy is in the Silicon-Germanium-On-

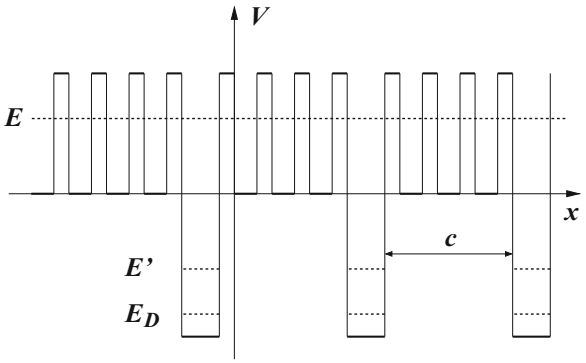


Fig. 18.16 Potential energy in the Kronig-Penney model modified to account for impurity atoms

Insulator (SGOI) technology. The difference in the lattice constants of germanium and silicon induces a strain in the material under the gate, that makes electron mobility to increase.

18.7.2 Qualitative Analysis of the Impurity Levels

A qualitative analysis of the impurity levels may be carried out basing on a modified version of the Kronig-Penney model discussed in Sect. 17.9.4. To this purpose, consider the case of donor atoms placed at equal distances in a one-dimensional lattice, like that of Fig. 18.16. The deeper wells are those introduced by the dopants, while the finer structure above the x axis is due to the semiconductor nuclei. Note that the relative distances in the figure are not realistic and are used only for illustrative purposes; assuming in fact that the structure represented a cross-section of a three-dimensional semiconductor, where a uniform donor concentration N_D is present, the distance between two neighboring impurity atoms is found to be $(1/N_D)^{1/3}$. If the semiconductor's concentration is N_{sc}, the ratio $(N_{\text{sc}}/N_D)^{1/3}$ indicates how many semiconductor atoms are present in the interval between two neighboring impurities. Considering silicon by way of example ($N_{\text{sc}} = 5 \times 10^{22}$ cm^{-3}) and taking $N_D = 5 \times 10^{16}$ cm^{-3} yield $(N_{\text{sc}}/N_D)^{1/3} = 100$.

With this provision, let c be the width of the barrier separating two dopant-induced wells, and consider a negative value E' of the electron's energy. If c is large, the probability that the electron tunnels from a well to an adjacent, empty one is negligibly small; in this case, the electron is confined within the well where the localization probability is the largest, and the energy states $E' < 0$ are similar to those of a single well (the energy states of the finite rectangular well are worked out in Sect. 11.5). The lowest state of the well is the ground state E_D shown in Fig. 18.16. If, instead, a positive energy state E is considered, the wave function is extended over the whole lattice like in the Kronig-Penney model. In summary, the addition of dopant atoms with a low or moderately high concentration, such that

their mutual distance c is large, provides a set of energy states that adds up to the band structure of the intrinsic semiconductor. The states introduced by the dopants are localized in space at the positions of the dopant atoms, and are distributed in energy as discrete levels whose mutual distances depend on the form of the well. In turn, the states with positive energy have a structure similar to that of the intrinsic semiconductor, because in this respect the dopant atoms have little effect on the lattice; in Fig. 18.16, the lower edge E_C of the conduction band coincides with $E = 0$. As said above, electrons belonging to the dopant atoms cannot move as long as they are confined within wells, because tunneling is precluded; on the other hand, they may be promoted to conduction-band states by absorbing energy in a collision, and become band electrons, that is, mobile. Conversely, a band electron may lose energy due to a collision, and be trapped in an empty well.

When, due to an increasing dopant concentration, the width c of the barrier becomes smaller, the transmission coefficient increases, and the electrons belonging to the wells have a non-negligible probability of moving from a well to another. When the structure of the wells becomes finer, the description of their energy states becomes more similar to that applicable to the intrinsic semiconductor: for an infinite structure one obtains a continuum of states that fill up the well and connect to those of the conduction band. This explains the band-gap narrowing phenomenon introduced in Sect. 18.6.

18.7.3 Position of the Impurity Levels

To determine the position of the impurity levels for the case of a low or moderately high impurity concentration of the donor type, one considers the dispersion relation $E(\mathbf{k})$ of the conduction band of the intrinsic semiconductor, in the parabolic-band approximation (17.57,17.58):

$$E(\mathbf{k}) \simeq E_C + \sum_{i=1}^{3} \frac{\hbar^2}{2\, m_{ia}} (k_i - k_{ia})^2 \,, \qquad \frac{1}{m_{ia}} = \frac{1}{\hbar^2} \left(\frac{\partial^2 E}{\partial k_i^2} \right)_a > 0 \,. \tag{18.72}$$

Now, assume that a donor-type impurity is added in the origin, and that the impurity is ionized. As a consequence, it gives rise to a hydrogenic-like potential energy[20] of the form $V = -q^2/(4\,\pi\,\varepsilon_{\rm sc}\,r)$. The effect of V may be considered as a local perturbation over the band structure, so that when the ionized impurity is present the total energy of an electron becomes $H = E(\mathbf{k}) + V(r)$, with $E(\mathbf{k})$ the same as in the unperturbed case. Shifting the origin of $\mathbf{k}$ to $\mathbf{k}_a$, and replacing k_i with $-\mathrm{i}\,\partial/\partial x_i$,

[20] Compare with Sect. 7.2. Like in Sect. 18.5, the semiconductor permittivity is used instead of that of vacuum, because the wave function of an electron subjected to the force due to V extends over several lattice cells.

one obtains the Hamiltonian operator[21]

$$\mathscr{H} \simeq E_C - \sum_{i=1}^{3} \frac{\hbar^2}{2\, m_{ia}} \frac{\partial^2}{\partial x_i^2} - \frac{q^2}{4\,\pi\,\varepsilon_{\rm sc}\, r}\,. \tag{18.73}$$

To estimate the eigenvalues of $\mathscr{H}$ one may replace each m_{ia} with the average $m_{ea} = (m_{1a}\, m_{2a}\, m_{3a})^{1/3}$ (Sect. 17.6.3), to obtain

$$-\frac{\hbar^2}{2\, m_{ea}} \nabla^2 v_n - \frac{q^2}{4\,\pi\,\varepsilon_{\rm sc}\, r}\, v_n = (E'_n - E_C)\, v_n\,, \tag{18.74}$$

with v_n the eigenfunction. Apart from the coefficients, (18.74) is identical to the Schrödinger equation for the Coulomb case treated in Sect. 13.5.2, whose eigenvalues are given by (13.56). It follows

$$E'_n = E_C - \frac{m_{ea}}{2\,\hbar^2} \left(\frac{q^2}{4\,\pi\,\varepsilon_{\rm sc}}\right)^2 \frac{1}{n^2}\,, \qquad n = 1, 2, \ldots \tag{18.75}$$

Thus, the donor impurity provides infinite levels from the minimum E'_1 to the maximum E_C. Considering the case of silicon by way of example, one has $m_{1a} = m_l = 0.97\, m_0$, $m_{2a} = m_{3a} = m_t = 0.19\, m_0$ (Table 17.4), with $m_0 \simeq 9.11 \times 10^{-31}$ kg the rest mass of the electron, whence $m_{ea} = 0.33\, m_0$. Letting $E_D = E'_1$ and using $\varepsilon_{\rm sc} \simeq 11.7\, \varepsilon_0$, with $\varepsilon_0 = 8.854 \times 10^{-14}$ F cm^{-1} the vacuum permittivity, one finds the ionization energy of the donor impurity in silicon:

$$E_C - E_D = \frac{m_{ea}}{2\,\hbar^2} \left(\frac{q^2}{4\,\pi\,\varepsilon_{\rm sc}}\right)^2 \simeq 32.8 \quad \text{meV}\,. \tag{18.76}$$

The analysis is similar for an ionized, acceptor-type impurity. The hydrogenic-like potential energy becomes $V = q^2/(4\,\pi\,\varepsilon_{\rm sc}\, r)$, and the dispersion relation around a maximum $\mathbf{k}_a$ of the valence band reads

$$E(\mathbf{k}) \simeq E_V - \sum_{i=1}^{3} \frac{\hbar^2}{2 m_{ia}} (k_i - k_{ia})^2\,, \qquad \frac{1}{m_{ia}} = -\frac{1}{\hbar^2} \left(\frac{\partial^2 E}{\partial k_i^2}\right)_a > 0\,. \tag{18.77}$$

When the impurity is present the Hamiltonian operator becomes

$$\mathscr{H} \simeq E_V + \sum_{i=1}^{3} \frac{\hbar^2}{2\, m_{ia}} \frac{\partial^2}{\partial x_i^2} + \frac{q^2}{4\,\pi\,\varepsilon_{\rm sc}\, r}\,. \tag{18.78}$$

[21] More comments about the procedure of obtaining an operator from a simplified form of the eigenvalues of a more general operator are made in Sect. 19.2.

Again, to estimate the eigenvalues one replaces each m_{ia} with the average m_{ha}. One finds

$$-\frac{\hbar^2}{2\,m_{ha}}\,\nabla^2 v_n - \frac{q^2}{4\,\pi\,\varepsilon_{\rm sc}\,r}\,v_n = (E_V - E_n'')\,v_n\,, \tag{18.79}$$

whence

$$E_n'' = E_V + \frac{m_{ha}}{2\,\hbar^2}\left(\frac{q^2}{4\,\pi\,\varepsilon_{\rm sc}}\right)^2 \frac{1}{n^2}\,, \tag{18.80}$$

$n = 1, 2, \ldots$. The acceptor impurity provides infinite levels from the minimum E_V to the maximum E_1''. Letting $E_A = E_1''$ one finds the ionization energy $E_A - E_V$ of the donor impurity. Taking again silicon by way of example, and using the values $m_{hh} = 0.5\,m_0$, $m_{hl} = 0.16\,m_0$ from Table 17.3, one finds $E_A - E_V = 49.7$ meV for the heavy holes and $E_A - E_V = 15.9$ meV for the light holes.

The ionization energies of phosphorus and boron *in vacuo* are about 10.5 eV and 8.3 eV, respectively, that is, much higher than those calculated here. The strong difference is ascribed to the presence of the silicon crystal: a comparison between (13.56) and (18.75) or (18.80) shows in fact that the coefficients of $1/n^2$ in the crystal case are much smaller than that *in vacuo*, due to the presence of the effective mass in the numerator and of the square of the material permittivity in the denominator. The small distance between the ground state of the impurity atoms and the edge of the band explains the ease with which the dopants ionize at room temperature.

Problem

18.1 Prove (18.21) by combinatorial calculus.

Part VI
Transport Phenomena in Semiconductors

Chapter 19
Mathematical Model of Semiconductor Devices

19.1 Introduction

This chapter describes the reasoning that leads from the single-particle Schrödinger equation for an electron in a crystal to the mathematical model of semiconductor devices. The latter is a set of equations describing the evolution in space and time of a number of average quantities of interest: with reference to the electrons of the conduction band or holes of the valence band, such quantities are the concentration, average velocity, current density, average kinetic energy, and so on. The model of semiconductor devices has different levels of complexity depending on the trade-off between the information that one needs to acquire about the physical behavior of the device under investigation and the computational cost of the system of differential equations to be solved. In fact, the possible models are hierarchically ordered from the drift-diffusion model, which is the simplest one, to the hydrodynamic model, and so on. In essence, these models are different approaches to the problem of solving, in a more or less simplified form, the Boltzmann Transport Equation. Those described in this chapter are the most widely adopted in the commercial simulation programs used by semiconductor companies. Other important methods, that are not addressed in this book, are the Monte Carlo method and the spherical-harmonics expansion.

The steps leading to the mathematical model of semiconductor devices start with a form of the single-particle Schrödinger equation based on the equivalent Hamiltonian operator, where it is assumed that the external potential energy is a small perturbation superimposed to the periodic potential energy of the nuclei; this leads to a description of the collisionless electron's dynamics in terms of canonically conjugate variables, that are the expectation values of the wave packet's position and momentum. The dynamics of the Hamiltonian type makes it possible to introduce the statistical description of a many-electron system, leading to the semiclassical Boltzmann Transport Equation. After working out the collision operator, the perturbative approximation is considered; the simplified form of the transport equation thus found is tackled by means of the moments method, whence the hydrodynamic and drift-diffusion versions of the model are derived. A detailed analysis of the

M. Rudan, *Physics of Semiconductor Devices*, DOI 10.1007/978-3-319-63154-7_19

derivation of the electron and hole mobility in the parabolic-band approximation is provided. Then, the semiconductor model is coupled with Maxwell's equation, and the applicability of the quasi-static approximation is discussed. The typical boundary conditions used in the analysis of semiconductor devices are shown, and an example of analytical solution of the one-dimensional Poisson equation is given.

The complements discuss the analogy between the equivalent Hamiltonian operator and the corresponding Hamiltonian function introduced in an earlier chapter, provide a detailed description of the closure conditions of the models, and illustrate the Matthiessen's rule for the relaxation times. Finally, a short summary of the approximations leading to the derivation of the semiconductor model is given.

19.2 Equivalent Hamiltonian Operator

The separation procedure outlined in Sects. 16.2 through 16.5 has led to the single-electron Schrödinger equation (17.40), namely, $[-\hbar^2/(2\,m)\,\nabla^2 + V]\,w = E\,w$, where the nuclei are kept fixed and the force acting onto the electron derives from a potential energy having the periodicity of the direct lattice: $V(\mathbf{r}+\mathbf{l}) = V(\mathbf{r})$, with $\mathbf{l}$ given by (17.1); as mentioned in Sects. 16.4,16.5, the external forces are absent ($U_{\text{ext}} = 0$).

Thanks to the periodicity of the Hamiltonian operator (17.40) it is possible to recast the time-independent Schrödinger equation into a different form as shown below. The procedure is based on the analogy with the Schrödinger equation for a free particle, $-\nabla^2 w = k^2\,w$ with $k^2 = 2\,m\,E/\hbar^2$. One notes in fact that the left-hand side is obtained by replacing $\mathbf{k}$ with $-\mathrm{i}\,\mathrm{grad}$ in the right-hand side; this is just another form of the transformation of momentum into the momentum operator (Sect. 8.5), which in the present case yields the whole Hamiltonian operator because for a free particle the energy is purely kinetic. This type of transformation can be pursued also for the Schrödinger equation with a periodic potential energy by observing that the eigenvalues $E_i(\mathbf{k})$, for each branch i, are periodic within the reciprocal, scaled lattice, $E_i(\mathbf{k}+\mathbf{g}) = E_i(\mathbf{k})$ (Sect. 17.6), hence they can be expanded in terms of the direct-lattice translation vectors $\mathbf{l}$ (Sect. 17.3):

$$E_i(\mathbf{k}) = \sum_{\mathbf{l}} E_{i\mathbf{l}}\,\exp(\mathrm{i}\,\mathbf{l}\cdot\mathbf{k})\,, \qquad E_{i\mathbf{l}} = E_i(\mathbf{l}) = \frac{1}{\tau_g}\int_{\tau_g} E_i(\mathbf{k})\,\exp(-\mathrm{i}\,\mathbf{l}\cdot\mathbf{k})\,\mathrm{d}^3k\,. \tag{19.1}$$

The eigenvalue $E_i(\mathbf{k})$ is now transformed into an operator by letting $\mathbf{k} \leftarrow -\mathrm{i}\,\mathrm{grad}$; remembering (17.174), this yields

$$E_i(\mathbf{k}) \leftarrow E_i(-\mathrm{i}\,\mathrm{grad}) = \sum_{\mathbf{l}} E_{i\mathbf{l}}\,\exp(\mathbf{l}\cdot\mathrm{grad}) = \sum_{\mathbf{l}} E_{i\mathbf{l}}\,\mathscr{T}_{\mathbf{l}}\,. \tag{19.2}$$

The form of operator (19.2) is purely kinetic, as the space coordinates do not appear in it. The shape of the potential energy V whence $E_i(\mathbf{k})$ originates is embedded in the

coefficients $E_{i\mathbf{l}}$ of expansion (19.1). To complete the procedure one must show that operator (19.2) has the same eigenvalues and eigenfunctions as the original operator $[-\hbar^2/(2\,m)\,\nabla^2 + V]\,w = E\,w$. Applying $E_i(-\mathrm{i}\,\mathrm{grad})$ to a Bloch function $w_{i\mathbf{k}}$ yields, using the first relation in (19.1) and the periodicity[1] of $\zeta_{i\mathbf{k}}$ (Sect. 17.6),

$$\sum_{\mathbf{l}} E_{i\mathbf{l}}\, \mathscr{T}_{\mathbf{l}}\, \zeta_{i\mathbf{k}}(\mathbf{r})\, \exp(\mathrm{i}\,\mathbf{k}\cdot\mathbf{r}) = \zeta_{i\mathbf{k}}(\mathbf{r})\, \sum_{\mathbf{l}} E_{i\mathbf{l}}\, \exp[\mathrm{i}\,\mathbf{k}\cdot(\mathbf{r}+\mathbf{l})] =$$

$$= \zeta_{i\mathbf{k}}(\mathbf{r})\, \exp(\mathrm{i}\,\mathbf{k}\cdot\mathbf{r})\, \sum_{\mathbf{l}} E_{i\mathbf{l}}\, \exp(\mathrm{i}\,\mathbf{k}\cdot\mathbf{l}) = w_{i\mathbf{k}}(\mathbf{r})\, E_i(\mathbf{k})\,. \tag{19.3}$$

The result

$$E_i(-\mathrm{i}\,\mathrm{grad})\, w_{i\mathbf{k}} = E_i(\mathbf{k})\, w_{i\mathbf{k}} \tag{19.4}$$

shows that, for each branch i, the purely kinetic operator $E_i(-\mathrm{i}\,\mathrm{grad})$ has the same eigenvalues and eigenfunctions as the Hamiltonian whence $E_i(\mathbf{k})$ originates. For this reason, operator $E_i(-\mathrm{i}\,\mathrm{grad})$ is called *equivalent Hamiltonian*. In summary, when the potential energy is periodic it is possible to directly reconstruct an equivalent operator by letting $\mathbf{k} \leftarrow -\mathrm{i}\,\mathrm{grad}$, in the same way as for a free particle. In this respect, the latter case may be viewed as a limiting condition of the former one, obtained by extending the period of the potential energy to infinity.[2] Another similarity between the free-particle case and that of a periodic potential energy is that the Hamiltonian operator is purely kinetic (albeit, in the latter case, at the cost of a more complicate form of the kinetic term).

19.2.1 Electron Dynamics

As illustrated in Sect. 17.6.1, the general solution of the time-dependent Schrödinger equation

$$\left[-\frac{\hbar^2}{2\,m}\,\nabla^2 + V(\mathbf{r})\right]\,\psi = \mathrm{i}\,\hbar\,\frac{\partial\psi}{\partial t}\,, \tag{19.5}$$

for a particle subjected to a periodic potential energy V, is (17.51), or its approximation using the wave packet of branch $i = n$,

[1] Like in Sect. 17.6.1, the periodic part of the Bloch function is indicated with $\zeta_{i\mathbf{k}}$ to avoid confusion with the group velocity.

[2] This method of reconstructing the operator from the eigenvalues was anticipated in Sect. 18.7.3.

$$\psi(\mathbf{r},t) \simeq \zeta_{n0}\, \exp(\mathrm{i}\, \Phi_{n0}) \sum_{\mathbf{k}} c_{n\mathbf{k}}\, \exp\left[\mathrm{i}\ (\mathbf{r} - \mathbf{u}_n\, t) \cdot (\mathbf{k} - \mathbf{k}_0)\right] , \tag{19.6}$$

with $\Phi_{n0} = \mathbf{k}_0 \cdot \mathbf{r} - \omega_{n0}\, t$.

Now, consider the case where an external,[3] non-periodic potential energy $U(\mathbf{r})$ is added to the periodic one; the Hamiltonian operator becomes $-\hbar^2/(2m)\, \nabla^2 + V + U$, yielding the eigenvalue equation

$$\left[-\frac{\hbar^2}{2\,m}\, \nabla^2 + V(\mathbf{r}) + U(\mathbf{r})\right] w'_{\mathbf{q}} = E'_{\mathbf{q}}\, w'_{\mathbf{q}} , \tag{19.7}$$

with $\mathbf{q}$ the label of the new eigenvalues. As the set $w'_{\mathbf{q}}$ is complete, a possible expansion of ψ is

$$\psi = \sum_{\mathbf{q}} c'_{\mathbf{q}}\, w'_{\mathbf{q}}\, \exp(-\mathrm{i}\, E'_{\mathbf{q}}\, t/\hbar) . \tag{19.8}$$

However, expansion (19.8) is inconvenient because the Hamiltonian operator in (19.7) is not periodic; as a consequence, the properties of the eigenvalues and eigenfunctions typical of the periodic case are lost. A more suitable expansion[4] is found by using the eigenfunctions of the Hamiltonian operator corresponding to $U = 0$, namely, the Bloch functions; in this case the coefficients of the expansion depend on time:

$$\psi = \sum_{i\mathbf{k}} a_{i\mathbf{k}}(t)\, w_{i\mathbf{k}} = \sum_{i\mathbf{k}} c_{i\mathbf{k}}(t)\, w_{i\mathbf{k}}\, \exp(-\mathrm{i}\, \omega_{i\mathbf{k}}\, t) , \tag{19.9}$$

where $c_{i\mathbf{k}} = a_{i\mathbf{k}}\, \exp(\mathrm{i}\, \omega_{i\mathbf{k}}\, t) \to$ const as $U \to 0$. This form is more convenient because it holds also in the case where the external potential energy depends on time, $U = U(\mathbf{r}, t)$. The approximate expression (19.6) of the wave function becomes

$$\psi(\mathbf{r},t) \simeq \zeta_{n0}\, \exp(\mathrm{i}\, \Phi_{n0})\, A , \qquad A = \sum_{\mathbf{k}} c_{n\mathbf{k}}(t)\, \exp\left[\mathrm{i}\ (\mathbf{r} - \mathbf{u}_n\, t) \cdot (\mathbf{k} - \mathbf{k}_0)\right] , \tag{19.10}$$

with $|\psi|^2 = |\zeta_{n0}|^2\, |A|^2$ and $\int_\Omega |\psi|^2\, \mathrm{d}^3 r = 1$. As ζ_{n0} is a rapidly varying function of $\mathbf{r}$, the physical information about the dynamics of the wave packet is given by $|A|^2$.

So far, the only approximation in (19.10) is the use of a single branch n of the dispersion relation. On the other hand, it must be observed that in the sum $V + U$ the first term has the periodicity of the lattice, namely, it varies rapidly in space, whereas the external potential energy is typically a slowly varying function; in fact, it is due

[3] For the sake of simplicity, suffix "ext" is dropped from the symbol of the external energy.

[4] The approach is the same as that used for treating the time-dependent perturbation theory (compare with 14.4).

to the application of external generators and/or to the presence of a nonuniform distribution of charge within the material.[5] Also, the field associated with U is weak, so that it does not influence the form of V. This leads to the idea of treating U as a perturbation superimposed to the periodic Hamiltonian operator $-\hbar^2/(2m)\,\nabla^2 + V$. Using this approximation, the Hamiltonian operator of the perturbed problem is rewritten as

$$-\frac{\hbar^2}{2\,m}\,\nabla^2 + V(\mathbf{r}) + U(\mathbf{r},t) \simeq E_n(-\mathrm{i}\,\mathrm{grad}) + U(\mathbf{r},t)\,, \tag{19.11}$$

where index n reminds one that the eigenfunctions of only the nth branch are used in the expansion. The approximation inherent in (19.11) consists in using the properties of the unperturbed problem in the perturbed case; in fact, the functional dependence of $E_n(-\mathrm{i}\,\mathrm{grad})$ on $-\mathrm{i}\,\mathrm{grad}$ derives from the unperturbed eigenvalues $E_n(\mathbf{k})$. Remembering that $E_n(-\mathrm{i}\,\mathrm{grad})$ is purely kinetic, (19.11) is similar to the Hamiltonian operator of a particle subjected only to the external potential U. The approximate form of the time-dependent Schrödinger equation then reads

$$[E_n(-\mathrm{i}\,\mathrm{grad}) + U]\;\psi = \mathrm{i}\,\hbar\frac{\partial\psi}{\partial t}\,. \tag{19.12}$$

19.2.2 Expectation Values—Crystal Momentum

The solution of (19.12) consists in determining the coefficients $c_{n\mathbf{k}}(t)$ of (19.10); this can be tackled by the method illustrated in Sect. 14.2, namely, by reducing (19.12) to a system of coupled differential equations in the unknowns $c_{n\mathbf{k}}$. More interesting it is to use (19.12) for calculating the expectation values of position and momentum; remembering that $\psi \simeq \exp(\mathrm{i}\,\Phi_{n0})\,\zeta_{n0}\,A$ is normalized to unity, one readily finds $\langle\mathbf{r}\rangle = \langle\psi|\mathbf{r}|\psi\rangle = \mathbf{r}_0$, where $\mathbf{r}_0$ denotes the center of the wave packet in the position space. As for momentum, it is $\langle\mathbf{p}\rangle = \langle\psi|-\mathrm{i}\,\hbar\mathrm{grad}|\psi\rangle$, namely, using $\Phi_{n0} = \mathbf{k}_0\cdot\mathbf{r} - \omega_{n0}\,t$,

$$\langle\mathbf{p}\rangle = -\mathrm{i}\,\hbar\int_\Omega \psi^*\;[\mathrm{i}\,\mathbf{k}_0\,\psi + \exp(\mathrm{i}\,\Phi_{n0})\;\mathrm{grad}(\zeta_{n0}\,A)]\;\mathrm{d}^3r =$$

$$= \hbar\,\mathbf{k}_0\int_\Omega |\psi|^2\,\mathrm{d}^3r - \mathrm{i}\,\hbar\int_\Omega (\zeta_{n0}A)^*\,\mathrm{grad}(\zeta_{n0}\,A)\,\mathrm{d}^3r\,. \tag{19.13}$$

[5]The field produced by nonuniformities in the local charge density, which is present also in an equilibrium condition if the dopant distribution is not spatially constant (compare with Sect. 18.5), is classified as "external" because it can be treated as a perturbation. Instead, rapid variations of the physical properties of the material, like those that typically occur at interfaces, cannot be treated using the perturbative method and require the solution of the Schrödinger equation without approximations.

The first term at the right-hand side of (19.13) yields $\hbar\,\mathbf{k}_0$ due to normalization. Letting $\zeta_{n0}\,A = a + \mathrm{i}\,b$, one finds in the second term $(\zeta_{n0}\,A)^*\,\mathrm{grad}(\zeta_{n0}\,A) = (1/2)\,\mathrm{grad}(a^2+b^2) + \mathrm{i}\,(a\,\mathrm{grad}b - b\,\mathrm{grad}a)$. The contribution of $\mathrm{grad}(a^2+b^2)$ to the integral is zero due to normalization; since A is slowly varying and normalizable, while ζ_{n0} oscillates rapidly, it follows

$$\langle \mathbf{p} \rangle = \hbar\,\mathbf{k}_0 + \hbar \int_\Omega (a\,\mathrm{grad}b - b\,\mathrm{grad}a)\;\mathrm{d}^3 r \simeq \hbar\,\mathbf{k}_0\,, \tag{19.14}$$

with $\mathbf{k}_0$ the center of the wave packet in the $\mathbf{k}$ space. The product $\hbar\,\mathbf{k}_0$ is called *crystal momentum*. As for the time derivatives of $\langle \mathbf{r} \rangle$ and $\langle \mathbf{p} \rangle$ one finds, from (17.52),

$$\dot{\mathbf{r}}_0 = \mathbf{u}_n = \frac{1}{\hbar}\,(\mathrm{grad}_{\mathbf{k}} E_n)_{\mathbf{k}_0} = \frac{\mathrm{i}}{\hbar} \sum_{\mathbf{l}} \mathbf{l}\,E_{i\mathbf{l}}\,\exp(\mathrm{i}\,\mathbf{l}\cdot\mathbf{k}_0)\,, \tag{19.15}$$

where the last expression derives from (19.1). For the time derivative of momentum one preliminarily observes that $E_n(-\mathrm{i}\,\mathrm{grad})$ commutes with the gradient operator; in fact,

$$E_n(-\mathrm{i}\,\mathrm{grad})\,\mathrm{grad}\psi = \sum_{\mathbf{l}} E_{n\mathbf{l}}\,\mathscr{T}_{\mathbf{l}}\,\mathrm{grad}\psi(\mathbf{r},t) = \sum_{\mathbf{l}} E_{n\mathbf{l}}\,\mathrm{grad}\psi(\mathbf{r}+\mathbf{l},t)\,, \tag{19.16}$$

$$\mathrm{grad}\,E_n(-\mathrm{i}\,\mathrm{grad})\psi = \mathrm{grad} \sum_{\mathbf{l}} E_{n\mathbf{l}}\,\psi(\mathbf{r}+\mathbf{l},t) = \sum_{\mathbf{l}} E_{n\mathbf{l}}\,\mathrm{grad}\psi(\mathbf{r}+\mathbf{l},t)\,. \tag{19.17}$$

Then, using definition (10.24) of the time derivative of an expectation value,[6] and remembering that the operator associated with $\mathbf{p}$ is $-\mathrm{i}\,\hbar\,\mathrm{grad}$, one finds

$$\hbar\,\dot{\mathbf{k}}_0 = \frac{\mathrm{d}\langle \mathbf{p} \rangle}{\mathrm{d}t} = \langle \psi |\,[E_n(-\mathrm{i}\,\mathrm{grad}) + U]\;\mathrm{grad} - \mathrm{grad}\;[E_n(-\mathrm{i}\,\mathrm{grad}) + U]\,|\psi\rangle\,. \tag{19.18}$$

Moreover it is $U\,\mathrm{grad}\psi - \mathrm{grad}(U\,\psi) = -\psi\,\mathrm{grad}U$, so that (19.18) eventually reduces to

$$\hbar\dot{\mathbf{k}}_0 = \frac{\mathrm{d}\langle \mathbf{p} \rangle}{\mathrm{d}t} = -\int_\Omega |\psi|^2\,\mathrm{grad}U\,\mathrm{d}^3 r\,. \tag{19.19}$$

As U is slowly varying in space, Ehrenfest approximation (10.33) applies, whence

$$\hbar\,\dot{\mathbf{k}}_0 = \frac{\mathrm{d}\langle \mathbf{p} \rangle}{\mathrm{d}t} \simeq -\,(\mathrm{grad}U)_{r_0}\,. \tag{19.20}$$

[6] Definition (10.24) could be used also for deriving (19.15).

Introducing the function $H_n(\mathbf{r}_0, \mathbf{k}_0, t) = E_n(\mathbf{k}_0) + U(\mathbf{r}_0, t)$, one finds that (19.15) and (19.20) are equivalent, respectively, to

$$\dot{x}_{i0} = \frac{\partial H_n}{\partial(\hbar\, k_{i0})}\,, \qquad \hbar\, \dot{k}_{i0} = -\frac{\partial H_n}{\partial x_{i0}}\,, \qquad i = 1, 2, 3\,. \tag{19.21}$$

Relations (19.21) are of paramount importance in solid-state theory. They show in fact that within a periodic lattice the dynamics of the expectation values of a wave packet, subjected to an external potential energy that varies slowly in space, is described by Hamilton equations (compare with (1.42)), where $\mathbf{r}_0 = \langle \mathbf{r} \rangle$ and $\hbar\, \mathbf{k}_0 = \langle \mathbf{p} \rangle$ play the role of position and momentum, respectively. It follows that H_n is a Hamiltonian function proper. Another important observation is that the time variations of the wave packet's momentum are due to the external force only; as a consequence, if $U = \text{const}$ one has $\hbar\, \dot{\mathbf{k}}_0 = 0$, namely, the crystal momentum is a constant of motion.

A further insight into the structure of H_n is obtained by calculating the work exerted onto the wave packet by the external force $-\text{grad}_{\mathbf{r}_0} U = \hbar\, \dot{\mathbf{k}}_0$ during an elementary time dt:

$$\mathrm{d}W = \hbar\, \dot{\mathbf{k}}_0 \cdot \mathrm{d}\mathbf{r}_0 = \hbar\, \dot{\mathbf{k}}_0 \cdot \mathbf{u}_n\, \mathrm{d}t = \hbar\, \mathbf{u}_n \cdot \mathrm{d}\mathbf{k}_0 = (\text{grad}_{\mathbf{k}} E_n)_{\mathbf{k}_0} \cdot \mathrm{d}\mathbf{k}_0 = \mathrm{d}E_n\,. \tag{19.22}$$

The work equals the variation of E_n; it follows that E_n, apart from an additive constant, is the kinetic energy of the wave packet. In turn, U is the potential energy which, as mentioned above, derives from the external force only. If the force acting on the electron is due to an electric field, then $U = -q\,\varphi$; this justifies the modified form (18.54) of the Fermi-Dirac statistics to be used when an electric field is present.[7] In the more general case where a magnetic field is also acting on the electron, $\hbar\, \delta\dot{\mathbf{k}}$ is given by the Lorentz force

$$\hbar\, \delta\dot{\mathbf{k}} = \mathbf{F} = -q\,(\mathbf{E} + \mathbf{u}_n \wedge \mathbf{B})\,, \tag{19.23}$$

and the Hamiltonian operator in (19.12) must be modified accordingly (compare with (9.19)).

It is important to remark again that the description of the wave packet's dynamics given in this section holds when the force is a weak perturbation with respect to the unperturbed situation. As a consequence, the description does not apply when the electron undergoes a collision; in fact, the force acting during a collision is strong and cannot be treated as a perturbation.

[7] More comments about the analogy with the perturbation theory in the classical case are made in Sect. 19.6.1.

19.2.3 Dynamics in the Parabolic-Band Approximation

When the wave packet is centered onto a wave vector $\mathbf{k}_0$ near the ath minimum of the conduction band, the diagonal expansion of $E_n(\mathbf{k})$ yields (17.57). Dropping the branch index n and letting $\mathbf{k} = \mathbf{k}_0$, $\delta k_i = k_{i0} - k_{ia}$ yield

$$E_e = E(\mathbf{k}_0) - E_C \simeq \frac{1}{2} \sum_{i=1}^{3} \frac{\hbar^2}{m_{ia}} (k_{i0} - k_{ia})^2 = \frac{1}{2} \hbar\, \delta\mathbf{k} \cdot (\hat{m}_a)^{-1}\, \hbar\, \delta\mathbf{k} \geq 0\,, \qquad (19.24)$$

with $(\hat{m}_a)^{-1}$ given by (17.58). Expression (19.24) bears a strong similarity with the kinetic energy of the classical case. The same applies to the expression of group velocity given by (17.61), namely,

$$\mathbf{u} = (\hat{m}_a)^{-1}\, \hbar\, \delta\mathbf{k}\,. \qquad (19.25)$$

Replacing (19.25) into (19.24) yields $E_e = (1/2)\, \hat{m}_a \mathbf{u} \cdot \mathbf{u}$. When the expectation value $\hbar\, \mathbf{k}_0$ of momentum coincides with $\hbar\, \mathbf{k}_a$, corresponding to an absolute minimum E_C of the conduction band, it is $E_e = 0$. Such a value is also the minimum of the positive-definite quadratic form at the right-hand side of (19.24). This shows that E_e is the kinetic energy of the electron, and allows one to identify E_C as the additive constant mentioned above.

In general, the relation between force and acceleration within a crystal is anisotropic. For the sake of simplicity consider the case of the parabolic-band approximation; the time derivative of (19.25) then yields

$$\dot{\mathbf{u}} = (\hat{m}_a)^{-1}\, \hbar\, \delta\dot{\mathbf{k}} = (\hat{m}_a)^{-1}\, \mathbf{F}\,. \qquad (19.26)$$

If the entries of the mass tensor are different from each other, the acceleration is not parallel to the force; the physical reason for this is easily understood if one thinks that the forces due to the crystal structure are embedded in the mass tensor through the second derivatives of $E(\mathbf{k})$. The mass tensor becomes a scalar only if the branch E is isotropic: $\hat{m}_a = m_a\, \mathscr{I}$, with $\mathscr{I}$ the identity tensor. More comments about this issue are made in Sect. 19.6.2.

The analysis for the valence band is similar. Again, the branch index n is dropped and symbols $\mathbf{k} = \mathbf{k}_0$, $\delta k_i = k_{i0} - k_{ia}$ are used,[8] to find (17.64), namely,

$$E_h = E_V - E(\mathbf{k}_0) \simeq \frac{1}{2} \sum_{i=1}^{3} \frac{\hbar^2}{m_{ia}} (k_{i0} - k_{ia})^2 = \frac{1}{2} \hbar\, \delta\mathbf{k} \cdot (\hat{m}_a)^{-1}\, \hbar\, \delta\mathbf{k} \geq 0\,, \qquad (19.27)$$

with $(\hat{m}_a)^{-1}$ given by (17.63), $m_{ia} > 0$. For the group velocity one finds

[8] For Si, Ge, and GaAs it is $k_{ia} = 0$ (Sect. 17.6.5).

$$\mathbf{u} = \frac{1}{\hbar}\,(\mathrm{grad}_{\mathbf{k}} E)_{\mathbf{k}_0} = -\,(\hat{m}_a)^{-1}\,\hbar\,\delta\mathbf{k}\,. \tag{19.28}$$

The work exerted onto the wave packet by the external force $-\mathrm{grad}_{\mathbf{r}_0} U = \hbar\,\dot{\mathbf{k}}_0$ during an elementary time $\mathrm{d}t$ is

$$\mathrm{d}W = \hbar\,\dot{\mathbf{k}}_0 \cdot \mathbf{u}\,\mathrm{d}t = \mathrm{d}\left[-\sum_{i=1}^{3} \frac{\hbar^2}{2\,m_{ia}}\,\delta k_i^2\right] = \mathrm{d}E\,, \tag{19.29}$$

which, again, shows that E is the kinetic energy of the electron apart from an additive constant. The negative signs in (19.27) and (19.28) make the discussion of the valence-band case somewhat awkward; however, the difficulty is readily eliminated if one refers to holes instead of electrons. For example, consider the case of an electron whose expectation value of momentum, initially equal to $\hbar\,\mathbf{k}_a$, is brought by the action of an external field to some other value $\hbar\,\mathbf{k}_0'$ in the vicinity of $\hbar\,\mathbf{k}_a$. For this transition to occur it is implied that the initial state $\mathbf{k}_a$ is occupied[9] and the final state $\mathbf{k}_0'$ is empty. As a consequence of (19.27), E changes from E_V to $E(\mathbf{k}_0') < E_V$, namely, it decreases during the time interval Δt during which the energy variation occurs; hence, the external field has exerted in Δt a negative work onto the electron, in fact, energy has been absorbed from the electron by the field. If a hole is considered instead, the initial and final states of the transition exchange roles; however, from the standpoint of the energy balance nothing changes, namely, the field still absorbs energy from the particle. It follows that the hole's energy must decrease due to the transition: this is possible only if the energy axis associated with the hole is reversed with respect to that of the electron, so that, apart from an additive constant, the hole's kinetic energy is $-E$. From this point on, the reasoning becomes identical to that outlined above for the electron of the conduction band: using (19.27), when the expectation value $\hbar\,\mathbf{k}_0$ of momentum coincides with $\hbar\,\mathbf{k}_a$, corresponding to an absolute maximum E_V of the valence band, it is $E_h = 0$. Such a value is also the minimum of the positive-definite quadratic form at the right-hand side of (19.27). This shows that E_h is the kinetic energy of the hole, and allows one to identify E_V as the additive constant.

19.3 Dynamics in the Phase Space

The theory outlined in Sects. 19.2, 19.2.1, and 19.2.2 has led to the conclusion that the dynamics of the expectation values of a wave packet describing an electron's motion, subjected to an external potential energy that varies slowly in space, is described by the Hamilton equations (19.21) where $\langle\mathbf{r}\rangle$ and $\langle\mathbf{p}\rangle$ play the role of position and momentum.

[9] For the sake of simplicity, spin is not considered here.

For a system made of a large number of electrons, the description of the dynamics of the individual wave packets is impossible from the practical standpoint. In this case one resorts to the same device as that used in Sect. 6.2 for a system of classical particles, namely, the distribution function. Being the formal apparatus identical to that of Sect. 6.2, only the relevant differences will be remarked. The μ-type phase space is defined here by the variables

$$\mathbf{s} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ k_1 \\ k_2 \\ k_3 \end{bmatrix}, \qquad \mathbf{e} = \begin{bmatrix} \partial H/\partial k_1 \\ \partial H/\partial k_2 \\ \partial H/\partial k_3 \\ -\partial H/\partial x_1 \\ -\partial H/\partial x_2 \\ -\partial H/\partial x_3 \end{bmatrix}, \tag{19.30}$$

(compare with (1.57)) so that the distribution function[10] reads $f = f(\mathbf{r}, \mathbf{k}, t)$. Note that the units of f are different from those of the classical distribution function. For the latter, in fact, it is $[f_\mu] = (\mathrm{J\ s})^{-3}$, so that the product $f_\mu\, \mathrm{d}^3 r\, \mathrm{d}^3 p$ is dimensionless (compare with (6.1)); in the present case, instead, both $f\, \mathrm{d}^3 r\, \mathrm{d}^3 k$ and $\mathrm{d}^3 r\, \mathrm{d}^3 k$ are dimensionless, hence the distribution function itself is dimensionless.

The system considered for the investigation is that of the electrons belonging to the conduction band. Remembering the first relation in (6.3), the concentration and average velocity of such electrons are given by

$$n(\mathbf{r}, t) = \iiint_{-\infty}^{+\infty} f(\mathbf{r}, \mathbf{k}, t)\, \mathrm{d}^3 k\,, \qquad \mathbf{v}(\mathbf{r}, t) = \frac{1}{n} \iiint_{-\infty}^{+\infty} \mathbf{u}(\mathbf{k}) f(\mathbf{r}, \mathbf{k}, t)\, \mathrm{d}^3 k\,, \tag{19.31}$$

with $\mathbf{u}$ the electron's group velocity. In the equilibrium condition it is $f^{\mathrm{eq}} = Q P$, where the Fermi-Dirac statistics P depends on $\mathbf{k}$ only through $E(\mathbf{k})$, namely, it is even with respect to $\mathbf{k}$. In turn, $\mathbf{u} = (1/\hbar)\, \mathrm{grad}_{\mathbf{k}} E$ is odd, so that the whole integrand in the second definition of (19.31) is odd. As the integration domain is symmetric with respect to $\mathbf{k} = 0$, it is $\mathbf{v}^{\mathrm{eq}} = 0$ as should be. In a nonequilibrium condition it is $f = Q \Phi$, with $\Phi(\mathbf{r}, \mathbf{k})$ the occupation probability of a state. If the band is completely filled, then $\Phi = 1$; the electron flux $n\mathbf{v}$ then becomes proportional to the integral of $\mathbf{u}$; as the latter is odd, the flux vanishes: this explains why a completely filled band does not contribute to the material's conduction, as anticipated in Sect. 17.6.5.

The Boltzmann collisionless equation in the $\mathbf{r}, \mathbf{k}$ space is derived in the same manner as for (6.28); it reads

$$\frac{\partial f}{\partial t} + \dot{\mathbf{r}} \cdot \mathrm{grad}_{\mathbf{r}} f + \dot{\mathbf{k}} \cdot \mathrm{grad}_{\mathbf{k}} f = 0\,. \tag{19.32}$$

[10] Suffix μ is dropped to distinguish this distribution function from that of the classical case.

The effects of collisions may be grouped into two classes: the collisions of the first class induce transitions that change the number of electrons of the band. Such transitions are the generations and recombinations introduced in Sect. 17.6.6, where the initial state of the electron belongs to the conduction band and the final one belongs to the valence band, or vice versa.[11] The transitions of this class are collectively called *inter-band transitions*.

The collisions of the second class are those where the initial and final state belong to the same band, and are called *intra-band transitions*; they do not change the number of electrons of the band. The distinction between the two classes is useful because the inter-band transitions exhibit characteristic times that are much larger than those of the intra-band transitions. In turn, the intra-band transitions are further divided into two subclasses: the *intra-valley transitions*, where the initial and final states are in the vicinity of the same extremum of the band, and the *inter-valley transitions*, where the initial and final state are in the vicinity of different extrema.[12]

Within each class, the transitions are further grouped depending on the entity with which the collision occurs; typical examples of collisions are those with phonons, impurities, defects, and photons. Like in the classical case, collisions are not accounted for in the derivation of (19.32), where the effect of only the slowly varying external potential energy is present; the further time change of f due to collisions is more conveniently kept separate from that of the external potential energy. Also, it is assumed that the system under consideration is dilute, so that each wave packet spends a relatively large fraction of time without suffering any collision; in other terms, the time during which an electron is subjected to the external field is much longer than that involved in a collision. For this reason it is preferable to write the Boltzmann Transport Equation (BTE), when the collisions are accounted for, as

$$\frac{\partial f}{\partial t} + \mathbf{u} \cdot \mathrm{grad}_{\mathbf{r}} f - \frac{q}{\hbar} \left(\mathbf{E} + \mathbf{u} \wedge \mathbf{B}\right) \cdot \mathrm{grad}_{\mathbf{k}} f = C \tag{19.33}$$

(compare with (6.29) and (6.31)). To derive (19.33), the expression (19.23) of the Lorentz force acting on the electron is used, after dropping index n from the group velocity. Term C embeds the forces acting during the collisions; such forces are short ranged and much more intense than those due to the external field; as a consequence, the Ehrenfest approximation (10.33) does not apply, so that a full quantum-mechanical approach is necessary to treat the collision term.

The relations involving the expectation values at the left-hand side of (19.32) and (19.33) are formally identical to those of the classical case, despite the fact

[11] In addition to this one must also consider the trapping-detrapping phenomena involving localized states. So far, only the localized states due to dopants have been considered (Sect. 18.4); other types of localized states are introduced in Chap. 20.

[12] Here the extrema are the minima of the conduction band. The region near an extremum of a band is also called *valley*.

the dynamic variables are actually derived from a quantum-mechanical description; in addition, the electromagnetic field appears in the non-quantized form. Therefore, the form of such equations is also called *semiclassical*.

19.3.1 Collision Term

The left-hand side of (19.33) equals $\mathrm{d}f/\mathrm{d}t$, whence the right-hand side is the rate of change of f due to collisions. As $f(\mathbf{r}, \mathbf{k}, t)\,\mathrm{d}^3r\,\mathrm{d}^3k$ is the number of electrons of the conduction band that at time t belong to the elementary volume $\mathrm{d}^3r\,\mathrm{d}^3k$ centered on $(\mathbf{r}, \mathbf{k})$, the rate of change can be expressed as

$$C = C_{\mathrm{in}} - C_{\mathrm{out}}\,, \tag{19.34}$$

where $C_{\mathrm{in}}\,\mathrm{d}^3r\,\mathrm{d}^3k$ is the number of electrons entering $\mathrm{d}^3r\,\mathrm{d}^3k$ per unit time, due to collisions, and $C_{\mathrm{out}}\,\mathrm{d}^3r\,\mathrm{d}^3k$ is the number of electrons leaving $\mathrm{d}^3r\,\mathrm{d}^3k$ per unit time, due to collisions. To illustrate the reasoning it is convenient to refer to Fig. 19.1, where a one-dimensional case is illustrated using the x, k coordinates. Instead of an elementary volume $\mathrm{d}x\,\mathrm{d}k$, a finite, rectangular cell is considered, whose position at time t_1 is fixed by the vertices A_1, B_1, C_1, and D_1. For simplicity it is assumed that no external force is present ($U = \text{const}$), so that the crystal momentum is conserved. In particular, the vertices' momenta at $t = t_1$ are $\hbar\, k_m = \hbar\, k(A_1) = \hbar\, k(B_1)$ and $\hbar\, k_M = \hbar\, k(C_1) = \hbar\, k(D_1)$. The corresponding positions are $x(A_1) = x(C_1) = 0$, $x_0 = x(B_1) = x(D_1)$. Letting m^* indicate the effective mass, the position of the vertices at a subsequent time $t_2 = t_1 + \Delta t$ is

$$x(A_2) = \frac{\hbar\, k_m}{m^*}\,\Delta t\,, \quad x(B_2) = x_0 + x(A_2)\,, \quad x(C_2) = \frac{\hbar\, k_M}{m^*}\,\Delta t\,, \quad x(D_2) = x_0 + x(C_2)\,,$$

this giving rise to the parallelogram also shown in Fig. 19.1. If no collisions occur, the electrons inside the parallelogram at $t = t_2$ are the same as those that were inside the rectangle at $t = t_1$; in contrast, when collisions occur, some electrons leave the rectangle without reaching the parallelogram (hence, $C_{\mathrm{out}} \neq 0$), while the parallelogram is reached by other electrons that originally did not belong to the rectangle ($C_{\mathrm{in}} \neq 0$). This is schematically indicated by the arrows in Fig. 19.1. In general it is $C_{\mathrm{out}} \neq C_{\mathrm{in}}$, so that $\mathrm{d}f/\mathrm{d}t \neq 0$. The description is the same also in the case when the external force is present, the difference being that the trajectories in the phase space are not rectilinear and the deformation of the domain is more complicate.

To give the analysis a more formal aspect it is necessary to determine how the population of the elementary domain $\mathrm{d}^6s = \mathrm{d}^3r\,\mathrm{d}^3k$ evolves in the elementary time interval $\mathrm{d}t$. To begin, one introduces the *scattering probability per unit time and unit*

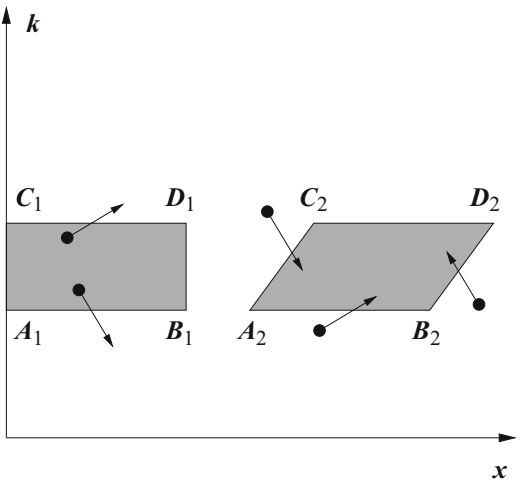


Fig. 19.1 Example of the time evolution of a phase-space domain in a one-dimensional case. The situation with no external force is considered

phase volume, S, from an initial state to a final state of the phase space.[13] The initial (final) state is indicated by the first (second) pair of arguments of S, namely,

$$S\left(\mathbf{r},\mathbf{k}\to\mathbf{r}',\mathbf{k}'\right)\,\mathrm{d}^3r'\,\mathrm{d}^3k' \tag{19.35}$$

is the probability per unit time that an electron scatters from $(\mathbf{r},\mathbf{k})$ to the elementary volume $\mathrm{d}^3r'\,\mathrm{d}^3k'$ centered at $(\mathbf{r}',\mathbf{k}')$. Then, let $\mathrm{d}N_{\mathrm{in}} = C_{\mathrm{in}}\,\mathrm{d}^6s$, $\mathrm{d}N_{\mathrm{out}} = C_{\mathrm{out}}\,\mathrm{d}^6s$, and $\mathbf{s} = (\mathbf{r},\mathbf{k})$, $\mathbf{s}' = (\mathbf{r}',\mathbf{k}')$. The number $\mathrm{d}N_{\mathrm{in}}$ is determined by observing that the electrons contributing to it are those that initially belong to elementary phase-space volumes, say, d^6s', different from d^6s. The population of d^6s' at time t is $f(\mathbf{s}',t)\,\mathrm{d}^6s'$; if the latter is multiplied by the scattering probability per unit time from $\mathbf{s}'$ to d^6s, given by $S(\mathbf{s}'\to\mathbf{s})\,\mathrm{d}^6s$, the unconditional number of transitions from d^6s' to d^6s is obtained. The actual number of such transitions is then found by remembering that electrons are fermions, so that transitions towards d^6s are possible only if the final states are empty; in other terms, the unconditional number of $\mathbf{s}'\to\mathbf{s}$ transitions must be multiplied by $1-\Phi(\mathbf{s},t)$, where $\Phi(\mathbf{s},t)$ is the probability that the final state is full. Finally, the contributions of all elementary volumes d^6s' must be summed up, to find[14]

$$\mathrm{d}N_{\mathrm{in}} = \int_{\mathbf{s}'}\left[f(\mathbf{s}',t)\,\mathrm{d}^6s'\right]\,\left[S\left(\mathbf{s}'\to\mathbf{s}\right)\,\mathrm{d}^6s\right]\,\left[1-\Phi(\mathbf{s},t)\right]\,. \tag{19.36}$$

The derivation of $\mathrm{d}N_{\mathrm{out}}$ is similar; one obtains

[13] The units of S are $[S]=\mathrm{s}^{-1}$.

[14] For the sake of conciseness, in Sects. 19.3.1, 19.3.2, and 19.3.3 the six-fold integrals over $\mathrm{d}^3r'\,\mathrm{d}^3k'$ and the three-fold integrals over d^3k' are indicated with $\int_{\mathbf{s}'}$ and $\int_{\mathbf{k}'}$, respectively.

$$\mathrm{d}N_{\mathrm{out}} = \int_{\mathbf{s}'} \left[f(\mathbf{s},t)\,\mathrm{d}^6 s\right]\ \left[S\left(\mathbf{s}\to\mathbf{s}'\right)\ \mathrm{d}^6 s'\right]\ \left[1-\Phi(\mathbf{s}',t)\right]\,. \tag{19.37}$$

The collision term $C = C_{\mathrm{in}} - C_{\mathrm{out}}$ is now determined by subtracting (19.37) from (19.36) and dividing the result by $\mathrm{d}^6 s$. This shows that C is the sum of two terms; the first one is linear with respect to f and reads

$$\int_{\mathbf{s}'} \left[f(\mathbf{s}',t)\,S\left(\mathbf{s}'\to\mathbf{s}\right) - f(\mathbf{s},t)\,S\left(\mathbf{s}\to\mathbf{s}'\right)\right]\,\mathrm{d}^6 s'\,. \tag{19.38}$$

As for the second term, one must preliminarily observe that $f = Q\,\Phi$, with $Q = 1/(4\,\pi^3)$ the density of states in the phase space, (17.50); then, the second term of C turns out to be quadratic with respect to Φ or f:

$$Q\int_{\mathbf{s}'} \Phi(\mathbf{s},t)\,\Phi(\mathbf{s}',t)\,\left[S\left(\mathbf{s}\to\mathbf{s}'\right) - S\left(\mathbf{s}'\to\mathbf{s}\right)\right]\,\mathrm{d}^6 s'\,. \tag{19.39}$$

19.3.2 Point-Like Collisions

The two summands (19.38), (19.39) in the expression of C are substantially simplified thanks to the property that the collisional forces, albeit very strong, are short ranged; as a consequence, whereas the momentum of the colliding electron may undergo a large change due to the collision, the electron's position changes little. The issue is illustrated with the aid of Fig. 19.2, that schematically describes an electron collision with a negatively ionized dopant. The latter is represented by the black circle, whereas the gray region around it indicates the positive charge attracted by the negative ion; such a positive charge acts like an *electric screen* that tends to neutralize the ion. As a consequence of the screen, the decay of the electrostatic potential acting between the ion and the incoming electron, when the relative distance increases, is much stronger than in the pure Coulomb case.[15] In practice, one can assume that the electron-ion repulsion is non-negligible only when the electron is inside the screen. This makes the dynamics of the interaction rather different from that of the pure Coulomb case, treated in Sect. 3.8. As shown in the figure, the final momentum $\mathbf{k}'$ may differ largely from the initial one, $\mathbf{k}$; in contrast, considering the atomic scale of the phenomenon, the final position $\mathbf{r}'$ may be thought of as coinciding with the initial one, $\mathbf{r}$. To vest this observation with mathematical

[15] The derivation and treatment of the screened Coulomb interaction are carried out in Sects. 20.6.4 and 14.7, respectively.

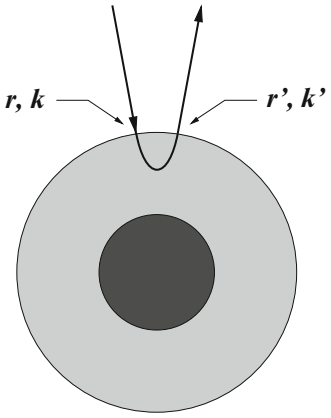


Fig. 19.2 Qualitative picture of a collision between an electron and a negatively ionized impurity. The latter is schematically represented by the black circle, whereas the gray area indicates the screening region. The initial and final state of the electron are indicated with $(\mathbf{r}, \mathbf{k})$ and $(\mathbf{r}', \mathbf{k}')$, respectively

form, considering that the scattering probability in (19.38) and (19.39) undergoes an integration over $\mathbf{r}'$, one lets[16]

$$S\left(\mathbf{s} \to \mathbf{s}'\right) = S_0\left(\mathbf{r}, \mathbf{k} \to \mathbf{k}'\right)\, \delta(\mathbf{r}' - \mathbf{r})\,. \tag{19.40}$$

Another important consequence of the above discussion is that although the duration of the interaction is very short, the force acting on the electron due the interaction is much stronger than the external forces, to the extent that the effects of the latter can be neglected during the interaction itself. It follows that S and S_0 do not depend on the external forces; this greatly simplifies the analysis of S_0. Inserting (19.40) into (19.38) yields, for the first part of C,

$$\int_{\mathbf{k}'} \left[f(\mathbf{r}, \mathbf{k}', t)\, S_0\left(\mathbf{r}, \mathbf{k}' \to \mathbf{k}\right) - f(\mathbf{r}, \mathbf{k}, t)\, S_0\left(\mathbf{r}, \mathbf{k} \to \mathbf{k}'\right)\right] \mathrm{d}^3 k'\,; \tag{19.41}$$

in turn, the second part (19.39) becomes

$$Q \int_{\mathbf{k}'} \Phi(\mathbf{r}, \mathbf{k}, t)\, \Phi(\mathbf{r}, \mathbf{k}', t)\, \left[S_0\left(\mathbf{r}, \mathbf{k} \to \mathbf{k}'\right) - S_0\left(\mathbf{r}, \mathbf{k}' \to \mathbf{k}\right)\right] \mathrm{d}^3 k'\,. \tag{19.42}$$

The term S_0 is typically calculated using the first-order perturbation theory (Sect. 14.3), which shows that the transition probability is invariant upon reversal of the initial and final states. It follows that the quantity in brackets in (19.42) vanishes, so that (19.41) is in fact the only contribution to C. The latter is recast in a more compact form by defining the *relaxation time* $\tau(\mathbf{r}, \mathbf{k})$ and the *collision operator* $\tilde{f}(\mathbf{r}, \mathbf{k})$ such that

[16] Note that the units of S_0 are different from those of S: in fact, $[S_0] = \mathrm{cm}^3/\mathrm{s}$. Examples of calculations of phonon scattering and ionized-impurity scattering are given in Sects. 20.5.1 and 20.5.2, respectively.

$$\frac{1}{\tau} = \int_{\mathbf{k}'} S_0\left(\mathbf{r}, \mathbf{k} \to \mathbf{k}'\right) \mathrm{d}^3k' , \qquad \tilde{f} = \frac{\int_{\mathbf{k}'} f(\mathbf{r}, \mathbf{k}', t)\, S_0\left(\mathbf{r}, \mathbf{k}' \to \mathbf{k}\right) \mathrm{d}^3k'}{\int_{\mathbf{k}'} S_0\left(\mathbf{r}, \mathbf{k} \to \mathbf{k}'\right) \mathrm{d}^3k'} , \tag{19.43}$$

to find $C = C_{\text{in}} - C_{\text{out}} = (\tilde{f} - f)/\tau$. The BTE (19.33) thus becomes

$$\frac{\partial f}{\partial t} + \mathbf{u} \cdot \mathrm{grad}_{\mathbf{r}} f - \frac{q}{\hbar} \left(\mathbf{E} + \mathbf{u} \wedge \mathbf{B}\right) \cdot \mathrm{grad}_{\mathbf{k}} f = -\frac{f - \tilde{f}}{\tau} . \tag{19.44}$$

In the derivation of (19.44) no distinction is made between the inter-band and intra-band transitions. For a dilute system one can assume that the transitions of the two types are uncorrelated, so that the corresponding probabilities are additive: $S_0 = S_{0b} + S_{0v}$, where index b (v) stands for "inter-band" ("intra-band"). As a consequence,

$$\frac{1}{\tau} = \frac{1}{\tau_b} + \frac{1}{\tau_v} , \qquad \frac{1}{\tau_b} = \int_{\mathbf{k}'} S_{0b}\, \mathrm{d}^3k' , \qquad \frac{1}{\tau_v} = \int_{\mathbf{k}'} S_{0v}\, \mathrm{d}^3k' . \tag{19.45}$$

For the semiconductors of interest, the relaxation times defined in (19.45) differ by several orders of magnitude (e.g., in electronic-grade silicon[17] is $\tau_b > 10^{-6}$ s, $\tau_v < 10^{-12}$ s). This makes the intra-band transitions dominant ($\tau = \tau_b\, \tau_v/(\tau_b + \tau_v) \simeq \tau_v$); one exception exists though, where the effects of the intra-band transitions cancel each other exactly, so that the inter-band transitions only are left. Such an exception is discussed in Sect. 19.4.

19.3.3 Perturbative Form of the BTE

The Boltzmann Transport Equation (19.44) is an integral-differential equation in the phase space and time, in the unknown f. The kernel of the integral part is S_0, while the equation coefficients are $\mathbf{E}(\mathbf{r}, t)$, $\mathbf{B}(\mathbf{r}, t)$, $\tau(\mathbf{r}, \mathbf{k})$, and $\mathbf{u}(\mathbf{k}) = (1/\hbar)\, \mathrm{grad}_{\mathbf{k}} H$. In equilibrium f becomes $f^{\text{eq}} = Q\, P$, with $P(\mathbf{r}, \mathbf{k})$ the Fermi-Dirac statistics; as shown in Sect. 19.2.2, P depends on position if the semiconductor is not uniform. Moreover it is $\mathrm{d}f^{\text{eq}}/\mathrm{d}t = 0$; hence, to make the collision term to vanish at equilibrium, it must be $\tilde{f}^{\text{eq}} = f^{\text{eq}}$ (detailed-balance principle, Sect. 6.5).

In general, the solution of the BTE is quite a difficult task. The issue of effective solution methods for this equation will be introduced later; however, a solution procedure is outlined here which, although seldom used in practical cases, has the advantage of providing a simplified form of (19.44), upon which a number of models for semiconductor-device analysis are based. The procedure consists in setting up

[17] The semiconductor's purification degree necessary for manufacturing integrated circuit is called *electronic grade*; it indicates that the ratio between the concentration of impurities and that of the semiconductor atoms is smaller than 10^{-9}. Lower-quality materials, with a ratio smaller than 10^{-6}, are used in the fabrication of solar cells; in this case the purification degree is called *solar grade*.

the iterative scheme

$$\frac{\mathrm{d}f^{(m+1)}}{\mathrm{d}t} = -\frac{f^{(m+1)} - \tilde{f}^{(m)}}{\tau}\,, \qquad \tilde{f}^{(m)} = \frac{\int_{\mathbf{k}'} f^{(m)}\, S_0\, \mathrm{d}^3k'}{\int_{\mathbf{k}'} S_0\, \mathrm{d}^3k'}\,. \tag{19.46}$$

with m the iteration index. In this way the $\tilde{f}^{(m)}$ term at the right-hand side of the first equation in (19.46) is known from the previous iteration, so that the integral-differential equation is transformed into a number of differential-only equations. If convergence occurs, the iterations are brought to an end when a suitable norm $||f^{(m+1)} - f^{(m)}||$ is smaller than a prescribed value.

To start the procedure it is reasonable to choose, for the approximation of order zero, the equilibrium distribution: $f^{(0)} = f^{\mathrm{eq}}$; from the detailed-balance principle it follows $\tilde{f}^{(0)} = f^{\mathrm{eq}}$. The first step of the iteration procedure then yields $f^{(1)}$, called *first-perturbation solution*. In many cases of practical interest, the material or device under investigation is sufficiently close to the equilibrium condition to make $f^{(1)}$ an acceptable solution; one then stops the iterations at the first-perturbation solution and takes $f \simeq f^{(1)}$. This is equivalent to solving the *perturbative form of the BTE*

$$\frac{\partial f}{\partial t} + \mathbf{u} \cdot \mathrm{grad}_{\mathbf{r}} f - \frac{q}{\hbar}\,(\mathbf{E} + \mathbf{u} \wedge \mathbf{B}) \cdot \mathrm{grad}_{\mathbf{k}} f = -\frac{f - f^{\mathrm{eq}}}{\tau}\,. \tag{19.47}$$

It is interesting to comment on the form of (19.47). The first term at the left-hand side differs from zero only if the distribution function depends explicitly on time; hence, it vanishes in a nonequilibrium condition if the latter is of the steady-state type. The second term vanishes if the distribution function is independent of the spatial coordinates, hence it describes a contribution to electron transport that originates from a spatial nonuniformity. For this reason this term is called *diffusive term* (compare with Sect. 23.3). The third term vanishes if the external force is missing, hence it originates from the action of the external forces on the electrons and, for this reason, is called *drift term*. At the right-hand side of (19.47), the magnitude of the relaxation time influences the amount by which the distribution function departs from equilibrium; to better show this, one recasts (19.47) as

$$f = f^{\mathrm{eq}} - \tau\, \mathscr{L} f\,, \qquad \mathscr{L} = \frac{\partial}{\partial t} + \mathbf{u} \cdot \mathrm{grad}_{\mathbf{r}} - \frac{q}{\hbar}\,(\mathbf{E} + \mathbf{u} \wedge \mathbf{B}) \cdot \mathrm{grad}_{\mathbf{k}}\,, \tag{19.48}$$

with $\mathscr{L}$ the *Liouvillian operator*. If $\tau \to 0$, then $f \to f^{\mathrm{eq}}$; this shows that the perturbative solution is in fact acceptable if the relaxation time is sufficiently small.[18] It is worth adding that in Classical Mechanics a simplified form of (19.48) is widely used to investigate transport problems. A brief outline of the procedure is given in Sect. 19.6.7.

[18] On the other hand, in a collisionless case it is $S_0 \to 0$ whence, from (19.43), it follows $\tau \to \infty$. In this situation there is no limit to the departure of f from f^{eq}.

A final comment refers to the spatially uniform case, where $f = f(\mathbf{k}, t)$, $\tau = \tau(\mathbf{k})$; then, (19.47) simplifies to

$$\frac{\partial f}{\partial t} - \frac{q}{\hbar} \, (\mathbf{E} + \mathbf{u} \wedge \mathbf{B}) \cdot \mathrm{grad}_{\mathbf{k}} f = -\frac{f - f^{\mathrm{eq}}}{\tau} \, . \tag{19.49}$$

If the fields are set to zero at some instant of time, say, $t = 0$, (19.49) reduces to a differential equation with respect to time only, with $\mathbf{k}$ a parameter; thus, for each $\mathbf{k}$ the solution approaches the equilibrium distribution[19] according to the law

$$f = f^{\mathrm{eq}} + (f_{t=0} - f^{\mathrm{eq}}) \, \exp(-t/\tau) \, . \tag{19.50}$$

In passing, this result explains why τ is called "relaxation time."

19.4 Moments Expansion of the BTE

The *moments expansion of the BTE* is a transformation method that has successfully been applied to the analysis of semiconductor devices. It reduces the original equation to a set of partial-differential equations in the $\mathbf{r}, t$ space; the number of such equations can be adapted to the type of information to be acquired about the device under investigation. More specifically, applying the moments method to the BTE and truncating the series of moments at a suitable order, one extracts from the BTE a hierarchically ordered set of models, ranging from the simplest to the more complicate ones ([112–114], and references therein).

The BTE is an equation in the $\mathbf{r}, \mathbf{k}, t$ space. The basic idea of the moments method is that for the description of carrier-transport phenomena in semiconductor devices, it is often sufficient to rely on equations defined over the $\mathbf{r}, t$ space alone; in fact, for practical applications the information about the distribution of the crystal momentum is less important. The equations in the $\mathbf{r}, t$ space are extracted from the BTE by multiplying the latter by suitable functions $\alpha(\mathbf{k})$ and integrating the result over the $\mathbf{k}$ space. The integration saturates the $\mathbf{k}$ coordinates and, as shown in Sect. C.6, provides an equation in the $\mathbf{r}, t$ space.[20] Remembering that the electron dynamics using the equivalent Hamiltonian operator is described by expanding the electron's wave function in terms of the Bloch functions (Sect. 19.2.1), the integration over the $\mathbf{k}$ space is in fact limited to the first Brillouin zone. On the other hand, the typical behavior of the distribution function at the boundary Γ of the first Brillouin zone is such that $\alpha(\mathbf{k}) f(\mathbf{r}, \mathbf{k}, t) \to 0$ when $\mathbf{k} \to \mathbf{k}_{\Gamma}$. This amount to

[19] Compare with the discussion carried out in Sect. 6.6.3.

[20] As indicated in Sect. C.6, term "moment" is specifically used when α is a polynomial in $\mathbf{k}$. As the dependence of α on $\mathbf{k}$ is not specified yet, it is implied that the form of α is such that the integrals in (19.51) converge.

assuming that there are no electrons at the boundary.[21] From a practical standpoint, the hypothesis has the same effect as that of replacing the first Brillouin zone with an infinite domain and assuming that the distribution function in a nonequilibrium condition vanishes at infinity in the same exponential-like fashion as it does at equilibrium, where it becomes proportional to the Fermi-Dirac statistics [108, 111]. With these premises, using the general form (19.33) of the equation, the moment of the BTE with respect to α reads

$$\iiint_{-\infty}^{+\infty} \alpha \left[\frac{\partial f}{\partial t} + \mathbf{u} \cdot \operatorname{grad}_{\mathbf{r}} f - \frac{q}{\hbar} \left(\mathbf{E} + \mathbf{u} \wedge \mathbf{B} \right) \cdot \operatorname{grad}_{\mathbf{k}} f \right] \mathrm{d}^3 k = \iiint_{-\infty}^{+\infty} \alpha \, C \, \mathrm{d}^3 k \, . \tag{19.51}$$

As the BTE is a continuity equation of the distribution function in the phase space, (19.51) is expected to be a continuity equation in the **r** space; in fact, as shown below, it is the continuity equation of the product $n\overline{\alpha}$, where n is the concentration of the electrons in the conduction band, given by the first relation in (19.31), and

$$\overline{\alpha}(\mathbf{r}, t) = \frac{1}{n} \iiint_{-\infty}^{+\infty} \alpha(\mathbf{k}) f(\mathbf{r}, \mathbf{k}, t) \, \mathrm{d}^3 k \tag{19.52}$$

is the average of α over the **k** space. The continuity equation is derived below by working out separately the different terms appearing in (19.51).

19.4.1 Time Derivative

The derivation of this term is readily accomplished by observing that the $\alpha \, \partial f / \partial t = \partial(\alpha f)/\partial t$, whence

$$\iiint_{-\infty}^{+\infty} \alpha \frac{\partial f}{\partial t} \, \mathrm{d}^3 k = \frac{\partial}{\partial t} \iiint_{-\infty}^{+\infty} \alpha f \, \mathrm{d}^3 k = \frac{\partial}{\partial t} \left(n \overline{\alpha} \right) \, . \tag{19.53}$$

This shows that (19.51) is the continuity equation of $n\overline{\alpha}$, as anticipated.

[21] In the case of the conduction band of germanium, the minima are at the boundary (Sect. 17.6.5), which makes the hypothesis inconsistent as it stands; to perform the integration one must shift the origin of the **k** space and exploit the periodicity of the band structure. The hypothesis that the distribution function vanishes at the boundary of the first Brillouin zone is made also in the application of the moments method to the holes of the valence band.

19.4.2 Diffusion Term

To calculate this terms one starts with the relation $\alpha\,\mathbf{u}\cdot\mathrm{grad}_{\mathbf{r}} f = \mathrm{div}_{\mathbf{r}}(\alpha\,\mathbf{u} f)$, that derives from the second identity in (A.16) and from the fact that $\alpha\,\mathbf{u}$ does not depend on $\mathbf{r}$. Thus,

$$\iiint_{-\infty}^{+\infty} \alpha\,\mathbf{u}\cdot\mathrm{grad}_{\mathbf{r}} f\,\mathrm{d}^3k = \mathrm{div}_{\mathbf{r}} \iiint_{-\infty}^{+\infty} \alpha\,\mathbf{u} f\,\mathrm{d}^3k = \mathrm{div}_{\mathbf{r}}\,(n\,\overline{\alpha\,\mathbf{u}})\,. \tag{19.54}$$

19.4.3 Drift Term

The term containing the electric field is treated starting from the identity

$$\alpha\,\mathbf{E}\cdot\mathrm{grad}_{\mathbf{k}} f = \mathbf{E}\cdot\mathrm{grad}_{\mathbf{k}}(\alpha f) - f\,\mathbf{E}\cdot\mathrm{grad}_{\mathbf{k}}\alpha\,. \tag{19.55}$$

The integral of the first term at the right-hand side of (19.55) vanishes due to identity (A.26) and to the asymptotic behavior of f described earlier. As a consequence,

$$-\iiint_{-\infty}^{+\infty} \alpha\,\mathbf{E}\cdot\mathrm{grad}_{\mathbf{k}} f\,\mathrm{d}^3k = \mathbf{E}\cdot\iiint_{-\infty}^{+\infty} f\,\mathrm{grad}_{\mathbf{k}}\alpha\,\mathrm{d}^3k = \mathbf{E}\cdot n\,\overline{\mathrm{grad}_{\mathbf{k}}\alpha}\,. \tag{19.56}$$

The term containing the magnetic induction is treated more easily by rewriting the mixed product as $\mathrm{grad}_{\mathbf{k}} f\cdot\alpha\,\mathbf{u}\wedge\mathbf{B} = \mathrm{grad}_{\mathbf{k}} f\wedge\alpha\,\mathbf{u}\cdot\mathbf{B}$ (compare with (A.31)) and using the first identity in (A.35), to find

$$\mathrm{grad}_{\mathbf{k}} f\wedge\alpha\,\mathbf{u}\cdot\mathbf{B} = \mathrm{rot}_{\mathbf{k}}(f\,\alpha\,\mathbf{u})\cdot\mathbf{B} - f\,\mathrm{rot}_{\mathbf{k}}(\alpha\,\mathbf{u})\cdot\mathbf{B}\,. \tag{19.57}$$

The integral of $\mathrm{rot}_{\mathbf{k}}(f\,\alpha\,\mathbf{u})\cdot\mathbf{B}$ over $\mathbf{k}$ vanishes due to identity (A.38) and to the asymptotic behavior of f; this yields

$$-\iiint_{-\infty}^{+\infty} \alpha\,\mathbf{u}\wedge\mathbf{B}\cdot\mathrm{grad}_{\mathbf{k}} f\,\mathrm{d}^3k = \mathbf{B}\cdot\iiint_{-\infty}^{+\infty} f\,\mathrm{rot}_{\mathbf{k}}(\alpha\,\mathbf{u})\,\mathrm{d}^3k\,. \tag{19.58}$$

In turn, identity (A.35) transforms the integrand at the right-hand side of (19.58) as $f\,\mathrm{rot}_{\mathbf{k}}(\alpha\,\mathbf{u}) = f\,\alpha\,\mathrm{rot}_{\mathbf{k}}\mathbf{u} + f\,\mathrm{grad}_{\mathbf{k}}\alpha\wedge\mathbf{u}$ where, thanks to the definition (17.52) of the group velocity, it is $\mathrm{rot}_{\mathbf{k}}\mathbf{u} = (1/\hbar)\,\mathrm{rot}_{\mathbf{k}}\mathrm{grad}_{\mathbf{k}} H = 0$. Thus, (19.58) becomes

$$-\iiint_{-\infty}^{+\infty} \alpha\,\mathbf{u}\wedge\mathbf{B}\cdot\mathrm{grad}_{\mathbf{k}} f\,\mathrm{d}^3k = \mathbf{B}\cdot\iiint_{-\infty}^{+\infty} f\,\mathrm{grad}_{\mathbf{k}}\alpha\wedge\mathbf{u}\,\mathrm{d}^3k =$$

$$= \iiint_{-\infty}^{+\infty} f\,\mathrm{grad}_{\mathbf{k}}\alpha\cdot\mathbf{u}\wedge\mathbf{B}\,\mathrm{d}^3k = n\,\overline{\mathrm{grad}_{\mathbf{k}}\alpha\cdot\mathbf{u}\wedge\mathbf{B}}\,. \tag{19.59}$$

In (19.59), the term containing the magnetic induction does not contribute to the moment if α or f depends on $\mathbf{k}$ through energy alone, $\alpha = \alpha(H)$ or $f = f(H)$. In fact, in the first case it is $\mathrm{grad}_{\mathbf{k}}\alpha = (\mathrm{d}\alpha/\mathrm{d}H)\,\mathrm{grad}_{\mathbf{k}}H$, whence

$$\mathrm{grad}_{\mathbf{k}}\alpha \cdot \mathbf{u} \wedge \mathbf{B} = \frac{\mathrm{d}\alpha}{\mathrm{d}H}\,\hbar\,\mathbf{u} \cdot \mathbf{u} \wedge \mathbf{B} = 0\,. \tag{19.60}$$

The same calculation holds when $f = f(H)$, starting from the integrand at the left-hand side of (19.59).

19.4.4 Collision Term

Here it is convenient to distinguish between the inter-band and intra-band transitions, introduced in Sects. 19.3, 19.3.2. Thus, the collision term is written $C = C_b + C_v$, where as above suffix b (v) stands for "inter-band" ("intra-band"). This yields

$$\iiint_{-\infty}^{+\infty} \alpha\, C\, \mathrm{d}^3k = W_b + W_v\,, \qquad W_{b(v)}[\alpha] = \iiint_{-\infty}^{+\infty} \alpha\, C_{b(v)}\, \mathrm{d}^3k\,, \tag{19.61}$$

where the functional symbol reminds one that $W_{b(v)}$ is determined by the form of α.

19.4.5 Moment Equations

Adding up (19.53), (19.54), (19.56), (19.59), and (19.61) after multiplying the drift terms by $q/\hbar$ provides the explicit form of (19.51), that reads

$$\frac{\partial}{\partial t}\,(n\,\overline{\alpha}) + \mathrm{div}_{\mathbf{r}}\,(n\,\overline{\alpha\,\mathbf{u}}) + \frac{q}{\hbar}\,n\,\overline{\mathrm{grad}_{\mathbf{k}}\alpha \cdot (\mathbf{E} + \mathbf{u} \wedge \mathbf{B})} = W_b[\alpha] + W_v[\alpha]\,. \tag{19.62}$$

A simple reasoning shows that the equations of the form (19.62) that are obtained from different choices of α are coupled with each other. To show this one takes for simplicity the one-dimensional case and let $\alpha = c\,k^m$, with m a positive integer and c a constant. Also, the parabolic-band approximation is assumed to hold, so that $\mathbf{u}$ is a linear function of k. It follows that the time derivative in (19.62) contains the moment of order m of f, while the diffusion term (due to the product $\alpha\,\mathbf{u}$) contains the moment of order $m + 1$; in turn, the summand proportional to $\mathbf{E}$ in the drift term contains the moment of order $m - 1$ due to the derivative of α, while the summand proportional to $\mathbf{B}$ contains the moment of order m. These considerations are sufficient to show that the equation whose unknown is the moment of order m is coupled with those whose unknowns are the moment of order $m - 1$ and $m + 1$.

In the typical applications of the moments expansion a finite set of equations is considered, starting from the lowest-order moment $m = 0$ up to some order m_0. The system of equations thus obtained is indeterminate because, due to the coupling mentioned above, the number of equations is m_0 whereas the number of unknown moments appearing in them is $m_0 + 1$. To make the system determinate it is then necessary to add an extra condition, that is found by prescribing an approximate form of the $(m_0 + 1)$th moment.[22] Such a prescription reduces the number of unknown moments to m_0 and makes the system of differential equations determinate; for this reason it is called *closure condition*. The typical choice for the closure condition is to approximate the $(m_0 + 1)$th moment using the equilibrium distribution.

Moment of Order Zero

The *moment of order zero* is obtained by letting $\alpha = 1$ in (19.62), whose left-hand side becomes $\partial n/\partial t + \mathrm{div}_{\mathbf{r}}\,(n\,\overline{\mathbf{u}})$; the electron concentration n is the moment of order zero and $\overline{\mathbf{u}} = \mathbf{v}$ is the average velocity, as defined in (19.31). If the zero-order moment of the collision term does not introduce further unknowns, the equation's unknowns are two: n and $\mathbf{v}$. The form of the left-hand side shows that integrating the zero-order moment of the BTE over an arbitrary volume Ω of the $\mathbf{r}$ space provides the balance equation for the number of electrons of the conduction band (compare with Sect. 23.2):

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\Omega} n\,\mathrm{d}\Omega + \int_{\Sigma} n\,\mathbf{v}\cdot\mathbf{s}\,\mathrm{d}\Sigma = \int_{\Omega} W_b[1]\,\mathrm{d}\Omega + \int_{\Omega} W_v[1]\,\mathrm{d}\Omega\,, \tag{19.63}$$

where Σ is the boundary of Ω and $\mathbf{s}$ the unit vector normal to Σ, oriented in the outward direction. The second integral at the right-hand side of (19.63) does not contribute to the electrons' balance; in fact, it describes transitions that do not influence the number of electrons of the band because both the initial and final state belong to it. On the other hand, due to the arbitrariness of Ω, the integral vanishes only if $W_v[1] = 0$. This result shows that the zero-order moment of the intra-band transitions vanishes,[23] so that the only transitions of importance for the zero-order moment, despite being dominated by much larger relaxation times, are the inter-band ones. This is the exception anticipated in Sect. 19.3.2. The zero-order moment for the electrons of the conduction band then reads

[22] The choice of the highest-order moment as the function to be approximated is reasonable in view of the analysis of the moments method carried out in Sect. C.6. In fact, as the moments are the coefficients of a converging Taylor series, they become smaller and smaller as the order increases; thus, the error due to approximating the highest-order coefficient is expected to be the smallest.

[23] A similar reasoning is used to explain (20.16).

$$\frac{\partial n}{\partial t} + \mathrm{div}_{\mathbf{r}}\,(n\,\mathbf{v}) = W_b[1]\,. \tag{19.64}$$

The form of the inter-band term $W_b[1]$, which is not relevant for the analysis in hand, is worked out in Chap. 20.

General Form of the Higher-Order Moments

As shown above, the contribution of the intra-band transitions vanishes for $\alpha = 1$. In contrast, it becomes dominant for other choices of α; in fact, in such cases the intra-band transitions do not cancel out any more and their scattering rates turn out to be much higher than those of the inter-band transitions. This allows one to adopt an approximation for $W_b[\alpha]$, namely,

$$W_b[\alpha] = \iiint_{-\infty}^{+\infty} \alpha\, C_b\, \mathrm{d}^3k \simeq \overline{\alpha} \iiint_{-\infty}^{+\infty} C_b\, \mathrm{d}^3k\,. \tag{19.65}$$

In other terms it is assumed that since the contribution of $W_b[\alpha]$ to the collision term is small when $\alpha \neq 1$, the error introduced by (19.65) is negligible. Expanding the time derivative in (19.62) and using (19.64), (19.65) yield

$$n\,\frac{\partial \overline{\alpha}}{\partial t} + \mathrm{div}_{\mathbf{r}}\,(n\,\overline{\alpha}\,\overline{\mathbf{u}}) - \overline{\alpha}\,\mathrm{div}_{\mathbf{r}}\,(n\,\mathbf{v}) + \frac{q}{\hbar}\,n\,\overline{\mathrm{grad}_{\mathbf{k}}\alpha \cdot (\mathbf{E} + \mathbf{u} \wedge \mathbf{B})} = W_v[\alpha]\,, \tag{19.66}$$

where only the intra-band transitions appear. Due to its simpler form, (19.66) will be used in the following to derive the balance equations with $\alpha \neq 1$.

Moments of Order One, Two, and Three

The *moment of order one* of the BTE is found by letting $\alpha = u_i$ with $i = 1, 2, 3$ in (19.66); this yields the continuity equation for the ith component of the average velocity of the electrons, $\overline{u}_i = v_i$:

$$n\,\frac{\partial v_i}{\partial t} + \mathrm{div}_{\mathbf{r}}\,(n\,\overline{u_i\,\mathbf{u}}) - v_i\,\mathrm{div}_{\mathbf{r}}\,(n\,\mathbf{v}) + \frac{q}{\hbar}\,n\,\overline{\mathrm{grad}_{\mathbf{k}}u_i \cdot (\mathbf{E} + \mathbf{u} \wedge \mathbf{B})} = W_v[u_i]\,. \tag{19.67}$$

To proceed it is necessary to introduce the definition of average kinetic energy and average flux of the electrons' kinetic energy,[24]

[24] In the equilibrium condition the product $E_e f^{\mathrm{eq}}$ is even with respect to $\mathbf{k}$. In turn, $\mathbf{u} = (1/\hbar)\,\mathrm{grad}_{\mathbf{k}}E$ is odd, so that $\mathbf{b}^{\mathrm{eq}} = 0$. Compare with the similar comment made about the average velocity in (19.31).

$$w(\mathbf{r},t) = \frac{1}{n} \iiint_{-\infty}^{+\infty} E_e(\mathbf{k})\, f(\mathbf{r},\mathbf{k},t)\, \mathrm{d}^3k\,, \tag{19.68}$$

$$\mathbf{b}(\mathbf{r},t) = \frac{1}{n} \iiint_{-\infty}^{+\infty} E_e(\mathbf{k})\, \mathbf{u}(\mathbf{k})\, f(\mathbf{r},\mathbf{k},t)\, \mathrm{d}^3k\,, \tag{19.69}$$

with $E_e = E(\mathbf{k}) - E_C$. Then, the *moment of order two* of the BTE is found by letting $\alpha = E_e$ in (19.66); this yields the continuity equation for the average kinetic energy of the electrons, (19.68). In the derivation, the term containing the magnetic induction vanishes due to (19.60); using the definition (17.52) of the group velocity, the equation reads

$$n\,\frac{\partial w}{\partial t} + \mathrm{div}_{\mathbf{r}}\,(n\,\mathbf{b}) - w\,\mathrm{div}_{\mathbf{r}}\,(n\,\mathbf{v}) + q\,n\;\mathbf{v}\cdot\mathbf{E} = W_v[E_e]\,. \tag{19.70}$$

The *moment of order three* of the BTE is found by letting $\alpha = E_e\,u_i$ with $i = 1, 2, 3$ in (19.66); this yields the continuity equation for the average flux of the electrons' kinetic energy, (19.69); the equation reads

$$n\,\frac{\partial b_i}{\partial t} + \mathrm{div}_{\mathbf{r}}\left(n\,\overline{E_e\,u_i\,\mathbf{u}}\right) - b_i\,\mathrm{div}_{\mathbf{r}}\,(n\,\mathbf{v}) + \frac{q}{\hbar}\,n\;\overline{\mathrm{grad}_{\mathbf{k}}(E_e\,u_i)\cdot(\mathbf{E}+\mathbf{u}\wedge\mathbf{B})} = W_v[E_e\,u_i]\,. \tag{19.71}$$

The choices $\alpha = 1$, $\alpha = u_i$, $\alpha = E_e$, and $\alpha = E_e\,u_i$ are such that each moment equation provides the balance relation of a dynamic quantity of interest: number of electrons, average velocity, average kinetic energy, average flux of the kinetic energy; the even-order moments yield a scalar equation, whereas the odd-order moments yield a vector equation.

19.4.6 Hierarchical Models

The order-one moment (19.67) contains the new unknown $\overline{u_i\,\mathbf{u}}$ besides n and $\mathbf{v}$ already present in (19.64); the order-two moment (19.70) contains again n and $\mathbf{v}$, and the new unknowns w, $\mathbf{b}$. The order-three moment contains n, $\mathbf{v}$, $\mathbf{b}$, and the new unknown $\overline{E_e\,u_i\,\mathbf{u}}$. The drift terms and the collision terms, depending on their form, may or may not introduce extra unknowns; even if they don't, the number of unknowns listed above exceeds that of the equations. It is worth anticipating that the finite set of balance equations indicated in Sect. 19.4.5 is obtained by taking the equations in pairs: specifically, the first pair is made of the balance equations of order zero and one, (19.64) and (19.66), that are collectively termed *drift-diffusion model*; in this case, the three unknowns n, $\mathbf{v}$, and $\overline{u_i\,\mathbf{u}}$ are reduced to two by the closure condition (Sect. 19.4.5), that consists in replacing the highest-order moment $\overline{u_i\,\mathbf{u}}$ with its equilibrium expression. A more elaborate model is obtained by taking the first two pairs, namely, the balance equations of order zero

through three, (19.64), (19.66), (19.71), and (19.71), that are collectively termed *hydrodynamic model*; the six unknowns n, $\mathbf{v}$, and $\overline{u_i\,\mathbf{u}}$, w, $\mathbf{b}$, $\overline{E_e\,u_i\,\mathbf{u}}$ are reduced to five by prescribing the closure condition, then to four by determining a relation between the second-order moments $\overline{u_i\,\mathbf{u}}$ and w.

By this procedure one constructs a set of hierarchically ordered models of increasing complexity. The type of model adopted in practical applications depends on the trade-off between the information that one needs to acquire about the physical behavior of the device under investigation and the computational cost of the system of differential equations to be solved. To date, the moments method has been investigated up to order 6 [60], and has been extended to order 21 using a scheme based on Legendre polynomial expansion [75]; the standard implementations in the commercial simulation programs used by semiconductor Companies adopt the hydrodynamic model.[25]

The balance equations derived so far are still rather cumbersome in view of the application to the analysis of semiconductor devices. A number of simplifications are illustrated below, which eventually lead to the standard form of the hydrodynamic model [50, 54, 111, 116]. To begin, one considers the time derivatives at the left-hand side of the balance equations. Such derivatives differ from zero only if the distribution function depends explicitly on time, which typically happens when time-dependent boundary conditions are imposed to the device under investigation. In the practical cases, the maximum frequency of the electric signals applied to a device or an integrated circuit is lower by many orders of magnitude than the inverse relaxation times associated with the intra-band transitions; this makes it possible to neglect the time derivatives of v_i, w, and b_i. A quasi-static approximation[26] is thus assumed in the continuity equations (19.67), (19.70), and (19.71). The argument leading to this approximation does not apply to the case of (19.64) because only the inter-band transitions take place there, whose relaxation times are much longer than those of the intra-band transitions and, in many cases, also than the inverse maximum frequency of the external signal. As a consequence, the term $\partial n/\partial t$ in (19.64) must be retained when the boundary conditions depend on time.

As a second approximation, one adopts the parabolic-band approximation; this implies that in a nonequilibrium condition the electrons of the conduction band still occupy energy states in the vicinity of the absolute minima. Such a condition is in general fulfilled as shown below.[27] Letting a indicate one of the absolute minima of the conduction band and using (19.24) after dropping suffix "0" yield

[25] Comprehensive reviews of the solution methods for the BTE are in [71, 72] as far as the Monte Carlo method is concerned, and in [64] for deterministic methods.

[26] A similar reasoning is used to treat the time derivative of the vector potential when the semiconductor equations are coupled with the Maxwell equations (Sect. 19.5.4).

[27] The adoption of the parabolic-band approximation may be avoided at the cost of redefining the carrier temperature and introducing more relaxation times [134].

$$u_i = \frac{\hbar\,(k_i - k_{ia})}{m_{ia}}\,, \quad \mathrm{grad}_{\mathbf{k}} u_i = \frac{\hbar}{m_{ia}}\,\mathbf{i}_i\,, \quad \mathrm{grad}_{\mathbf{k}}(E_e\,u_i) = \hbar\left(\frac{E_e}{m_{ia}}\,\mathbf{i}_i + u_i\,\mathbf{u}\right), \tag{19.72}$$

with $\mathbf{i}_i$ the unit vector of the ith axis. From now on, the equations derived from the parabolic-band approximation refer to the ath valley. In principle, the electron concentration, average velocity, and the other averages should be indicated with n_a, $\mathbf{v}_a$, and so on; this is not done here to avoid complicacies in the notation. The suffix will be introduced in Sect. 19.5.2, where the contributions of the valleys are summed up. This comment does not apply to the moment of order zero, (19.64), because its derivation does not entail any simplifying hypothesis.

With these premises, one manipulates the terms $\overline{u_i\,\mathbf{u}}$, $E_e\,\overline{u_i\,\mathbf{u}}$ by introducing the auxiliary quantity $\mathbf{c} = \mathbf{u} - \mathbf{v}$, called *random velocity*. Clearly it is $\overline{c}_i = \overline{u}_i - v_i = 0$, so that $\overline{u_i\,\mathbf{u}} = v_i\,\mathbf{v} + \overline{u_i\,\mathbf{u}}$; it follows

$$\mathrm{div}_{\mathbf{r}}(n\,\overline{u_i\,\mathbf{u}}) = \mathrm{div}_{\mathbf{r}}(n\,\overline{c_i\,\mathbf{c}}) + v_i\,\mathrm{div}_{\mathbf{r}}(n\,\mathbf{v}) + n\,\mathbf{v}\cdot\mathrm{grad}_{\mathbf{r}} v_i\,. \tag{19.73}$$

The last term at the right-hand side of (19.73) is called *convective term*. In the typical operating conditions of the semiconductor devices this term can be neglected (refer to [108] and the comments below). Replacing into (19.67) the simplified form of (19.73) along with the second relation of (19.72) yields

$$\mathrm{div}_{\mathbf{r}}\,(n\,m_{ia}\,\overline{c_i\,\mathbf{c}}) + q\,n\;\;(\mathbf{E} + \mathbf{v}\wedge\mathbf{B})_i = m_{ia}\,W_v[u_i]\,. \tag{19.74}$$

The latter equation contains the unknowns n and $\mathbf{v}$ already present in (19.64), and the new unknown $\overline{c_i\,\mathbf{c}}$. The drift term does not introduce extra unknowns. Note that $\overline{c_i\,\mathbf{c}}$ is actually made of three vectors, so that it may be thought of as a symmetric 3×3 tensor with components $\overline{c_i\,c_j}$, $i,j = 1,2,3$. To give it a more compact form, after observing that $m_{ia}\,\overline{c_i\,\mathbf{c}}$ has the units of an energy, one defines the *electron-temperature tensor* of components T_{ij} such that

$$k_B\,T_{ij} = m_{ia}\,\overline{c_i\,c_j}\,, \qquad n\,k_B\,T_{ij} = \iiint_{-\infty}^{+\infty} m_{ia}\,c_i\,c_j\,f\,\mathrm{d}^3k\,, \tag{19.75}$$

with k_B the Boltzmann constant. Letting $\mathbf{T}_i$ be the vector of entries T_{i1}, T_{i2}, T_{i3}, so that $k_B\,\mathbf{T}_i = m_{ia}\,\overline{c_i\,\mathbf{c}}$, one finds for the ith component of the moment of order one

$$\mathrm{div}_{\mathbf{r}}\,(n\,k_B\,\mathbf{T}_i) + q\,n\;\;(\mathbf{E} + \mathbf{v}\wedge\mathbf{B})_i = m_{ia}\,W_v[u_i]\,. \tag{19.76}$$

In the equilibrium condition the electron-temperature tensor reduces to the product of a scalar coefficient times the identity tensor; in the limit of the Boltzmann distribution, the scalar coefficient identifies with the lattice temperature (Sect. 19.6.4), thus providing an estimate of the modulus $|\mathbf{c}|$ of the random velocity. The modulus $|\mathbf{v}|$ of the average velocity in a nonequilibrium condition can be estimated as well, basing upon the current density and carrier concentration of the devices' operating conditions. It is found that in typical situations it is $|\mathbf{v}| \ll |\mathbf{c}|$, so that the average

motion of the carriers in a nonequilibrium condition can be thought of as that of a slowly drifting fluid. This justifies the neglect of the convective term in (19.73), and also allows one to neglect $v_i\,\mathbf{v}$ with respect to $\overline{c_i\,\mathbf{c}}$ when these terms appear in the same expression.

The simplifications used in (19.67) apply in the same manner to the moment of order three, (19.71); in fact it is $\overline{E_e\,u_i\,\mathbf{u}} = \overline{E_e\,u_i\,\mathbf{c}} + \overline{E_e\,u_i}\;\mathbf{v}$, whence

$$\mathrm{div}_{\mathbf{r}}(n\,\overline{E_e\,u_i\,\mathbf{u}}) = \mathrm{div}_{\mathbf{r}}(n\,\overline{E_e\,u_i\,\mathbf{c}}) + b_i\,\mathrm{div}_{\mathbf{r}}(n\,\mathbf{v}) + n\,\mathbf{v}\cdot\mathrm{grad}_{\mathbf{r}}\,b_i. \tag{19.77}$$

Using the third relation of (19.72) in the drift term of (19.71) transforms the latter into $(q/m_{ia})\,n\;[(w\,\mathbf{i}_i + m_{ia}\,v_i\,\mathbf{v} + m_{ia}\,\overline{c_i\,\mathbf{c}})\cdot\mathbf{E} + \mathbf{b}\wedge\mathbf{B}\cdot\mathbf{i}_i + m_{ia}\,\overline{u_i\,\mathbf{u}\cdot\mathbf{u}\wedge\mathbf{B}}]$, where the mixed product vanishes due to the repeated factor, and $m_{ia}\,v_i\,\mathbf{v}$ is negligible as shown above. Replacing (19.77) into (19.71) after neglecting the time derivative and the convective term yields

$$\mathrm{div}_{\mathbf{r}}\left(n\,m_{ia}\,\overline{E_e\,u_i\,\mathbf{c}}\right) + q\,n\;\;[(w\,\mathbf{i}_i + k_B\,\mathbf{T}_i)\cdot\mathbf{E} + \mathbf{b}\wedge\mathbf{B}\cdot\mathbf{i}_i] = m_{ia}\,W_v[E_e\,u_i]. \tag{19.78}$$

The relation between the second-order moments necessary to reduce the number of unknowns is now determined starting from the expression of E_e in the parabolic-band approximation. Using the first relation in (19.72) yields $E_e = (1/2)\,\sum_{i=1}^{3} m_{ia}\,u_i^2$ whence, from (19.68),

$$w = \frac{1}{2}\sum_{i=1}^{3} m_{ia}\,v_i^2 + \frac{3}{2}\,k_B\,T_e\,, \qquad T_e = \frac{T_{11}+T_{22}+T_{33}}{3}, \tag{19.79}$$

with T_e the *electron temperature*. The two summands of w in (19.79) are also called *convective part* and *thermal part* of the average kinetic energy, respectively. The same reasoning that has led to the neglect of $m_{ia}\,v_i\,\mathbf{v}$ with respect to $m_{ia}\,c_i\,\mathbf{c}$ also shows that the thermal part is dominant with respect to the convective part, so that $w \simeq (3/2)\,k_B\,T_e$. Moreover, it can be assumed that the electron-temperature tensor retains the same structure of the equilibrium case, so that

$$\begin{bmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \end{bmatrix} \simeq \begin{bmatrix} T_{11} & 0 & 0 \\ 0 & T_{22} & 0 \\ 0 & 0 & T_{33} \end{bmatrix} \simeq T_e(\mathbf{r},t)\,\mathscr{I}, \tag{19.80}$$

with $\mathscr{I}$ the identity tensor. As a consequence, $(w\,\mathbf{i}_i + k_B\,\mathbf{T}_i)\cdot\mathbf{E} = (5/2)\,k_B\,T_e\,\mathbf{i}_i\cdot\mathbf{E}$ and $\mathrm{div}_{\mathbf{r}}(n\,k_B\,\mathbf{T}_i) = \partial(n\,k_B\,T_e)/\partial x_i$. The latter is the ith component of $\mathrm{grad}_{\mathbf{r}}\,(n\,k_B\,T_e)$. In summary, the balance equations for the moments of order one, two, and three read

$$\frac{\partial\,(n\,k_B\,T_e)}{\partial x_i} + q\,n\;\;(\mathbf{E}+\mathbf{v}\wedge\mathbf{B})_i = m_{ia}\,W_v[u_i]\,, \tag{19.81}$$

$$\mathrm{div}_{\mathbf{r}}\,(n\,\mathbf{b}) - (3/2)\,k_B\,T_e\,\mathrm{div}_{\mathbf{r}}\,(n\,\mathbf{v}) + q\,n\;\mathbf{v}\cdot\mathbf{E} = W_v[E_e]\,, \tag{19.82}$$

$$\mathrm{div}_{\mathbf{r}}\left(n\,m_{ia}\,\overline{E_e\,u_i\,\mathbf{c}}\right) + q\,n\;\;[(5/2)\,k_B\,T_e\,\mathbf{E} + \mathbf{b}\wedge\mathbf{B}]_i = m_{ia}\,W_v[E_e\,u_i]\,. \tag{19.83}$$

Macroscopic Relaxation Times of the Higher-Order Moments

As remarked in Sect. (19.4.5), the collision terms of the moments of order higher than zero account for the intra-band transitions only. These terms are worked out here using the perturbative form of the BTE (Sect. 19.3.3); this approach is coherent with the other approximations from which the balance equations (19.81–19.83) derive. The collision term of (19.82) then becomes

$$W_v[E_e] = -\iiint_{-\infty}^{+\infty} E_e \frac{f - f^{\mathrm{eq}}}{\tau} \, \mathrm{d}^3k \,, \tag{19.84}$$

with $\tau \simeq \tau_v$ (Sect. 19.3.2). The equilibrium part is worked out by defining the *energy-relaxation time* τ_w such that

$$\iiint_{-\infty}^{+\infty} E_e \frac{f^{\mathrm{eq}}}{\tau_v} \, \mathrm{d}^3k = \frac{1}{\tau_w} \iiint_{-\infty}^{+\infty} E_e f^{\mathrm{eq}} \, \mathrm{d}^3k = \frac{n^{\mathrm{eq}} w^{\mathrm{eq}}}{\tau_w} \simeq \frac{3}{2} \frac{n^{\mathrm{eq}} k_B T_e^{\mathrm{eq}}}{\tau_w} \,, \tag{19.85}$$

where the definitions (19.68), (19.79) of the electrons' average kinetic energy and temperature are used. The left-hand side of (19.85) does not vanish because the integrand is positive definite. The nonequilibrium part of $W_v[E_e]$ is approximated as

$$\iiint_{-\infty}^{+\infty} E_e \frac{f}{\tau_v} \, \mathrm{d}^3k \simeq \frac{1}{\tau_w} \iiint_{-\infty}^{+\infty} E_e f \, \mathrm{d}^3k = \frac{n w}{\tau_w} \simeq \frac{3}{2} \frac{n k_B T_e}{\tau_w} \,, \tag{19.86}$$

based on the observation that due to the smallness of the intra-band relaxation time τ_v, the distribution function departs little from the equilibrium one (Sect. 19.3.3).

The derivation of the analogues of τ_w for the collision terms of (19.81) and (19.83) is somewhat more complicate. In fact, in most semiconductors, among which Si, Ge, and GaAs, the relaxation time τ_v is even with respect to $\mathbf{k}$ [73], which makes the integrals of $u_i f^{\mathrm{eq}}/\tau_v$ and $E_e u_i f^{\mathrm{eq}}/\tau_v$ to vanish because the integrand is odd. To overcome the difficulty one expands $f - f^{\mathrm{eq}}$ into a Taylor series with respect to a parameter and truncates the series in such a way as to retain the first summand which is odd with respect to $\mathbf{k}$. For instance, letting λ be the parameter[28] and assuming that the first-order term of the expansion is odd, one lets $f - f^{\mathrm{eq}} \simeq (\mathrm{d}f/\mathrm{d}\lambda)^{\mathrm{eq}} \, \lambda$ whence

$$\iiint_{-\infty}^{+\infty} u_i \frac{(\mathrm{d}f/\mathrm{d}\lambda)^{\mathrm{eq}}}{\tau_v} \, \mathrm{d}^3k = \frac{1}{\tau_{pi}} \iiint_{-\infty}^{+\infty} u_i \, (\mathrm{d}f/\mathrm{d}\lambda)^{\mathrm{eq}} \, \mathrm{d}^3k \,, \tag{19.87}$$

$$\iiint_{-\infty}^{+\infty} E_e \, u_i \frac{(\mathrm{d}f/\mathrm{d}\lambda)^{\mathrm{eq}}}{\tau_v} \, \mathrm{d}^3k = \frac{1}{\tau_{bi}} \iiint_{-\infty}^{+\infty} E_e \, u_i \, (\mathrm{d}f/\mathrm{d}\lambda)^{\mathrm{eq}} \, \mathrm{d}^3k \,, \tag{19.88}$$

[28] Typically the parameter used in this procedure is the electric field [73]. An expansion truncated to the first order is coherent with the first-order perturbation approach.

with τ_{pi} and τ_{bi} the *momentum-relaxation time* and *relaxation time of the energy flux*, respectively.[29] Due to their definitions, τ_{pi} and τ_{bi} are diagonal tensors. However, as their degree of anisotropy is small, they are approximated by scalar quantities, $\tau_{pi} \simeq \tau_p$ and $\tau_{bi} \simeq \tau_b$. Investigations about the relaxation times have been carried out by different techniques, specifically, the spherical-harmonics expansion method to determine the dependence on the average energy [112, 116], and the Monte Carlo method to study the anisotropy properties [17, 18, 57].

Using (19.87), (19.88) along with the definitions (19.31), (19.69) of the average velocity and average flux of kinetic energy finally yields

$$\iiint_{-\infty}^{+\infty} \frac{u_i}{\tau_v} \left(\frac{\mathrm{d}f}{\mathrm{d}\lambda}\,\lambda\right)^{\mathrm{eq}} \mathrm{d}^3k = \frac{n\,v_i}{\tau_p}\,, \qquad \iiint_{-\infty}^{+\infty} \frac{E_e\,u_i}{\tau_v} \left(\frac{\mathrm{d}f}{\mathrm{d}\lambda}\,\lambda\right)^{\mathrm{eq}} \mathrm{d}^3k = \frac{n\,b_i}{\tau_b}\,. \tag{19.89}$$

19.5 Hydrodynamic and Drift-Diffusion Models

In Sect. 19.4 the moments method has been applied to derive a set of balance equations; the general form of the latter has successively been modified by introducing a number of simplifications: among them is the parabolic-band approximation, due to which, as indicated in Sect. 19.4.6, a set of equations restricted to the ath valley of the conduction band is obtained. In order to recover the equations for the whole band, it is necessary to add up the single-valley contributions. The procedure is the same for the hydrodynamic and drift-diffusion models; it will be worked out explicitly only for the simpler case of the drift-diffusion model.

19.5.1 HD Model

As anticipated in Sect. 19.4.6, the hydrodynamic (HD) model is obtained by taking the balance equations of order zero through three, (19.64), (19.81), (19.82), and (19.83), and imposing the closure condition onto the fourth-order moment. For simplicity, the latter is considered in the nondegenerate case whence, from (19.159), it is $m_{ia}\,\left(\overline{E_e\,u_i\,\mathbf{c}}\right)^{\mathrm{eq}} \simeq \mathbf{i}_i\,(5/2)\,(k_B\,T)^2$. Letting $W = W_b[1]$ and using (19.85), (19.86), (19.89) yield

$$\frac{\partial n}{\partial t} + \mathrm{div}_{\mathbf{r}}\,(n\,\mathbf{v}) = W\,, \qquad \frac{\partial\,(n\,k_B\,T_e)}{\partial x_i} + q\,n\;\left(\mathbf{E} + \mathbf{v} \wedge \mathbf{B}\right)_i = -\frac{m_{ia}}{\tau_p}\,n\,v_i\,, \tag{19.90}$$

[29] The term "momentum" for τ_{pi} derives from the observation that the continuity equation for the ith component of the average velocity v_i of the electrons (19.67) may also be thought of as the continuity equation for the ith component of the average momentum, $m_{ia}\,v_i$. In turn, τ_{bi} is also called *heat-relaxation time*.

$$\mathrm{div}_{\mathbf{r}}\,(n\,\mathbf{b}) - \frac{3}{2}\,k_B\,T_e\,\mathrm{div}_{\mathbf{r}}\,(n\,\mathbf{v}) + q\,n\;\mathbf{v}\cdot\mathbf{E} = -\frac{3}{2}\,\frac{k_B}{\tau_w}\,[n\,T_e - (n\,T_e)^{\mathrm{eq}}]\;, \tag{19.91}$$

$$\frac{5}{2}\,(k_B\,T)^2\,\frac{\partial n}{\partial x_i} + q\,n\,\left(\frac{5}{2}\,k_B\,T_e\,\mathbf{E} + \mathbf{b}\wedge\mathbf{B}\right)_i = -\frac{m_{ia}}{\tau_b}\,n\,b_i\,. \tag{19.92}$$

which constitute a system of first-order, partial-differential equations in the unknowns n, $\mathbf{v}$, T_e, and $\mathbf{b}$. In general, the model's equations are to be solved over a volume that encloses the device under investigation; the boundary conditions that typically apply are discussed in Sect. 19.5.6. Two of the equations are scalar (namely, (19.91) and the first one in (19.90)), while the other two are vector equations. The system is nonlinear because the unknowns are multiplied by each other.[30] Note, however, that the second equation in (19.90) is linear with respect to the components of $\mathbf{v}$; the latter can be extracted and replaced into the two scalar equations. The same procedure is applicable to (19.92), which is linear with respect to the components of $\mathbf{b}$. After the replacements are completed, the system reduces to two scalar equations of the second order. An example is given in Sect. 19.5.5, with reference to the simpler case of the drift-diffusion model. Due to the components m_{ia} of the effective-mass tensor, the vector equations are anisotropic; however, when the contributions of the different valleys are combined together, the anisotropy cancels out (the explicit calculation is provided for the drift-diffusion model below).

The qualitative analysis of the model carried out above implies that the electric field and magnetic induction are known, so that they are embedded in the model's coefficients. In fact, this is not true, because the fields are influenced by the distribution of electric charge and current density that are some of the model's unknowns. For this reason, as shown below, the hydrodynamic equations, and the drift-diffusion ones as well, must be coupled with the Maxwell equations.

19.5.2 DD Model

The drift-diffusion (DD) model is obtained by taking the balance equations of order zero and one, (19.90), and imposing the closure condition onto the second-order moment. For simplicity, the latter is considered in the nondegenerate case, whence $T_e^{\mathrm{eq}} = T$ (Sect. 19.6.4); the model thus reads

$$\frac{\partial n}{\partial t} + \mathrm{div}_{\mathbf{r}}\,(n\,\mathbf{v}) = W\,, \qquad k_B\,T\,\frac{\partial n_a}{\partial x_i} + q\,n_a\;\;(\mathbf{E} + \mathbf{v}_a\wedge\mathbf{B})_i = -\frac{m_{ia}}{\tau_p}\,n_a\,v_{ia}\,. \tag{19.93}$$

[30] Also, the generation-recombination term W embeds nonlinear dependencies on some of the unknowns, Chap. 20.

As indicated in Sect. 19.4.6, the first equation in (19.93) refers to the whole conduction band because its derivation did not entail any simplifying hypothesis; in contrast, the second equation refers to the ath minimum of the band due to the parabolic-band approximation. This explains the index attached to n and to the average velocity; the momentum-relaxation time, instead, does not depend on the valley [112]. As noted above, the dependence on the components v_{ia} is linear, which makes it possible to express them in terms of the other functions. In fact, it is more convenient to extract, instead of $\mathbf{v}_a$, the electron-current density of the ath minimum; remembering (4.21) and (4.22), the latter is given by $\mathbf{J}_a = -q\, n_a\, \mathbf{v}_a$. Then, the second equation in (19.93) is recast as

$$J_{ia} = J'_{ia} - \frac{q\,\tau_p}{m_{ia}}\,(\mathbf{J}_a \wedge \mathbf{B})_i\,, \qquad J'_{ia} = k_B\,T\,\frac{q\,\tau_p}{m_{ia}}\,\frac{\partial n_a}{\partial x_i} + \frac{q\,\tau_p}{m_{ia}}\,q\,n_a\,(\mathbf{E})_i\,, \tag{19.94}$$

with $J'_{ia} = J_{ia}(\mathbf{B} = 0)$. Letting $\mu_{ia} = q\,\tau_p/m_{ia}$, the matrix form of (19.94) reads

$$\begin{bmatrix} J_{1a} \\ J_{2a} \\ J_{3a} \end{bmatrix} = \begin{bmatrix} J'_{1a} \\ J'_{2a} \\ J'_{3a} \end{bmatrix} - \begin{bmatrix} \mu_{1a} & 0 & 0 \\ 0 & \mu_{2a} & 0 \\ 0 & 0 & \mu_{3a} \end{bmatrix} \begin{bmatrix} J_{2a}\,B_3 - J_{3a}\,B_2 \\ J_{3a}\,B_1 - J_{1a}\,B_3 \\ J_{1a}\,B_2 - J_{2a}\,B_1 \end{bmatrix}, \tag{19.95}$$

equivalent to

$$\begin{bmatrix} 1 & \mu_{1a}\,B_3 & -\mu_{1a}\,B_2 \\ -\mu_{2a}\,B_3 & 1 & \mu_{2a}\,B_1 \\ \mu_{3a}\,B_2 & -\mu_{3a}\,B_1 & 1 \end{bmatrix} \begin{bmatrix} J_{1a} \\ J_{2a} \\ J_{3a} \end{bmatrix} = \begin{bmatrix} J'_{1a} \\ J'_{2a} \\ J'_{3a} \end{bmatrix}. \tag{19.96}$$

The diagonal tensor $\hat{\mu}_a$ of entries μ_{ia} is called *mobility tensor* of the ath valley. Note that the product of a mobility by a magnetic induction is dimensionless. Letting $\mathbf{M}_a = \mu_{1a}\,\mu_{2a}\,\mu_{3a}\,(\hat{\mu}_a)^{-1}\,\mathbf{B}$, the components of the current density are found by solving the algebraic system (19.96), where the determinant of the matrix is

$$D_M = 1 + \mu_{1a}\,\mu_{2a}\,\mu_{3a}\left(\frac{B_1^2}{\mu_{1a}} + \frac{B_2^2}{\mu_{2a}} + \frac{B_3^2}{\mu_{3a}}\right) = 1 + \mathbf{B}\cdot\mathbf{M}_a\,. \tag{19.97}$$

The components of $\mathbf{J}_a$ are finally found to be

$$D_M\,J_{ia} = J'_{ia} + \mu_{ia}\,\left(\mathbf{B}\wedge\mathbf{J}'_a\right)_i + \left(\mathbf{M}_a\cdot\mathbf{J}'_a\right)\,B_i\,. \tag{19.98}$$

In typical situations the modulus of the magnetic induction is small, so that terms that are quadratic in the components of $\mathbf{B}$ may be neglected. This yields the approximate form

$$J_{ia} \simeq J'_{ia} + \mu_{ia}\,\left(\mathbf{B}\wedge\mathbf{J}'_a\right)_i\,. \tag{19.99}$$

The electron current density of the whole conduction band is thus found as

$$\mathbf{J} = \sum_{a=1}^{M_C} \mathbf{J}_a = \sum_{a=1}^{M_C} \left[k_B\, T\, \hat{\mu}_a\, \mathrm{grad} n_a + \hat{\mu}_a\, q\, n_a\, \mathbf{E} + \hat{\mu}_a\, \left(\mathbf{B} \wedge \mathbf{J}'_a \right) \right] . \tag{19.100}$$

In the perturbative approach followed here, it can be assumed that the total electron concentration n equally distributes[31] over the valleys, $n_a = n/M_C$. The first two summands at the right-hand side of (19.100) then yield

$$\mathbf{J}' = \sum_{a=1}^{M_C} \mathbf{J}'_a = q\, \hat{\mu}_n\, n\, \mathbf{E} + q\, \hat{D}_n\, \mathrm{grad} n\,, \tag{19.101}$$

where the diagonal tensors $\hat{\mu}_n$, $\hat{D}_n$ are defined as

$$\hat{\mu}_n = \frac{1}{M_C} \sum_{a=1}^{M_C} \hat{\mu}_a\,, \qquad \hat{D}_n = \frac{k_B\, T}{q}\, \hat{\mu}_n\,. \tag{19.102}$$

They are called *electron-mobility tensor* and *electron-diffusivity tensor*, respectively. The second relation in (19.102), that states that diffusivity and mobility are proportional through $k_B\, T/q$, is called *Einstein relation*.[32] The form of $\hat{\mu}_n$ is specified on a case-by-case basis, depending on the semiconductor under consideration. Taking silicon by way of example ($M_C = 6$), the mass tensor is obtained from (17.82); thus, the mobility tensor $\hat{\mu}_a$ has one of the following forms:

$$\begin{bmatrix} \mu_l & 0 & 0 \\ 0 & \mu_t & 0 \\ 0 & 0 & \mu_t \end{bmatrix}, \qquad \begin{bmatrix} \mu_t & 0 & 0 \\ 0 & \mu_l & 0 \\ 0 & 0 & \mu_t \end{bmatrix}, \qquad \begin{bmatrix} \mu_t & 0 & 0 \\ 0 & \mu_t & 0 \\ 0 & 0 & \mu_l \end{bmatrix}, \tag{19.103}$$

with $\mu_l = q\, \tau_p/m_l$, $\mu_t = q\, \tau_p/m_t$. The first form in (19.103) applies to the two minima belonging to axis k_1, and so on; thus, the electron-mobility tensor (19.102) is found to be

$$\hat{\mu}_n = \frac{1}{6} \left(2 \begin{bmatrix} \mu_l & 0 & 0 \\ 0 & \mu_t & 0 \\ 0 & 0 & \mu_t \end{bmatrix} + 2 \begin{bmatrix} \mu_t & 0 & 0 \\ 0 & \mu_l & 0 \\ 0 & 0 & \mu_t \end{bmatrix} + 2 \begin{bmatrix} \mu_t & 0 & 0 \\ 0 & \mu_t & 0 \\ 0 & 0 & \mu_l \end{bmatrix} \right) = \mu_n\, \mathscr{I}\,, \tag{19.104}$$

[31] From this assumption and from (19.100) it also follows $\mathbf{J} = -q \sum_{a=1}^{M_C} n_a\, \mathbf{v}_a = -q\, (n/M_C) \sum_{a=1}^{M_C} \mathbf{v}_a$, whence $\mathbf{J} = -q\, n\, \mathbf{v}$ with $\mathbf{v} = (1/M_C) \sum_{a=1}^{M_C} \mathbf{v}_a$.

[32] The relation derives from Einstein's investigation on the Brownian motion [46] and has therefore a broader application. In a semiconductor it holds within the approximations of parabolic bands and nondegenerate conditions.

with $\mathscr{I}$ the identity tensor and $\mu_n = (\mu_l + 2\,\mu_t)/3$ the *electron mobility*. The second definition in (19.102) then yields $\hat{D}_n = D_n\,\mathscr{I}$, with $D_n = (k_B\,T/q)\mu_n$ the *electron diffusivity* or *electron-diffusion coefficient*. From these results and (19.101) one derives

$$\mathbf{J}' = q\,\mu_n\,n\,\mathbf{E} + q\,D_n\,\mathrm{grad}n\,, \qquad \mathbf{J}'_a = \frac{1}{\mu_n\,M_C}\,\hat{\mu}_a\,\mathbf{J}'\,. \tag{19.105}$$

From this, after a somewhat lengthy calculation, the last term at the right-hand side of (19.100) is found to be

$$\sum_{a=1}^{M_C} \hat{\mu}_a\,\left(\mathbf{B}\wedge\mathbf{J}'_a\right) = a_n\,\mu_n\,\mathbf{B}\wedge\mathbf{J}'\,, \qquad a_n = \frac{\mu_t\,(\mu_t + 2\,\mu_l)}{3\,\mu_n^2}\,. \tag{19.106}$$

As anticipated in the qualitative discussion about the HD model, despite the fact that each vector equation is anisotropic, when the contributions of the different valleys are combined together the anisotropy cancels out. From the definition of the electron mobility $\mu_n = (\mu_l + 2\,\mu_t)/3$ one may also extract a scalar effective mass $m_n = q\,\tau_p/\mu_n$ that fulfills $1/m_n = (2/m_t + 1/m_l)/3$. Using the room-temperature values taken from Table 17.4 yields, for silicon, $m_n/m_0 \simeq 0.26$. By the same token one finds $a_n \simeq 2.61$.

In the next sections, the current density of the electrons in the conduction band will be used in equations involving also the current density of the holes in the valence band; for this reason it is necessary to use different symbols. Specifically, $\mathbf{J}_n$ for the former and $\mathbf{J}_p$ for the latter; with this provision, the above calculation yields

$$\mathbf{J}_n = q\,\mu_n\,n\,\mathbf{E} + q\,D_n\,\mathrm{grad}n + q\,a_n\,\mu_n\,\mathbf{B}\wedge(\mu_n\,n\,\mathbf{E} + D_n\,\mathrm{grad}n)\,, \tag{19.107}$$

which is called *drift-diffusion transport equation*. Thus the DD model for the electrons of the conduction band is given by (19.107) along with the first equation in (19.93); the latter is rewritten here as

$$\frac{\partial n}{\partial t} - \frac{1}{q}\,\mathrm{div}(\mathbf{J}_n) = W_n\,, \tag{19.108}$$

where a specific symbol for the generation-recombination term has been introduced as well.

19.5.3 DD Model for the Valence Band

The transport models illustrated so far are applicable to the valence band as well; here, the DD model will be worked out. Remembering the discussion of Sect. 19.2.3 about the dynamics in the parabolic-band approximation, the model is described

in terms of the concentration and current density of holes. The two quantities are defined by adapting the corresponding expression for electrons, (19.31), as shown below. Letting $f = Q\,\Phi$, with $Q = 1/(4\,\pi^3)$ the density of states in the phase space and Φ the occupation probability, the hole concentration is

$$p(\mathbf{r}, t) = \iiint_{-\infty}^{+\infty} Q\ (1 - \Phi)\ \mathrm{d}^3 k\,. \tag{19.109}$$

In turn, the hole current density is defined starting from the definition of the electron current density of the valence band. The latter is similar to (19.31), the difference being that the integration in (19.31) is restricted to the branch of $E(\mathbf{k})$ belonging to the conduction band, whereas the integration in (19.110) below is restricted to one of the two branches of the valence band:

$$\mathbf{J}_a = -q \iiint_{-\infty}^{+\infty} \mathbf{u}(\mathbf{k})\, Q\, \Phi(\mathbf{r}, \mathbf{k}, t)\, \mathrm{d}^3 k\,. \tag{19.110}$$

Letting $\Phi = 1 - (1 - \Phi)$ transforms (19.110) into

$$\mathbf{J}_a = q \iiint_{-\infty}^{+\infty} \mathbf{u}\, Q(1 - \Phi)\, \mathrm{d}^3 k - q \iiint_{-\infty}^{+\infty} \mathbf{u}\, Q\, \mathrm{d}^3 k\,, \tag{19.111}$$

where the second integral vanishes because $\mathbf{u}$ is odd with respect to $\mathbf{k}$. As a consequence, the current density of the branch under consideration may also be thought of as given by the motion of the empty states (holes), having the group velocity $\mathbf{u}(\mathbf{k})$ and the positive charge q. Moreover, one defines the average velocity of holes using $Q\,(1 - \Phi)$ as weighing function, to find

$$\mathbf{J}_a = q\, p_a\, \frac{\int_{\mathbf{k}} \mathbf{u}\, Q\, (1 - \Phi)\, \mathrm{d}^3 k}{\int_{\mathbf{k}} Q\, (1 - \Phi)\, \mathrm{d}^3 k} = q\, p_a\, \mathbf{v}_a \tag{19.112}$$

where the definition (19.109) of the hole concentration has been specified for the branch under consideration, and the short-hand notation $\int_{\mathbf{k}}$ has been used.

Given the above definitions, the derivation of the drift-diffusion model for holes follows the same pattern as for the electrons. Remembering the description of the band structure given in Sect. 17.6.5, for the valence band index a ranges over h and l; moreover, due to the isotropy of each branch deriving from the parabolic-band approximation (compare with 17.78), the effective mass is scalar. Then, the equivalent of (19.94) read, in vector form,

$$\mathbf{J}_a = \mathbf{J}'_a + \frac{q\,\tau_{pa}}{m_a}\,\mathbf{J}_a \wedge \mathbf{B}\,, \qquad \mathbf{J}'_a = -k_B\,T\,\frac{q\,\tau_{pa}}{m_a}\,\mathrm{grad}\ p_a + \frac{q\,\tau_{pa}}{m_a}\,q\,p_a\,\mathbf{E}\,, \tag{19.113}$$

where index a is attached also to the momentum-relaxation time because the two branches are different. Still due to such a difference, the holes do not distribute

equally over the branches; the contribution of the drift and diffusion components then read, respectively,

$$\frac{q\,\tau_{ph}}{m_{hh}}\,q\,p_h\,\mathbf{E}+\frac{q\,\tau_{pl}}{m_{hl}}\,q\,p_l\,\mathbf{E}=q\,(\mu_{ph}+\mu_{pl})\,p\,\mathbf{E}\,, \tag{19.114}$$

$$-k_B\,T\,\frac{q\,\tau_{ph}}{m_{hh}}\,\mathrm{grad}\;p_h-k_B\,T\,\frac{q\,\tau_{pl}}{m_{hl}}\,\mathrm{grad}\;p_l=-k_B\,T\,(\mu_{ph}+\mu_{pl})\,\mathrm{grad}\;p\,, \tag{19.115}$$

where

$$\mu_{ph}=\frac{q\,\tau_{ph}}{m_{hh}}\,\frac{p_h}{p}\,,\qquad \mu_{pl}=\frac{q\,\tau_{pl}}{m_{hl}}\,\frac{p_l}{p}\,. \tag{19.116}$$

An approximate expression of μ_{ph}, μ_{ph} is obtained by replacing the concentrations with the corresponding equilibrium values $p_h = N_{Vh}\,\Phi_{1/2}(\xi_h)$, $p_l = N_{Vl}\,\Phi_{1/2}(\xi_h)$, with

$$N_{Vh}=2\,M_V\left(\frac{m_{hh}}{2\,\pi\,\hbar^2}\,k_B\,T\right)^{3/2}\,,\qquad N_{Vl}=2\,M_V\left(\frac{m_{hl}}{2\,\pi\,\hbar^2}\,k_B\,T\right)^{3/2} \tag{19.117}$$

(compare with (18.8)), whence, using $p = p_h + p_l$,

$$\mu_{ph}\simeq\frac{q\,\tau_{ph}\,m_{hh}^{1/2}}{m_{hh}^{3/2}+m_{hl}^{3/2}}\,,\qquad \mu_{pl}\simeq\frac{q\,\tau_{pl}\,m_{hl}^{1/2}}{m_{hh}^{3/2}+m_{hl}^{3/2}}\,. \tag{19.118}$$

Then, $\mathbf{J}_a$ is extracted from the first relation in (19.113), whose matrix form is

$$\begin{bmatrix} 1 & \mu_a\,B_3 & -\mu_a\,B_2 \\ -\mu_a\,B_3 & 1 & \mu_a\,B_1 \\ \mu_a\,B_2 & -\mu_a\,B_1 & 1 \end{bmatrix}\begin{bmatrix} J_{1a} \\ J_{2a} \\ J_{3a} \end{bmatrix}=\begin{bmatrix} J'_{1a} \\ J'_{2a} \\ J'_{3a} \end{bmatrix}\,, \tag{19.119}$$

$\mu_a = q\,\tau_{pa}/m_a$. The determinant of the matrix in (19.119) is $D_M = 1 + \mu_a^2\,B^2$. Still considering the case where $\mathbf{B}$ is weak, one finds

$$\mathbf{J}_a\simeq\mathbf{J}'_a-\mu_a\,\mathbf{B}\wedge\mathbf{J}'_a\,. \tag{19.120}$$

In turn, the contribution of the last term at the right-hand side of the above yields

$$-a_p\,\mathbf{B}\wedge\left(q\,\mu_p\,p\,\mathbf{E}-q\,D_p\,\mathrm{grad}p\right)\,,\qquad \mu_p=\mu_{ph}+\mu_{pl}\,, \tag{19.121}$$

with μ_p the *hole mobility*. In turn, the *hole diffusivity* (or *hole-diffusion coefficient*) and the dimensionless parameter a_p are given by

$$D_p=\frac{k_B\,T}{q}\,\mu_p\,,\qquad a_p=\frac{1}{\mu_p^2}\left(\frac{q\,\tau_{ph}}{m_{hh}}\,\mu_{ph}+\frac{q\,\tau_{pl}}{m_{hl}}\,\mu_{pl}\right)\,. \tag{19.122}$$

Putting (19.114), (19.115), and (19.121) together finally provides the *drift-diffusion transport equation for the holes*,

$$\mathbf{J}_p = q\,\mu_p\,p\,\mathbf{E} - q\,D_p\,\mathrm{grad}p - q\,a_p\,\mu_p\,\mathbf{B}\wedge\left(\mu_p\,p\,\mathbf{E} - D_p\,\mathrm{grad}p\right)\,. \tag{19.123}$$

Thus, the DD model for the holes of the valence band is given by (19.123) along with the balance equation for the holes' number (compare with (19.93)), that reads

$$\frac{\partial p}{\partial t} + \frac{1}{q}\,\mathrm{div}(\mathbf{J}_p) = W_p\,. \tag{19.124}$$

19.5.4 Coupling with Maxwell's Equations

As anticipated in Sect. 19.5.1, as the electromagnetic field is influenced by the distribution of charge and current density, it is necessary to couple the equations describing the charge transport (in the form, e.g., of the hydrodynamic or drift-diffusion model) with the Maxwell equations. For this, one inserts the total charge density ϱ and current density $\mathbf{J}$ into the right-hand sides of (4.19); considering that there are different groups of charges and currents, one uses (4.22), where the charge density is given by (18.53), namely,[33]

$$\varrho = q\,(p - n + N)\,, \qquad N = N_D^+ - N_A^-\,. \tag{19.125}$$

In turn, the current density reads

$$\mathbf{J} = \mathbf{J}_p + \mathbf{J}_n = \varrho_p\,\mathbf{v}_p + \varrho_n\,\mathbf{v}_n = q\,p\,\mathbf{v}_p - q\,n\,\mathbf{v}_n\,, \tag{19.126}$$

with $\mathbf{J}_n$ and $\mathbf{J}_p$ given by (19.107) and (19.123), respectively. As noted in Sect. 18.5, the material's permittivity must be used here instead of vacuum's; as a consequence, the relation between electric displacement and field reads $\mathbf{D} = \varepsilon_{\mathrm{sc}}\,\mathbf{E}$.

One notes that the $\mathbf{E}$ and $\mathbf{B}$ fields are the sum of two contributions: the first one derives from the internal charge and current-density distribution as mentioned above, while the second one derives from external sources, e.g., voltage or current generators connected to the device or integrated circuits, or electric and magnetic fields present in the environment. In general, the internal contribution to $\mathbf{B}$ is negligible and is not considered in semiconductor devices or integrated circuit;

[33] Equation (18.53) is the definition of charge density in a semiconductor; as a consequence it holds in general, not only in the equilibrium condition considered in Sect. 18.5. In fact, it can readily be extended to account for charges trapped in energy states different from those of the dopants (Sect. 20.2.2).

it follows that **B** is to be accounted for in (19.107) and (19.123) only when it derives from external sources[34] and, due to this, it must be thought of as a prescribed function of **r** and t. With these premises, the analysis will continue here after letting $\mathbf{B} = 0$. Despite this simplification, the continuity and transport equations (19.108), (19.107) and (19.123), (19.124) must be coupled with the whole set of Maxwell equations: in fact, the expression of the electric field in terms of the potentials is given by the second relation in (4.26), namely, $\mathbf{E} = -\text{grad}\,\varphi - \partial\mathbf{A}/\partial t$; as a consequence, in a dynamic condition both the scalar and vector potential must be determined. In a steady-state or equilibrium condition, instead, the expression of the electric field reduces to $\mathbf{E} = -\text{grad}\,\varphi$; in this case it is sufficient to couple it with the first equation in (4.19) only.

The presence of the vector potential **A** makes the model more complicate; thus, it is useful to ascertain whether, in the typical operating conditions, the time derivative $\partial\mathbf{A}/\partial t$ in the expression of **E** should be kept or not. As noted above, the derivative differs from zero only if the boundary conditions (e.g., the applied voltages) vary with time. To associate a characteristic time with a boundary condition one takes the period associated with the maximum frequency of the boundary condition's spectrum, $\tau_{\min} = 1/\nu_{\max}$; then, one compares $\tau_{\min}$ with the time Δt necessary for the electromagnetic perturbation produced by the boundary condition to propagate to a position internal to the semiconductor. If d is the distance between a point on the boundary and an internal point, the propagation time can be estimated to be $\Delta t = d/u_f$, with u_f the radiation's phase velocity corresponding to $\nu_{\max}$. If it happens that $\Delta t \ll \tau_{\min}$, the propagation is practically instantaneous, namely, the electromagnetic field at the internal point is consistent with the boundary condition existing at the same instant of time; as a consequence, the boundary condition is thought of as stationary, and the $\partial\mathbf{A}/\partial t$ derivative is neglected. This is called *quasi-static approximation*; the condition of its applicability is summarized as[35]

$$\Delta t = \frac{d}{u_f} \ll \tau_{\min} = \frac{1}{\nu_{\max}}\,, \qquad \nu_{\max} \ll \frac{u_f}{d}\,. \tag{19.127}$$

To estimate the condition one must fix the value of d; as a conservative choice one takes the channel length of the MOSFET transistors of the old generations, $d \approx 10^{-7}$ m. Using $u_f \approx 10^8$ m s^{-1} yields $\nu_{\max} \ll 10^{15}$ Hz, which is amply fulfilled in the present state-of-the-art integrated circuits. Note that the condition is even better verified in the last-generation devices, whose channel length is shorter than 10^{-7} m.

The choice of the channel length in the above estimate is dictated by the fact that the channel is the active region of the device. Choosing, instead, d as the (much

[34] A typical example is found when a semiconductor device or circuit is used as a magnetic-field sensor or in specific measurement setups, like in the Hall-voltage measurement (Sect. 25.4).

[35] A similar reasoning is used to treat the time derivative of v_i, w, and b_i in the derivation of the BTE's moments of order larger than zero (Sect. 19.4.6).

larger) thickness of the silicon wafer would not make sense, because the phenomena taking place in the wafer's bulk are relatively unimportant. Other distances within an integrated circuit are larger by orders of magnitude than the value of d considered in the estimate; for instance, the diameter of the integrated circuit itself is of the order of 10^{-2} m, hence the quasi-static approximation is not applicable to the case of two devices placed, e.g., at opposite corners of a chip and connected by a line. In fact, the propagation of signals along the lines connecting different devices on a chip is modeled using the whole set of Maxwell equations.[36]

19.5.5 Semiconductor-Device Model

Thanks to the quasi-static approximation, the equations describing the semiconductor, in the drift-diffusion case and with $\mathbf{B} = 0$, read

$$\mathrm{div}\mathbf{D} = q\,(p - n + N)\,, \qquad \mathbf{D} = -\varepsilon_{\mathrm{sc}}\,\mathrm{grad}\varphi, \tag{19.128}$$

$$\frac{\partial n}{\partial t} - \frac{1}{q}\,\mathrm{div}\mathbf{J}_n = W_n\,, \qquad \mathbf{J}_n = q\,\mu_n\,n\,\mathbf{E} + q\,D_n\,\mathrm{grad}n\,, \tag{19.129}$$

$$\frac{\partial p}{\partial t} + \frac{1}{q}\,\mathrm{div}\mathbf{J}_p = W_p\,, \qquad \mathbf{J}_p = q\,\mu_p\,p\,\mathbf{E} - q\,D_p\,\mathrm{grad}p\,. \tag{19.130}$$

As outlined in Chap. 24, insulating layers play an essential role in the fabrication of integrated circuits; it is then necessary to extend the model to incorporate also the description of such layers. This is easily accomplished by observing that mobile charges are absent in an insulator, so that the balance equations for the number of particles and the transport equations reduce to identities, $0 = 0$. The model for the insulators then reduces to Poisson's equation only. The right-hand side of the latter does not necessarily vanish because, as indicated in Sect. 24.1, contaminants may enter the insulator during the fabrication steps; some of these contaminants may ionize and act as fixed charges so that, letting N_{ox} be their density, the model for the insulator reads[37]

$$\mathrm{div}\mathbf{D} = q\,N_{\mathrm{ox}}\,, \qquad \mathbf{D} = -\varepsilon_{\mathrm{ox}}\,\mathrm{grad}\varphi\,, \qquad n = p = 0\,, \qquad \mathbf{J}_n = \mathbf{J}_p = 0\,. \tag{19.131}$$

The set of equations (19.128–19.131) is commonly called *semiconductor-device model*. It is made of partial-differential equations of the first order, in the unknowns φ, n, p and $\mathbf{D}$, $\mathbf{J}_n$, $\mathbf{J}_p$. The coefficients are μ_n, μ_p, $D_n = (k_B\,T/q)\,\mu_n$,

[36] The progressive device scaling from one generation to the next is in general associated with an increase in the size of the chips. Due to this, the constraints on the circuit's speed are rather imposed by the lines connecting the devices than by the devices themselves.

[37] The insulator's permittivity is indicated with $\varepsilon_{\mathrm{ox}}$ because, in the examples used later, silicon dioxide (SiO_2) is used as the reference insulator.

$D_p = (k_B\,T/q)\,\mu_p$. In turn, N, W_n, W_p are either known functions or are expressed in terms of the unknowns themselves. Some of the equations are nonlinear because the unknowns are multiplied by each other. Each equation on the left in (19.128–19.131) contains the divergence of a vector; in turn, the expression of the vector is given by the corresponding equation on the right, in terms of the scalar unknowns (in fact it is $\mathbf{E} = -\mathrm{grad}\varphi$). It follows that by introducing the expressions of $\mathbf{D}$, $\mathbf{J}_n$, and $\mathbf{J}_p$ into the divergence operator, each pair of first-order equation is transformed into a single, second-order equation. This observation is useful in view of the application of numerical methods to the solution of the semiconductor-device model.

Remembering the derivation of the transport model, the terms W_n, W_p in (19.129), (19.130) are due to the generation-recombination phenomena. Specifically, W_n is the difference between the number of electrons entering the conduction band, and of those leaving it, per unit volume and time; in turn, W_p is the difference between the number of holes entering the valence band, and of those leaving it, per unit volume and time.[38] For this reason, they are also called *net generation rates*. As mentioned in Sect. 19.3, the transitions of a given class are further grouped depending on the entity with which the particle's collision occurs. As far as the net generation rates are concerned, it is customary to separate the contribution of the phonon collisions from those of the other types (e.g., electron-electron collisions, electron-photon collisions, and so on); in fact, unless the device is kept at a very low temperature, the phonon collisions are the most important ones. Thus, the net generation rates are recast as

$$W_n = G_n - U_n\,, \qquad W_p = G_p - U_p\,, \tag{19.132}$$

where U_n, U_p describe the transitions due to phonon collisions, while G_n, G_p describe those of the other types. The minus signs in (19.132) come from the fact that U_n is defined as the difference between the number of electrons leaving the conduction band, and of those entering it, because of phonon collisions, per unit volume and time; similarly, U_p is defined as the difference between the number of holes leaving the valence band, and of those entering it, because of phonon collisions, per unit volume and time. The terms used for U_n, U_p are *net thermal recombination rates*, those for G_n, G_p are *net nonthermal generation rates*.

Another comment about the semiconductor-device model concerns the drift terms in (19.129), (19.130). The latter can be recast as $\mathbf{J}_n^{\mathrm{dr}} = \sigma_n\,\mathbf{E}$ and $\mathbf{J}_p^{\mathrm{dr}} = \sigma_p\,\mathbf{E}$, where

$$\sigma_n = q\,\mu_n\,n\,, \qquad \sigma_p = q\,\mu_p\,p \tag{19.133}$$

are the *electron conductivity* and *hole conductivity*, respectively. From $\mathbf{J} = \mathbf{J}_n + \mathbf{J}_p$, in a uniform material one obtains $\mathbf{J} = \sigma\,\mathbf{E}$, that is, *Ohm's law*, with

$$\sigma = \sigma_n + \sigma_p = q\,\left(\mu_n\,n + \mu_p\,p\right). \tag{19.134}$$

[38] The units are $[W_n, W_p] = \mathrm{m}^{-3}\ \mathrm{s}^{-1}$.

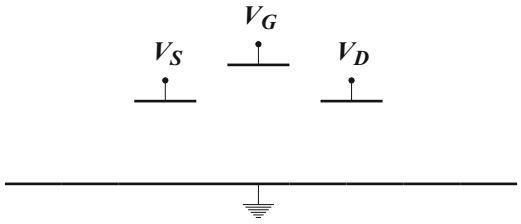


Fig. 19.3 MOS structure used to discuss the boundary conditions for the mathematical model of semiconductor devices. Only the conducting boundaries are shown. Note that the vertical scale of the drawing is not realistic

19.5.6 Boundary Conditions

In practical applications, the equations of the semiconductor-device model, (19.128) through (19.131), are solved over a closed domain whose boundary[39] is indicated here with Γ. The boundary is partitioned into portions, some of which, indicated with $\Gamma_{i1}, \Gamma_{i2}, \ldots$, are *insulating boundaries*, namely, they cannot be crossed by electrons or holes; the remaining portions, $\Gamma_{c1}, \Gamma_{c2}, \ldots$, can be crossed by the carriers and are termed *conducting boundaries*.

Considering that the domain over which the equations are solved is in general a part of a much larger domain enclosing an integrated circuit, some flexibility exists as for the choice of Γ. Thanks to this, it is possible to select $\Gamma_{i1}, \Gamma_{i2}, \ldots$ such that the normal component of the vector unknowns vanishes there; in other terms, letting $\mathbf{s}$ be the unit vector normal to an insulating boundary at some point $\mathbf{r}$, it is

$$\mathbf{E} \cdot \mathbf{s} = 0\,, \qquad \frac{\partial \varphi}{\partial s} = 0\,, \qquad \mathbf{r} \in \Gamma_{i1}, \Gamma_{i2}, \ldots\,. \tag{19.135}$$

where $\partial/\partial s$ indicates the derivative normal to the boundary at $\mathbf{r}$. An example of how this can be accomplished is given in Figs. 19.3 and 19.4, representing the schematic cross-section of a MOSFET. In Fig. 19.3 only the conducting boundaries are shown, with V_S, V_G, and V_D indicating the voltage applied to the source, gate, and drain contact, respectively. The bulk contact is grounded, as shown by the line below; such a contact is the ground reference for the whole chip and, for this reason, extends laterally beyond the region occupied by the MOSFET under consideration. The insulating boundaries must now be selected in order to form a closed line that completes the boundary Γ; in principle, such a completion may be accomplished in different ways. Consider, however, the completion shown in the upper part of Fig. 19.4: as in general it is $V_S \neq V_G$, it is likely that one of the field lines in the region between the source and gate contacts coincides with the segment ab. As a

[39] A two- or three-dimensional case is considered. In the one-dimensional case the boundary reduces to the two points enclosing the segments over which the equations are to be solved.

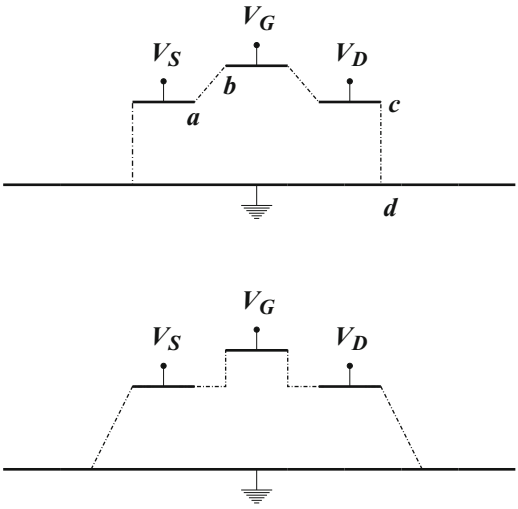


Fig. 19.4 The same structure as in Fig. 19.3, to which the insulating boundaries have been added (dash-dotted lines). The upper part of the figure shows the correct placement of the insulating boundaries, the lower part shows a wrong placement

consequence, choosing *ab* as the insulating boundary in that region guarantees that the component of **E** normal to such a boundary is zero, thus achieving the condition sought. The same reasoning applies to line *cd*. By this procedure one succeeds in prescribing the boundary conditions for the Poisson equation along the insulating boundaries. If, instead, the insulating boundaries were placed differently, like, e.g., in the lower part of Fig. 19.4, the component of **E** normal to the boundary would be different from zero; besides that, it would be impossible to determine it *a priori*, and Poisson's equation would become ill-posed.

Once the insulating boundaries are completed, the same condition as (19.135) is prescribed onto the current densities, namely, $\mathbf{J}_n \cdot \mathbf{s} = \mathbf{J}_p \cdot \mathbf{s} = 0$ whence, using the second equation in (19.129),

$$-q\,\mu_n\,n\frac{\partial\varphi}{\partial s} + q\,D_n\,\frac{\partial n}{\partial s} = 0\,, \qquad \mathbf{r} \in \Gamma_{i1}, \Gamma_{i2}, \ldots\,. \tag{19.136}$$

Combining the above with (19.135) yields $\partial n/\partial s = 0$; repeating the calculation for the holes finally shows that at the insulating boundaries the boundary condition is the same for all scalar unknowns φ, n, p:

$$\frac{\partial\varphi}{\partial s} = 0\,, \qquad \frac{\partial n}{\partial s} = 0\,, \qquad \frac{\partial p}{\partial s} = 0\,, \qquad \mathbf{r} \in \Gamma_{i1}, \Gamma_{i2}, \ldots\,, \tag{19.137}$$

i.e., a boundary condition of the *homogeneous Neumann type*.

The conducting boundaries $\Gamma_{c1}, \Gamma_{c2}, \ldots$ are typically covered with metal layers or heavily doped polycrystalline layers that provide the electric contacts to the device. Unless the operating condition of the device departs strongly from equilibrium, a contact is able to supply the amount of charge necessary to keep the equilibrium and charge-neutrality conditions in the semiconductor layer adjacent to it. Thus at

each point $\mathbf{r}$ of this layer one takes $\varrho = 0$ or, more specifically,

$$\varrho_c = q\,(p_c - n_c + N_c) = 0\,, \qquad \mathbf{r} \in \Gamma_{c1}, \Gamma_{c2}, \ldots, \tag{19.138}$$

where index c indicates the conducting boundary. In most cases the metal or polycrystalline layer is connected to a voltage generator, so that the electric potential is prescribed, or to another part of the integrated circuit, so that the electric potential can be determined from a separate calculation. In these cases, the electric potential of the contact is known. From it, one derives the electric potential φ_c of the conducting boundary adjacent to the contact; in fact, when the departure from the equilibrium condition is not too strong, the difference between the electric potential of the conducting boundary and that of the contact does not depend on the current density that crosses the boundary, and equals the contact's work function.[40] The latter is experimentally known, thus yielding φ_c. In conclusion, at a conducting boundary where the voltage is prescribed, the boundary condition is the same for all scalar unknowns: $\varphi = \varphi_c$, $n = n_c$, $p = p_c$, $\mathbf{r} \in \Gamma_{c1}, \Gamma_{c2}, \ldots$, i.e., a boundary condition of the *Dirichlet type*. Note that the quantities in parenthesis in (19.138) depend on φ_c (compare with (18.56) and (18.57)); as a consequence, they can be calculated only after the electric potential has been determined.

In some instances a current generator is connected to a contact; as a consequence, the voltage is not prescribed at the corresponding conducting boundary. However, such a voltage can be determined by observing that the flux of the current density across the boundary equals the generator's current. This provides the extra relation that keeps the well-posedness of the mathematical problem.[41]

19.5.7 Quasi-Fermi Potentials

The drift-diffusion transport equations, given by the second relation in (19.129), (19.130), can be recast in a monomial form by defining two auxiliary functions

$$\varphi_n(\mathbf{r}, t) = \varphi - \frac{k_B\,T}{q}\,\log\left(\frac{n}{n_i}\right), \qquad \varphi_p(\mathbf{r}, t) = \varphi + \frac{k_B\,T}{q}\,\log\left(\frac{p}{n_i}\right), \tag{19.139}$$

whose inversion yields

$$n = n_i\,\exp\left[\frac{q\,(\varphi - \varphi_n)}{k_B\,T}\right], \qquad p = n_i\,\exp\left[\frac{q\,(\varphi_p - \varphi)}{k_B\,T}\right]. \tag{19.140}$$

[40] Examples of application of this concept are given in Sects. 21.2.2 and 22.2.

[41] This outcome becomes immediately clear by applying a numerical-discretization method to the problem. In fact, the component of the current density normal to the contact depends on the electric potential of the contact itself; thus, the extra relation provided by the flux-conservation equation embeds the extra unknown φ_c.

In the equilibrium limit, (19.140) must coincide with (18.62), namely, $\varphi_n \to \varphi_F$, $\varphi_p \to \varphi_F$. It follows that the auxiliary functions (19.139) are a formal generalization of the concept of Fermi potential; they have the advantage of keeping the exponential form of the expressions of n and p in the nonequilibrium case. For this reason, φ_n and φ_p are called *Quasi-Fermi potentials* for electrons and holes, respectively.[42] From (19.140) one finds $(k_B\,T/q)\,\mathrm{grad}n = n\,\mathrm{grad}(\varphi - \varphi_n)$ and $(k_B\,T/q)\,\mathrm{grad}p = p\,\mathrm{grad}(\varphi_p - \varphi)$ which, replaced into the second relation of (19.129), (19.130), respectively, yield

$$\mathbf{J}_n = -q\,\mu_n\,n\,\mathrm{grad}\varphi + q\,\frac{k_B T_L}{q}\,\mu_n\,\mathrm{grad}n = -q\,\mu_n\,n\,\mathrm{grad}\varphi_n, \tag{19.141}$$

$$\mathbf{J}_p = -q\,\mu_p\,p\,\mathrm{grad}\varphi - q\,\frac{k_B\,T}{q}\,\mu_p\,\mathrm{grad}p = -q\,\mu_p\,p\,\mathrm{grad}\varphi_p. \tag{19.142}$$

One notes that the monomial forms (19.142), (19.142) thus achieved are similar to drift-diffusion equations where the drift term only is present. This result allows one to interpret $-\mathrm{grad}\varphi_n$ and $-\mathrm{grad}\varphi_p$ as effective fields acting on the electrons (or, respectively, holes) and incorporating both drift and diffusion effects. The monomial form is useful for describing unipolar devices, where one of the current densities $\mathbf{J}_n$, $\mathbf{J}_p$ dominates over the other and is essentially solenoidal (Sect. 22.8).

The definition of the quasi-Fermi potential given above is applicable only to drift-diffusion equations of the form (19.129), (19.130), that are valid within the approximations of parabolic band and nondegenerate conditions. However, the concept of quasi-Fermi potential can be generalized by disposing, e.g., of the nondegeneracy hypothesis. For this, one starts from the equilibrium expressions of n and p, given by (18.56) and (18.57), respectively, and replaces E_F in the definitions of ξ_e, ξ_h with $E_F = E_{Fi} - q\,\varphi_F$ (compare with (18.26)); then, the Fermi potential φ_F is replaced with φ_n in the definition of ξ_e, and with φ_p in that of ξ_h. The nonequilibrium concentrations then read

$$n = N_C\,\Phi_{1/2}\,(\xi_e)\ , \qquad \xi_e = -\frac{E_C - E_{Fi}}{k_B\,T} + \frac{q\,(\varphi - \varphi_n)}{k_B\,T}\ , \tag{19.143}$$

and

$$p = N_V\,\Phi_{1/2}\,(\xi_h)\ , \qquad \xi_h = -\frac{E_{Fi} - E_V}{k_B\,T} + \frac{q\,(\varphi_p - \varphi)}{k_B\,T}\ . \tag{19.144}$$

[42] By some authors, φ_n and φ_p are called *Imref potentials*, where "Imref" is "Fermi" read from right to left [128].

19.5.8 Poisson Equation in a Semiconductor

In the equilibrium condition the concentrations of electrons (19.143) and holes (19.144) depend on the electric potential only. In turn, the ionized donor and acceptor concentrations depend on the electric potential and, possibly, on position if the dopant distributions are position dependent. In summary, the general form of the equilibrium charge concentration is $\varrho = \varrho(\varphi, \mathbf{r})$. The semiconductor-device model reduces to the Poisson equation alone because at equilibrium it is $\partial/\partial t = 0$, $\mathbf{J}_n = \mathbf{J}_p = 0$; the equation reads $-\varepsilon_{\rm sc} \nabla^2 \varphi = q\,(p - n + N) = \varrho(\varphi, \mathbf{r})$, which is a semi-linear partial differential equation (PDE).[43] If the explicit dependence on $\mathbf{r}$ is absent, $\varrho = \varrho(\varphi)$, and the problem is one dimensional, say, in the x direction, Poisson's equation can be solved analytically over a domain I where $\mathrm{d}\varphi/\mathrm{d}x \neq 0$. In fact, multiplying by $\mathrm{d}\varphi/\mathrm{d}x$ both sides of $-\varepsilon_{\rm sc}\,\mathrm{d}^2\varphi/\mathrm{d}x^2 = \varrho(\varphi)$, one obtains

$$-\varepsilon_{\rm sc}\,\frac{\mathrm{d}^2\varphi}{\mathrm{d}x^2}\,\frac{\mathrm{d}\varphi}{\mathrm{d}x} = \varrho(\varphi)\,\frac{\mathrm{d}\varphi}{\mathrm{d}x}\,, \qquad \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\mathrm{d}\varphi}{\mathrm{d}x}\right)^2 = \frac{\mathrm{d}}{\mathrm{d}x} S(\varphi)\,, \tag{19.145}$$

where $\mathrm{d}S/\mathrm{d}\varphi = -2\,\varrho(\varphi)/\varepsilon_{\rm sc}$. Integrating (19.145) from x_0 to x, where both points belong to I, and letting $\varphi_0 = \varphi(x_0)$ yield

$$\left(\frac{\mathrm{d}\varphi}{\mathrm{d}x}\right)^2 = G^2(\varphi)\,, \qquad G(\varphi) = \left[\left(\frac{\mathrm{d}\varphi}{\mathrm{d}x}\right)_0^2 + S(\varphi) - S(\varphi_0)\right]^{1/2}. \tag{19.146}$$

Separating the variables in (19.146) then provides

$$\mathrm{d}x = \pm\frac{\mathrm{d}\varphi}{G(\varphi)}\,, \qquad x = x_0 \pm \int_{\varphi_0}^{\varphi} \frac{\mathrm{d}\tilde{\varphi}}{G(\tilde{\varphi})}\,, \tag{19.147}$$

namely, the inverse relation $x = x(\varphi)$. The choice of the sign is made on case-by-case basis (an example is given in Sect. 21.2). The approach can be extended to the nonequilibrium case if it happens that the electric potential depends on one independent variable, say, x, while the Poisson equation contains terms that depend also on independent variables different from x. In fact, the other variables can be considered as parameters during the integration with respect to x. An example of this case is given in Sect. 22.11.1.

[43] A PDE of order s in the unknown φ is called *quasi-linear* if it is linear in the order-s derivatives of φ and its coefficients depend on the independent variables and the derivatives of φ of order $m < s$. A quasi-linear PDE where the coefficients of the order-s derivatives are functions of the independent variables alone is called *semi-linear*. A PDE which is linear in the unknown function and all its derivatives, with coefficients depending on the independent variables alone, is called *linear*. PDEs not belonging to the classes above are *fully nonlinear*.

19.6 Complements

19.6.1 Comments on the Equivalent Hamiltonian Operator

It has been observed in Sect. 19.2.2 that in the description of the wave-packet dynamics based on the equivalent Hamiltonian operator and crystal momentum, that the time variations of the latter are due to the external force only; as a consequence, if $U = \text{const}$ one has $\hbar \dot{\mathbf{k}}_0 = 0$, namely, the crystal momentum is a constant of motion. The periodic part of the potential energy is incorporated with the kinetic part to form an equivalent operator. Thus, the expression of the Hamiltonian operator is eventually made of two terms, the equivalent-kinetic part, where the space coordinates do not explicitly appear, and the potential part due to the external force only; from this standpoint it is similar to that of a particle *in vacuo*.

The same results were found in the analysis of the motion of a classical particle subjected to a periodic potential onto which a weak perturbation is superimposed (Sect. 3.11). The classical investigation makes it clear that the crystal momentum is in fact different from the actual momentum of the particle; also, the subsequent elaboration carried out in Sect. 3.12 shows that the concept of effective mass is not distinctive of Quantum Mechanics.

19.6.2 Special Cases of Anisotropy

It is interesting to note that for an anisotropic branch the acceleration $\dot{\mathbf{u}}$ may still be parallel to $\mathbf{F}$; this happens when the force is parallel to one of the coordinate axes. Taking by way of example a force parallel to the first coordinate axis, $\mathbf{F} = F\,\mathbf{i}_1$, it follows in fact

$$\dot{u}_{1g} = F/m_{1a}\,, \qquad \dot{u}_{2g} = 0\,, \qquad \dot{u}_{3g} = 0\,. \tag{19.148}$$

This seems to imply that a suitable choice of the reference is sufficient to make the acceleration parallel to the force. However, this is not so: as noted in Sect. 17.6.2, for (19.24) to hold, the reference in the $\mathbf{k}$ space must be chosen in such a way as to make the Hessian matrix of $E_n(\mathbf{k})$ diagonal; this, in turn, fixes the reference in the $\mathbf{r}$ space, because the two references are reciprocal to each other. As a consequence, if one rotates the $\mathbf{r}$ reference to align one of its axes with the force, the $\mathbf{k}$ reference rotates as well; as the mass tensor in the new reference is not necessarily diagonal, the anisotropy present in the old reference is maintained.

19.6.3 α-Moment at Equilibrium

The derivation of the moment with respect to α of the BTE has been shown in Sect. 19.4.5, leading to (19.62). In the equilibrium condition the distribution function f^{eq} is independent of t and depends on $\mathbf{k}$ through the Hamiltonian function H only; since the latter is even with respect to $\mathbf{k}$, the distribution function is even as well. In turn, as the transitions balance each other, the right-hand side of (19.62) vanishes. The term with the magnetic induction vanishes as well (compare with (19.60)); in conclusion, in the equilibrium condition (19.62) reduces to

$$\mathrm{div}_{\mathbf{r}} \left(n\,\overline{\alpha\,\mathbf{u}}\right)^{\mathrm{eq}} + \frac{q}{\hbar}\left(n\,\overline{\mathrm{grad}_{\mathbf{k}}\alpha}\cdot\mathbf{E}\right)^{\mathrm{eq}} = 0\,. \tag{19.149}$$

It is easily found that (19.149) yields the identity $0 = 0$ if α is even with respect to $\mathbf{k}$. This, on the contrary, is not true when α is odd; in this case the equilibrium condition consists in the balance between the diffusion term, due to the spatial nonuniformity of $(n\,\overline{\alpha\,\mathbf{u}})^{\mathrm{eq}}$, and the second term, proportional to the carrier concentration and linearly dependent on the electric field.

19.6.4 Closure Conditions

The closure conditions for the drift-diffusion and hydrodynamic model are derived in this section. The former consists in calculating (19.75) using the equilibrium distribution $f^{\mathrm{eq}} = Q\,P$, with $Q = 1/(4\,\pi^3)$ the density of states in the $\mathbf{r}, \mathbf{k}$ space and P the Fermi-Dirac statistics. In the equilibrium case it is $\mathbf{v} = 0$, whence $\mathbf{c} = \mathbf{u}$, and

$$n^{\mathrm{eq}}\,k_B\,T_{ij}^{\mathrm{eq}} = \frac{m_{ia}}{4\,\pi^3}\iiint_{-\infty}^{+\infty}\frac{u_i\,u_j}{\exp[(E_e+\zeta_e)/(k_B\,T)]+1}\,\mathrm{d}^3k\,, \tag{19.150}$$

with $E_e = E - E_C$ and $\zeta_e = E_C - q\,\varphi - E_F$ (compare with (18.54)). In the above, E_e is even with respect to all components of $\mathbf{k}$. For $j \neq i$ the integrand is odd with respect to k_i because $u_i = (1/\hbar)\,\partial E_e/\partial k_i$, and with respect to k_j because $u_j = (1/\hbar)\,\partial E_e/\partial k_j$; as a consequence it is $T_{ij} = 0$ for $j \neq i$, while

$$n^{\mathrm{eq}}\,k_B\,T_{ii}^{\mathrm{eq}} = \frac{m_{ia}}{4\,\pi^3}\iiint_{-\infty}^{+\infty}\frac{u_i^2}{\exp[(E_e+\zeta_e)/(k_B\,T)]+1}\,\mathrm{d}^3k\,, \tag{19.151}$$

namely, in the equilibrium condition the electron-temperature tensor is diagonal. For this result to hold, the parabolic-band approximation is not necessary.

In the parabolic-band approximation, (19.24) and (19.72) hold, and the temperature tensor in equilibrium is evaluated at the ath minimum of the conduction band by letting

$$\eta_i = \frac{\hbar\,\delta k_i}{\sqrt{2\,m_{ia}}}\,, \qquad \mathrm{d}^3k = 2\,\frac{\sqrt{2}}{\hbar^3}\,m_{ea}^{3/2}\,\mathrm{d}^3\eta\,, \qquad m_{ia}\,u_i^2 = 2\,\eta_i^2\,, \tag{19.152}$$

where the first relation is the Herring-Vogt transformation (17.66) and m_{ea} is defined in (17.68). Adding up over the minima and using (17.72) yield

$$n^{\mathrm{eq}}\,k_B\,T_{ii}^{\mathrm{eq}} = \frac{\sqrt{2}}{\pi^3\,\hbar^3}\iiint_{-\infty}^{+\infty}\frac{M_C\,m_e^{3/2}\,\eta_i^2}{\exp[(\eta^2+\zeta_e)/(k_B\,T)]+1}\,\mathrm{d}^3\eta\,. \tag{19.153}$$

As the value of the integral in (19.153) does not depend[44] on index i, it follows that the diagonal entries T_{ii}^{eq} are equal to each other; as a consequence, the common value of the three integrals can be replaced with $T_e^{\mathrm{eq}} = (T_{11}^{\mathrm{eq}} + T_{22}^{\mathrm{eq}} + T_{33}^{\mathrm{eq}})/3$, namely,

$$n^{\mathrm{eq}}\,k_B\,T_{ii}^{\mathrm{eq}} = \frac{\sqrt{2}}{3\,\pi^3\,\hbar^3}\iiint_{-\infty}^{+\infty}\frac{M_C\,m_e^{3/2}\,\eta^2}{\exp[(\eta^2+\zeta_e)/(k_BT_L)]+1}\,\mathrm{d}^3\eta\,. \tag{19.154}$$

Turning to spherical coordinates (B.1) yields $\mathrm{d}^3\eta = \eta^2\,\mathrm{d}\eta\,\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\phi$, with $\eta^2 = \eta_1^2+\eta_2^2+\eta_3^2 = E_e$ and $\mathrm{d}\eta = 1/(2\sqrt{E_e})\,\mathrm{d}E_e$, $\eta^2\,\mathrm{d}\eta = \sqrt{E_e}\,\mathrm{d}E_e/2$. The integral over the angles equals $4\,\pi$ whence, using (C.111) with $\Gamma(1+3/2) = (3/2)\,\sqrt{\pi}/2$,

$$n^{\mathrm{eq}}\,k_B\,T_{ii}^{\mathrm{eq}} = k_B\,T\,N_C\,\Phi_{3/2}(\xi_e)\,, \tag{19.155}$$

with $N_C = 2\,M_C\,[m_e\,k_B\,T/(2\,\pi\,\hbar^2)]^{3/2}$ the effective density of states (18.4), and $\xi_e = -\zeta_e/(k_BT)$. On the other hand, from (18.17) it is $n^{\mathrm{eq}} = N_C\,\Phi_{1/2}(\xi_e)$, whence

$$T_e^{\mathrm{eq}}(\mathbf{r}) = T\,\frac{\Phi_{3/2}(\xi_e)}{\Phi_{1/2}(\xi_e)}\,. \tag{19.156}$$

The dependence on position is due to $\xi_e = (q\,\varphi - E_C + E_F)/(k_B\,T)$. However, in the nondegenerate case the approximation (C.112) holds, $\Phi_\alpha(\xi_e) = \exp(\xi_e)$, and the electron-temperature tensor at equilibrium reduces to $T_e^{\mathrm{eq}} = T$.

Coming now to the hydrodynamic model, and remembering (19.83), the closure condition is found by calculating the equilibrium value of $n\,m_{ia}\,\overline{E_e\,u_i}\,\mathbf{c}$, that reads

$$n^{\mathrm{eq}}\,m_{ia}\,\left(\overline{E_e\,u_i\,\mathbf{c}}\right)^{\mathrm{eq}} = \frac{m_{ia}}{4\,\pi^3}\iiint_{-\infty}^{+\infty}E_e\,u_i\,\mathbf{c}\,P\,\mathrm{d}^3k\,, \tag{19.157}$$

[44] Let $i = 1$, whence (19.153) takes the form $\int_{-\infty}^{+\infty}\eta_1\,S(\eta_1^2+\eta_2^2+\eta_3^2)\,\mathrm{d}\eta_1\,\mathrm{d}\eta_2\,\mathrm{d}\eta_3$. A cyclic permutation $1 \leftarrow 2 \leftarrow 3 \leftarrow 1$ of the axes transforms the integral into $\int_{-\infty}^{+\infty}\eta_2\,S(\eta_1^2+\eta_2^2+\eta_3^2)\,\mathrm{d}\eta_1\,\mathrm{d}\eta_2\,\mathrm{d}\eta_3$, where the integration limits of each coordinate are unchanged. As the value of an integral does not depend on the names of the axes, the two integrals above are equal to each other.

with P the Fermi-Dirac statistics. The procedure is similar to that of the drift-diffusion model. At equilibrium one can replace $\mathbf{c}$ with $\mathbf{u}$; then, out of the three components of $\mathbf{u}$, only that of index i contributes to the integral, because those of index $j \neq i$ make the integrand odd with respect to both k_i and k_j. In other terms, the integrand in (19.157) is replaced with $E_e\, u_i^2\, P\, \mathbf{i}_i$, which differs from the integrand of (19.151) because of factor $E_e \mathbf{i}_i$; this shows that the tensor defined by (19.157) is diagonal as well.[45] When the transformation (19.152) is used, the right-hand side of (19.157) becomes similar to (19.153), the only difference being the additional factor $\eta^2\, \mathbf{i}_i$ that derives from $E_e \mathbf{i}_i$. The next step, replacing the common value of the three integrals with one third of their sum, yields an expression similar to (19.154), the only difference being that η^2 is replaced with $\eta^4\, \mathbf{i}_i$. Transforming into spherical coordinates and inserting (C.111) with $\Gamma(1 + 5/2) = (5/2)\,\Gamma(1 + 3/2) = (15/4)\,\sqrt{\pi}/2$ finally yield

$$n^{\mathrm{eq}}\, m_{ia}\, \left(\overline{E_e\, u_i\, \mathbf{c}}\right)^{\mathrm{eq}} = \mathbf{i}_i\, \frac{5}{2}\, (k_B\, T)^2\, N_C\, \Phi_{5/2}(\xi_e)\,. \tag{19.158}$$

It follows

$$m_{ia}\, \left(\overline{E_e\, u_i\, \mathbf{c}}\right)^{\mathrm{eq}} = \mathbf{i}_i\, \frac{5}{2}\, (k_B\, T)^2\, \frac{\Phi_{5/2}(\xi_e)}{\Phi_{1/2}(\xi_e)} \simeq \mathbf{i}_i\, \frac{5}{2}\, (k_B\, T)^2\,, \tag{19.159}$$

where the approximation holds in the nondegenerate case.

19.6.5 Matthiessen's Rule

The effects of the inter-band and intra-band transitions have been separated in Sect. 19.3.2 under the assumption that the two types are uncorrelated. The separation may further be pursued within each class, depending on the entity with which the collision occurs. Here the collisions leading to the intra-band transitions only are considered. With reference to (19.45), and assuming that the intra-band transitions are uncorrelated, one lets $S_{0v} = S_{0v}^{(1)} + S_{0v}^{(2)} + \ldots$, whence

$$\frac{1}{\tau_v} = \frac{1}{\tau_{v1}} + \frac{1}{\tau_{v2}} + \ldots\,, \qquad \frac{1}{\tau_{vj}} = \iiint_{-\infty}^{+\infty} S_{0v}(j)(\mathbf{r}, \mathbf{k} \to \mathbf{k}')\, \mathrm{d}^3 k'\,. \tag{19.160}$$

This way of combining the relaxation times, also called *Matthiessen's rule*, still holds in the definitions of the macroscopic relaxation times τ_p, τ_w, and τ_b (Sect. 19.4.6), and in the definitions (19.104), (19.118) of electron and hole mobilities through τ_p.

[45] As above, for this result to hold the parabolic-band approximation is not necessary.

19.6.6 Order of Magnitude of Mobility and Conductivity

As shown in Sect. 19.5.2, the electron and hole mobilities are expressed by relations of the form $\mu = q\,\tau_p/m^*$, where τ_p is the momentum-relaxation time and m^* an effective mass. To the purpose of estimating the order of magnitude it is not necessary to distinguish between electron and hole cases. Considering the $T = 300$ K case and taking $\tau_p \approx 0.25 \times 10^{-12}$ s, $m^* \approx 0.4 \times 10^{-30}$ kg yield[46]

$$\mu = q\,\frac{\tau_p}{m^\star} \approx 1.60 \times 10^{-19}\ \text{C} \times \frac{0.25 \times 10^{-12}\ \text{s}}{0.4 \times 10^{-30}\ \text{kg}} = 10^3 \quad \frac{\text{cm}^2}{\text{V s}}\,. \tag{19.161}$$

The diffusion coefficient at $T = 300$ K is estimated from the Einstein relation (19.102), $D = (k_B\,T/q)\,\mu$, where

$$\frac{k_B\,T}{q} \approx \frac{1.38 \times 10^{-23}\ (\text{J/K}) \times 300\ \text{K}}{1.60 \times 10^{-19}\ \text{C}} = 26 \times 10^{-3} \quad \text{V}\,. \tag{19.162}$$

One finds

$$D \approx 26 \times 10^{-3}\ \text{V} \times 10^3\ \frac{\text{cm}^2}{\text{V s}} = 26 \quad \frac{\text{cm}^2}{\text{s}}\,. \tag{19.163}$$

To estimate the conductivity one takes by way of example the expression for electrons (19.133), $\sigma_n = q\,\mu_n\,n$, where, due to the estimates above, it is $q\,\mu_n \approx 1.60 \times 10^{-19}\ \text{C} \times 10^3\ \text{cm}^2/(\text{V s}) = 1.60 \times 10^{-16}\ \text{cm}^2/\Omega$. For silicon at $T = 300$ K it is $n = n_i \approx 10^{10}\ \text{cm}^{-3}$ (Table 18.2)[47]; in comparison, when silicon is doped with a uniform donor concentration equal to, say, $N'_D = 10^{16}\ \text{cm}^{-3}$, it is $n \simeq N'_D = 10^{16}\ \text{cm}^{-3}$ (compare with (18.30)). In conclusion,

$$\sigma_n[\text{intrinsic}] \approx 10^{-6}\ (\Omega\,\text{cm})^{-1}\,, \qquad \sigma_n[N'_D] \approx 1\ (\Omega\,\text{cm})^{-1}\,. \tag{19.164}$$

As a further comparison, the estimates for an insulator and, respectively, a conductor are made with $n \simeq 10^4\ \text{cm}^{-3}$ and $n \simeq 10^{22}\ \text{cm}^{-3}$ to find

$$\sigma_n[\text{insulator}] \approx 10^{-12}\ (\Omega\,\text{cm})^{-1}\,, \qquad \sigma_n[\text{conductor}] \approx 10^{6}\ (\Omega\,\text{cm})^{-1}\,. \tag{19.165}$$

[46] Mobility is traditionally expressed in $\text{cm}^2/(\text{V s})$ instead of $\text{m}^2/(\text{V s})$.

[47] The equilibrium concentrations are used in the estimates.

19.6.7 Onsager Relations

A method widely used in Classical Mechanics to treat transport problems is based on a simplified form of (19.48), which is illustrated below. To avoid a change in notation, the method will be described as if it were applied to the conduction band of a semiconductor. To this purpose, consider the equilibrium statistics for the electrons of the conduction band in a uniform case, given by (18.54) with $\varphi = 0$, and recast the numerator of the exponent as $E - E_F = E_e + E_C - E_F$, with $E_e = E(\mathbf{k}) - E_C$ the kinetic energy of the electrons (compare with (19.24)). Taking the nondegenerate form of (18.54), the equilibrium distribution function $f^{\mathrm{eq}} = Q\,P$ becomes $f^{\mathrm{eq}} = f_0 \exp[-E_e(\mathbf{k})/(k_B\,T)]$. The simplified form of (19.48) is obtained, first, by letting $\mathbf{B} = 0$ and considering the steady-state case; then, the distribution function acted upon by the Liouvillian operator is replaced with $f_a = f_0 \exp[-E_e(\mathbf{k})/(k_B\,T_a)]$, with $T_a(\mathbf{r})$ a function to be determined. The transport equation (19.48) then becomes $f = f^{\mathrm{eq}} - \tau\,\mathscr{L} f_a$; observing that

$$\mathrm{grad}_k f_a = -\frac{\hbar\,\mathbf{u}}{k_B\,T_a} f_a\,, \qquad \mathrm{grad}_r f_a = \frac{E_e}{k_B\,T_a}\,\frac{\mathrm{grad}_r T_a}{T_a} f_a\,, \tag{19.166}$$

its explicit form reads

$$f = f^{\mathrm{eq}} - \frac{\tau}{k_B\,T_a}\left[q\,\mathbf{E}\cdot\mathbf{u} + E_e\,\mathbf{u}\cdot\frac{\mathrm{grad}_r T_a}{T_a}\right] f_a\,. \tag{19.167}$$

Then, one calculates the moments of order one and three of (19.167); using the symbols of Sect. 19.4.5 yields, for the moment of order one,

$$-n\,v_i = \frac{q\,\mathbf{E}}{k_B\,T_a}\cdot\iiint_{-\infty}^{+\infty} u_i\,\mathbf{u}\,\tau\,f_a\,\mathrm{d}^3k + \frac{\mathrm{grad}_r T_a}{k_B\,T_a^2}\cdot\iiint_{-\infty}^{+\infty} E_e\,u_i\,\mathbf{u}\,\tau\,f_a\,\mathrm{d}^3k\,. \tag{19.168}$$

Similarly, the moment of order three yields

$$-n\,b_i = \frac{q\,\mathbf{E}}{k_B\,T_a}\cdot\iiint_{-\infty}^{+\infty} E_e\,u_i\,\mathbf{u}\,\tau\,f_a\,\mathrm{d}^3k + \frac{\mathrm{grad}_r T_a}{k_B\,T_a^2}\cdot\iiint_{-\infty}^{+\infty} E_e^2\,u_i\,\mathbf{u}\,\tau\,f_a\,\mathrm{d}^3k\,. \tag{19.169}$$

Note that the coefficient of the second term at the right-hand side of (19.168) is identical to the coefficient of the first term at the right-hand side of (19.169); this identity is also called *Onsager relation*. Similar relations hold for the next moments; in fact, the moment of order five is obtained by integrating (19.167) over $\mathbf{k}$ after multiplying it by $E_e^2\,u_i$, and so on.

The above finding holds irrespective of the dependence of τ on $\mathbf{k}$ because it is essentially based on the form of f_a. On the other hand, the analysis of the coefficients of (19.168) and (19.169) is greatly simplified when the dependence is of the form $\tau = \tau(E_e)$; thanks to spherical symmetry one can in fact follow the same reasoning as in Sect. 19.6.4: the only even component of $u_i\,\mathbf{u}$ is in this case $u_i^2\,\mathbf{i}_i$, whence $\mathbf{E}\cdot u_i\,\mathbf{u}$

is replaced with $E_i\,u_i^2$ and $\mathrm{grad}_r T_a \cdot u_i\,\mathbf{u}$ is replaced with $(\partial T_a/\partial x_i)\,u_i^2$. The integrals of the other components vanish. The relation between f_0 and the electron concentration n is given by

$$n = \iiint_{-\infty}^{+\infty} f_0\,\exp[-E_e/(k_B\,T_a)]\,\mathrm{d}^3k\,. \tag{19.170}$$

In the spherically symmetric case considered here it is $E_e = \hbar^2\,(k_1^2+k_2^2+k_3^2)/(2\,m)$ whence, letting $\eta_i = \hbar\,k_i/\sqrt{2\,m\,k_B\,T_a}$ and turning to spherical coordinates, $\mathrm{d}\eta^3 = \eta^2\,\mathrm{d}\eta\,\sin\vartheta\,\mathrm{d}\vartheta\,\mathrm{d}\varphi$, one finds, after an integration over the angles that yields a factor $4\,\pi$,

$$n = 4\,\pi\,f_0\left(\frac{2\,m\,k_B\,T_a}{\hbar^2}\right)^{3/2}\int_0^{+\infty}\eta^2\,\exp(-\eta^2)\,\mathrm{d}\eta\,. \tag{19.171}$$

The integral in (19.171) is equal to $\sqrt{\pi}/4$ due to C.33. From now on it is convenient to use the components of $\mathbf{u} = (\hbar/m)\,\mathbf{k}$ as integration variables in (19.168) and (19.169); thus, $f_0\,\mathrm{d}^3k$ must be replaced with $F_0\,\mathrm{d}^3u$, with $F_0 = f_0\,(m/\hbar)^3$ and[48]

$$n = \left(\frac{2\,\pi\,k_B\,T_a}{m}\right)^{3/2} F_0\,. \tag{19.172}$$

To proceed one observes that, apart from factor $F_0\,\tau$, all integrands in (19.168) and (19.169) have the form $E_e^s\,u_i^2\,\exp[-E_e/(k_B\,T_a)]\,\mathrm{d}^3u$ with $s = 0, 1, 2$; thus, they can be manipulated in the same way. First, like in Sect. 19.6.4, one replaces u_i with $u^2/3$; this makes the coefficients independent of the coordinate index. Second, one turns to spherical coordinates and integrates over the angles, thus yielding a factor $4\,\pi$. Third, one lets $\lambda = E_e/(k_B\,T_a)$; the integrand thus becomes $(4/3)\,\pi\,(k_B\,T_a)^s\,\lambda^s\,u^4\,\exp(-\lambda)\,\mathrm{d}u$. On the other hand, it is $m\,u^2/2 = k_B\,T_a\,\lambda$, whence $u^4\,\mathrm{d}u = (2\,k_B\,T_a/m)^{5/2}\,(1/2)\,\lambda^{3/2}\,\mathrm{d}\lambda$. Combining with the previous one and using (19.172) yield the general form of the integrands of (19.168) and (19.169):

$$n\,\frac{\tau}{m}\,\frac{4}{3\,\sqrt{\pi}}\,(k_B\,T_a)^{s+1}\,\lambda^{s+3/2}\,\exp(-\lambda)\,\mathrm{d}\lambda\,. \tag{19.173}$$

A simple approximation for the relaxation time is $\tau = \tau_0\,(E_e/E_0)^{-\alpha}$, where τ_0, E_0, and α are positive parameters independent of E_e (compare with Sect. 20.5.1). The approximation transforms (19.173) into

$$n\,\frac{\tau_0}{m}\,\frac{4}{3\,\sqrt{\pi}}\,(k_B\,T_a)^{s+1}\left(\frac{E_0}{k_B\,T_a}\right)^{\alpha}\lambda^{s-\alpha+3/2}\,\exp(-\lambda)\,\mathrm{d}\lambda\,. \tag{19.174}$$

[48] Note that f_0 is dimensionless, while the units of F_0 are $(s/m^2)^3$.

Replacing (19.174) into (19.168) after multiplying by q both sides of the latter and integrating with respect to λ from zero to infinity yield

$$-q\,n\,\mathbf{v} = \mathbf{J}_n = \sigma\,\mathbf{E} + \chi\,\frac{\mathrm{grad}_r T_a}{T_a}\,, \tag{19.175}$$

where, thanks to (C.95),

$$\sigma = q\,\frac{\tau_0}{m}\,n\,\frac{4}{3\,\sqrt{\pi}}\left(\frac{E_0}{k_B\,T_a}\right)^{\alpha}\,\Gamma(5/2-\alpha)\,, \qquad \chi = \frac{k_B\,T_a}{q}\,\frac{\Gamma(7/2-\alpha)}{\Gamma(5/2-\alpha)}\,\sigma\,. \tag{19.176}$$

By the same token one finds from (19.169)

$$-n\,\mathbf{b} = \chi\,\mathbf{E} + \kappa\,\mathrm{grad}_r T_a\,, \tag{19.177}$$

with

$$\kappa = \frac{k_B^2\,T_a}{q^2}\,\frac{\Gamma(9/2-\alpha)}{\Gamma(5/2-\alpha)}\,\sigma\,. \tag{19.178}$$

The σ and κ coefficients are the *electric conductivity* and *thermal conductivity* of the electrons, respectively.[49] The ratio κ/σ is proportional to T_a; this property has been observed first in metals and is called *Wiedemann-Franz law* [79, Chap. 10]. The ratio $\kappa/(\sigma\,T_a)$ is called *Lorenz number*. If the flow of electrons is prevented, letting $\mathbf{J}_n = 0$ in (19.175) shows that a temperature gradient produces an electric field $\mathbf{E} = -[\chi/(\sigma\,T_a)]\,\mathrm{grad}_r T_a$; this is called *thermoelectric effect*, with $\chi/(\sigma\,T_a)$ the *thermoelectric coefficient*. Eliminating $\mathbf{E}$ from (19.175) and (19.177) yields

$$-n\,\mathbf{b} = \frac{\chi}{\sigma}\,\mathbf{J}_n + \left(\kappa - \frac{\chi^2}{\sigma\,T_a}\right)\,\mathrm{grad}_r T_a\,, \tag{19.179}$$

where the coefficient of $\mathrm{grad}_r T_a$ is called *Peltier coefficient*.

The approach depicted in this section prescribes a form of the distribution function and, for this reason, is less general than the hydrodynamic model illustrated in section 19.5.1. It is interesting to check whether the Onsager relation is found also in the hydrodynamic equations describing the particle and energy flux, namely, the second of (19.90) and (19.92). Apparently, the Onsager relation is not fulfilled because $\tau_b \neq \tau_p$; however, it can be shown that when the perturbation with respect to equilibrium is small it is $\tau_b \to \tau_p$ [132]. Thus, the Onsager relation can also be deduced from the limiting case of the hydrodynamic model.

[49] These definitions hold also in the general case where the form of τ has no special symmetry. In this case, however, all coefficients of (19.168) and (19.169) are tensors.

19.6.8 A Resumé of the Transport Model's Derivation

The number of steps that lead to the hydrodynamic or drift-diffusion model for semiconductors is quite large; thus, a brief summary is of use. The starting point is the single-particle Schrödinger equation for an electron in a crystal. To reach this stage of the theory a considerable amount of work has already been spent, necessary to reduce the many-particle problem to a tractable form (Chap. 16). When the external forces are absent, the single-particle equation is recast in a form based on the equivalent Hamiltonian operator, that exploits the periodicity of the lattice (this implies in turn that the nuclei are kept fixed in the equilibrium positions). Finally, the external forces are added, assuming that the external potential energy is a small perturbation; thanks to this hypothesis, it is possible to describe the collisionless motion of a single electron by means of a Hamiltonian function whose canonical variables are the expectation values of the wave packet's position and momentum (Sects. 19.2.1 and 19.2.2).

Basing on the Hamiltonian function thus found, the analysis shifts from the description of the single-particle to the statistical treatment of a system made of a large number of electrons, following the same pattern as in the classical case; this leads to the semiclassical BTE (Sect. 19.3), for which the general form of the collision term is worked out (Sect. 19.3.1). The latter is simplified, first, by considering point-like collisions, then by taking the perturbative form of the collision operator (Sects. 19.3.2 and 19.3.3).

The perturbative form of the BTE is treated with the moments method, that provides a hierarchical set of models, e.g., the drift-diffusion and the hydrodynamic model. Important approximations at this stage are, for all moments of order larger than zero, the neglect of the inter-band transitions, of the time derivatives, and of the convective terms. The models reach the final form thanks to the hypothesis of parabolic bands and the approximation of the relaxation-time tensors with scalar quantities.

Problem

19.1 In the expressions (19.115), (19.118) defining the hole mobility μ_p, assume that $\tau_{ph} \simeq \tau_{pI}$. Letting τ_p be the common value, determine the value of the normalized effective mass $\overline{m}_h/m_0$ to be used in $\mu_p = q\,\tau_p/\overline{m}_h$ for silicon at room temperature. Also, determine the value of parameter a_p in (19.122) in the same conditions.

Chapter 20
Generation-Recombination and Mobility

20.1 Introduction

This chapter illustrates the main contributions to the transitions of the inter-band type, that give rise to the generation-recombination terms in the continuity equations for electrons and holes, and to those of the intra-band type, that give rise to the electron and hole mobilities in the current-density equations. The inter-band transitions that are considered are the net thermal recombinations (of the direct and trap-assisted type), Auger recombinations, impact-ionization generations, and net-optical recombinations. The model for each type of event is first given as a closed-form function of the semiconductor-device model's unknowns, like carrier concentrations, electric field, or current densities. Such functions contain a number of coefficients, whose derivation is successively worked out in the complements by means of a microscopic analysis. The case of semiconductors having a distribution of traps within the gap, like, e.g., polycrystalline silicon, is treated as well. Some discussion is devoted to the optical-generation and recombination events to show how the concepts of semiconductor laser, solar cell, and optical sensor may be derived as particular cases of nonequilibrium interaction between the material and an electromagnetic field. The intra-band transitions are treated in a similar manner: two examples, the collisions with acoustic phonons and ionized impurities, are worked out in some detail; the illustration then follows of how the contributions from different scattering mechanisms are combined together in the macroscopic mobility models. The material is supplemented with a brief discussion about advanced modeling methods.

M. Rudan, *Physics of Semiconductor Devices*, DOI 10.1007/978-3-319-63154-7_20

20.2 Net Thermal Recombinations

As anticipated in Sect. 19.5.5, it is customary to separate the net generation rates W_n, W_p into two contributions, namely, those deriving from the phonon collisions and those of the other types (e.g., electron-electron collisions, electron-photon collisions, and so on). The separate contributions are defined in (19.132); this section deals with the net thermal recombination rates U_n, U_p.

In the calculations carried out below, the nonequilibrium carrier concentrations are derived by integrating over the bands' energy. This is consistent with the general definitions (19.31) and (19.109). In fact, considering the nonequilibrium electron concentration n as defined in (19.31), one introduces the variable transformation illustrated in Sect. B.5 and replaces the quantities appearing in it as follows:

$$(u, v, w) \leftarrow (k_1, k_2, k_3)\,, \qquad \sigma \leftarrow (\mathbf{r}, t)\,, \qquad \eta \leftarrow E\,, \tag{20.1}$$

$$S \leftarrow n\,, \qquad s \leftarrow f = Q\,\Phi\,, \qquad b \leftarrow \gamma\,, \qquad \bar{s} \leftarrow P\,, \tag{20.2}$$

where Q, $\gamma(E)$ are, respectively, the density of states in the phase space $\mathbf{r}, \mathbf{k}$ and the combined density of states in energy and $\mathbf{r}$ space, while $\Phi(\mathbf{r}, \mathbf{k}, t)$, $P(\mathbf{r}, E, t)$ are the nonequilibrium occupation probabilities in the phase space and, respectively, in energy; the integration in energy is carried out over the range corresponding to the conduction band's branch. The hole concentration is treated in the same manner. In conclusion,

$$n(\mathbf{r}, t) = \iiint_{-\infty}^{+\infty} Q\,\Phi\,\mathrm{d}^3k = \int_{E_C}^{E_{CU}} \gamma\,P\,\mathrm{d}E\,, \tag{20.3}$$

$$p(\mathbf{r}, t) = \iiint_{-\infty}^{+\infty} Q\,(1-\Phi)\,\mathrm{d}^3k = \int_{E_{VL}}^{E_V} \gamma\,(1-P)\,\mathrm{d}E\,. \tag{20.4}$$

20.2.1 Direct Thermal Recombinations

To begin, a graphic example of thermal transitions is shown in Fig. 20.1, where the edges of the conduction and valence bands are indicated with the same symbols used in Sect. 18.2; the transition marked with a is a recombination event, in which an electron belonging to an energy state of the conduction band transfers to an empty state of the valence band. The energy difference between the initial and final state is released to the lattice in the form of a phonon. The opposite transition, where the electron's energy increases due to phonon absorption, is an electron-hole generation and is marked with b in the figure. The transitions of type a and b are

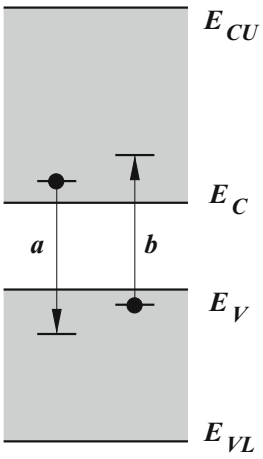


Fig. 20.1 A graphic example of direct thermal recombination (*a*) and generation (*b*). The edges of the conduction and valence bands are indicated with the same symbols used in Sect. 18.2. The same drawing applies also to the description of the direct optical recombinations and generations (Sect. 20.4)

called *direct thermal recombination* and *direct thermal generation*, respectively. Let r_a be the number of direct thermal recombination per unit volume and time, and r_b the analogue for the generations; considering the conduction band as a reference, the difference $r_a - r_b$ provides the contribution to the net thermal recombination rate U_n due to the direct thermal transitions. When the valence band is considered instead, the rates of electrons transitions reverse; however, for the valence band the transitions of holes must be considered: as consequence, the contribution to U_p is again $r_a - r_b$. In conclusion,

$$U_{DT} = U_{DTn} = U_{DTp} = r_a - r_b \,, \tag{20.5}$$

where D stands for "direct" and T for "thermal." The expressions of r_a, r_b are determined by a reasoning similar to that used in Sect. 19.3.1 to express the collision term of the BTE; here, however, the analysis is carried out directly in the energy space instead of the $\mathbf{k}$ space.[1] Let $P(\mathbf{r}, E, t)$ be the occupation probability of a state at energy E; then, let C be the probability per unit time and volume (in $\mathbf{r}$) of an electron transition from a state of energy E to a state of energy E' belonging to a different band, induced by the interaction with a phonon.[2] Such a probability depends on the phonon energy $\hbar\,\omega$ (Sect. 12.5), and also on the position in $\mathbf{r}$ if the semiconductor is nonuniform. Typically, the equilibrium distribution is assumed for the phonons, which makes C independent of time; as the collisions are point-like

[1] A more detailed example of calculations is given below, with reference to collisions with ionized impurities.

[2] The units of C are $[C] = \mathrm{m}^{-3}\,\mathrm{s}^{-1}$. Remembering that the phonon energy equals the change in energy of the electron due to the transition (Sect. 14.8.2), it is $C = 0$ for $\hbar\,\omega < E_C - E_V = E_G$ (refer also to Fig. 20.1).

(Sect. 19.3.2), the spatial positions of the initial and final states coincide, whence $C = C(\mathbf{r}, \hbar\,\omega, E \to E')$.

Indicating with $g(E)$ the density of states of the band where the initial state belongs, the product $g\,\mathrm{d}E\,P$ is the number of electrons within the elementary interval $\mathrm{d}E$ around the initial state; such a product is multiplied by C to find the number of unconditional $E \to E'$ transitions per unit time and volume. On the other hand, the transitions take place only if the final states around E' are empty; as the empty states in that interval are $g'\,\mathrm{d}E'\,(1 - P')$, the number of actual transitions per unit time and volume from $\mathrm{d}E$ to $\mathrm{d}E'$ turns out to be $g\;\mathrm{d}E\,P\,C\,g'\,\mathrm{d}E'\,(1 - P')$. Now, to calculate the r_a or r_b rate it is necessary to add up all transitions: for r_a one lets E range over the conduction band and E' over the valence band; the converse is done for r_b. As the calculation of the latter is somewhat easier, it is shown first:

$$r_b = \int_{E_{VL}}^{E_V} g\,\mathrm{d}E\,P \int_{E_C}^{E_{CU}} C\,g'\,\mathrm{d}E'\,\left(1 - P'\right)\,. \tag{20.6}$$

As in normal operating conditions the majority of the valence-band states are filled, while the majority of the conduction-band states are empty, one lets $P \simeq 1$ and $1 - P' \simeq 1$, whence, using symbol G_{DT} for r_b,

$$G_{DT}(\mathbf{r}, \hbar\,\omega) = \int_{E_{VL}}^{E_V} g\,\mathrm{d}E \int_{E_C}^{E_{CU}} C\,g'\,\mathrm{d}E'\,. \tag{20.7}$$

Thus, the generation rate is independent of the carrier concentrations. To proceed, one uses the relation $g = \Omega\,\gamma$, with γ the combined density of states in energy and volume, given by (15.65), and the definition (20.4) of the hole concentration. Thus, the recombination rate is found to be

$$r_a = \int_{E_C}^{E_{CU}} g\;\mathrm{d}E\,P \int_{E_{VL}}^{E_V} C\,g'\,\mathrm{d}E'\;(1 - P') = p \int_{E_C}^{E_{CU}} K\,g\,P\;\mathrm{d}E\,, \tag{20.8}$$

where $K(\mathbf{r}, \hbar\,\omega, E)$, whose units are $[K] = \mathrm{s}^{-1}$, is the average of $\Omega\,C$ over the valence band, weighed by $g'\,(1 - P')$:

$$K = \frac{\int_{E_{VL}}^{E_V} \Omega\,C\,g'\,(1 - P')\,\mathrm{d}E'}{\int_{E_{VL}}^{E_V} g'\,(1 - P')\,\mathrm{d}E'}\,. \tag{20.9}$$

Strictly speaking, K is a functional of P'; however, the presence of P' in both numerator and denominator of (20.9) makes such a dependence smoother, so that one can approximate K using the equilibrium distribution instead of P'. By the same token one uses the definition of the electron concentration (20.3) to find

$$r_a = \alpha_{DT}\, n\, p\,, \qquad \alpha_{DT}(\mathbf{r}, \hbar\omega) = \frac{\int_{E_C}^{E_{CU}} \Omega\, K\, g\, P\, \mathrm{d}E}{\int_{E_C}^{E_{CU}} g\, P\, \mathrm{d}E}\,, \tag{20.10}$$

where the integrals are approximated using the equilibrium probability. In conclusion,

$$U_{DT} = \alpha_{DT}\, n\, p - G_{DT}\,, \tag{20.11}$$

where α_{DT} is the *transition coefficient* of the direct thermal transitions, with units $[\alpha_{DT}] = \mathrm{m}^3\ \mathrm{s}^{-1}$, and G_{DT} their *generation rate* ($[G_{DT}] = \mathrm{m}^{-3}\ \mathrm{s}^{-1}$). As in the equilibrium case it is $r_a = r_b$, namely, $G_D = \alpha_D\, n^{\mathrm{eq}}\, p^{\mathrm{eq}}$, it follows $U_{DT} = \alpha_{DT}\, (np - n^{\mathrm{eq}} p^{\mathrm{eq}})$.

20.2.2 Trap-Assisted Thermal Recombinations

An important contribution to the thermal generation and recombination phenomena is due to the so-called *trap-assisted transitions*. As mentioned in Sect. 19.3, among the possible collisions undergone by electrons or holes are those with *lattice defects*. The latter may originate from lattice irregularities (e.g., dislocations of the material's atoms occurring during the fabrication process, Sect. 24.1), or from impurities that were not eliminated during the semiconductor's purification process, or were inadvertently added during a fabrication step. Some defects may introduce energy states localized in the gap; such states, called *traps*, may capture an electron from the conduction band and release it towards the valence band, or vice versa. The phenomena are illustrated in Fig. 20.2, where four traps located in the energy gap are shown in order to distinguish among the different transition events, that are: *a*) capture of a conduction-band electron by a trap, *b*) release of a trapped electron towards the conduction band, *c*) release of a trapped electron towards the valence band (more suitably described as the capture of a valence-band hole by the trap),

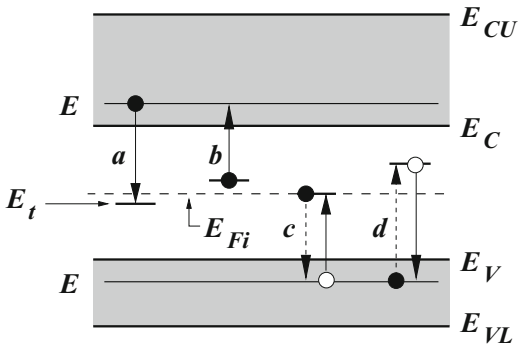


Fig. 20.2 Different types of trap-assisted transitions

and *d*) capture of a valence-band electron from the valence band (more suitably described as the release of a hole towards the valence band). Each transition is accompanied by the absorption or emission of a phonon. Thus, transitions of type *a* and *b* contribute to the net thermal recombination U_n of the conduction band, while those of type *c* and *d* contribute to the net thermal recombination U_p of the valence band. Also, a sequence of two transitions, one of type *a* involving a given trap, followed by one of type *c* involving the same trap, produces an electron-hole recombination and is therefore called *trap-assisted thermal recombination*; similarly, a sequence of two transitions, one of type *d* involving a given trap, followed by one of type *b* involving the same trap, produces an electron-hole generation and is therefore called *trap-assisted thermal generation*.

To calculate the contribution of the trap-assisted transitions to U_n and U_p it is necessary to distinguish between two kinds of traps: those of *donor type*, that are electrically neutral when the electron is present in the trap and become positively charged when the electron is released, and those of *acceptor type*, that are electrically neutral when the electron is absent from the trap and become negatively charged when the electron is captured. In this respect, the traps are similar to the dopants' atoms. Instead, a strong difference is made by the position of the traps' energy within the gap. Consider, for instance, traps localized near the gap's midpoint (the latter is indicated by the intrinsic Fermi level E_{Fi} in Fig. 20.2); the phonon energy necessary for the transition is about $E_G/2$ in all cases, to be compared with the value E_G necessary for a direct transition. On the other hand, the equilibrium-phonon distribution (Sect. 16.6) is the Bose-Einstein statistics (15.55); it follows that the number $\mathrm{d}N_{\mathrm{ph}}$ of phonons in the interval $\mathrm{d}\omega$ is

$$\mathrm{d}N_{\mathrm{ph}} = \frac{g_{\mathrm{ph}}(\omega)\,\mathrm{d}\omega}{\exp[\hbar\,\omega/(k_B\,T)] - 1}\,, \tag{20.12}$$

with $\hbar\,\omega$ the energy and g_{ph} the density of states of the phonons. Due to (20.12), $\mathrm{d}N_{\mathrm{ph}}/\mathrm{d}\omega$ rapidly decreases as the phonon energy increases, thus making the probability of an electron-phonon interaction much larger at lower energies. For this reason, even in an electronic-grade semiconductor, where the concentration of defects is very small (Sect. 19.3.2), the traps are able to act as a sort of "preferred path" in energy for the inter-band transitions, to the extent that the contribution to U_n, U_p of the trap-assisted transitions is largely dominant over that of the direct transitions. Therefore, in the continuity equations (20.13) below, and in the subsequent derivation of the trap-assisted, thermal-transition rates, symbols U_n, U_p refer only to the latter transitions, not any more to the sum of the trap-assisted and direct ones.

The net thermal-recombination terms U_n, U_p appear in (19.129) and (19.130) after replacing W_n, W_p with (19.132); this yields

$$\frac{\partial n}{\partial t} + U_n - \frac{1}{q}\,\mathrm{div}\mathbf{J}_n = G_n\,, \qquad \frac{\partial p}{\partial t} + U_p + \frac{1}{q}\,\mathrm{div}\mathbf{J}_p = G_p\,. \tag{20.13}$$

To introduce the trap-assisted transitions one formally duplicates (20.13) as if the acceptor and donor traps formed two additional bands; as the acceptor traps are either neutral or negatively charged, the charge and current densities of the band associated with them are thought of as due to electrons; instead, the charge and current densities of the band associated with the donor traps are thought of as due to holes. In summary, the two additional equations read

$$\frac{\partial n_A}{\partial t} + U_{nA} - \frac{1}{q}\,\mathrm{div}\mathbf{J}_{nA} = G_{nA}\,, \qquad \frac{\partial p_D}{\partial t} + U_{pD} + \frac{1}{q}\,\mathrm{div}\mathbf{J}_{pD} = G_{pD}\,, \tag{20.14}$$

with a and d standing for "acceptor" and "donor," respectively. To ease the calculation it is assumed that the nonthermal phenomena are absent, whence $G_n = G_p = G_{nA} = G_{pD} = 0$. Combining (20.13) with (20.14) and observing that $\mathbf{J} = \mathbf{J}_p + \mathbf{J}_{pD} + \mathbf{J}_n + \mathbf{J}_{nA}$ is the total current density of the semiconductor yield

$$\frac{\partial[q\,(p + p_D - n - n_A)]}{\partial t} + \mathrm{div}\mathbf{J} = q\,(U_n + U_{nA}) - q\,(U_p + U_{pD})\,. \tag{20.15}$$

As the net dopant concentration N is independent of time, it is $\partial[q\,(p + p_D - n - n_A)]/\partial t = \partial[q\,(p + p_D - n - n_A + N)]/\partial t = \partial\varrho/\partial t$; thus, the left-hand side of (20.15) vanishes due to (4.23), and[3]

$$U_n + U_{nA} = U_p + U_{pD}\,. \tag{20.16}$$

The two continuity equations (20.14) are now simplified by observing that in crystalline semiconductors the current densities $\mathbf{J}_{pD}$, $\mathbf{J}_{nA}$ of the traps are negligible. In fact, the trap concentration is so low that inter-trap tunneling is precluded by the large distance from a trap to another; the reasoning is the same as that used in Sect. 18.7.2 with respect to the impurity levels.[4] Letting $\mathbf{J}_{pD} = \mathbf{J}_{nA} = 0$ makes the two equations (20.14) local:

$$\frac{\partial n_A}{\partial t} = -U_{nA}\,, \qquad \frac{\partial p_D}{\partial t} = -U_{pD}\,. \tag{20.17}$$

In steady-state conditions the traps' populations are constant, thus yielding $U_{nA} = U_{pD} = 0$ and, from (20.16), $U_n = U_p$. In equilibrium all continuity equations reduce to the identity $0 = 0$, whence the net-recombination terms vanish independently, $U_n^{\mathrm{eq}} = U_{nA}^{\mathrm{eq}} = U_p^{\mathrm{eq}} = U_{pD}^{\mathrm{eq}} = 0$.

[3] The result expressed by (20.16) is intuitive if one thinks that adding up all continuity equations amounts to counting all transitions twice, the first time in the forward direction (e.g., using the electrons), the second time in the backward direction (using the holes). The reasoning is similar to that leading to the vanishing of the intra-band contribution in (19.63).

[4] In a polycrystalline semiconductor with a large spatial concentration of traps it may happen that the traps' current densities are not negligible; in fact, the whole system of equations (20.13) and (20.14) must be used to correctly model the material [26–28]. The conduction phenomenon associated with these current densities is called *gap conduction*.

20.2.3 Shockley-Read-Hall Theory

The *Shockley-Read-Hall theory* describes the trap-assisted, net thermal-recombination term in a crystalline semiconductor based upon the steady-state relation $U_n = U_p$. In fact, the outcome of the theory is used also in dynamic conditions; this approximation is acceptable because, due to the smallness of the traps' concentration, the contribution of the charge density stored within the traps is negligible with respect to that of the band and dopant states; the contribution of the time variation of the traps' charge density is similarly negligible. The theory also assumes that only one trap level is present, of energy E_t; with reference to Fig. 20.2, the trap levels must be thought of as being aligned with each other. If more than one trap level is present, the contributions of the individual levels are summed up at a later stage. In the theory it is not important to distinguish between acceptor-type or donor-type traps; however, one must account for the fact that a trap can accommodate one electron at most.

Still with reference to Fig. 20.2, let r_a be the number of type-a transitions per unit volume and time, and similarly for r_b, r_c, r_d. The derivation of these rates is similar to that of the direct transitions and is shown in the complements; here the expressions of the net thermal-recombination terms are given, that read

$$U_n = r_a - r_b = \alpha_n \, n \, N_t \, (1 - P_t) - e_n \, N_t \, P_t \,, \tag{20.18}$$

$$U_p = r_c - r_d = \alpha_p \, p \, N_t \, P_t - e_p \, N_t \, (1 - P_t) \,, \tag{20.19}$$

where N_t is the concentration of traps of energy E_t, P_t the trap-occupation probability, α_n, α_p the *electron- and hole-transition coefficients*, respectively, and e_n, e_p the *electron- and hole-emission coefficients*, respectively.[5] The ratios e_n/α_n, e_p/α_p are assumed to vary little from the equilibrium to the nonequilibrium case. From $U_n^{\text{eq}} = U_p^{\text{eq}} = 0$ one derives

$$\frac{e_n}{\alpha_n} = n^{\text{eq}} \left(\frac{1}{P_t^{\text{eq}}} - 1 \right) , \qquad \frac{e_p}{\alpha_p} = p^{\text{eq}} \left(\frac{1}{P_t^{\text{eq}}} - 1 \right)^{-1} . \tag{20.20}$$

The occupation probability at equilibrium is the modified Fermi-Dirac statistics (compare with (18.21) or (18.36))

$$P_t^{\text{eq}} = \left[\frac{1}{d_t} \exp\left(\frac{E_t - E_F}{k_B \, T} \right) + 1 \right]^{-1} , \qquad \frac{1}{P_t^{\text{eq}}} - 1 = \frac{1}{d_t} \exp\left(\frac{E_t - E_F}{k_B \, T} \right) , \tag{20.21}$$

with d_t the degeneracy coefficient of the trap. It follows, after introducing the shorthand notation $n_B = e_n/\alpha_n$, $p_B = e_p/\alpha_p$,

[5] It is $[\alpha_{n,p}] = \text{m}^3 \, \text{s}^{-1}$, $[e_{n,p}] = \text{s}^{-1}$.

$$n_B = \frac{n^{\mathrm{eq}}}{d_t}\,\exp\left(\frac{E_t - E_F}{k_B\,T}\right), \qquad p_B = p^{\mathrm{eq}}\,d_t\,\exp\left(\frac{E_F - E_t}{k_B\,T}\right). \tag{20.22}$$

Note that $n_B\,p_B = n^{\mathrm{eq}}\,p^{\mathrm{eq}}$. Replacing (20.22) into (20.18), (20.19) and letting $U_n = U_p$ yield

$$\alpha_n\,n\,(1 - P_t) - \alpha_n\,n_B\,P_t = \alpha_p\,p\,P_t - \alpha_p\,p_B\,(1 - P_t), \tag{20.23}$$

whence

$$P_t = \frac{\alpha_n\,n + \alpha_p\,p_B}{\alpha_n\,(n + n_B) + \alpha_p\,(p + p_B)}, \qquad 1 - P_t = \frac{\alpha_n\,n_B + \alpha_p\,p}{\alpha_n\,(n + n_B) + \alpha_p\,(p + p_B)}. \tag{20.24}$$

In this way one expresses the trap-occupation probability as a function of two of the unknowns of the semiconductor-device model, namely, n and p, and of a few parameters. Among the latter, n_B and p_B are known (given the trap's energy) because they are calculated in the equilibrium condition. In conclusion, replacing (20.24) into (20.18) or (20.19) yields, for the common value $U_{\mathrm{SRH}} = U_n = U_p$,

$$U_{\mathrm{SRH}} = \frac{n\,p - n^{\mathrm{eq}}\,p^{\mathrm{eq}}}{(n + n_B)/(N_t\,\alpha_p) + (p + p_B)/(N_t\,\alpha_n)}, \tag{20.25}$$

where the indices stand for "Shockley-Read-Hall." Eventually, the only unknown parameters turn out to be the products $N_t\,\alpha_p$ and $N_t\,\alpha_n$ which, as shown in Sect. 25.2, can be obtained from measurements.

The expression obtained so far, (20.25), has been derived considering a single trap level E_t. Before adding up over the levels it is convenient to consider how sensitive U_{SRH} is to variations of E_t; in fact, one notes that the numerator of (20.25) is independent of E_t, whereas the denominator D has the form

$$D = c + 2\,\lambda\,\cosh\eta, \qquad \eta = \frac{E_t - E_F}{k_B\,T} + \frac{1}{2}\,\log\mu, \tag{20.26}$$

where

$$c = \frac{1}{N_t}\left(\frac{n}{\alpha_p} + \frac{p}{\alpha_n}\right), \qquad \lambda = \frac{1}{N_t}\sqrt{\frac{n^{\mathrm{eq}}}{\alpha_p}\,\frac{p^{\mathrm{eq}}}{\alpha_n}}, \qquad \mu = \frac{1}{d_t^2}\,\frac{n^{\mathrm{eq}}/\alpha_p}{p^{\mathrm{eq}}/\alpha_n}. \tag{20.27}$$

The denominator has a minimum where $\eta = 0$; thus, U_{SRH} has a maximum there. Moreover, the maximum is rather sharp due to the form of the hyperbolic cosine. It follows that the trap level E_{tM} that most efficiently induces the trap-assisted transitions is found by letting $\eta = 0$. The other trap levels have a much smaller efficiency and can be neglected; in conclusion, it is not necessary to add up over the trap levels.[6] With this provision, one finds

[6] This simplification is not applicable in a polycrystalline or amorphous semiconductor.

$$E_{tM} = E_F + \frac{k_B\,T}{2}\,\log\left(d_t^2\,\frac{p^{\rm eq}/\alpha_n}{n^{\rm eq}/\alpha_p}\right)\,. \tag{20.28}$$

An estimate of E_{tM} is easily obtained by considering the nondegenerate condition, whence $n^{\rm eq} = N_C\,\exp[(E_F - E_C)/(k_B\,T)]$ and $p^{\rm eq} = N_V\,\exp[(E_V - E_F)/(k_B\,T)]$ (compare with (18.28)). It follows

$$E_{tM} \simeq \frac{E_C + E_V}{2} + \frac{k_B\,T}{2}\,\log\left(d_t^2\,\frac{N_V\,\alpha_p}{N_C\,\alpha_n}\right)\,. \tag{20.29}$$

Observing that the second term at the right-hand side of (20.29) is small, this result shows that the most efficient trap level is near the gap's midpoint which, in turn, is near the intrinsic Fermi level E_{Fi}. In fact, combining (20.29) with (18.16) yields

$$E_{tM} \simeq E_{Fi} + \frac{k_B\,T}{2}\,\log\left(d_t^2\,\frac{\alpha_p}{\alpha_n}\right) \simeq E_{Fi}\,. \tag{20.30}$$

Defining the *lifetimes*

$$\tau_{p0} = \frac{1}{N_t\,\alpha_p}\,, \qquad \tau_{n0} = \frac{1}{N_t\,\alpha_n}\,, \tag{20.31}$$

gives (20.25) the standard form

$$U_{\rm SRH} = \frac{n\,p - n^{\rm eq}\,p^{\rm eq}}{\tau_{p0}\,(n + n_B) + \tau_{n0}\,(p + p_B)}\,, \tag{20.32}$$

which is also called *Shockley-Read-Hall recombination function*. In equilibrium it is $U_{\rm SRH}^{\rm eq} = 0$; in a nonequilibrium condition, a positive value of $U_{\rm SRH}$, corresponding to an excess of the $n\,p$ product with respect to the equilibrium product $n^{\rm eq}\,p^{\rm eq}$, indicates that recombinations prevail over generations, and vice versa. In a nonequilibrium condition it may happen that $U_{\rm SRH} = 0$; this occurs at the boundary between a region where recombinations prevail and another region where generations prevail.

In a nondegenerate semiconductor (20.22) become, letting $E_t = E_{tM} = E_{Fi}$ and using (18.12),

$$n_B = \frac{n_i}{d_t}\,, \qquad p_B = d_t\,n_i\,, \tag{20.33}$$

whence $n_B\,p_B = n_i^2$. This result is useful also in a degenerate semiconductor for discussing possible simplifications in the form of $U_{\rm SRH}$.

Limiting Cases of the Shockley-Read-Hall Theory

The operating conditions of semiconductor devices are often such that the SRH recombination function (20.32) can be reduced to simpler forms. The first case is the so-called *full-depletion condition*, where both electron and hole concentrations are negligibly small with respect to n_B and p_B. Remembering that $n^{\mathrm{eq}}\, p^{\mathrm{eq}} = n_B\, p_B$ one finds

$$U_{\mathrm{SRH}} \simeq -\frac{n_B\, p_B}{\tau_{p0}\, n_B + \tau_{n0}\, p_B} = -\frac{\sqrt{n_B\, p_B}}{\tau_g}\,, \qquad \tau_g = \sqrt{\frac{n_B}{p_B}}\,\tau_{p0} + \sqrt{\frac{n_B}{p_B}}\,\tau_{n0}\,. \tag{20.34}$$

In a nondegenerate condition n_B, p_B take the simplified form (20.33), whence $\sqrt{n_B/p_B} = n_i$ and $\tau_g = \tau_{p0}/d_t + \tau_{n0}\, d_t$. In a full-depletion condition U_{SRH} is always negative, namely, generations prevail over recombinations; for this reason, τ_g is called *generation lifetime*.

The second limiting case of interest is the so-called *weak-injection condition*. This condition occurs when both inequalities below are fulfilled:

$$|n - n^{\mathrm{eq}}| \ll c^{\mathrm{eq}}\,, \qquad |p - p^{\mathrm{eq}}| \ll c^{\mathrm{eq}}\,, \tag{20.35}$$

where c^{eq} is the equilibrium concentration of the majority carriers in the spatial position under consideration. From the above definition it follows that the concept of weak injection is applicable only after specifying which carriers are the majority ones. Expanding the product $n\,p$ to first order in n and p around the equilibrium value yields $n\,p \simeq n^{\mathrm{eq}}\, p^{\mathrm{eq}} + n^{\mathrm{eq}}\,(p - p^{\mathrm{eq}}) + p^{\mathrm{eq}}\,(n - n^{\mathrm{eq}})$. As a consequence, the numerator of (20.32) becomes

$$n\,p - n^{\mathrm{eq}}\, p^{\mathrm{eq}} \simeq n^{\mathrm{eq}}\,(p - p^{\mathrm{eq}}) + p^{\mathrm{eq}}\,(n - n^{\mathrm{eq}})\,. \tag{20.36}$$

To proceed, it is necessary to distinguish between the n-type and p-type regions.

Weak-Injection Condition, *n*-Type Semiconductor

The weak-injection condition (20.35) reads $|n - n^{\mathrm{eq}}| \ll n^{\mathrm{eq}}$, $|p - p^{\mathrm{eq}}| \ll n^{\mathrm{eq}}$. As a consequence, one lets $n \simeq n^{\mathrm{eq}}$ in the denominator of (20.32) and neglects n_B with respect to n^{eq}; in fact, in a nondegenerate condition it is $n_B \simeq n_i \ll n^{\mathrm{eq}}$, and the same inequality is also applicable in a degenerate condition. As the lifetimes are similar to each other, the term $\tau_{n0}\,(p + p_B)$ in the denominator is negligible with respect to $\tau_{p0}\, n^{\mathrm{eq}}$, because p is a concentration of minority carriers and p_B is similar to n_B. In conclusion, the denominator of (20.32) simplifies to $\tau_{p0}\, n^{\mathrm{eq}}$, whence

$$U_{\mathrm{SRH}} \simeq \frac{p - p^{\mathrm{eq}}}{\tau_{p0}} + \frac{n - n^{\mathrm{eq}}}{(n^{\mathrm{eq}}/p^{\mathrm{eq}})\,\tau_{p0}}\,. \tag{20.37}$$

The second term at the right-hand side of (20.37) is negligible[7] because $n^{\mathrm{eq}}/p^{\mathrm{eq}} \gg 1$; letting $\tau_p = \tau_{p0}$ finally yields

$$U_{\mathrm{SRH}} \simeq \frac{p - p^{\mathrm{eq}}}{\tau_p} , \tag{20.38}$$

with τ_p the *minority-carrier lifetime* in an n-doped region.

Weak-Injection Condition, p-Type Semiconductor

The weak-injection condition (20.35) reads $|n - n^{\mathrm{eq}}| \ll p^{\mathrm{eq}}$, $|p - p^{\mathrm{eq}}| \ll p^{\mathrm{eq}}$. As a consequence, one lets $p \simeq p^{\mathrm{eq}}$ in the denominator of (20.32) and neglects p_B with respect to p^{eq}; the other term in the denominator of (20.35) is neglected as above, thus simplifying the denominator to $\tau_{n0}\, p^{\mathrm{eq}}$. In conclusion,

$$U_{\mathrm{SRH}} \simeq \frac{p - p^{\mathrm{eq}}}{(p^{\mathrm{eq}}/n^{\mathrm{eq}})\, \tau_{n0}} + \frac{n - n^{\mathrm{eq}}}{\tau_{n0}} . \tag{20.39}$$

The first term at the right-hand side of (20.39) is negligible because $p^{\mathrm{eq}}/n^{\mathrm{eq}} \gg 1$; letting $\tau_n = \tau_{n0}$ finally yields

$$U_{\mathrm{SRH}} \simeq \frac{n - n^{\mathrm{eq}}}{\tau_n} , \tag{20.40}$$

with τ_n the *minority-carrier lifetime* in a p-doped region.

The simplified expressions of U_{SRH} found here are particularly useful; in fact, in contrast to (20.32), the weak-injection limits (20.38) and (20.40) are linear with respect to p or n. Moreover, as (20.38) and (20.40) depend on one unknown only, they decouple the continuity equation of the minority carriers (the first one in (19.129) or in (19.130)) from the other equations of the semiconductor's model; thanks to this it is possible to separate the system of equations. The simplification introduced by the full-depletion condition is even stronger, because (20.34) is independent of the model's unknowns. On the other hand, all simplifications illustrated here are applicable only in the regions where the approximations hold; once the simplified model's equations have been solved locally, it is necessary to match the solutions at the boundaries between adjacent regions.

[7] Considering for instance the example in Sect. 18.4.1, one has $n^{\mathrm{eq}} \simeq 10^{15}$ cm^{-3}, $p^{\mathrm{eq}} \simeq 10^{5}$ cm^{-3}, whence $n^{\mathrm{eq}}/p^{\mathrm{eq}} \simeq 10^{10}$.

20.2.4 Thermal Recombination with Tail and Deep States

In recent years, *thin-film transistors* made of amorphous silicon (α-Si:H TFTs) or polycrystalline silicon (poly-TFTs) have acquired great importance in microelectronics; in fact, they can be fabricated with a low thermal budget and on large-area substrates and, at the same time, they are able to achieve performances adequate for the realization of complex circuits. Typical applications are in the area of solid-state image sensors, active-matrix liquid-crystal displays, charge-coupled devices, and static random-access memories. These materials are characterized by a large amount of defects, giving rise to localized states with a complex energy distribution within the gap; typically, the concentration of the acceptor-like states is larger in the upper half of the gap, while that of the donor-like states is larger in the lower half of the gap.

Traps in amorphous silicon are due to the irregular distribution of atoms and to defects in the material; the lack of long-range order in the atomic structure produces a distribution of localized states with energies near the conduction- and valence-band edges (*tail states*); in turn, the defects give rise to a distribution of states localized near midgap (*deep states*). As for the spatial localization, traps in amorphous silicon are uniformly distributed in the semiconductor's volume, while defects in polycrystalline silicon are located at the grain boundaries; for the latter material a simplifying hypothesis is used, that consists in describing the traps as uniformly distributed over the volume [51]. Given these premises, for both materials the energy distribution of traps in the gap can be modeled as the superposition of two distributions of acceptor and donor states; in turn, the densities of states per unit volume, γ_D and γ_A, for each group of donor- and acceptor-like states is approximated as the sum of two exponential functions, describing the deep and tail states, respectively [25, 103, 137]:

$$\gamma_D(E) = \Gamma_{TD} \exp\left(\frac{E_V - E}{\eta_{TD}}\right) + \Gamma_{DD} \exp\left(\frac{E_V - E}{\eta_{DD}}\right), \tag{20.41}$$

$$\gamma_A(E) = \Gamma_{TA} \exp\left(\frac{E - E_C}{\eta_{TA}}\right) + \Gamma_{DA} \exp\left(\frac{E - E_C}{\eta_{DA}}\right) \tag{20.42}$$

(compare with 18.20, 18.35). In (20.41), (20.42), $\Gamma_{TD}, \ldots$ and $\eta_{TD}, \ldots$ are constants, with suffixes *TD*, *DD* standing for "Tail-Donor," "Deep-Donor," respectively, and the like for *TA*, *DA*. When the number of energy states in the gap is large, distinguishing between bands and gap seems meaningless; the distinction, however, is kept as long as the gap states, although dense in energy, are still much less dense in space than those of the bands, so that the contribution of the gap states to current transport is negligible (in other terms, the mobility of the carriers belonging to the gap states is much smaller than that of the band carriers). This condition is assumed here; thus, proceeding in the same manner as in Sect. 20.2.3 yields (compare with (20.18) and (20.19))

$$U_n = \int_{E_V}^{E_C} \{\gamma_A \, [n\,\alpha_{nA}\,(1-P_A) - e_{nA}\,P_A] + \gamma_D \, [n\,\alpha_{nD}\,(1-P_D) - e_{nD}\,P_D]\} \, \mathrm{d}E \,, \tag{20.43}$$

$$U_p = \int_{E_V}^{E_C} \left\{\gamma_A \left[p\,\alpha_{pA}\,P_A - e_{pA}\,(1-P_A)\right] + \gamma_D \left[p\,\alpha_{pD}\,P_D - e_{pD}\,(1-P_D)\right]\right\} \mathrm{d}E \,, \tag{20.44}$$

where n, p are the concentration of electrons in the conduction band and that of holes in the valence band, (20.3) and (20.4) respectively. The meaning of the other symbols in (20.43), (20.44) is as follows:

$$n\,\alpha_{nA} = \int_{E_C}^{E_{CU}} S_{CA}^{(0)}(E', E, \mathbf{r})\,\gamma_C(E')\,P_C(E', \mathbf{r}, t)\,\mathrm{d}E' \,, \tag{20.45}$$

with γ_C, P_C the density of states per unit volume and the nonequilibrium occupation probability of the conduction band; in turn, letting $S_{CA}(E', \mathbf{r}', E, \mathbf{r})$ be the unconditional probability per unit time of a transition from a conduction-band state of energy E' and position $\mathbf{r}'$ to an acceptor trap of energy E and position $\mathbf{r}$, it is $S_{CA} = S_{CA}^{(0)}\,\delta(\mathbf{r}' - \mathbf{r})$. The Dirac symbol in the definition indicates that the change in position during the transition is negligible (compare with 19.40). Symbol $n\,\alpha_{nD}$ has a similar meaning. Similarly,

$$p\,\alpha_{pA} = \int_{E_{VL}}^{E_V} S_{AV}^{(0)}(E, E', \mathbf{r})\,\gamma_V(E')\,\left[1 - P_V(E', \mathbf{r}, t)\right] \mathrm{d}E' \,, \tag{20.46}$$

with γ_V, P_V the density of states per unit volume and the nonequilibrium occupation probability of the valence band; in turn it is $S_{AV} = S_{AV}^{(0)}\,\delta(\mathbf{r}' - \mathbf{r})$, where S_{AV} is the unconditional probability per unit time of a transition from an acceptor trap of energy E and position $\mathbf{r}$ to a valence-band state of energy E' and position $\mathbf{r}'$. Finally, the emission probability per unit time of an electron, from an acceptor state of energy E to the conduction band, is given by

$$e_{nA} = \int_{E_C}^{E_{CU}} S_{AC}^{(0)}(E, E', \mathbf{r}')\,\gamma_C(E')\,\mathrm{d}E' \,, \tag{20.47}$$

and the like for e_{nD}; in turn, the emission probability per unit time of a hole, from an acceptor state of energy E to the valence band, is given by

$$e_{pA} = \int_{E_{VL}}^{E_V} S_{VA}^{(0)}(E', E, \mathbf{r})\,\gamma_V(E')\,\mathrm{d}E' \,, \tag{20.48}$$

and the like for e_{pD}. The concentrations of electrons belonging to the acceptor traps and of holes belonging to the donor traps are, respectively (compare with (20.3) and (20.4)),

$$n_A(\mathbf{r},t) = \int_{E_V}^{E_C} \gamma_A(E',\mathbf{r})\, P_A(E',\mathbf{r},t)\, \mathrm{d}E' \,, \tag{20.49}$$

$$p_D(\mathbf{r},t) = \int_{E_V}^{E_C} \gamma_D(E',\mathbf{r})\, \left[1 - P_D(E',\mathbf{r},t)\right]\, \mathrm{d}E' \,. \tag{20.50}$$

Observing that (20.14), (20.15), (20.16), and (20.17) hold also for the distribution of traps considered here, and replacing (20.17) into (20.16), one finds

$$U_p - \frac{\partial p_D}{\partial t} = U_n - \frac{\partial n_A}{\partial t} \,. \tag{20.51}$$

The expressions of U_p and U_n are given by (20.43), (20.44), while those of $\partial p_D/\partial t$ and $\partial n_A/\partial t$ are obtained from (20.49), (20.50); introducing such expressions into (20.51) and letting

$$D_A = \alpha_{nA}\, n + e_{nA} + \alpha_{pA}\, p + e_{pA} \,, \qquad D_D = \alpha_{nD}\, n + e_{nD} + \alpha_{pD}\, p + e_{pD} \,, \tag{20.52}$$

yield

$$\int_{E_V}^{E_C} \{\gamma_A\, [(\alpha_{nA}\, n + e_{pA}) - D_A\, P_A - \dot{P}_A] + \gamma_D\, [(\alpha_{nD}\, n + e_{pD}) - D_D\, P_D - \dot{P}_D]\}\, \mathrm{d}E = 0 \,, \tag{20.53}$$

with $\dot{P}_A = \partial P_A/\partial t$, $\dot{P}_D = \partial P_D/\partial t$. As equality (20.53) holds for any distributions of states, it follows

$$\frac{\partial P_A}{\partial t} + D_A\, P_A = \alpha_{nA}\, n + e_{pA} \,, \qquad \frac{\partial P_D}{\partial t} + D_D\, P_D = \alpha_{nD}\, n + e_{pD} \,. \tag{20.54}$$

The expressions (20.52) of the denominators D_A, D_D generalize that of the denominator in (20.24), which refers to the case of a single trap level; in turn, the steady-state form of (20.54) generalizes the first expression in (20.24). In conclusion, the continuity equations for a semiconductor having a distribution of traps within the gap, neglecting gap conduction, are (20.13) and (20.54), supplemented with the definitions (20.43), (20.44) of the net recombination rates.

20.3 Auger Recombination and Impact Ionization

An important, nonthermal recombination mechanism is *Auger recombination*. The phenomenon is due to the electron-electron or hole-hole collision and is illustrated in Fig. 20.3. With reference to case *a*, two electrons whose initial state is the conduction band collide and exchange energy. The outcome of the collision is that one of the electrons suffers an energy loss equal or larger than the energy gap and

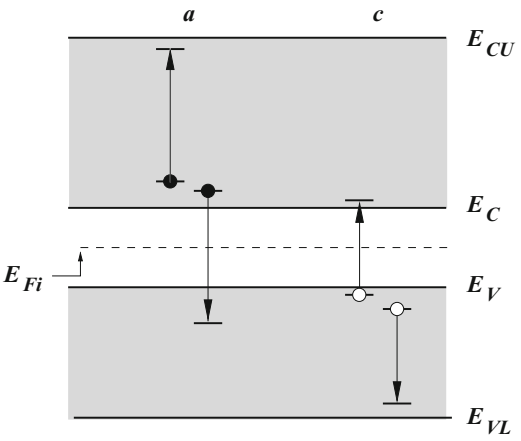


Fig. 20.3 Auger recombinations initiated by electrons (*a*) and holes (*c*)

makes a transition to an empty state of the valence band; the other electron absorbs the same amount of energy and makes a transition to a higher-energy state of the conduction band. The phenomenon is also indicated as an Auger recombination *initiated by electrons*. The analogue for holes is shown in case *c* of Fig. 20.3: two holes whose initial state is in the valence band collide and exchange energy. Remembering that hole energy increases in the opposite direction with respect to that of electrons (Sect. 19.2.3), the hole that suffers an energy loss equal or larger than the energy gap makes a transition to a filled state of the conduction band; the other hole absorbs the same amount of energy and makes a transition to a higher-energy state of the valence band. The phenomenon is indicated as an Auger recombination *initiated by holes*.

The phenomenon dual to Auger recombination is illustrated in Fig. 20.4 and is called *impact ionization*. With reference to case *b*, an electron whose initial state is in the conduction band at high energy collides and exchanges energy with an electron whose initial state is in the valence band. The initial energy E of the electron in the conduction band is such that $E - E_C$ is equal or larger than the energy gap, whereas the initial energy of the electron in the valence band is near E_V. The outcome of the collision is that although the high-energy electron suffers an energy loss equal or larger than the energy gap, its final state is still in the conduction band; the other electron absorbs the same amount of energy and makes a transition to the conduction band. The phenomenon is in fact an electron-hole pair generation and is also indicated as an impact-ionization event *initiated by electrons*. The analogue for holes is shown in case *d* of Fig. 20.4: a hole whose initial state is in the valence band at high energy collides and exchanges energy with a hole whose initial state is in the conduction band. The initial energy E of the hole in the valence band is such that $|E - E_V|$ is equal or larger than the energy gap, whereas the initial energy of the hole in the conduction band is near E_C. The outcome of the collision is that although the high-energy hole suffers an energy loss equal or larger than the energy gap, its final state is still in the valence band; the other hole absorbs the same amount of energy

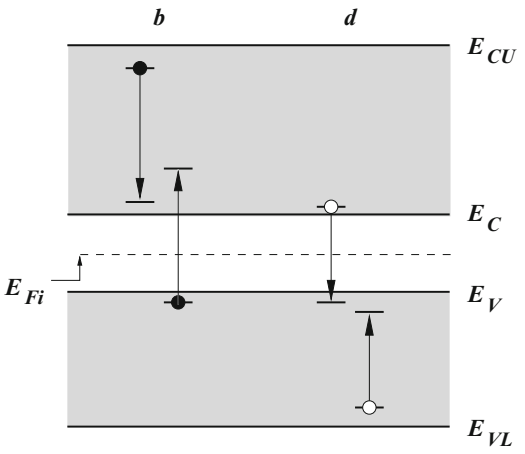


Fig. 20.4 Impact-ionization transitions initiated by electrons (*b*) and holes (*d*)

and makes a transition to the valence band. The phenomenon is in fact an electron-hole pair generation and is also indicated as an impact-ionization event *initiated by holes*.

The derivation of the Auger and impact-ionization rates is shown in the complements; here the expressions of the net recombinations due to the Auger and impact-ionization events are given, that read

$$U_n^{AI} = r_a - r_b = c_n\, n^2\, p - I_n\, n\,, \qquad U_p^{AI} = r_c - r_d = c_p\, p^2\, n - I_p\, p\,, \tag{20.55}$$

where U_n^{AI} refers to the electron-initiated transitions and U_p^{AI} to the hole-initiated ones. As usual, r_a indicates the number of transitions of type a per unit time and volume; the same holds for r_b, r_c, and r_d. In (20.55), c_n, I_n are the transition coefficients for the Auger recombination and impact ionization initiated by electrons, and c_p, I_p the analogue for holes; c_n, c_p are also called *Auger coefficients*.[8] In equilibrium it is $U_n^{AI} = U_p^{AI} = 0$, whence $I_n = c_n\, n^{\mathrm{eq}}\, p^{\mathrm{eq}}$, $I_p = c_p\, n^{\mathrm{eq}}\, p^{\mathrm{eq}}$. The above holds also in a nonequilibrium case as long as the operating conditions are not too far from equilibrium; with these premises it follows

$$U_n^{AI} = c_n\, n\ (n\, p - n^{\mathrm{eq}}\, p^{\mathrm{eq}})\ , \qquad U_p^{AI} = c_p\, p\ (n p - n^{\mathrm{eq}}\, p^{\mathrm{eq}})\ , \tag{20.56}$$

When the operating condition departs strongly from equilibrium, the simplification leading to (20.56) is no longer applicable and the general expressions (20.55) must be used. Referring to all recombinations as due to transitions of electrons, their rate is easily found to be $r_a + r_c$; similarly, the total generation rate is $r_b + r_d$. In conclusion, the net recombination rate due to the Auger and impact-ionization phenomena is given by

[8] The units are $[c_{n,p}] = \mathrm{cm}^6\ \mathrm{s}^{-1}$ and $[I_{n,p}] = \mathrm{s}^{-1}$.

$$U_{AI} = U_n^{AI} + U_p^{AI} . \tag{20.57}$$

For Auger recombination to occur it is necessary that an electron collides with another electron, or a hole collides with another hole. The probability of such an event is relatively small because in normal operating conditions and at room temperature there is a high probability that a carrier collides with a phonon; as a consequence, for the collisionless motion of an electron to be interrupted by a collision with another electron it is necessary that the electron concentration be very high. This situation occurs only in a heavily doped, *n*-type region; similarly, an Auger recombination initiated by holes can be significant only in a heavily doped, *p*-type region.[9]

Considering now the case of impact-ionization, for this phenomenon to occur it is necessary that an electron, or a hole, acquires a kinetic energy larger than the energy gap. This is a rare event as well,[10] because in general the carrier undergoes a phonon collision when its kinetic energy is still significantly lower than the energy gap. The impact-ionization event occurs only if the carrier acquires a substantial energy over a distance much shorter than the average collisionless path, which happens only in presence of a strong electric field.[11]

The qualitative reasoning outlined above explains why the conditions for a strong Auger recombination are incompatible with those that make impact-ionization dominant; in fact, a large charge density, like that imposed by a heavy dopant concentration, prevents the electric field from becoming strong. Vice versa, a strong electric field prevents a large carrier concentration from building up. It is therefore sensible to investigate situations where only one term dominates within U_{AI}.

20.3.1 Strong Impact Ionization

As indicated in Sect. 20.3, far from equilibrium the approximations $I_n = c_n n^{\text{eq}} p^{\text{eq}}$, $I_p = c_p n^{\text{eq}} p^{\text{eq}}$ are not valid, and the general expressions (20.55) must be used. Here the situation where impact ionization dominates over the other generation-recombination mechanisms is considered, using the steady-state case. If impact ionization is dominant, it is $U_n - G_n = U_p - G_p \simeq U_{AI} \simeq -I_n n - I_p p$. The continuity equations (the first ones in (19.129) and (19.130)) then become

[9] In fact, Auger recombination becomes significant in the source and drain regions of MOSFETs and in the emitter regions of BJTs, where the dopant concentration is the highest.

[10] In principle, high-energy electrons or hole exists also in the equilibrium condition; however, their number is negligible because of the exponentially vanishing tail of the Fermi-Dirac statistics.

[11] The high-field conditions able to produce a significant impact ionization typically occur in the reverse-biased *p*-*n* junctions like, e.g., the drain junction in MOSFETs and the collector junction in BJTs.

$$\text{div}\mathbf{J}_n = -q\,I_n\,n - q\,I_p\,p\,, \qquad \text{div}\mathbf{J}_p = q\,I_n\,n + q\,I_p\,p\,. \tag{20.58}$$

As outlined in Sect. 20.3, impact-ionization dominates if the electric field is high. For this reason, the transport equations in (19.129) and (19.130) are simplified by keeping the ohmic term only, to yield $\mathbf{J}_n \simeq q\,\mu_n\,n\,\mathbf{E}$ and $\mathbf{J}_p \simeq q\,\mu_p\,p\,\mathbf{E}$. As a consequence, the electron and hole current densities are parallel to the electric field. Let $\mathbf{e}(\mathbf{r})$ be the unit vector of the electric field, oriented in the direction of increasing field, $\mathbf{E} = |\mathbf{E}|\,\mathbf{e}$; it follows $\mathbf{J}_n = J_n\,\mathbf{e}$ and $\mathbf{J}_p = J_p\,\mathbf{e}$, with J_n and J_p strictly positive. Extracting n, p from the above and replacing them into (20.58) yield

$$-\,\text{div}\mathbf{J}_n = k_n\,J_n + k_p\,J_p\,, \qquad \text{div}\mathbf{J}_p = k_n\,J_n + k_p\,J_p\,, \tag{20.59}$$

where the ratios

$$k_n = \frac{I_n}{\mu_n\,|\mathbf{E}|}\,, \qquad k_p = \frac{I_p}{\mu_p\,|\mathbf{E}|}\,, \tag{20.60}$$

whose units are $[k_{n,p}] = \text{m}^{-1}$, are the *impact-ionization coefficients* for electrons and holes, respectively. Equations (20.59) form a system of differential equations of the first order, whose solution in the one-dimensional case is relatively simple if the dependence of the coefficients on position is given (Sect. 21.5).

20.4 Optical Transitions

The description of the optical transitions is similar to that of the direct thermal transitions given in Sect. 20.2.1; still with reference to Fig. 20.1, the transition marked with a can be thought of as an optical-recombination event if the energy difference between the initial and final state is released to the environment in the form of a photon. The opposite transition (b), where the electron's energy increases due to photon absorption from the environment, is an optical electron-hole generation. The expression of the net optical-recombination rate is similar to (20.11) and reads

$$U_O = \alpha_O\,n\,p - G_O\,, \tag{20.61}$$

whose coefficients are derived in the same manner as those of U_{DT} (Sect. 20.2.1).

In normal operating conditions the similarity between the direct-thermal and optical generation-recombination events extends also to the external agent that induces the transitions. In fact, the distribution of the phonon energies is typically the equilibrium one, given by the Bose-Einstein statistics (15.55) at the lattice temperature; as for the photons, the environment radiation in which the semiconductor is immersed can also be assimilated to the equilibrium one, again given by the Bose-Einstein statistics at the same temperature.

The conditions of the optical generation-recombination events drastically change if the device is kept far from equilibrium. Consider for instance the case where the electron concentration of the conduction band is artificially increased with respect to the equilibrium value at the expense of the electron population of the valence band, so that both n and p in (20.61) increase. This brings about an excess of recombinations; if the probability of radiative-type generation-recombination events is high,[12] the emission of a large number of photons follows. The angular frequencies of the emitted photons is close to $(E_C - E_V)/\hbar$, because the majority of the electrons in the conduction band concentrate near E_C, and the final states of the radiative transitions concentrate near E_V. In this way, the energy spent to keep the artificially high concentration of electron-hole pairs is transformed into that of a nearly monochromatic optical emission. In essence, this is the description of the operating principle of a *laser*.[13] Another method for keeping the device far from equilibrium is that of artificially decreasing both the concentration of electrons of the conduction band and the concentration of holes of the valence band. The outcome is opposite with respect to that described earlier: the decrease of both n and p in (20.61) brings about an excess of generations, which in turn corresponds to the absorption of photons from the environment. The absorption may be exploited to accumulate energy (thus leading to the concept of *solar cell*), or to provide an electrical signal whose amplitude depends on the number of absorbed photons (thus leading to the concept of *optical sensor*).

In a nonequilibrium condition the amount of energy exchanged between the semiconductor and the electromagnetic field is not necessarily uniform in space. Consider, by way of example, the case of an optical sensor on which an external radiation impinges; as the nonequilibrium conditions are such that the absorption events prevail, the radiation intensity within the material progressively decreases at increasing distances from the sensor's surface. Therefore, it is important to determine the radiation intensity as a function of position.

It is acceptable to assume that the absorption events are uncorrelated from each other. Thus, one can limit the analysis to a monochromatic radiation; the effect of the whole spectrum is recovered at a later stage by adding up over the frequencies. When absorption prevails, (20.61) simplifies to $U_O \simeq -G_O$, where G_O is a function of the radiation's frequency ν and possibly of position. If the radiation's intensity varies with time, G_O depends on time as well.[14] When the radiation interacts with the external surface of the material, part of the energy is reflected; moreover, the radiation is refracted at the boundary, so that the propagation direction outside the material differs in general from that inside. Letting ξ be the propagation direction inside the material, consider an elementary volume with a side $\mathrm{d}\xi$ aligned with ξ

[12] As indicated in Sect. 17.6.6, among semiconductors this is typical of the direct-gap ones.

[13] In fact, LASER is the acronym of Light Amplification by Stimulated Emission of Radiation.

[14] In principle, a time dependence of the intensity is incompatible with the hypothesis that the radiation is monochromatic. However, the frequency with which the intensity may vary is extremely small with respect to the optical frequencies.

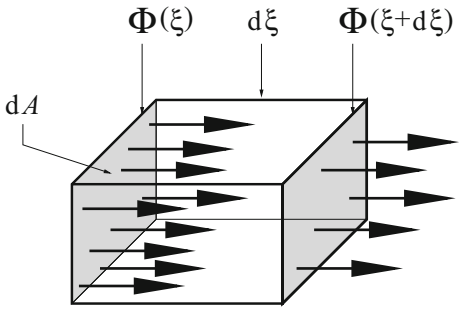


Fig. 20.5 Sketch of photon absorption in a material layer

and a cross-section dA normal to it (Fig. 20.5). The monochromatic radiation can be described as a flux of photons of equal energy $h\nu$, with h the Planck constant, and a momentum's direction parallel to ξ. Let $\Phi(\xi)$ be the flux density of photons entering the volume from the face corresponding to ξ, and $\Phi(\xi + \mathrm{d}\xi)$ the flux density leaving it at $\xi + \mathrm{d}\xi$; the following holds, $\Phi = K u_f$, where $K(\xi)$ is the concentration of the photons and u_f their constant phase velocity. Then,

$$\frac{\partial \Phi}{\partial \xi} = \frac{\partial K}{\partial (\xi / u_f)} = \frac{\partial K}{\partial t} \,. \tag{20.62}$$

The derivatives in (20.62) are negative because the photon concentration decreases in time due to absorption; as the loss of each photon corresponds to the loss of an energy quantum $h\nu$, the loss of electromagnetic energy per unit volume and time is $-h\nu\,(\partial\Phi/\partial\xi)$. By a similar token one finds[15] that the energy absorbed by the optical-generation events per unit time and volume is $h\nu\, G_O$. The latter is not necessarily equal to $-h\nu\,(\partial\Phi/\partial\xi)$; in fact, some photons crossing the elementary volume may be lost due to collisions with nuclei (this, however, is a rare event), or with electrons that are already in the conduction band, so that no electron-hole pair generation occurs. To account for these events one lets

$$G_O = -\eta \frac{\partial \Phi}{\partial \xi} > 0 \,, \tag{20.63}$$

with $0 < \eta < 1$ the *quantum efficiency*. In moderately doped semiconductors η is close to unity because the concentration of the conduction-band electrons is small; instead, the efficiency degrades in degenerate semiconductors. The spatial dependence of the generation term can be derived from (20.63) if that of the photon flux is known. To proceed, one defines the *absorption coefficient* as

[15] It is implied that $h\nu \geq E_C - E_V$, and that two-particle collisions only are to be considered.

$$k = -\frac{1}{\Phi}\,\frac{\partial \Phi}{\partial \xi} > 0\,, \tag{20.64}$$

with $[k] = \mathrm{m}^{-1}$. In general it is $k = k(\Phi, \xi, \nu)$; however, as the absorption effects are uncorrelated, the flux density lost per unit path $\mathrm{d}\xi$ is proportional to the flux density available at ξ. Then, k is independent of Φ; neglecting momentarily the dependence on ξ as well, one finds

$$\Phi(\xi) = \Phi_B \, \exp[-k(\nu)\,\xi]\,, \tag{20.65}$$

with $\Phi_B = \Phi(\xi = 0^+)$ on account of the fact that due to the reflection at the interface, the flux density on the inside edge of the boundary is different from that on the outside edge. When k is independent of position, its inverse $1/k$ is called *average penetration length of the radiation*. When k depends on position, (20.64) is still separable and yields

$$\Phi(\xi) = \Phi_B \, \exp(-k_m\,\xi)\,, \qquad k_m = \frac{1}{\xi}\int_0^{\xi} k(\xi';\nu)\,\mathrm{d}\xi'\,. \tag{20.66}$$

Combining (20.66) with (20.63), the optical-generation term is found to be

$$G_O = \eta\,\Phi_B\,k(\xi,\nu)\,\exp\left[-\int_0^{\xi} k(\xi',\nu)\,\mathrm{d}\xi'\right]\,. \tag{20.67}$$

20.5 Macroscopic Mobility Models

It has been shown in Sect. 19.5.2 that the carrier mobilities are defined in terms of the momentum-relaxation times. Specifically, in the parabolic-band approximation it is, for the electrons of the conduction band, $\mu_n = (\mu_l + 2\,\mu_t)/3$, with $\mu_l = q\,\tau_p/m_l$, $\mu_t = q\,\tau_p/m_t$, where τ_p is the electron momentum-relaxation time (19.87); similarly, for the holes of the valence band the carrier mobility is given by inserting (19.118) into the second relation of (19.121), namely, a linear combination of the heavy-hole and light-hole momentum-relaxation times. As, in turn, the inverse momentum-relaxation time is a suitable average of the inverse intra-band relaxation time, the Matthiessen rule follows (Sect. 19.6.5); in conclusion, the electron and hole mobilities are calculated by combining the effects of the different types of collisions (e.g., phonons, impurities, and so on) suffered by the carrier.[16] In the case of electrons, the application of the Matthiessen rule is straightforward, leading to

[16] As mentioned in Sect. 19.6.5, it is assumed that the different types of collisions are uncorrelated.

$$\frac{1}{\mu_n} = \frac{m_n}{q}\left(\frac{1}{\tau_p^{\text{ph}}} + \frac{1}{\tau_p^{\text{imp}}} + \ldots\right), \tag{20.68}$$

where the index refers to the type of collision, and $1/m_n = (1/m_l + 2/m_t)/3$. For holes a little more algebra is necessary, which can be avoided if the approximation $\tau_{ph} \simeq \tau_{pl}$ is applicable.

In the typical operating conditions of semiconductor devices the most important types of collisions are those with phonons and ionized impurities. For devices like surface-channel MOSFETs, where the flow lines of the current density are near the interface between semiconductor and gate insulator, a third type is also very important, namely, the collisions with the interface. The *macroscopic mobility models* are closed-form expressions in which mobility is related to a set of macroscopic parameters (e.g., temperature) and to some of the unknowns of the semiconductor-device model; the concept is similar to that leading to the expressions of the generation-recombination terms shown in earlier sections.

20.5.1 Example of Phonon Collision

By way of example, a simplified analysis of the contribution to mobility of the electron-phonon collision is outlined below, starting from the definition of the ith component of the momentum-relaxation tensor τ_{pi} given by (19.87); the simplifications are such that the first-order expansion $f - f^{\text{eq}} \simeq (\mathrm{d}f/\mathrm{d}\lambda)^{\text{eq}}\,\lambda$ is not used here. Starting from the perturbative form (19.47) one considers the steady-state, uniform case and lets $\mathbf{B} = 0$, $\tau = \tau_v$, to find

$$\frac{q}{\hbar}\,\mathbf{E}\cdot\mathrm{grad}_{\mathbf{k}} f = \frac{f - f^{\text{eq}}}{\tau_v}\,. \tag{20.69}$$

Replacing f with f^{eq} at the left-hand side of (20.69) and using the definition (17.52) of the group velocity yield $\mathrm{grad}_{\mathbf{k}} f^{\text{eq}} = (\mathrm{d}f^{\text{eq}}/\mathrm{d}H)\,\hbar\,\mathbf{u}$, with H the Hamiltonian function defined in Sect. 19.2.2. Inserting into (19.87) yields

$$\tau_{pi} \iiint_{-\infty}^{+\infty} u_i\,\mathbf{E}\cdot\mathbf{u}\,(\mathrm{d}f^{\text{eq}}/\mathrm{d}H)\,\mathrm{d}^3k = \iiint_{-\infty}^{+\infty} u_i\,\mathbf{E}\cdot\mathbf{u}\,(\mathrm{d}f^{\text{eq}}/\mathrm{d}H)\,\tau_v\,\mathrm{d}^3k\,. \tag{20.70}$$

As the derivative $\mathrm{d}f^{\text{eq}}/\mathrm{d}H$ is even with respect to $\mathbf{k}$, the integrals involving velocity components different from u_i vanish because the corresponding integrand is odd; as a consequence, only the ith component of the electric field remains, and cancels out. A further simplification is obtained by replacing the Fermi-Dirac statistics, with the Maxwell-Boltzmann distribution law, $f^{\text{eq}} \simeq Q\,\exp[(-E_e + q\,\varphi - E_C + E_F)/(k_B\,T)]$, to find

$$\tau_{pi} \iiint_{-\infty}^{+\infty} u_i^2 \, \exp[-E_e/(k_B\,T)] \, \mathrm{d}^3k = \iiint_{-\infty}^{+\infty} u_i^2 \, \exp[-E_e/(k_B\,T)] \, \tau_v \, \mathrm{d}^3k \,. \tag{20.71}$$

To proceed it is necessary to make an assumption about τ_v. Remembering the definition of the relaxation time given by the first relation in (19.43), it is reasonable to assume that the scattering probability S_0 increases with the kinetic energy E_e of the electron, so that the relaxation time decreases; a somewhat stronger hypothesis is that the relaxation time depends on E_e only, namely, the collision is isotropic.[17] In this case, (20.71) is readily manipulated by a Herring-Vogt transformation. Following the same procedure as in Sect. 19.6.4, one finds that all numerical factors cancel out; as a consequence, one may replace the auxiliary coordinate η_i^2 with $\eta^2/3 = E_e/3$, thus showing that $\tau_{pi} = \tau_p$ is isotropic as well. One eventually finds

$$\tau_p = \frac{\int_0^{+\infty} \tau_v(E_e) \; E_e^{3/2} \, \exp[-E_e/(k_B\,T)] \, \mathrm{d}E_e}{\int_0^{+\infty} E_e^{3/2} \, \exp[-E_e/(k_B\,T)] \, \mathrm{d}E_e} \,. \tag{20.72}$$

A simple approximation for the relaxation time is $\tau_v = \tau_{v0} \, (E_e/E_0)^{-\alpha}$, where τ_{v0}, E_0, and α are positive parameters independent of E_e (compare with Sect. 19.6.7). From (C.95) it follows

$$\tau_p = \tau_{v0} \, \frac{\Gamma(5/2-\alpha)}{\Gamma(5/2)} \left(\frac{E_0}{k_B\,T} \right)^{\alpha} \,. \tag{20.73}$$

When the electron-phonon interaction is considered, $\tau_{v0} = \tau_{v0}^{\mathrm{ph}}$ is found to be inversely proportional to $k_B\,T$ and to the concentration N_{sc} of semiconductor's atoms; moreover, for acoustic phonons[18] it is $\alpha = 1/2$ [78, Sects. 61,62], whence

$$\tau_p^{\mathrm{ap}} = \tau_{v0}(N_{\mathrm{sc}}, T) \, \frac{4}{3\,\sqrt{\pi}} \left(\frac{E_0}{k_B\,T} \right)^{1/2} \,, \qquad \mu_n^{\mathrm{ap}} \propto N_{\mathrm{sc}}^{-1} \, (k_B\,T)^{-3/2} \,, \tag{20.74}$$

where "ap" stands for "acoustic phonon." More elaborate derivations, including also the contribution of optical phonons, still show that carrier-phonon collisions make mobility to decrease when temperature increases.

[17] The first-principle derivation of the scattering probabilities is carried out by applying Fermi's Golden Rule (Sect. 14.8.3) to each type of perturbation, using the Bloch functions for the unperturbed states [73]. Examples are given in Sect. 14.8.6 for the case of the harmonic perturbation in a periodic structure, and in this chapter (Sect. 20.5.2) for the case of ionized-impurity scattering.

[18] Acoustic phonons are those whose momentum and energy belong to the acoustic branch of the lattice-dispersion relation (Sect. 17.9.5); a similar definition applies to optical phonons (Sect. 17.9.6).

20.5.2 Example of Ionized-Impurity Collision

As a second example one considers the collisions with ionized impurities. The interaction with a single ionized impurity is a perturbation of the Coulomb type; due to the presence of the crystal, the more suitable approach is the screened Coulomb perturbation, an example of which is shown in Sect. 14.7, leading to the perturbation-matrix element (14.34):

$$h^{(0)}_{\mathbf{kg}} = \frac{A/(2\,\pi)^3}{q_c^2 + q^2}\,, \qquad A = \frac{\kappa\, Z\, e^2}{\varepsilon_0}\,. \tag{20.75}$$

In (20.75), $e > 0$ is the elementary electric charge, Z a positive integer, ε_0 the vacuum permittivity, $q_c > 0$ the inverse screening length,[19] $q = |\mathbf{q}| = |\mathbf{k} - \mathbf{g}|$ and, finally, $\kappa = 1\,(-1)$ in the repulsive (attractive) case. The wave vectors $\mathbf{k}$ and $\mathbf{g}$ correspond to the initial and final state of the transition, respectively. In principle, (20.75) should not be used as is because it holds *in vacuo*; in fact, the eigenfunctions of the unperturbed Hamiltonian operator used to derive (20.75) are plane waves. Inside a crystal, instead, one should define the perturbation matrix $h_{\mathbf{kg}}(t)$ using the Bloch functions $w_{\mathbf{k}} = u_{\mathbf{k}} \exp(\mathrm{i}\,\mathbf{k}\cdot\mathbf{r})$ in an integral of the form (14.24). However, it can be shown that the contribution of the periodic part $u_{\mathbf{k}}$ can suitably be averaged and extracted from the integral, in the form of a dimensionless coefficient, whose square modulus G is called *overlap factor*.[20] For this reason, the collisions with ionized impurities is treated starting from the definition (20.75) to calculate the perturbation matrix, with the provision that the result is to be multiplied by G and the permittivity $\varepsilon_{\mathrm{sc}}$ of the semiconductor replaces ε_0 in the second relation of (20.75).

Like in Sect. 14.6, a Gaussian wave packet (14.27) centered on some wave vector $\mathbf{b} \neq \mathbf{g}$ is used as initial condition. In this case the perturbation is independent of time, $h_{\mathbf{bg}} = h^{(0)}_{\mathbf{bg}} = \text{const} \neq 0$; as a consequence, the infinitesimal probability $\mathrm{d}P_{\mathbf{b}}$ that such a perturbation induces a transition, from the initial condition (14.27), to a final state whose energy belongs to the range $\mathrm{d}E_{\mathbf{g}}$, is given by (14.32). In turn, the integral (14.31) providing $H^{(0)}_{\mathbf{b}}(E_{\mathbf{g}})$ is calculated in Prob. 14.1. Assuming that the duration t_P of the interaction is large enough to make Fermi's Golden Rule (14.44) applicable, and inserting the overlap factor, one finally obtains

$$\mathrm{d}P_{\mathbf{b}} \approx G \left(\frac{2\,\pi\, m}{\hbar^2}\right)^{3/2} \frac{8\,\pi\, t_P\, \delta(E_{\mathbf{b}} - E_{\mathbf{g}})\, A^2}{\lambda^3\, \hbar\, (2\,\pi)^5\, q_c^2\, (q_c^2 + 8\, m\, E_{\mathbf{g}}/\hbar^2)}\, \sqrt{E_{\mathbf{g}}}\; \mathrm{d}E_{\mathbf{g}}\,. \tag{20.76}$$

[19] An example of derivation of the screening length is given in Sect. 20.6.4.

[20] An example of this procedure is given in Sect. 14.8.6 with reference to the case where the spatial part of the perturbation has the form of a plane wave.

where the relation $E_{\mathbf{g}} = \hbar^2\, g^2/(2\,m)$ has been used. Integrating over $E_{\mathbf{g}}$ and dividing by t_P provides the probability per unit time of a transition from the initial energy $E_{\mathbf{b}}$ to any final energy; letting $E_c = \hbar^2\, q_c^2/(2\,m)$, one finds

$$\dot{P}(E_{\mathbf{b}}) = \frac{1}{\tau_{vc}}\,\frac{\sqrt{4\,E_{\mathbf{b}}/E_c}}{1+4\,E_{\mathbf{b}}/E_c}\,. \qquad \frac{1}{\tau_{vc}} = \frac{G\,A^2/\sqrt{2\,\pi\,m}}{8\,\pi^2\,(\lambda^2\,E_c)^{3/2}}\,. \tag{20.77}$$

The above expression provides the contribution to the intra-band relaxation time of the scattering due to a single impurity. One notes that since A is squared, the effect onto (20.77) of a positive impurity is the same as that of a negative one. If the effect of each impurity is uncorrelated with that of the others,[21] the probabilities add up; letting $N_I = N_D^+ + N_A^-$ be the total concentration of ionized impurities, the product $N_I\,\mathrm{d}^3r$ is the total number of ionized impurities in the elementary volume d^3r; it follows that the probability per unit time and volume is given by $\dot{P}(E_{\mathbf{b}})\,N_I$. Considering that N_I depends on position only, mobility inherits the inverse proportionality with N_I; letting "ii" indicate "ionized impurity," one finds $\mu_n^{\mathrm{ii}} \propto 1/N_I$.

The derivation of the dependence on N_I shown above is in fact oversimplified, and the resulting model does not reproduce the experimental results with sufficient precision. One of the reasons for this discrepancy is that the inverse screening length q_c depends on the dopant concentration as well, as is apparent, for instance, from the second relation of (20.105). In order to improve the model, while still keeping an analytical form, the expression is modified by letting $1/\mu_n^{\mathrm{ii}} \propto N_I^{\alpha}$, with α a dimensionless parameter to be extracted from the comparison with experiments. One then lets

$$\frac{1}{\mu_n^{\mathrm{ii}}(N_I)} = \frac{1}{\mu_n^{\mathrm{ii}}(N_R)}\left(\frac{N_I}{N_R}\right)^{\alpha}\,, \tag{20.78}$$

with N_R a reference concentration.

20.5.3 Bulk and Surface Mobilities

Combining the phonon and ionized-impurity contributions using the Matthiessen rule yields $1/\mu_n^{\mathrm{B}}(T,N_I) = 1/\mu_n^{\mathrm{ph}}(T) + 1/\mu_n^{\mathrm{ii}}(N_I)$, namely,

$$\mu_n^{\mathrm{B}}(T,N_I) = \frac{\mu_n^{\mathrm{ph}}(T)}{1+c(T)\,(N_I/N_R)^{\alpha}}\,, \tag{20.79}$$

[21] In silicon, this assumption is fulfilled for values of the concentration up to about 10^{19} cm^{-3} [80, 106].

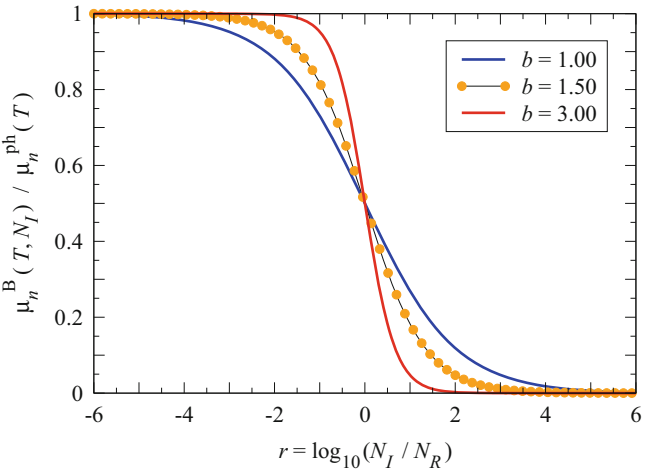


Fig. 20.6 Graph of the theoretical mobility curve (20.80), normalized to its maximum, for different values of b, with $b_0 = 0$. Each curve has a flex at $r = r_{\mathrm{flex}} = -b_0/b$ and takes the value 0.5 there. The slope at the flex is $-b/4$

with $c(T) = \mu_n^{\mathrm{ph}}(T)/\mu_n^{\mathrm{ii}}(N_R)$. In practical cases the doping concentration ranges over many orders of magnitude; for this reason, (20.79) is usually represented in a semilogarithmic scale: letting $r = \log_{10}(N_I/N_R)$, $b = \alpha \log_e 10$, and $b_0 = \log_e c$, (20.79) becomes

$$\mu_n^{\mathrm{B}}(T, N_I) = \frac{\mu_n^{\mathrm{ap}}(T)}{1 + \exp(b\, r + b_0)}\,. \tag{20.80}$$

The curves corresponding to $b = 1,\ 1.5,\ 3$ and $b_0 = 0$ are drawn in Fig. 20.6, using r as independent variable at a fixed T. Index "B" in the mobility defined in (20.79) or (20.80) stands for "bulk." More generally, the term *bulk mobility* is ascribed to the combination of all contributions to mobility different from surface collisions.

As mentioned at the beginning of this section, in surface-channel devices the degradation of mobility produced by the interaction of the carriers with the interface between channel and gate insulator is also very important. The macroscopic models of this effect are built up by considering that the carrier-surface interaction is more likely to occur if the flow lines of the current density are closer to the interface itself; such a closeness is in turn controlled by the intensity of the electric field's component normal to the interface, $E_\perp$. In conclusion, the model describes the contribution to mobility due to surface scattering as a decreasing function of $E_\perp$, e.g., for electrons,

$$\frac{1}{\mu_n^{\mathrm{s}}(E_\perp)} = \frac{1}{\mu_n^{\mathrm{s}}(E_R)} \left(\frac{E_\perp}{E_R}\right)^{\beta}, \tag{20.81}$$

with E_R a reference field and β a dimensionless parameter to be extracted from experiments. Combining the bulk and surface contributions using the Matthiessen rule yields $1/\mu_n(T, N_I, E_\perp) = 1/\mu_n^{\mathrm{B}}(T, N_I) + 1/\mu_n^{\mathrm{s}}(E_\perp)$, namely,

$$\mu_n(T, N_I, E_\perp) = \frac{\mu_n^{\mathrm{B}}(T, N_I)}{1 + d(T, N_I)\,(E_\perp/E_R)^\beta}\,, \tag{20.82}$$

with $d(T, N_I) = \mu_n^{\mathrm{B}}(T, N_I)/\mu_n^{\mathrm{s}}(E_R)$.

20.5.4 Beyond Analytical Modeling of Mobility

In general the analytical approaches outlined above do not attain the precision necessary for applications to realistic devices. For this reason, one must often resort to numerical-simulation methods; in this way, the main scattering mechanisms are incorporated into the analysis (e.g., for silicon: acoustic phonons, optical phonons, ionized impurities, and impact ionization), along with the full-band structure of the semiconductor, which is included in the simulation through the density of states and group velocity defined in the energy space. The latter, in turn, are obtained directly from the corresponding functions in the momentum space by integrating the full-band system over the angles. The energy range considered to date allows for the description of carrier dynamics up to 5 eV.

As mentioned above, the ionized-impurity collisions can be treated as interactions between the carrier and a single impurity as long as the impurity concentration is below some limit. When the limit is exceeded, impurity clustering becomes relevant and must be accounted for [80]. In fact, at high doping densities the carrier scatters with a cluster of K ions, where K is a function of the impurity concentration. Finally, different outcomes are found for majority- or minority-mobility calculations: e.g., minority-hole mobility is found to be about a factor 2 higher than the majority-hole mobility for identical doping levels.

Figures 20.7 and 20.8 show the outcome of electron- and hole-mobility calculations for bulk silicon, obtained from the spherical-harmonics method illustrated in [140]. The method incorporates the models for the scattering mechanisms listed above. The electron and hole mobility have been calculated as a function of the total ionized-dopant concentration N_I, using the lattice temperature T as a parameter; in the figures, they are compared with measurements taken from the literature.

To include the surface effects in the analysis it is necessary to account for the fact that in modern devices the thickness of the charge layer at the interface with the gate insulator is so small that quantum confinement and formation of subbands must be considered. The typical collisions mechanisms to be accounted for at the semiconductor–insulator interface are surface roughness, scattering with ionized impurities trapped at the interface, and surface phonons. Figures 20.9 and 20.10 show the outcome of electron and hole surface-mobility calculations in silicon,

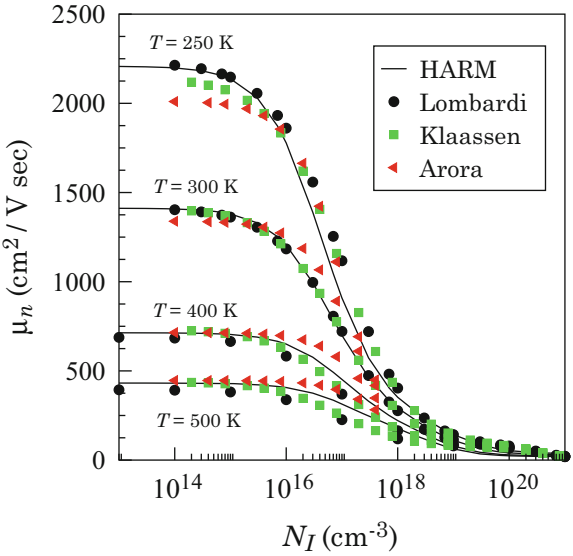


Fig. 20.7 Electron mobility in silicon calculated with the spherical-harmonics expansion method (HARM) as a function of the total ionized-dopant concentration N_I, using the lattice temperature T as parameter. The calculations are compared with measurements by Lombardi [91], Klaassen [80], and Arora [2] (courtesy of S. Reggiani)

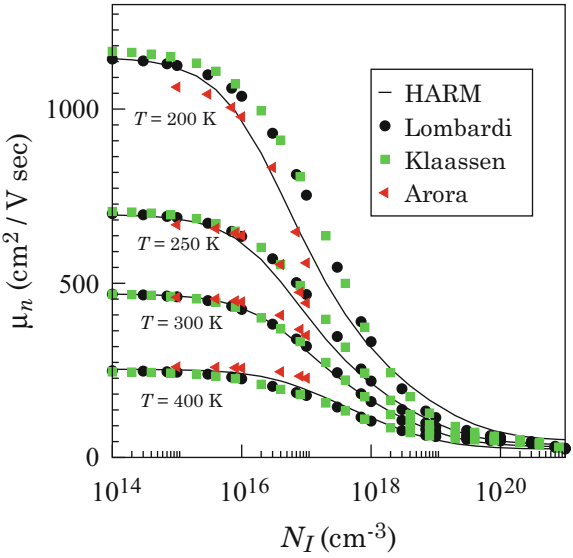


Fig. 20.8 Hole mobility in silicon calculated with the spherical-harmonics expansion method (HARM) as a function of the total ionized-dopant concentration N_I, using the lattice temperature T as parameter. The calculations are compared with measurements by Lombardi [91], Klaassen [80], and Arora [2] (courtesy of S. Reggiani)

also obtained from the spherical-harmonics method [106]. The electron and hole mobility have been calculated as functions of the dopant concentration (N_A and N_D, respectively), at room temperature; in the figures, they are compared with measurements taken from the literature.

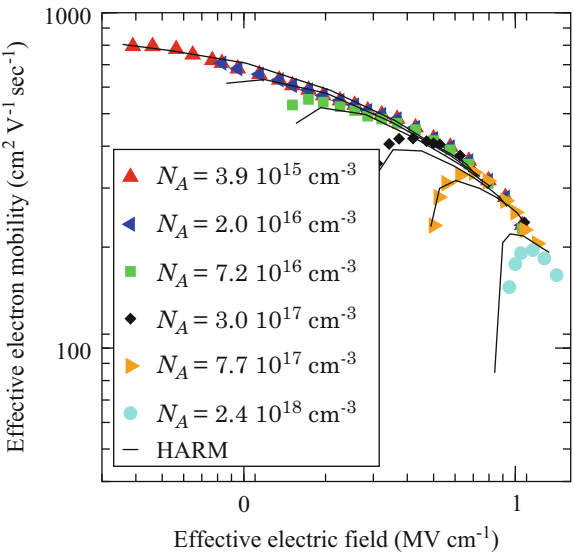


Fig. 20.9 Electron surface mobility in silicon calculated with the spherical-harmonics expansion method (HARM) method at room temperature, using the acceptor concentration N_A as parameter. The calculations are compared with measurements by Takagi [131] (courtesy of S. Reggiani)

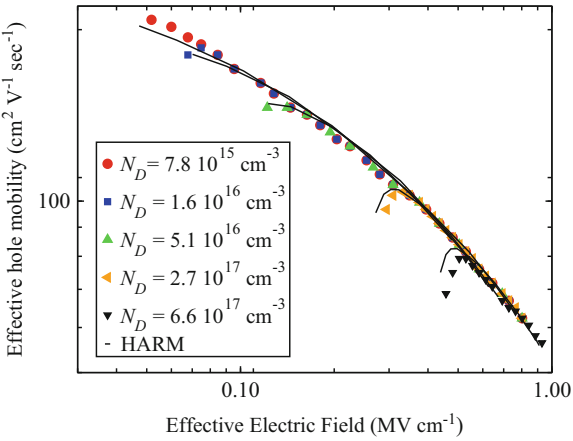


Fig. 20.10 Hole surface mobility in silicon calculated with the spherical-harmonics expansion method (HARM) at room temperature, using the donor concentration N_D as parameter. The calculations are compared with measurements by Takagi [131] (courtesy of S. Reggiani)

20.6 Complements

20.6.1 Transition Rates in the SRH Recombination Function

The expressions of the transition rates r_a, r_b, r_c, r_d to be used in the calculation of the Shockley-Read-Hall recombination function (20.32) are determined by the same reasoning as that used in Sect. 20.2.1 for the direct thermal transitions. Let $P(\mathbf{r}, E, t)$ be the occupation probability of a state at energy E, and $C(E \to E')$ the probability per unit time and volume (in $\mathbf{r}$) of a transition from a filled state of energy E to an empty state of energy E'. Such a probability is independent of time; it depends on

the energy of the phonon involved in the transition, and possibly on position. Then, define $P' = P(\mathbf{r}, E = E', t)$, $P_t = P(\mathbf{r}, E = E_t, t)$, where E_t is the energy of the trap. Finally, let $\gamma(E)$ be the combined density of states in energy and volume of the bands, and $\gamma_t(\mathbf{r}, E)$ the same quantity for the traps (the latter depends on position if the traps' distribution is nonuniform). The number of transitions per unit volume and time, from states in the interval $\mathrm{d}E$ belonging to a band, to states in the interval $\mathrm{d}E'$ belonging to the trap distribution, is obtained as the product of the number $\Omega\,\gamma(E)\,\mathrm{d}E\,P$ of filled states in the interval $\mathrm{d}E$, times the transition probability per unit volume and time C, times the number $\Omega\,\gamma_t(\mathbf{r}, E')\,\mathrm{d}E'\,(1 - P')$ of empty states in the interval $\mathrm{d}E'$. Thus, letting ΔE_t be an energy interval belonging to the gap and containing the traps, the transition rate from the conduction band to the traps is given by

$$r_a = \int_{E_C}^{E_{CU}} \int_{\Delta E_t} \Omega\,\gamma(E)\,\mathrm{d}E\,P\,C(E \to E')\,\Omega\gamma_t(\mathbf{r}, E')\,\mathrm{d}E'\,(1 - P')\,. \tag{20.83}$$

By the same token, the transition rate from the valence band to the traps is

$$r_d = \int_{E_{VL}}^{E_V} \int_{\Delta E_t} \Omega\,\gamma(E)\,\mathrm{d}E\,P\,C(E \to E')\,\Omega\gamma_t(\mathbf{r}, E')\,\mathrm{d}E'\,(1 - P')\,. \tag{20.84}$$

In turn, the number of transitions per unit volume and time, from states in the interval $\mathrm{d}E'$ belonging the trap distribution, to states in the interval $\mathrm{d}E$ belonging to a band, is obtained as the product of the number $\Omega\,\gamma_t(\mathbf{r}, E')\,\mathrm{d}E'\,P'$ of filled states in the interval $\mathrm{d}E'$, times $C(\mathbf{r}, E' \to E)$, times the number $\Omega\,\gamma(E)\,\mathrm{d}E\,(1 - P)$ of empty states in the interval $\mathrm{d}E$. Thus, the transition rates from the traps to conduction or valence band are respectively given by

$$r_b = \int_{E_C}^{E_{CU}} \int_{\Delta E_t} \Omega\,\gamma_t(\mathbf{r}, E')\,\mathrm{d}E'\,P'\,C(\mathbf{r}, E' \to E)\,\Omega\,\gamma(E)\,\mathrm{d}E\,(1 - P)\,, \tag{20.85}$$

$$r_c = \int_{E_{VL}}^{E_V} \int_{\Delta E_t} \Omega\,\gamma_t(\mathbf{r}, E')\,\mathrm{d}E'\,P'\,C(\mathbf{r}, E' \to E)\,\Omega\,\gamma(E)\,\mathrm{d}E\,(1 - P)\,. \tag{20.86}$$

The combined density of states of the traps is treated in the same manner as that of the dopant atoms (compare with (18.20) and (18.35)) by letting

$$\gamma_t(\mathbf{r}, E') = N_t(\mathbf{r})\,\delta(E' - E_t)\,, \tag{20.87}$$

where $N_t(\mathbf{r})$ is the trap concentration. Thanks to this, the integrals over ΔE_t are easily evaluated to yield

$$r_a = N_t\,(1 - P_t)\,\Omega^2 \int_{E_C}^{E_{CU}} \gamma\,P\,C(\mathbf{r}, E \to E_t)\,\mathrm{d}E = N_t\,(1 - P_t)\,\alpha_n\,n\,, \tag{20.88}$$

$$r_c = N_t\, P_t\, \Omega^2 \int_{E_{VL}}^{E_V} \gamma\, (1-P)\, C(\mathbf{r}, E_t \to E)\, \mathrm{d}E = N_t\, P_t\, \alpha_p\, p\,, \tag{20.89}$$

where the definitions (20.3), (20.4) of the electron and hole concentrations are used, and the transition coefficients for electrons and holes are defined as the weighed averages

$$\alpha_n = \Omega^2\, \frac{\int_{E_C}^{E_{CU}} \gamma\, P\, C\, \mathrm{d}E}{\int_{E_C}^{E_{CU}} \gamma\, P\, \mathrm{d}E}\,, \qquad \alpha_p = \Omega^2\, \frac{\int_{E_{VL}}^{E_V} \gamma\, (1-P)\, C\, \mathrm{d}E}{\int_{E_{VL}}^{E_V} \gamma\, (1-P)\, \mathrm{d}E}\,. \tag{20.90}$$

Like in the case of (20.10), the integrals in (20.90) are approximated using the equilibrium probability. The remaining transition rates r_b, r_d are determined in a similar manner, using also the approximation $1 - P \simeq 1$ in (20.85) and $P \simeq 1$ in (20.84). Like in Sect. 20.2.1, the approximation is justified by the fact that in normal operating conditions the majority of the valence-band states are filled, while the majority of the conduction-band states are empty. In conclusion,

$$r_b = N_t\, P_t\, \Omega^2 \int_{E_C}^{E_{CU}} \gamma\, (1-P)\, C(\mathbf{r}, E_t \to E)\, \mathrm{d}E \simeq N_t\, P_t\, e_n\,, \tag{20.91}$$

$$r_d = N_t\, (1-P_t)\, \Omega^2 \int_{E_{VL}}^{E_V} \gamma\, P\, C(\mathbf{r}, E \to E_t)\, \mathrm{d}E \simeq N_t\, (1-P_t)\, e_p\,, \tag{20.92}$$

with the emission coefficients defined by

$$e_n = \Omega^2 \int_{E_C}^{E_{CU}} \gamma\, C\, \mathrm{d}E\,, \qquad e_p = \Omega^2 \int_{E_{VL}}^{E_V} \gamma\, C\, \mathrm{d}E\,. \tag{20.93}$$

20.6.2 Coefficients of the Auger and Impact-Ionization Events

The expression of the coefficients c_n, c_p and I_n, I_p, to be used in the calculation of the net recombination rates (20.55) due to the Auger and impact-ionization phenomena, is found in the same way as the transition rates of the SRH recombination function (Sect. 20.2.3) or the direct thermal recombinations (Sect. 20.2.1). Let $P(\mathbf{r}, E, t)$ be the occupation probability of a state of energy E, and $C_n(E_1, E_2 \to E_1', E_2')$ the combined probability per unit time and volume (in $\mathbf{r}$) of an electron transition from a filled state of energy E_1 in the conduction band to an empty state of energy E_1' in the conduction band, and of another electron from a filled state of energy E_2 to an empty state of energy E_2', where E_2 and E_2' belong to different bands.

Auger Coefficients

In an Auger recombination it is $E_1' > E_1$; also, E_2 belongs to the conduction band while E_2' belongs to the valence band. Due to energy conservation it is[22]

$$C_n = C_{n0}\,\delta\left[(E_1 - E_1') + (E_2 - E_2')\right], \tag{20.94}$$

where $E_2 - E_2' \simeq E_G$; it follows $E_1' \simeq E_1 + E_G$. Then, define $P_i = P(\mathbf{r}, E = E_i, t)$, $P_i' = P(\mathbf{r}, E = E_i', t)$, with $i = 1, 2$, and let $\gamma(E)$ be the combined density of states in energy and volume for the bands; in particular, let $g_i = \Omega\,\gamma(E_i)$ and $g_i' = \Omega\,\gamma(E_i')$. From the above definitions one finds, for the rate r_a of the Auger recombinations initiated by electrons,

$$r_a = \int g_1\ \mathrm{d}E_1\,P_1\ g_2\ \mathrm{d}E_2\,P_2\ C_n\ g_1'\ \mathrm{d}E_1'\,(1 - P_1')\ g_2'\ \mathrm{d}E_2'\,(1 - P_2'), \tag{20.95}$$

where $\int$ indicates a fourfold integral that extends thrice over the conduction band and once over the valence band. Observing that $P_1' \ll 1$ and integrating over E_1' with $C_n = C_{n0}\,\delta(E_1 + E_G - E_1')$ yield

$$r_a = \int_{E_C}^{E_{CU}} g_1\ \mathrm{d}E_1\,P_1\ C_{n0}\,g_G \int_{E_C}^{E_{CU}} g_2\ \mathrm{d}E_2\,P_2 \int_{E_{VL}}^{E_V} g_2'\ \mathrm{d}E_2'\,(1 - P_2'), \tag{20.96}$$

where $g_G = g(E_1 + E_G)$ and $[C_{n0}\,g_G] = \mathrm{s}^{-1}\,\mathrm{m}^{-3}$. Thanks to (20.3) and (20.4), the second integral in (20.96) equals $\Omega\,n$ and the third one equals $\Omega\,p$. Letting

$$c_n = \Omega^3\,\frac{\int_{E_C}^{E_{CU}} C_{n0}\,g_G\,g_1\,P_1\ \mathrm{d}E_1}{\int_{E_C}^{E_{CU}} g_1\,P_1\ \mathrm{d}E_1}, \tag{20.97}$$

finally yields $r_a = c_n\,n^2\,p$. The derivation of $r_c = c_p\,p^2\,n$ is similar.

Impact Ionization's Transition Coefficients

Using the same symbols introduced at the beginning of Sect. 20.6.2, for an impact-ionization event induced by an electron it is $E_1 > E_1'$; in turn, E_2 belongs to the valence band and E_2' belongs to the conduction band. It follows

$$r_b = \int g_1\ \mathrm{d}E_1\,P_1\ g_2\ \mathrm{d}E_2\,P_2\ C_n\ g_1'\ \mathrm{d}E_1'\,(1 - P_1')\ g_2'\ \mathrm{d}E_2'\,(1 - P_2'), \tag{20.98}$$

[22] The units of C_{n0} are $[C_{n0} = \mathrm{J}\ \mathrm{s}^{-1}\ \mathrm{m}^{-3}]$.

where the fourfold integral extends thrice over the conduction band and once over the valence band. From the energy-conservation relation $E_1 + E_2 = E_1' + E_2'$ and from $E_2' - E_2 \simeq E_G$ it follows $E_1' \simeq E_1 - E_G$. Observing that $P_2 \simeq 1$, $P_1' \ll 1$, $P_2' \ll 1$, and integrating over E_1' with $C_n = C_{n0}\,\delta(E_1 - E_G - E_1')$ yield

$$r_b = \int_{E_C}^{E_{CU}} C_{n0}\, g_G\, g_1\, P_1\, \mathrm{d}E_1 \int_{E_{VL}}^{E_V} g_2\, \mathrm{d}E_2 \int_{E_C}^{E_{CU}} g_2'\, \mathrm{d}E_2' \,, \tag{20.99}$$

where $g_G = g(E_1 - E_G)$, and the product of the second and third integral is a dimensionless quantity that depends only on the semiconductor's structure. Indicating such a quantity with ν_n, and letting

$$I_n = \nu_n \frac{\int_{E_C}^{E_{CU}} C_{n0}\, g_G\, g_1\, P_1\, \mathrm{d}E_1}{\int_{E_C}^{E_{CU}} g_1\, P_1\, \mathrm{d}E_1} \,, \tag{20.100}$$

finally yields $r_b = I_n\, n$. The derivation of $r_d = I_p\, p$ is similar.

20.6.3 Total Recombination-Generation Rate

The expressions for the most important generation-recombination terms have been worked out in this chapter. Only one of them, the SRH recombination function U_{SRH}, involves energy states different from those of the conduction and valence bands; in principle, such states would require additional continuity equations to be added to the semiconductor-device model. However, as discussed in Sect. 20.2.3, this is not necessary in crystalline semiconductors. The other mechanisms (direct thermal recombination-generation U_{DT}, Auger recombination and impact ionization U_{AI}, and optical recombination-generation U_O) do not involve intermediate states. As a consequence, with reference to (20.13) the generation-recombination terms of the electron-continuity equation are equal to those of the hole continuity equation. Finally, assuming that the different generation-recombination phenomena are uncorrelated and neglecting U_{DT} with respect to U_{SRH} (Sect. 20.2.2) yield

$$U_n - G_n = U_p - G_p \simeq U_{SRH} + U_{AI} + U_{DO} \,. \tag{20.101}$$

20.6.4 Screened Coulomb Potential

In the context of physics, the general meaning of *screening* is the attenuation in the electric field intensity due to the presence of mobile charges; the effect is treated here using the Debye-Hückel theory [39], which is applicable to a nondegenerate semiconductor where the dopants are completely ionized. For a

medium of permittivity ε, with charge density ϱ, the electric potential in the equilibrium condition is found by solving Poisson's equation

$$-\varepsilon \nabla^2 \varphi = \varrho \,. \tag{20.102}$$

One starts by considering a locally neutral material, to which a perturbation is added due, for instance, to the introduction of a fixed charge $Z_c\, e_c$ placed in the origin; this, in turn, induces a variation in ϱ. The corresponding perturbation of φ is calculated to first order by replacing φ with $\varphi + \delta\varphi$ and ϱ with $\varrho + (\partial\varrho/\partial\varphi)\,\delta\varphi$, where the derivative is calculated at $\delta\varphi = 0$; the perturbed form of Poisson's equation reads:

$$-\varepsilon \nabla^2 \varphi - \varepsilon \nabla^2 \delta\varphi = \varrho + \frac{\partial \varrho}{\partial \varphi}\, \delta\varphi \,. \tag{20.103}$$

As the unperturbed terms cancel out due to (20.102), a Poisson equation in the perturbation is obtained,

$$\nabla^2 \delta\varphi = q_c^2\, \delta\varphi \,, \qquad q_c^2 = -\frac{\partial\varrho/\partial\varphi}{\varepsilon} \,, \tag{20.104}$$

where $1/q_c$ is the *screening length* or *Debye length*. The definition implies that $\partial\varrho/\partial\varphi < 0$; this is in fact true, as shown below with reference to a nondegenerate semiconductor with completely ionized dopants.[23] Letting $N_D^+ = N_D$, $N_A^- = N_A$ in (19.125), and using the nondegenerate expressions (18.60), (18.61), of the equilibrium concentrations, one finds that $N = N_D - N_A$ is left unaffected by the perturbation, while the electron concentration[24] n transforms into $n \exp[e\,\delta\varphi/(k_B\,T)]$ and the hole concentration p transforms into $p \exp[-e\,\delta\varphi/(k_B\,T)]$. From $\varrho = e\,(p - n + N)$ one obtains, to first order,

$$\frac{\partial \varrho}{\partial \varphi} = -\frac{e^2}{k_B\,T}\,(n + p) \,, \qquad q_c^2 = \frac{e^2\,(n + p)}{\varepsilon\, k_B\, T} > 0 \,. \tag{20.105}$$

The left-hand side of the Poisson equation in (20.104) is conveniently recast using a set of spherical coordinates r, θ, ϕ whose origin coincides with the center of symmetry of the perturbation; using (B.25) one finds

$$\nabla^2 \delta\varphi = \frac{1}{r}\frac{\partial^2}{\partial r^2}(r\,\delta\varphi) + \frac{r^{-2}}{\sin\theta}\,\frac{\partial}{\partial\theta}\left(\sin\theta\,\frac{\partial \delta\varphi}{\partial\theta}\right) + \frac{r^{-2}}{\sin^2\theta}\,\frac{\partial^2 \delta\varphi}{\partial\phi^2} \,. \tag{20.106}$$

Considering a perturbation with a spherical symmetry, only the first term at the right-hand side of (20.106) is left, whence (20.104) becomes an equation in the unknown $r\,\delta\varphi$:

[23] As shown by (A.118), the property $\partial\varrho/\partial\varphi < 0$ holds true also in the degenerate case.

[24] The electron charge is indicated here with e to avoid confusion with q_c.

$$\frac{\mathrm{d}^2}{\mathrm{d}r^2}(r\,\delta\varphi) = q_c^2\,(r\,\delta\varphi)\,. \tag{20.107}$$

The general solution of (20.107) is $r\,\delta\varphi = A_1\,\exp(-q_c\,r) + A_2\,\exp(q_c\,r)$, where it must be set $A_2 = 0$ to prevent the solution from diverging as r becomes large. In conclusion,

$$\delta\varphi = \frac{A_1}{r}\,\exp(-q_c\,r)\,. \tag{20.108}$$

The remaining constant is found by observing that for very small r the pure Coulomb case $\delta\varphi \simeq A_1/r$ is recovered, whence $A_1 = Z_c e_c e/(4\,\pi\,\varepsilon)$. This makes (20.108) to coincide with (14.33).

Part VII
Basic Semiconductor Devices

Chapter 21
Bipolar Devices

21.1 Introduction

The mathematical model of semiconductor devices, derived in Chap. 19, is applied here to the description of the fundamental bipolar device, the *p-n* junction. The term *bipolar* indicates that both electrons and holes contribute to the current. The analysis is carried out using the simple example of a one-dimensional abrupt junction in steady state, with the hypotheses of nondegeneracy and complete ionization, that lend themselves to an analytical treatment. The equilibrium condition is considered first, and the solution of Poisson's equation is tackled, showing that the structure can be partitioned into space-charge and quasi-neutral regions. Then, the Shockley theory is illustrated, leading to the derivation of the ideal $I(V)$ characteristic. The semiconductor model is then applied to illustrating two features of the reverse-bias condition, namely, the depletion capacitance and the avalanche due to impact ionization. Next, the model is used to work out the features of the photodiode, both in the continuous and storage mode, of the solar cell, and of the bipolar junction transistor. The complements justify the simplification of considering only the diffusive transport for the minority carriers in a quasi-neutral region, and provide the derivation of the Shockley boundary conditions. Finally, the expression of the depletion capacitance is worked out for the case of an arbitrary charge-density profile. The typical parameters of the *p-n* junction and of the solar cell are also illustrated in the complements, along with the single-generator equivalent circuit for the bipolar junction transistor.

M. Rudan, *Physics of Semiconductor Devices*, DOI 10.1007/978-3-319-63154-7_21

21.2 *P-N* Junction in Equilibrium

A very simple, yet fundamental, semiconductor device is the *p-n junction*, whose one-dimensional version is sketched in Fig. 21.1. The device is fabricated by thermally diffusing (Chap. 23), or ion implanting *p*-type dopant atoms into an *n*-type substrate, or vice versa. As a consequence, the diffused or implanted profile is not spatially uniform. The substrate profile may in turn result from a similar process, so that in general it is not uniform either. The locus of points where the ionized dopant concentrations are equal to each other, $N_D^+ = N_A^-$, is a surface called *metallurgical junction*.[1] The theory of the *p-n* junction is carried out with reference to a simplified structure, where the device is one dimensional and aligned with the x axis; in this case the metallurgical junction is a plane normal to x and, as shown in Fig. 21.1, its position is made to coincide with the reference's origin. Also, the nonuniform dopant concentrations $N_A(x)$ and $N_D(x)$ are approximated by piecewise-constant functions, $N_A = \text{const}$ for $x < 0$ and $N_D = \text{const}$ for $x > 0$. The device obtained from this approximation is called *abrupt p-n junction*. Considering the actual form of the dopant distribution, the approximation is not realistic; however, the much simpler model based on it is still able to capture the essential features of the device characteristics. Moreover, the model assumes that the conditions of nondegeneracy and complete ionization hold; this assumption makes the analytical approach possible.

Within an integrated circuit the *p-n* junction is supplemented with contacts that connect it to the rest of the circuit. Such contacts are typically made of metals, metal silicides, or heavily doped polycrystalline semiconductors; as a consequence, two more junctions are present: the first one is between the contact and the *p*-doped semiconductor, the other one between the contact and the *n*-doped semiconductor. It is implied that the contacts are made of the same material; if it is not so, more junctions must be considered as shown below.

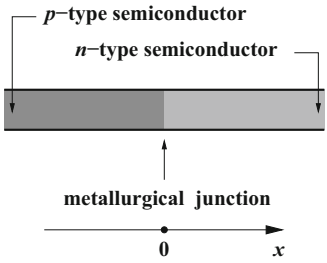


Fig. 21.1 Schematic example of a one-dimensional *p-n* junction

[1] The metallurgical junction is often indicated with the same term used for the whole device, namely, *p-n junction* or simply *junction*.

21.2.1 Built-In Potential

A qualitative description of the device in the equilibrium condition starts from the assumption that the extension of the p-doped and n-doped regions along the x axis is large, so that, far away from the junction, the semiconductor can be considered as uniformly doped of the p or n type, respectively. This fixes the boundary conditions for the electron and hole concentrations:[2] in fact, remembering that in the nondegeneracy and complete-ionization conditions the equilibrium concentrations in a uniform semiconductor are given by (18.42) for the p type and by (18.30) for the n type, one finds

$$p_{p0} = p(-\infty) \simeq N_A\,, \qquad n_{p0} = n(-\infty) \simeq \frac{n_i^2}{N_A}\,, \tag{21.1}$$

$$n_{n0} = n(+\infty) \simeq N_D\,, \qquad p_{n0} = p(+\infty) \simeq \frac{n_i^2}{N_D}\,. \tag{21.2}$$

The above concentrations are also called *asymptotic concentrations*; the last approximations are derived from the assumption $N_A, N_D \gg n_i$ which, as outlined in Sect. 18.4.1, has a vast range of validity. The distance between the conduction-band edge and the Fermi level is found from $n = N_C \exp[-(E_C - E_F)/(k_B T)]$ (compare with (18.28)); combining with (21.1) and (21.2) yields

$$\frac{n_i^2}{N_A} \simeq N_C \exp\left[-\frac{E_C(-\infty) - E_F}{k_B T}\right], \quad N_D \simeq N_C \exp\left[-\frac{E_C(+\infty) - E_F}{k_B T}\right], \tag{21.3}$$

whence

$$E_C(-\infty) - E_C(+\infty) = k_B T \log\left(\frac{N_A N_D}{n_i^2}\right). \tag{21.4}$$

An identical expression is found for $E_V(-\infty) - E_V(+\infty)$. These findings show that E_C, E_V are functions of position; their explicit form is determined below. Alternatively, one may use (18.60) and (18.61), to find

$$\psi_0 = \frac{k_B T}{q} \log\left(\frac{N_A N_D}{n_i^2}\right), \qquad \psi_0 = \varphi(+\infty) - \varphi(-\infty)\,, \tag{21.5}$$

where ψ_0 is called *built-in potential*.[3] One notes that so far the values of the constants $n^{(0)}$, $p^{(0)}$ in (18.60) and (18.61) have been left unspecified; remembering

[2] The use of asymptotic conditions is not applicable to shallow junctions like, e.g., those used for the fabrication of solar cells. In this case, the theory is slightly more involved (Sect. 21.7).

[3] The same quantity is also called *barrier potential* and is sometimes indicated with ψ_B.

that $n^{(0)}$ is the value of n in the position(s) where $\varphi = 0$, and the same for $p^{(0)}$, the numerical values sought are determined by specifying the zero point of φ. Here such a point is fixed by letting $\varphi(+\infty) = 0$ whence, using (21.3) and (21.4), one finds

$$p^{(0)} = p(+\infty) = p_{n0}\,, \qquad n^{(0)} = n(+\infty) = n_{n0}\,. \tag{21.6}$$

The expressions of the carrier concentrations in terms of the band energies E_C, E_V must be coherent with those expressed in terms of φ. In fact, the relation between $E_C(x)$ and $\varphi(x)$ is found by combining $n = n^{(0)} \exp[q\,\varphi/(k_B\,T)]$ with $n = N_C \exp[(E_F - E_C)/(k_B\,T)]$ and using $n^{(0)} = N_C \exp\{[E_F - E_C(+\infty)]/(k_B\,T)\}$; a similar procedure is applied to E_V, to eventually find

$$E_C(x) = E_C(+\infty) - q\,\varphi(x)\,, \qquad E_V(x) = E_V(+\infty) - q\,\varphi(x)\,. \tag{21.7}$$

Letting $N(x) = -N_A$ for $x < 0$ and $N(x) = +N_D$ for $x > 0$, the electric potential is found by solving the Poisson equation

$$\frac{\mathrm{d}^2\varphi}{\mathrm{d}x^2} = \frac{q}{\varepsilon_{\mathrm{sc}}}\left[n_{n0}\,\exp\left(\frac{q\,\varphi}{k_B\,T}\right) - p_{n0}\,\exp\left(\frac{-q\,\varphi}{k_B\,T}\right) - N(x)\right] \tag{21.8}$$

with boundary conditions $\varphi(-\infty) = -\psi_0$ and $\varphi(+\infty) = 0$. One notes that within each half domain the charge density in (21.8) has the form $\rho = \rho(\varphi)$, which makes the theory of Sect. 19.5.8 applicable. Therefore, it is convenient to separately solve (21.8) in each half space, and apply suitable matching conditions at $x = 0$ afterwards. When the regional-solution method is used, the boundary conditions $\varphi(-\infty) = -\psi_0$ and $\varphi(+\infty) = 0$ apply separately in each half domain, as shown below.

In the n-doped region the charge density reads $\rho = q\,(p - n + N_D)$; when $x \to +\infty$ the latter becomes $0 = p_{n0} - n_{n0} + N_D$: in fact, as at large distances from the origin the material behaves like a uniformly doped semiconductor, local charge neutrality is fulfilled at infinity. Using the dimensionless potential $u = q\,\varphi/(k_B\,T)$ and indicating the derivatives with primes give the equation the form

$$u'' = \frac{1}{L_D^2}\,A_D(u)\,, \quad A_D = \exp(u) - 1 + \frac{p_{n0}}{n_{n0}}\,[1 - \exp(-u)]\,, \quad L_D^2 = \frac{\varepsilon_{\mathrm{sc}}\,k_B\,T}{q^2\,n_{n0}}\,, \tag{21.9}$$

with L_D the *Debye length for the electrons*. The normalized charge density A_D vanishes for $u = 0$, and is positive (negative) when u is positive (negative). Note that, by letting $x \to +\infty$ in $n = n_i \exp[(\varphi - \varphi_F)/(k_B\,T)]$, $p = n_i \exp[(\varphi_F - \varphi)/(k_B\,T)]$, and using the normalized Fermi potential $u_F = q\,\varphi_F/(k_B\,T)$, one finds $p_{n0}/n_{n0} = (n_i/N_D)^2 = \exp(2\,u_F)$, where $u_F < 0$ on account of the fact that here an n-doped region is considered. Following the method illustrated in Sect. 19.5.8 transforms the left-hand side of the first equation in (21.9) into $u''\,u' = (1/2)\,[(u')^2]'$. This term is then integrated from $x = 0$ to $x = +\infty$; the

result is simplified by observing that the region far from the junction is substantially uniform, whence the electric potential is constant there. As u' is proportional to the electric field, it follows that $u'(+\infty) = 0$; in conclusion, the integration of (21.9) from $x \geq 0$ to $+\infty$ yields

$$\left(u'\right)^2 = \frac{2}{L_D^2} B_D(u)\,, \quad B_D = \exp(u) - 1 - u + \frac{n_i^2}{N_D^2}\left[u + \exp(-u) - 1\right]. \tag{21.10}$$

It is easily found that B_D is nonnegative; in fact, from (21.10) one derives $B_D = 0$ for $u = 0$, and at the same time $\mathrm{d}B_D/\mathrm{d}u = A_D(u)$ is positive for $u > 0$, negative for $u < 0$; as a consequence, B_D grows from zero in either direction when u departs from the origin. Letting $F_D^2 = B_D$ one finds $|u'| = \sqrt{2}\,F_D/L_D$; as the condition $u = 0$ holds only asymptotically, the modulus of u' always grows as u departs from the origin, showing that u is monotonic. Considering that u must fulfill the other boundary condition $u(-\infty) = -q\,\psi_0/(k_B\,T) < 0$, one concludes that u is a monotonically growing function, whence, choosing the positive sign and separating the variables, one finds

$$\frac{\mathrm{d}u}{F_D(u)} = \frac{\sqrt{2}}{L_D}\,\mathrm{d}x \tag{21.11}$$

for $x \geq 0$. The above must be tackled numerically because it has no analytical solution.[4] To integrate (21.11) one needs the value of the potential in the origin, $u(x = 0^+)$; for the time being the latter is not known, and must be calculated from the matching conditions as shown below. Such a calculation provides also the derivative in the origin, $u'(x = 0^+) = \sqrt{2}\,F_D[u(x = 0^+)]/L_D$.

The solution for $x < 0$ follows the same pattern, where the asymptotic neutrality condition reads $n_{n0}\exp(-\chi_0) - p_{n0}\exp(\chi_0) + N_A = 0$, with $\chi_0 = q\,\psi_0/(k_B\,T)$. Letting $v = u + \chi_0$, $n_{p0}/p_{p0} = (n_i/N_A)^2 = \exp(-2\,u_F) \ll 1$, and using (21.5) provide

$$v'' = \frac{1}{L_A^2} A_A(v)\,, \quad A_A = \frac{n_i^2}{N_A^2}\left[\exp(v) - 1\right] + 1 - \exp(-v)\,, \quad L_A^2 = \frac{\varepsilon_{\mathrm{sc}}\,k_B\,T}{q^2\,p_{p0}}\,, \tag{21.12}$$

where L_A is the *Debye length for the holes*. Like in the *n*-doped region, the normalized charge density A_A vanishes for $u = 0$ and is positive (negative) when u is positive (negative). Also in this case, one transforms the normalized Poisson equation (21.12) into $v''\,v' = (1/2)\,[(v')^2]'$; integrating the latter from $x = -\infty$, where $v(-\infty) = u(-\infty) + \chi_0 = 0$, to $x \leq 0$ and letting $v'(-\infty) = 0$ yield

[4] The numerical procedure is outlined in the note of Sect. 22.2.1.

$$\left(v'\right)^2 = \frac{2}{L_A^2} B_A(v)\,, \quad B_A = \frac{n_i^2}{N_A^2}\left[\exp(v) - 1 - v\right] + v + \exp(-v) - 1\,. \tag{21.13}$$

Following the same reasoning as for B_D, one shows that B_A is nonnegative, so u is a monotonically growing function also for $x < 0$. Letting $F_A^2 = B_A$ and choosing the positive sign when separating the variables, one finally finds

$$\frac{\mathrm{d}v}{F_A(v)} = \frac{\sqrt{2}}{L_A}\,\mathrm{d}x \tag{21.14}$$

for $x \leq 0$. From (21.13) one also obtains $v'(x = 0^-) = u'(x = 0^-) = \sqrt{2}\,F_A[v(x = 0^-)]/L_A$. On the other hand, as the electric potential and field are continuous, it is $u'(x = 0^-) = u'(x = 0^+) = u'(x = 0)$ and $u(x = 0^-) = u(x = 0^+) = u(x = 0)$; equating the derivatives and using the continuity of the potential yield the matching condition

$$\frac{1}{L_A} F_A[u(x = 0) + \chi_0] = \frac{1}{L_D} F_D[u(x = 0)]\,, \tag{21.15}$$

namely, a transcendental equation in the unknown $u(x = 0)$. Once found, $u(x = 0)$ is used as integration limit in (21.11) and (21.14).

21.2.2 Space-Charge and Quasi-Neutral Regions

The form of the electric potential φ is shown in Fig. 21.2 for a p-n junction at equilibrium with $N_A = 10^{16}$ cm^{-3}, $N_D = 10^{15}$ cm^{-3}. The form of the bands is shown in Fig. 21.3 for the same device. It is interesting to note that the device can be thought of as made of three regions: in the intermediate region, whose boundaries are marked by dashed vertical lines in Fig. 21.2, the electric potential has a non-negligible curvature, thus showing that the charge density is large. The region is called *space-charge region* and contains the metallurgical junction, marked by the continuous vertical line. In the two regions on the sides of the space-charge region, the electric potential is nearly constant,[5] whence the electric field $-\mathrm{d}\varphi/\mathrm{d}x$ is negligibly small. As a consequence, the charge density $\rho = -\varepsilon_{\mathrm{sc}}\,\mathrm{d}^2\varphi/\mathrm{d}x^2$ is negligible as well; for this reason, the two regions under consideration are called *quasi-neutral regions*.[6]

[5] The electric potential cannot be exactly constant, because the solution of (21.8) is an analytical function; as a consequence, if φ was constant in a finite interval, it would be constant everywhere.

[6] The inverse reasoning would not be correct: in fact, $\rho = 0$ may yield $\mathrm{d}\varphi/\mathrm{d}x = \text{const} \neq 0$, which makes φ a linear function of x; deducing $\varphi = \text{const}$ from $\rho = 0$ is correct only if the additional condition of spatial uniformity holds.

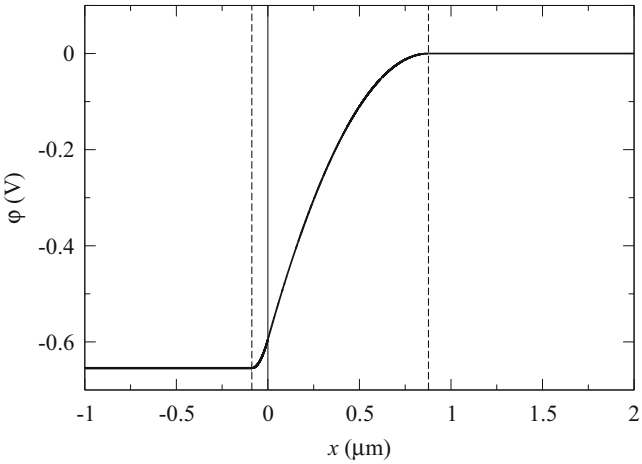


Fig. 21.2 Solution of the one-dimensional Poisson equation (21.8) in an abrupt *p-n* junction at equilibrium, with $N_A = 10^{16}$ cm^{-3}, $N_D = 10^{15}$ cm^{-3}. The *continuous vertical line* marks the position of the metallurgical junction, the *dashed vertical lines* mark the edges of the space-charge region

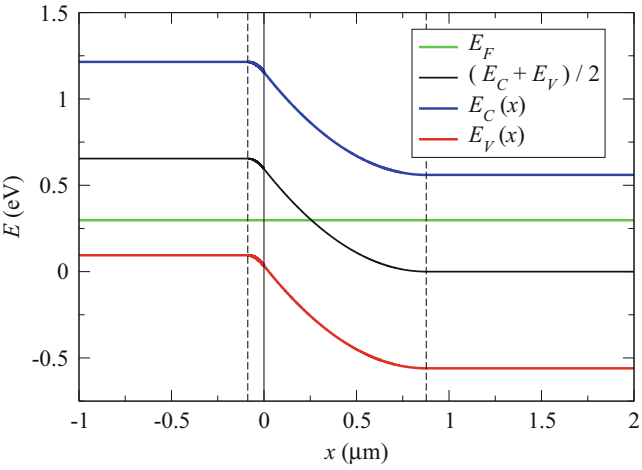


Fig. 21.3 Form of the bands for the same device as in Fig. 21.2

The transition from the space-charge region and one or the other quasi-neutral region is sharp. Thanks to this, it is possible to identify the width l of the space-charge region and correlate it with other parameters of the device; such a correlation is worked out in Sect. 21.4 in a specific operating regime. For convenience, the width of the space-charge region is expressed as $l = l_p + l_n$, where l_p is the extension of the space-charge region on the p side of the metallurgical junction, and l_n the analogue

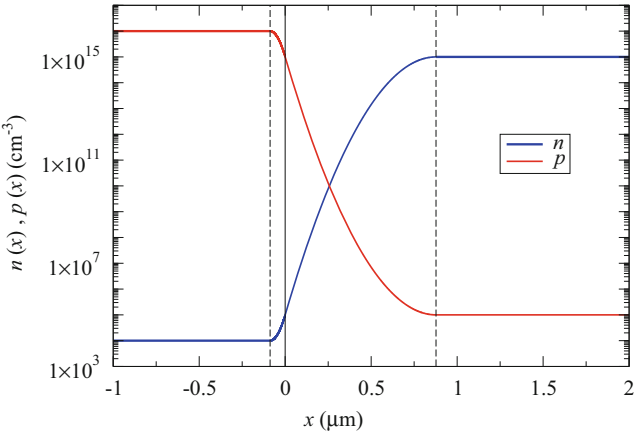


Fig. 21.4 Electron and hole concentrations in a one-dimensional, abrupt p-n junction at equilibrium, with $N_A = 10^{16}$ cm^{-3}, $N_D = 10^{15}$ cm^{-3}. The figure is drawn in a logarithmic scale. The *continuous vertical line* marks the position of the metallurgical junction, the *dashed vertical lines* mark the edges of the space-charge region

on the n side. As shown in Fig. 21.2, it is $l_n > l_p$; as explained in Sect. 21.4, this is due to the global charge neutrality and to the fact that $N_D < N_A$.

If the equilibrium carrier concentrations corresponding to the electric potential of Fig. 21.2 are drawn in a logarithmic scale, the curves look similar to that of Fig. 21.2, apart from scaling factors and from the inversion due to the negative sign in $p = p_{n0} \exp[-q\,\varphi/(k_B\,T)]$. This is shown in Fig. 21.4. A more realistic representation, shown in Fig. 21.5, uses a linear scale. The hole concentration p ranges from $p_{p0} \simeq N_A = 10^{16}$ cm^{-3} in the p-type quasi-neutral region to $p_{n0} \simeq n_i^2/N_D = 10^5$ cm^{-3} in the n-type quasi-neutral region; similarly, the electron concentration n ranges from $n_{p0} \simeq n_i^2/N_A = 10^4$ cm^{-3} in the p-type quasi-neutral region to $n_{n0} \simeq N_D = 10^{15}$ cm^{-3} in the n-type quasi-neutral region. This shows that the two concentrations vary by several orders of magnitude over the space-charge region, whose length is about 1 μm; for this reason, if these gradients existed alone, they would make holes (electrons) to diffuse in the positive (negative) direction of the x axis. Such diffusions in fact do not occur, because in the equilibrium condition they are balanced by the electric field. The latter is negative: in fact, the electric potential increases with x, so that the force associated with the electric field E is negative (positive) for holes (electrons); the equations for the current densities of the semiconductor device model in (19.129) and (19.130) yield in this case

$$-q\,\mu_p\,p\,E = -q\,D_p\,\frac{\mathrm{d}p}{\mathrm{d}x} > 0\,, \qquad -q\,\mu_n\,n\,E = +q\,D_n\,\frac{\mathrm{d}n}{\mathrm{d}x} > 0\,, \tag{21.16}$$

whence $J_p = J_n = 0$. The description of the p-n junction in the equilibrium condition is completed by adding the contacts; as indicated above, this amounts to

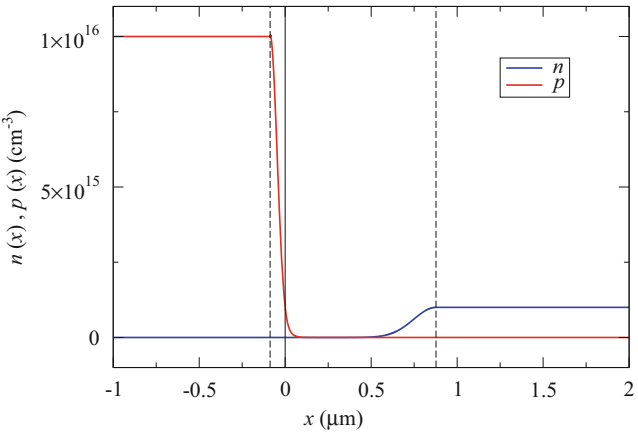


Fig. 21.5 The same concentrations as in Fig. 21.4, drawn in a linear scale

introducing two more junctions. The contacts are made of materials different from the semiconductor of which the *p-n* junction is made, hence the atomic structure of the contact's material must adapt to that of the semiconductor when the contact is deposited on it; for this reason, the structure of the contact-semiconductor junction must be described on a case-by-case basis. From the qualitative standpoint, one can use the analogy with the *p-n* junction to deduce the existence of a built-in potential Φ_{mp} between the contact and the *p*-type semiconductor, and of another built-in potential Φ_{mn} between the contact and the *n*-type semiconductor. The built-in potentials are influenced by the dopant concentration of the semiconductor, namely, $\Phi_{mp} = \Phi_{mp}(N_A)$ and $\Phi_{mn} = \Phi_{mn}(N_D)$. Assume that the contacts are made of the same material; if they are short-circuited, a closed loop is formed where, from Kirchhoff's voltage law, the built-in potentials fulfill the relation

$$\psi_0 + \Phi_{mn} - \Phi_{mp} = 0\,. \tag{21.17}$$

This situation is schematically illustrated in Fig. 21.6, where it is assumed that the material of the contacts is the same.[7] In the figure, the built-in potentials at the contacts are represented by discontinuities; in the practical cases, in fact, to prevent the contact-semiconductor junction from behaving like a rectifying device, a heavy dose of dopant is preliminarily introduced into the semiconductor region onto which the contact is to be deposited. For this reason, the spatial extension where Φ_{mp} or Φ_{mn} occurs is negligibly small with respect to the typical scale length of the device. Regardless of this, another important outcome of the fabrication process mentioned above is that the concentration of carriers available in the materials forming a contact is very large; for this reason, as mentioned in Sect. 19.5.6, a

[7] If it is not so, one must add to the left-hand side of (21.17) the barrier between the two materials.

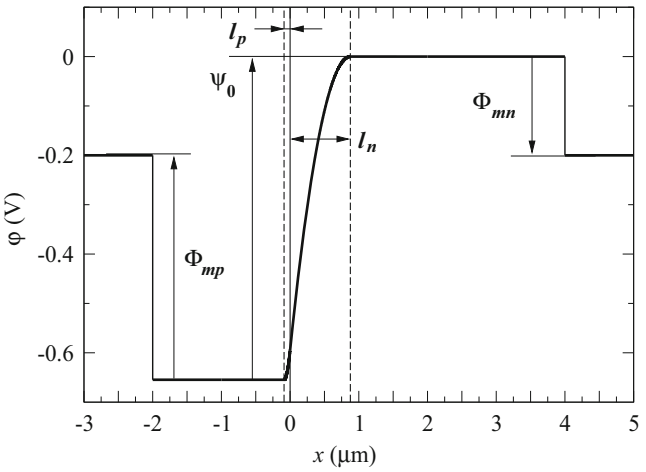


Fig. 21.6 Electric potential for the same device as in Fig. 21.2, including the built-in potentials of the contacts

contact is able to supply the amount of charge necessary to keep the equilibrium and charge-neutrality conditions in the semiconductor layer adjacent to it. This also implies that within some limits to be specified later, in a nonequilibrium condition the built-in potential is practically the same as in equilibrium. In circuit theory, a contact whose built-in potential is independent of the current that crosses it is called *ideal Ohmic contact*; this condition is equivalent to that of a vanishing differential resistivity of the contact.

21.3 Shockley Theory of the *P-N* Junction

The analytical derivation of the current-voltage characteristic of the *p-n* junction is based on the hypothesis that the device is not too far from equilibrium, so that the weak-injection condition (20.35) is fulfilled in the quasi-neutral regions. Within this limit, the approximation that the contacts are ideal is acceptable. A nonequilibrium condition is obtained by applying, e.g., a bias voltage V between the contacts; if V is such that the electric potential at the contact of the p region is higher than that of the n region, the condition is called *forward bias*; when it is lower, the condition is called *reverse bias*. Figure 21.7 shows the symbol of the *p-n* junction used in circuit theory, along with the standard references for the applied voltage V and current I. In a one-dimensional case the current is given by $I = A_e J$, with A_e the device cross-sectional area and J the total current density.

Numerical solutions of the semiconductor-device model show that in the weak-injection condition the partitioning of the device into space-charge and quasi-neutral regions still holds; this implies that when a bias voltage is applied between the

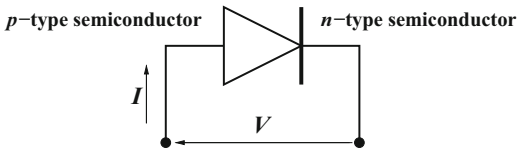


Fig. 21.7 Symbol and typical I, V reference for the p-n junction

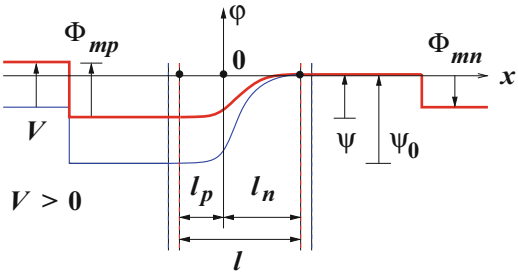


Fig. 21.8 Electric potential φ in a p-n junction in forward-biased ($V > 0$, *thick-red lines*) and equilibrium (*thin-blue lines*) conditions. When $V > 0$, the extension l of the space-charge region is smaller than in equilibrium. The drawing is not in the same scale as that of Fig. 21.6

contacts, such a voltage adds algebraically to the built-in potential. In fact, the discontinuities at the contacts are the same as in the equilibrium condition due to the contacts' ideality, and the electric potential in the quasi-neutral regions is nearly constant; the extension of the space-charge region, instead, changes due to the application of the external voltage, $l_p = l_p(V)$, $l_n = l_n(V)$. When a forward bias is applied, l_p and l_n slightly decrease, as qualitatively shown in Fig. 21.8; the drawing is not in the same scale as Fig. 21.6 and is meant only to show the change in l. The same applies to Fig. 21.9 that refers to the reverse bias and shows that in this case l_p and l_n increase. Using (21.17), the application of Kirchhoff's voltage law to either case yields, for the voltage drop ψ across the space-charge region, the expression

$$\psi + \Phi_{mn} + V - \Phi_{mp} = 0\,, \qquad \psi + V - \psi_0 = 0\,. \tag{21.18}$$

In reverse bias ($V < 0$) it is always $\psi > \psi_0 > 0$; in forward bias ($V > 0$) a sufficiently large value of V in (21.18) could make ψ to become negative. However, when V becomes large the weak-injection condition does not hold any more, and (21.18) does not apply; in conclusion, the range of forward biases to be considered here is such that the condition $\psi_0 > \psi > 0$ is always fulfilled.

When a forward bias is applied, due to $\psi < \psi_0$ the electric field within the space-charge region decreases with respect to the equilibrium case;[8] thus, the drift term in the drift-diffusion equations of (19.129) and (19.130) becomes weaker. The diffusion term, in contrast, becomes slightly stronger, because the values of the electron concentrations in the quasi-neutral regions are fixed by the asymptotic conditions, and the width of the space-charge region slightly decreases with respect

[8]The width of the space-charge region decreases as well (Fig. 21.8); such a decrease, however, is small and does not compensate for the decrease in the potential drop.

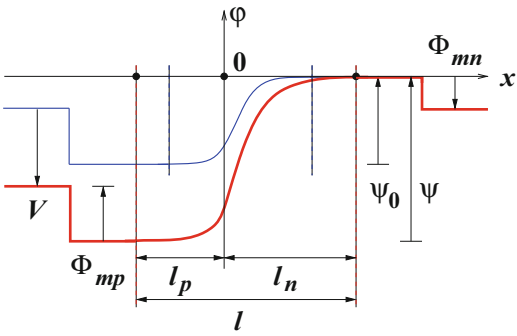


Fig. 21.9 Electric potential φ in a *p-n* junction in reverse-biased ($V < 0$, *thick-red lines*) and equilibrium (*thin-blue lines*) conditions. When $V < 0$, the extension l of the space-charge region is larger than in equilibrium. The drawing is not in the same scale as that of Fig. 21.6

to the equilibrium case. In conclusion, the diffusion term prevails and the current-density equations yield

$$-q\,D_p\,\frac{\mathrm{d}p}{\mathrm{d}x} > -q\,\mu_p\,p\,E > 0\,, \qquad q\,D_n\,\frac{\mathrm{d}n}{\mathrm{d}x} > -q\,\mu_n\,n\,E > 0\,, \tag{21.19}$$

so that $\mathbf{J}_p \cdot \mathbf{i} = q\,p\,v_p > 0$ and $\mathbf{J}_n \cdot \mathbf{i} = -q\,n\,v_n > 0$. The total current density $(\mathbf{J}_p + \mathbf{J}_n) \cdot \mathbf{i}$ is positive as well.

When a reverse bias is applied, due to $\psi > \psi_0$ the voltage drop across the space-charge region increases with respect to the equilibrium case. The region's width increases as well; however, the increase in l is relatively weak, whence the electric field within the space-charge region increases and the drift term in the drift-diffusion equations of (19.129) and (19.130) becomes stronger. The diffusion term, in contrast, becomes weaker, because the values of the electron concentrations in the quasi-neutral regions are fixed by the asymptotic conditions. In conclusion, the drift term prevails and the current-density equations yield

$$-q\,\mu_p\,p\,E > -q\,D_p\,\frac{\mathrm{d}p}{\mathrm{d}x} > 0\,, \qquad -q\,\mu_n\,n\,E > q\,D_n\,\frac{\mathrm{d}n}{\mathrm{d}x} > 0\,, \tag{21.20}$$

so that $\mathbf{J}_p \cdot \mathbf{i} = q\,p\,v_p < 0$, $\mathbf{J}_n \cdot \mathbf{i} = -q\,n\,v_n < 0$. The total current density is negative as well.

The $I(V)$ relation of the *p-n* junction is worked out here in the one-dimensional and steady-state case, thus leading to the *Shockley equations* [121, 122]. The steady-state form $\mathrm{div}\mathbf{J} = 0$ of the continuity equation (4.23) reduces in one dimension to $\mathrm{d}J/\mathrm{d}x = 0$, whence

$$J = J_p(x) + J_n(x) = \text{const}\,. \tag{21.21}$$

The hole and electron current densities depend on position and fulfill the steady-state continuity equations of (19.129) and (19.130); in the latter, only the net thermal recombination term is considered to find

$$\frac{\mathrm{d}J_p}{\mathrm{d}x} = -q\,U_{\mathrm{SRH}}\,, \qquad \frac{\mathrm{d}J_n}{\mathrm{d}x} = q\,U_{\mathrm{SRH}}\,. \tag{21.22}$$

Once J_p, J_n are determined from (21.22), they are specified at a suitable position and added up. As shown below, such positions are the boundaries $-l_p$ and l_n between the space-charge and quasi-neutral regions, for instance, $J = J_p(-l_p) + J_n(-l_p)$. Observing that $-l_p$ is the boundary of the p-type region, $J_p(-l_p)$ is a majority-carrier current density, whereas $J_n(-l_p)$ is a minority-carrier current density. The opposite happens if the other boundary is chosen, to yield $J = J_p(l_n) + J_n(l_n)$. To proceed, it is convenient to seek for an expression of J where both current densities refer to minority carriers; this is achieved by integrating (21.22) over the space-charge region to define the *recombination current density*

$$J_U = \int_{-l_p}^{l_n} q\,U_{\mathrm{SRH}}\;\mathrm{d}x = J_p(-l_p) - J_p(l_n) = J_n(l_n) - J_n(-l_p)\,. \tag{21.23}$$

Combining (21.23) with the expression of the total current density at, e.g., l_n provides

$$J = J_p(l_n) + J_n(l_n) = J_p(l_n) + J_n(-l_p) + J_U\,, \tag{21.24}$$

which has the desired form.

21.3.1 Derivation of the I(V) Characteristic

Remembering the simplified form (20.38) or (20.40) of the net thermal-recombination term in the weak-injection condition, and using $p^{\mathrm{eq}} = p_{n0}$ in the n-type region and $n^{\mathrm{eq}} = n_{p0}$ in the p-type region, transforms (21.22) into, respectively,

$$\frac{\mathrm{d}J_p}{\mathrm{d}x} = -q\,\frac{p - p_{n0}}{\tau_p}\,, \qquad x > l_n\,, \tag{21.25}$$

$$\frac{\mathrm{d}J_n}{\mathrm{d}x} = q\,\frac{n - n_{p0}}{\tau_n}\,, \qquad x < -l_p\,. \tag{21.26}$$

In this way, the hole- and electron-continuity equations are decoupled from each other; also, they are to be solved over disjoint intervals. To the current-continuity equations one associates the corresponding drift-diffusion equation taken from (19.129) or (19.130); it follows that in each quasi-neutral region the drift-diffusion equation is that of the minority carriers. It can be shown that in a quasi-neutral region, when the weak-injection condition holds, the diffusion term of the minority carriers dominates over the drift term (the details are worked out in Sect. 21.9.1). Thus, for $x < -l_p$ (p-type region), $J_n = q\,\mu_n\,n\,E + q\,D_n\,\mathrm{d}n/\mathrm{d}x \simeq$

$q\,D_n\,\mathrm{d}n/\mathrm{d}x$. Inserting the latter into (21.26) yields $D_n\,\mathrm{d}^2n/\mathrm{d}x^2 = (n - n_{p0})/\tau_n$. In this derivation the diffusion coefficient $D_n = k_B\,T\,\mu_n/q$ is not subjected to the derivative; in fact, the two parameters that influence bulk mobility, T and N_A (Sect. 20.5), are independent of position. The equation then reads,

$$\frac{\mathrm{d}^2(n - n_{p0})}{\mathrm{d}x^2} = \frac{n - n_{p0}}{L_n}\,, \qquad L_n = \sqrt{\tau_n\,D_n}\,, \tag{21.27}$$

with L_n the *diffusion length* of the minority carriers in the p-type region. It is a second-order, linear equation in the unknown n; it is decoupled from the rest of the semiconductor-device model: in fact, the simplified form of the net-recombination term contains only the electron concentration, and the neglect of the drift term eliminates the coupling with the Poisson equation. The boundary conditions must be fixed at $x \to -\infty$ and $x = -l_p$; the former is $n(-\infty) = n_{p0}$, whereas the latter needs a more elaborate derivation, given in Sect. 21.9.2, whose outcome (also called *Shockley's boundary condition*) is $n(-l_p) = n_{p0}\,\exp[q\,V/(k_B\,T)]$. The general solution of (21.27) is

$$n = n_{p0} + A_n\,\exp(x/L_n) + B_n\,\exp(-x/L_n)\,, \tag{21.28}$$

whence the asymptotic boundary condition yields $B_n = 0$. The other boundary condition provides $n(-l_p) = n_{p0} + A_n^-$, with $A_n^- = A_n\,\exp(-l_p/L_n)$. The electron current density is then found from

$$J_n = q\,D_n\,\frac{\mathrm{d}n}{\mathrm{d}x} = q\,\frac{D_n}{L_n}\,A_n\,\exp(x/L_n) = J_n(-l_p)\,\exp[(x + l_p)/L_n]\,, \tag{21.29}$$

where, using $n(-l_p) = n_{p0} + A_n^-$ and the boundary condition at $x = -l_p$,

$$J_n(-l_p) = \frac{q\,D_n\,A_n^-}{L_n} = \frac{q\,D_n\,n_{p0}}{L_n}\,F\,, \qquad F(V) = \exp\left[q\,V/(k_B\,T)\right] - 1\,. \tag{21.30}$$

In the same manner one finds

$$J_p(l_n) = \frac{q\,D_p\,p_{n0}}{L_p}\,F\,, \qquad L_p = \sqrt{\tau_p\,D_p}\,, \tag{21.31}$$

where F is the same as in (21.30), and L_p is the diffusion length of the minority carriers in the n-type region. Inserting (21.30) and (21.31) into (21.24) yields the total current density,

$$J = J_p(l_n) + J_n(-l_p) + J_U = q\left(\frac{D_p\,p_{n0}}{L_p} + \frac{D_n\,n_{p0}}{L_n}\right)F + J_U\,. \tag{21.32}$$

Multiplying (21.32) by the cross-sectional area A_e and defining the *saturation current density*

$$J_s = q\left(\frac{D_p\, p_{n0}}{L_p} + \frac{D_n\, n_{p0}}{L_n}\right) = q\, n_i^2 \left(\frac{\sqrt{D_p/\tau_p}}{N_D} + \frac{\sqrt{D_n/\tau_n}}{N_A}\right), \tag{21.33}$$

yield the expression of the $I(V)$ characteristic of the *p-n* junction:

$$I = I_s\left[\exp\left(\frac{q\,V}{k_B\,T}\right) - 1\right] + I_U\,, \qquad I_s = A_e\, J_s\,, \quad I_U = A_e\, J_U\,. \tag{21.34}$$

The characteristic fulfills the equilibrium condition $I(0) = 0$; in fact, at equilibrium it is $U_{\mathrm{SRH}} = 0$. When $q\,V/k_B\,T \gg 1$, the exponential term in (21.34) prevails over the other terms and the characteristic becomes $I \simeq I_s \exp\left[qV/(k_B\,T)\right]$, namely, the well-known exponential form of the forward-bias case. Finally, when $q\,V/k_B\,T \ll -1$, the current becomes $I \simeq -I_s + I_U$. As the order of magnitude of I_s may be similar to that of I_U, it is necessary to calculate the latter explicitly. The analysis is made easier by the observation that the electric field (which in the reverse-bias condition prevails over diffusion) drains the holes from the space-charge region to the *p*-type quasi-neutral region; similarly, the electrons of the space-charge region are drained towards the *n*-type quasi-neutral region. As a consequence, the carrier concentrations in the space-charge region are negligible, and the full-depletion condition (20.34) applies there; using the nondegenerate expression one finds, for the reverse-bias current,

$$I \simeq -I_s + I_U \simeq -I_s + A_e \int_{-l_p}^{l_n} q\,\frac{-n_i}{\tau_g}\,\mathrm{d}x = -I_s - q\,A_e\,\frac{n_i}{\tau_g}\; l(V) < 0\,, \tag{21.35}$$

where the expression (21.23) of the recombination current density has been used, and $l = l(V)$ is the width of the space-charge region. As shown in Sect. 21.4, in the reverse-bias condition and for an abrupt junction it is $l \propto \sqrt{\psi_0 + |V|}$; as a consequence, $|I_U|$ increases with $|V|$.

The approximations that have been introduced to derive the $I(V)$ characteristic are many; in fact, (21.34) is referred to as the *ideal characteristic*. However, it captures quite well the general behavior of the device as long as the applied voltage is within the limit of the weak-injection approximation. When the forward bias exceeds such a limit, the drift term is not negligible anymore and the effect of the electric field in the quasi-neutral regions must be accounted for. Considering in turn the reverse-bias condition, at large values of $|V|$ the electric field in the space-charge region becomes sufficiently strong to induce impact ionization (Sect. 20.3) and, possibly, the junction breakdown due to avalanche (Sect. 21.5).

The dependence of the $I(V)$ characteristic on temperature is due, besides that of $q\,V/(k_B\,T)$, to the coefficients of J_s and J_U. To this purpose, the second form of (21.33) is more useful because it shows explicitly the term n_i^2, whose

temperature dependence is exponential (18.14); in fact, the temperature dependence of D_p/τ_p, D_n/τ_n is much weaker. The same considerations apply to J_U, whose main dependence on temperature is due to factor n_i. The dependence on n_i of the reverse current (at constant temperature) has been used in the considerations about the parasitic currents in integrated circuits made in Sect. 18.7.

21.4 Depletion Capacitance of the Abrupt *P-N* Junction

It has been anticipated in Sect. 21.3.1 that when a reverse bias is applied to the junction, the full-depletion condition holds in the space-charge region; as a consequence, the charge density ρ in the latter is essentially due to the dopant atoms. In the abrupt junction considered here, the dopants' concentration is piecewise constant; a simplified description of the charge density in the situation in hand is obtained from the *abrupt space-charge edge* (ASCE) approximation, which describes ρ with the form sketched in Fig. 21.10. In other terms, the approximation consists in replacing with a discontinuity the smooth change of ρ at $x = -l_p$, $x = 0$, and $x = l_n$; thus, the space charge is given by

$$\rho = -q N_A \,, \qquad -l_p < x < 0 \,, \tag{21.36}$$

$$\rho = q N_D \,, \qquad 0 < x < l_n \,, \tag{21.37}$$

and $\rho = 0$ elsewhere. The electric field and the electric potential are continuous because there are no charge layers or double layers; letting $E_0 = E(0)$, from $\mathrm{d}E/\mathrm{d}x = \rho/\varepsilon_{\mathrm{sc}}$ and (21.36), (21.37) one draws

$$\frac{E_0 - E(-l_p)}{l_p} = -\frac{q N_A}{\varepsilon_{\mathrm{sc}}} \,, \qquad \frac{E(l_n) - E_0}{l_n} = \frac{q N_D}{\varepsilon_{\mathrm{sc}}} \,, \tag{21.38}$$

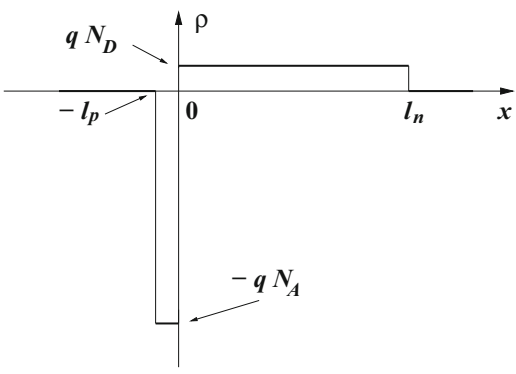


Fig. 21.10 Charge density in a reverse-biased *p-n* junction using the ASCE approximation, in arbitrary units. The ratio N_A/N_D is the same as in Fig. 21.6

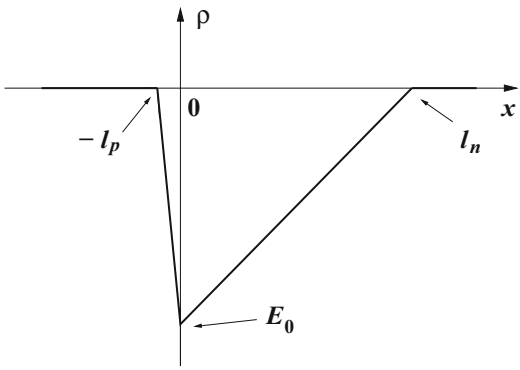


Fig. 21.11 Electric field consistent with the charge density of Fig. 21.10, in arbitrary units

the first of which holds for $-l_p \le x \le 0$, the second one for $0 \le x \le l_n$. In the quasi-neutral regions the field is negligible; in the order of approximation used here one lets $E = 0$ in such regions, whence $E(-l_p) = E(l_n) = 0$ and, from (21.38),

$$E_0 = -\frac{q\,N_D}{\varepsilon_{sc}}\,l_n = -\frac{q\,N_A}{\varepsilon_{sc}}\,l_p < 0\,, \qquad N_D\,l_n = N_A\,l_p\,. \tag{21.39}$$

In conclusion, the electric field is a piecewise-linear function whose form is shown in Fig. 21.11. Due to $\mathrm{d}\varphi/\mathrm{d}x = -E$, the integral of $-E$ over the space-charge region equals the potential drop ψ:

$$\psi = \varphi(l_n) - \varphi(-l_p) = -\int_{-l_p}^{+l_n} E\;\mathrm{d}x = -\frac{1}{2}\,E_0\,(l_n + l_p)\,. \tag{21.40}$$

Inserting into (21.40) one or the other form of E_0 from (21.39), one obtains two equivalent expressions for $l_p + l_n$:

$$l_n + l_p = l_n + \frac{N_D}{N_A}\,l_n = N_D\,l_n\left(\frac{1}{N_D} + \frac{1}{N_A}\right) = N_A\,l_p\left(\frac{1}{N_D} + \frac{1}{N_A}\right). \tag{21.41}$$

Then, combining (21.41) with (21.40) one finds

$$\psi = \frac{q}{2\,\varepsilon_{sc}}\left(\frac{1}{N_D} + \frac{1}{N_A}\right)(N_D l_n)^2 = \frac{q}{2\,\varepsilon_{sc}}\left(\frac{1}{N_D} + \frac{1}{N_A}\right)\left(N_A l_p\right)^2\,, \tag{21.42}$$

whence

$$l_n = \frac{1}{N_D}\left(\frac{2\,\varepsilon_{sc}\,\psi/q}{1/N_D + 1/N_A}\right)^{1/2}, \qquad l_p = \frac{1}{N_A}\left(\frac{2\,\varepsilon_{sc}\,\psi/q}{1/N_D + 1/N_A}\right)^{1/2}, \tag{21.43}$$

and

$$l = l_n + l_p = \left[\frac{2\,\varepsilon_{sc}}{q} \left(\frac{1}{N_D} + \frac{1}{N_A} \right) \psi \right]^{1/2} , \qquad \psi = \psi_0 - V \,. \tag{21.44}$$

Multiplying by $q\,A_e$ both sides of the second relation in (21.39) yields $q\,N_D\,A_e\,l_n = q\,N_A\,A_e\,l_p$, that represents the global charge conservation in the device. Such a conservation is implied by the assumption that $E = 0$ in the quasi-neutral regions, as is found by integrating $\mathrm{div}\mathbf{D} = \rho$ over the space-charge region. In the charge-conservation relation the widths l_p, l_n depend on V through (21.43); it follows that in the reverse-bias condition the device can be assimilated to a nonlinear capacitor where the charge per unit area of the two oppositely charged sides is, respectively, $Q_p = -q\,N_A\,l_p$ and $Q_n = q\,N_D\,l_n$. The differential capacitance per unit area is defined as $C = \mathrm{d}Q_p/\mathrm{d}V = -\mathrm{d}Q_n/\mathrm{d}V$; from the definition,[9] two equivalent expressions follow,

$$C = -q\,N_A \frac{\mathrm{d}l_p}{\mathrm{d}V} = q \frac{\mathrm{d}(N_A\,l_p)}{\mathrm{d}\psi} , \qquad C = -q\,N_D \frac{\mathrm{d}l_n}{\mathrm{d}V} = q \frac{\mathrm{d}(N_D\,l_n)}{\mathrm{d}\psi} . \tag{21.45}$$

Using (21.43), the differential capacitance per unit area of the abrupt *p-n* junction is found to be

$$C = \left[\frac{q\,\varepsilon_{sc}/(2\,\psi)}{1/N_D + 1/N_A} \right]^{1/2} = \frac{[(q\,\varepsilon_{sc}/2)/(1/N_D + 1/N_A)]^{1/2}}{[\psi_0\,(1 - V/\psi_0)]^{1/2}} , \tag{21.46}$$

which is given the more compact form[10]

$$C = C_0 \left(1 - \frac{V}{\psi_0} \right)^{-1/2} , \qquad C_0 = C(V = 0) = \left[\frac{q\,\varepsilon_{sc}/(2\,\psi_0)}{1/N_D + 1/N_A} \right]^{1/2} . \tag{21.47}$$

Combining (21.47) with (21.44) one derives the interesting relation

$$\frac{1}{C^2} = \frac{2}{q\,\varepsilon_{sc}} \left(\frac{1}{N_D} + \frac{1}{N_A} \right) \psi = \frac{l^2}{\varepsilon_{sc}^2} , \qquad C = \frac{\varepsilon_{sc}}{l} , \tag{21.48}$$

namely, the standard expression for the capacitance per unit area of the parallel-plate capacitor. Such an expression is not limited to the case where the dopant

[9] Definition $C = \mathrm{d}Q_p/\mathrm{d}V$ is coherent with the choice of the reference in Fig. 21.7. The units of C are $[C] = \mathrm{F\,cm^{-2}}$. Compare with the calculation of the MOS capacitance in Sect. 22.3.

[10] It is worth reminding that the result holds only in the reverse-bias condition. In the forward-bias condition the injection of carriers from the quasi-neutral regions into the space-charge region prevents one from neglecting the contribution of the carrier concentrations to the charge density and makes the use of (21.47) erroneous.

concentration is piecewise constant; as shown in Sect. 21.9.3, it applies in fact to all cases.

From the standpoint of circuit design, the capacitance associated with a *p-n* junction is a parasitic effect that hampers the circuit's speed. However, the effect is also exploited to manufacture voltage-controlled capacitors, called *variable capacitors* or *varactors*. In these devices, that are operated in reverse bias, the geometry is designed to maximize the capacitance; they are used, e.g., in voltage-controlled oscillators, parametric amplifiers, and frequency modulation.[11] The bias range of these devices is such that the reverse current is negligibly small; if the modulus of the reverse bias is made to increase and is eventually brought outside this range, carrier multiplication due to impact ionization (Sect. 20.3) takes place; this, as shown in Sect. 21.5, leads to a strong increase of the reverse current.

21.5 Avalanche Due to Impact Ionization

The situation where impact ionization dominates over the other generation-recombination mechanisms has been illustrated in Sect. 20.3.1, showing that in the steady-state case the continuity equations for electrons and holes reduce to (20.59). Such equations are applicable, for instance, to the space-charge region of a reverse-biased *p-n* junction; when the value of $|V|$ becomes large, the increase in the number of carriers due to impact ionization may give rise to an avalanche phenomenon that eventually leads to the junction's *avalanche breakdown*, namely, a strong increase in the current due to carrier multiplication.[12] The absolute value V_B of the voltage at which the phenomenon occurs is called *breakdown voltage*. To illustrate avalanche, the one-dimensional case is considered, so that (20.59) become

$$\frac{\mathrm{d}J_n}{\mathrm{d}x} = k_n J_n + k_p J_p \,, \qquad \frac{\mathrm{d}J_p}{\mathrm{d}x} = -k_n J_n - k_p J_p \,, \tag{21.49}$$

where the impact-ionization coefficients k_n, k_p depend on x through the electric field E and are determined experimentally.[13] To ease the notation the boundaries of the space-charge region are indicated with a, b; also, considering the reference's orientation (Fig. 21.7), it is $E, J_n, J_p < 0$, where the electric field is significant

[11] Varactors are also manufactured using technologies other than the bipolar one; e.g., with MOS capacitors or metal-semiconductor junctions.

[12] If the breakdown is accompanied by current crowding, the junction may be destroyed due to excessive heating. Special *p-n* junctions, called *avalanche diodes*, are designed to have breakdown uniformly spread over the surface of the metallurgical junction, to avoid current crowding. Such devices are able to indefinitely sustain the breakdown condition; they are used as voltage reference and for protecting electronic circuits against excessively high voltages.

[13] An example of model for k_n, k_p is that proposed by Chynoweth [23]: $k_n = k_{ns} \exp(-|E_{cn}/E|^{\beta_n})$, $k_p = k_{ps} \exp(-|E_{cp}/E|^{\beta_p})$, where the parameters depend on temperature [105, 138].

only for $a \leq x \leq b$. As in the one-dimensional and steady-state case it is $J = J_n(x) + J_p(x) = \text{const}$, eliminating J_p from the first equation in (21.49) yields

$$\frac{\mathrm{d}J_n}{\mathrm{d}x} = k_n J_n + k_p (J - J_n) = (k_n - k_p) J_n + k_p J \,, \tag{21.50}$$

namely, a first-order equation in J_n containing the yet undetermined parameter J. The equation is recast as

$$\frac{\mathrm{d}J_n}{\mathrm{d}x} - \frac{\mathrm{d}m}{\mathrm{d}x} J_n = k_p J = k_n J - \frac{\mathrm{d}m}{\mathrm{d}x} J \,, \qquad m = \int_a^x (k_n - k_p)\, \mathrm{d}x' \tag{21.51}$$

where $m(a) = 0$, $\mathrm{d}m/\mathrm{d}x = k_n - k_p$. Multiplying by the integrating factor $\exp(-m)$ and dividing by J transform (21.51) into

$$\frac{1}{J} \frac{\mathrm{d}}{\mathrm{d}x} [J_n \exp(-m)] = k_n \exp(-m) - \frac{\mathrm{d}m}{\mathrm{d}x} \exp(-m) \,, \tag{21.52}$$

where $-(\mathrm{d}m/\mathrm{d}x) \exp(-m) = \mathrm{d}\exp(-m)/\mathrm{d}x$. Integrating (21.52) from a to b and using $m(a) = 0$ yield

$$\frac{J_n(b)}{J} \exp[-m(b)] - \frac{J_n(a)}{J} = Y_n + \exp[-m(b)] - 1 \,, \tag{21.53}$$

where the *electron-ionization integral* is defined as

$$Y_n = \int_a^b k_n \exp(-m)\, \mathrm{d}x \,. \tag{21.54}$$

The above result is somewhat simplified by observing that due to impact ionization, the concentration of electrons in the conduction band increases from a to b, whereas that of holes increases from b to a; the concept is rendered in Fig. 21.12: consider a portion $x_1 < x < x_2$ of the space-charge region, where it is assumed for simplicity that the electric potential is linear. As the electric field is oriented toward the left, electrons (indicated by the black dots) are accelerated toward the right end of the space-charge region, holes (the white dots) are accelerated toward the left end. The vertical lines indicate the exchange of energies involved in impact-ionization events initiated by electrons or holes. An electron transited from the valence to the conduction band is accelerated by the field and may acquire a kinetic energy sufficient for initiating an impact-ionization event itself; the same applies to holes. As a consequence, the number of conduction-band electrons is multiplied from left to right, namely, the number of those exiting at x_2 is larger than the number entering at x_1; similarly, the number of valence-band holes is multiplied from right to left. Due to the multiplication mechanism, taking $x = b$ by way of example, the major contribution to $J = J_n(b) + J_p(b)$ is given by $J_n(b)$. Then, letting $J \simeq J_n(b)$ in (21.53) and canceling out some terms provide

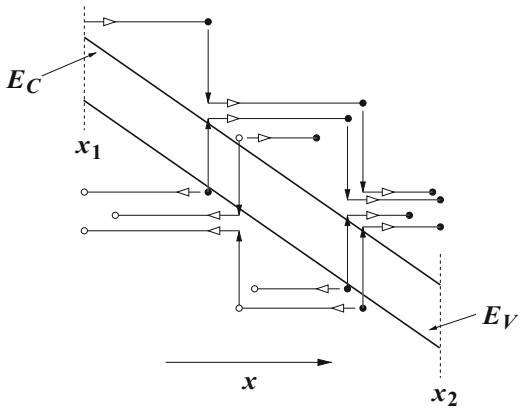


Fig. 21.12 Schematic description of the avalanche phenomenon. The details are given in the text

$$1 - \frac{1}{M_n} = Y_n\,, \qquad M_n = \frac{J_n(b)}{J_n(a)} \geq 1\,, \tag{21.55}$$

where M_n is the *electron-multiplication factor*. The latter is a measure of the impact ionization's level. As long as $Y_n < 1$, corresponding to a finite value of M_n, the avalanche condition does not occur; when $Y_n \to 1$, then $M_n \to \infty$: in this case, the injection of a negligibly small number of electrons at a produces a large electron current density at b. The operating conditions where M_n is large must be avoided because an excessive current may damage the device. Note that in the design stage of the device one first calculates Y_n from (21.54), then obtains M_n from (21.55). Thus, it may well happen that $Y_n > 1$, corresponding to $M_n < 0$; this outcome is not physically sound and simply indicates that the parameters used in the calculation of the ionization integral (21.54) are not consistent.[14]

The analysis of the impact-ionization condition can also be carried out starting with the elimination of J_n from the second equation in (21.49). The equation corresponding to (21.51) reads

$$\frac{\mathrm{d}J_p}{\mathrm{d}x} = -k_p\,J_p - k_n\,(J - J_p) = \frac{\mathrm{d}m}{\mathrm{d}x}\,J_p - k_n\,J\,; \tag{21.56}$$

in turn, the equation corresponding to (21.53) is

$$\frac{J_p(b)}{J}\,\exp[-m(b)] - \frac{J_p(a)}{J} = -\frac{Y_p}{\exp[m(b)]} + \exp[-m(b)] - 1\,, \tag{21.57}$$

with the *hole-ionization integral* given by

$$Y_p = \int_a^b k_p\,\exp[m(b) - m]\,\mathrm{d}x\,. \tag{21.58}$$

[14] This happens, for instance, if a value of $|V|$ larger than the breakdown voltage is used in (21.54).

Letting $J \simeq J_p(a)$ in (21.57) yields

$$1 - \frac{1}{M_p} = Y_p \,, \qquad M_p = \frac{J_p(a)}{J_p(b)} \geq 1 \,, \tag{21.59}$$

where M_p is the *hole-multiplication factor*. Using the definition (21.51) of m, the relation between the ionization integrals is found to be

$$Y_n = \exp[-m(b)]\, Y_p + 1 - \exp[-m(b)] \,. \tag{21.60}$$

The above shows that $Y_p = 1$ corresponds to $Y_n = 1$, namely, the avalanche condition $Y_p = 1$ for the holes coincides with that of the electrons, as should be.

21.6 Photodiode

The analysis of the *p-n* junction, depicted in Sects. 21.2 and 21.3, is readily extended to incorporate the description of a radiation detector, the *photodiode*. More specifically, the device described in this section is the *depletion-layer photodiode*, that consists of a reverse-biased *p-n* junction subjected to an impinging radiation. In the reverse-bias condition, the space-charge region is practically depleted from the mobile carriers; it follows that in such a region the radiation-induced generations prevail over the recombinations. The carriers thus generated modulate the device current: in this way, as outlined in the discussion of Sect. 20.4, the impinging radiation is exploited to provide an electric signal whose amplitude depends on the number of absorbed photons. Typical applications aim at detecting visible light; in this case the device is also termed *optical sensor*.[15]

In the description that follows, a one-dimensional, abrupt junction is considered, whose structure is analyzed with the help of Fig. 21.13; in the latter, the impinging photons are sketched by the wavy arrows, while the black dot indicates an electron and the white dot a hole. The following considerations are also worth attention: first, the approximation of considering the structure as one dimensional makes the upper contact to cover the semiconductor area completely; in real devices, the upper contact covers only a fraction of the sensor's area and, moreover, the area free from the contact is covered with an *anti-reflection coating* to minimize the loss of photons;[16] in the simplified structure considered here it is assumed that

[15] A reverse-biased *p-n* junction can also be used for detecting photons in a broader range of frequencies than the visible one, including x-rays or γ-rays, or high-energy particles, like electrons, α-particles, or neutrons. The mechanism is in all cases the creation of electron-hole pairs without significantly altering the junction's functioning.

[16] Anti-reflection coatings are structures made of the superposition of layers having different refraction indices. The thicknesses of the layers are such that a destructive interference is produced in the rays reflected at the interfaces, while a constructive interference is produced in the

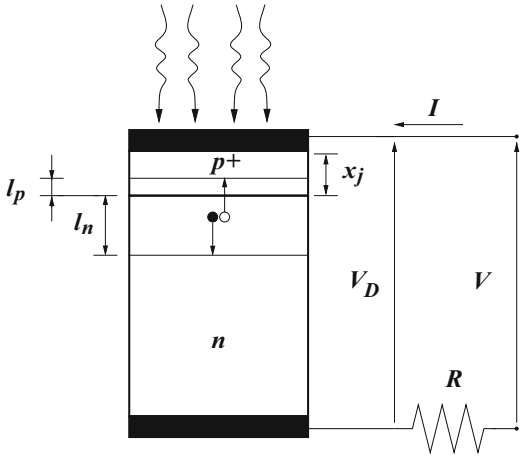


Fig. 21.13 Continuous mode photodiode. The applied bias V is negative and constant; the resistor mimics the input resistance of the circuit that measures the photodiode's current

the contact is partially transparent. Next, the radiation-induced generations occur in both the space-charge and quasi-neutral regions.[17] The electron-hole pairs generated in the space-charge region are separated by the electric field and move in opposite directions, as graphically sketched in Fig. 21.13. In contrast, the electrons and holes generated in the quasi-neutral regions, where the electric field is small, have a relatively high chance of recombining.[18] It follows that most of the electron-hole pairs generated within the quasi-neutral regions are useless as far as the functioning of the photodiode is concerned. This disadvantage is particularly severe in the quasi-neutral region close to the upper contact, which in the example of Fig. 21.13 is the p^+ region: in fact, remembering from the analysis of Sect. 20.4 that the optical generation decays exponentially along the propagation direction (x in this case), the number of generated pairs is maximum near the upper contact. To limit the loss it is then necessary to make the depth of the upper neutral region as small as possible, namely, the junction's width x_j must be minimized (*shallow junction*). At the same time, it is necessary to make the space-charge region's depth as large as possible, so that the majority of the pairs are generated in it; for this reason, the junction is strongly asymmetric, specifically, the dopant concentration of the upper region is much larger than that of the lower region. In the case considered here it is $N_A \gg N_D$ whence, remembering the second relation in (21.39), it follows $l_p \ll l_n$.

transmitted rays. The coating's design is optimized in order to obtain maximum transmission in a specific range of wavelengths and incidence angles.

[17] As shown in detail in Sect. 21.6.1, the photodiode is analyzed under the same hypotheses as those used for the *p-n* junction, among which is the partitioning of the structure into space-charge and quasi-neutral regions (Sect. 21.2.2).

[18] An exception is made for the pairs generated in the quasi-neutral regions near the edges of the space-charge region. The pairs move due to diffusion; if they succeed in reaching the edge before recombining, they are separated by the electric field and contribute to the radiation-generated current.

The minimization of x_j brings about the problem, already noted in Sect. 21.2.1, that the asymptotic boundary condition typical of the *p-n* junction's theory does not hold when the distance between the contact and the edge of the space-charge region is too short; a slightly more involved theory must be applied, whose outcome is that the expression of the saturation current density J_s differs from that given by (21.33). The derivation of the new expression is identical for the photodiode and the solar cell, and will therefore be postponed to the analysis of the latter (Sect. 21.7). As a final remark, in the derivation of the photodiode's characteristic it is acceptable to assume that the absorption events are uncorrelated from each other, so that one can limit the analysis to a monochromatic radiation. The more general approach in which one recovers the effect of the whole spectrum by adding up over the frequencies is shown in Sect. 21.7.

21.6.1 Photodiode—Continuous Mode

The *continuous mode* of operation of the photodiode corresponds to the case where the applied bias $V < 0$ is constant in time. Therefore, one uses the one-dimensional, steady-state form of the continuity equations for the electrons and holes. The analysis follows the same pattern as for the *p-n* junction (Sect. 21.3), the difference being that in addition to thermal recombination, one must consider optical generation as well. The continuity equations (21.22) then become

$$\frac{\mathrm{d}J_p}{\mathrm{d}x} = q\,G_O - q\,U_{\mathrm{SRH}}\,, \qquad \frac{\mathrm{d}J_n}{\mathrm{d}x} = -q\,G_O + q\,U_{\mathrm{SRH}}\,, \tag{21.61}$$

with G_O given by (20.67). Using the recombination current density J_U found in (21.23), and defining the *optical current density*

$$J_G = \int_{-l_p}^{l_n} q\,G_O \;\mathrm{d}x\,, \tag{21.62}$$

one integrates over the space-charge region one of the equations in (21.61), say, the first one, to find

$$J_p(l_n) - J_p(-l_p) = J_G - J_U\,. \tag{21.63}$$

Combining (21.63) with the expression of the total current density at $-l_p$, namely, $J = J_p(-l_p) + J_n(-l_p)$, yields

$$J = J_p(l_n) + J_n(-l_p) + J_U - J_G\,, \tag{21.64}$$

to be compared with (21.24). Like in Sect. 21.3, the rest of the calculation is based upon the weak-injection approximation, which allows one to decouple the continuity equations for the minority carriers in the quasi-neutral region and to assume that, still in such regions, the diffusion term in the transport equation for the minority carriers dominates over the drift term. Remembering that a reverse bias is applied to the photodiode, the recombination current is given by the same expression as in (21.35), namely, $I_U = -q A_e n_i l/\tau_g$, with $l = l(V)$ the width of the space-charge region and A_e the cross-sectional area of the device. In turn, the optical current $I_G = A_e J_G$ is found by inserting (20.67) into (21.62); assuming for simplicity that the absorption coefficient k is independent of position, one finds

$$I_G = A_e\, q\, \eta\, \Phi_B\, k \int_{-l_p}^{l_n} \exp(-k\,x)\ \mathrm{d}x = A_e\, q\, \eta\, \Phi_B\ \exp(k\, l_p)\ [1 - \exp(-k\, l)]\ , \tag{21.65}$$

where k depends on the photons' wavelength λ. In (21.65), $0 < \eta < 1$ is the quantum efficiency (compare with (20.63)), and $\Phi_B = \Phi(x = 0^+)$ is the photons' flux density on the inside edge of the interface between the top contact and the semiconductor. The expression of I_G simplifies when specific materials or applications are considered; taking for instance the case of silicon used for visible-light detection, $1/k$ ranges monotonically between $1/k \simeq 0.25\ \mu$m at $\lambda = 0.4\ \mu$m and $1/k \simeq 5\ \mu$m at $\lambda = 0.7\ \mu$m, with $1/k \simeq 1\ \mu$m at $\lambda = 0.5\ \mu$m [128, Sect. 12-4]. In turn, the order of magnitude of l can be determined with the aid of the expressions shown in Sect. 21.9.4. From the analysis one finds that the dopant concentrations and the applied voltage can be adapted to obtain $k\, l \gg 1$ and $k\, l_p \ll 1$, so that $I_G \simeq A_e\, q\, \eta\, \Phi_B$. This brings about the advantage that I_G becomes independent of the applied voltage.

The calculation of the term $J_p(l_n) + J_n(-l_p)$ in (21.64) follows the same pattern as for the solar cell and is postponed to Sect. 21.7. It leads to a constant saturation current whose expression, as anticipated above, differs from that given by (21.33) due to a modified boundary condition. Indicating the saturation current with I_s' and introducing the other simplifications discussed above provide the expression

$$I \simeq -q A_e \left(\eta\, \Phi_B + l\, \frac{n_i}{\tau_g}\right) - I_s' \tag{21.66}$$

for the current of a photodiode operating in the continuous mode. As the calculation of the current has been carried out starting from the steady-state form of the semiconductor equation, the result is applicable in principle only to the situations where the flux density of photons, Φ_B, is independent of time. However, the result may still be used in time-dependent cases, provided the period of the time variation of Φ_B is much larger than the other relevant time constants (the minority-carrier lifetimes in this case).

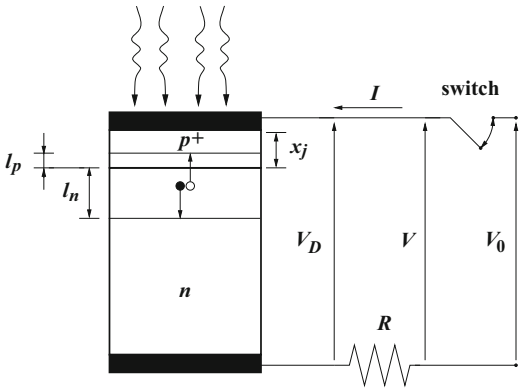


Fig. 21.14 Storage mode photodiode. The applied bias V_0 is negative and constant, while V, V_D vary with time depending on the action of the switch; the resistor mimics the input resistance of the circuit that measures the photodiode's current

21.6.2 Photodiode—Storage Mode

The current of a photodiode functioning in the continuous mode, whose simplified expression is (21.66), may be rather small under the typical illumination conditions; for this reason, the noise introduced by the external circuit (such a circuit is mimicked by resistor R in Fig. 21.13) may bring in an unacceptable degradation of the electric signal obtained from the device. To increase the amplitude of the output signal one may renounce the continuous mode operation, and sample the signal only at specific instants. This solution leads to the concept of *storage-mode* photodiode, which can be accomplished by introducing a suitable modification in the external circuit.

In the storage-mode operation of the photodiode, a constant voltage $V_0 < 0$ is applied (Fig. 21.14), while the switch is alternately opened and closed. When the switch is open, no current can flow in the circuit; thus, the effect of the generation-recombination processes is to modify the charge stored in the photodiode. This, in turn, changes the voltage drop V_D across the photodiode according to the relation $\mathrm{d}Q = C\,\mathrm{d}V_D$, where Q is the charge per unit area and $C = C(V_D)$ the differential capacitance per unit area. The latter is obtained from (21.47) in the case of an abrupt junction, and from (21.126) in the case of a diffused junction; here a general expression will be used, namely,

$$C = \frac{C_0}{(1 - V_D/\psi_0)^{1/m}}\,, \qquad m = 2, 3\,, \tag{21.67}$$

where the constant C_0 must be consistent with the choice of m. The expression of $\mathrm{d}Q$ is more easily derived with reference to the equivalent circuit shown in Fig. 21.15, where $I_G = q\,A_e\,\eta\,\Phi_B$ indicates the optical-generation current.[19]

[19]For the use of a current generator in the equivalent circuit, refer to the analysis of the solar cell carried out in Sect. 21.7.

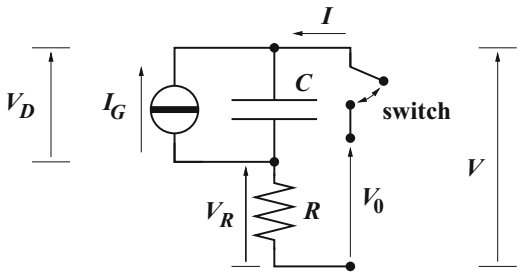


Fig. 21.15 Equivalent circuit used to analyze the functioning of the storage-mode photodiode

It is assumed that I_G is independent of time, and much larger than the other two summands at the right-hand side of (21.66); therefore, only I_G is considered in the equivalent circuit. When the switch is kept closed and possible earlier transients have vanished, it is $I = -I_G < 0$, so that the voltage drop across the photodiode is $V_D = V_0 + R\,I_G$; the choice of V_0 is such that $V_0 + R\,I_G < 0$. If the switch is opened at, say, $t = 0$, the external current I vanishes while I_G starts modifying the charge stored in the capacitor according to the relation $A_e\,\mathrm{d}Q = I_G\,\mathrm{d}t$. Combining the latter with $\mathrm{d}Q = C\,\mathrm{d}V_D$ and using (21.67) yield

$$\frac{\mathrm{d}t}{\tau_0} = \frac{\mathrm{d}(V_D/\psi_0)}{(1 - V_D/\psi_0)^{1/m}}\,, \qquad \tau_0 = \frac{A_e\,C_0\,\psi_0}{I_G}\,, \tag{21.68}$$

whose solution provides the time evolution of V_D when the switch is open. This condition is kept until a suitable time T_i, called *integration time*, has elapsed; then, the switch is closed again. To find the voltage drop across the photodiode right before the closing of the switch, one solves (21.68) and calculates the solution at $t = T_i$; this yields

$$\left(1 - \frac{1}{m}\right)\frac{T_i}{\tau_0} = \left(1 - \frac{V_0 + R\,I_G}{\psi_0}\right)^{1-1/m} - \left(1 - \frac{V_i}{\psi_0}\right)^{1-1/m}\,, \tag{21.69}$$

with $V_i = V_D(T_i)$ and $V_0 + R\,I_G = V_D(0)$. The left-hand side of (21.69) and I_G are positive, whence $V_i > V_0$. Also, T_i is selected in such a way as to keep V_i negative, whence $0 > V_i > V_0$. Note that during the integration time it is $V_R = 0$. The closing of the switch restores the voltage V_0 of the upper electrode of the photodiode, forcing a voltage variation, from V_i to V_0, virtually in zero time. As the capacitor does not allow for a time discontinuity of V_D, the same variation is forced in the lower electrode of the photodiode; as a consequence V_R changes, virtually in zero time, from 0 to $V_0 - V_i < 0$, making a current peak to flow in the load resistor. The ratio between the current peak and the optical-generation current,

$$\Gamma = \frac{(V_i - V_0)/R}{I_G} \tag{21.70}$$

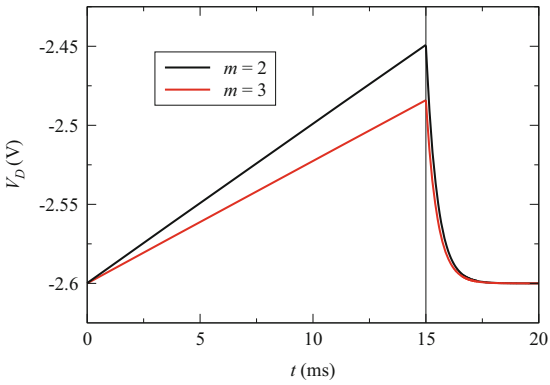


Fig. 21.16 Time evolution of the voltage V_D across a storage-mode photodiode. The integration time T_i ranges from $t = 0$ to $t = 15$ ms (*vertical line*); the sampling time T_s ranges from $t = 15$ ms to $t = 20$ ms. The calculations are shown in Probs. 21.10 and 21.11

is also called *current gain* of the storage-mode operation. In typical operating conditions the current peak is much larger than I_G, thus alleviating the noise problems discussed above (an example is in Prob. 21.7).

The switch is kept closed for a time interval T_s, called *sampling time*, long enough to restore the condition $V_D = V_0 + R I_G$. The differential equation describing the circuit during the sampling time is different from (21.68) because here it is $I \neq 0$. From $(I + I_G)\,\mathrm{d}t = A_e\,\mathrm{d}Q = A_e\,C(V_D)\,\mathrm{d}V_D$ and $R\,I = V_0 - V_D$ one finds

$$\frac{\mathrm{d}t}{R A_e\,C_0} = \frac{\mathrm{d}V_D}{(V_0 + R I_G - V_D)\,(1 - V_D/\psi_0)^{1/m}}\,, \tag{21.71}$$

whose initial condition is $V_D = V_i$. The solution of (21.71) for $m = 2$ and $m = 3$ is in Probs. 21.8 and 21.9, respectively. The time evolution of V_D in the integration time and in the sampling time is shown in Fig. 21.16 for $m = 2$ and $m = 3$. Observing that during the sampling time it is $V_R = V_0 - V_D$ and $I = V_R/R$, the time evolution of V_R and I within the sampling time is the same as that of V_D apart from scaling factors and/or additive constants; within the integration time, instead, it is $V_R = 0$, $I = 0$ as remarked above.

Typical applications of storage-mode devices are in the field of imaging, where individual sensors are organized into arrays that provide a signal to be displayed on a screen. The temporal sensitivity and resolution of human vision are such that an image is perceived as stable when the frame rate is higher than about 50 Hz.[20] Above such a limit, no flicker is perceived by the average observer despite the fact that static images are presented one after the other. The *interlaced technique* can effectively be used to double the perceived frame rate; in this technique, each frame is made of two fields, one of which contains the odd-numbered lines of the image, the other one contains the even-numbered lines. When the interlaced technique

[20]The *frame rate* is the frequency at which an imaging device (e.g., a computer display, a film, or a video camera) displays consecutive images.

is used, the frame rate can in fact be reduced to about 25 Hz, corresponding to 40 ms per frame. The estimates carried out in Probs. 21.8 and 21.9 show that the photodiode's performance complies with these specifications.

21.7 Solar Cell

In contrast to the photodiode, whose purpose is that of sensing the radiation, the solar cell exploits the radiation-induced generations to convert the radiation's energy into electric energy. Such a conversion is also termed *photovoltaic (PV) effect*.

The schematic cross-section of a solar cell is shown in Fig. 21.17; observing that there is no external bias, one cannot prescribe the sign of V *a priori*: such an information must be obtained by solving the full p-n junction's model first, and deducing the value of V and I afterwards. Thus, besides considering optical generation in addition to thermal recombination, the general approach of Sect. 21.3 will be followed. As remarked in Sect. 21.6, the whole spectrum of the radiation must be used; to this purpose, one takes the simplified form of the optical-generation term (20.67) in which, like in Sect. 21.6.1, it is assumed that the absorption coefficient k is independent of position:

$$G_O(\nu) = g_O \exp(-k\,x)\,, \qquad g_O(\nu) = \eta\,\Phi_B\,k\,. \tag{21.72}$$

The zero of x is made to coincide with the inside edge of the interface between the top contact and the semiconductor. Considering a discrete spectrum of the radiation, the symbols G_O, g_O are replaced with G_i, g_i, respectively, where index i refers to the ith monochromatic component. As noted in Sect. 21.6.2, it is acceptable to assume that the absorption events are uncorrelated from each other, so that the generations due to the whole spectrum are described by $\sum_i G_i$. With these premises,

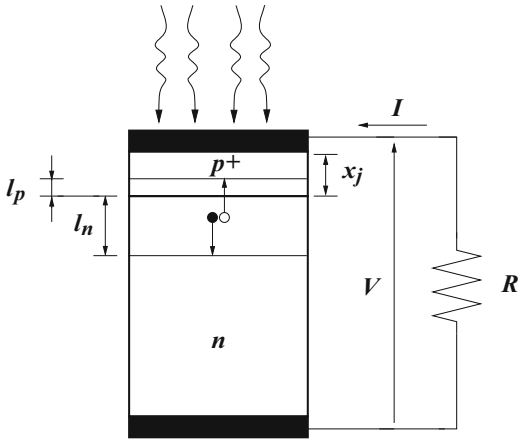


Fig. 21.17 Solar cell. The resistor mimics the input resistance of the load to which the energy is delivered

the continuity and transport equations for the minority carriers in the p-type, quasi-neutral region read

$$\frac{1}{q}\frac{\mathrm{d}J_n}{\mathrm{d}x} \simeq \frac{n-n_{p0}}{\tau_n} - \sum_i G_i\,, \qquad \frac{1}{q}J_n \simeq D_n\frac{\mathrm{d}n}{\mathrm{d}x} = D_n\frac{\mathrm{d}(n-n_{p0})}{\mathrm{d}x}\,, \tag{21.73}$$

where the weak-injection expression $U_{\mathrm{SRH}} \simeq (n-n_{p0})/\tau_n$ has been used for the thermal-recombination term.

21.7.1 Current of the Solar Cell

The equation to be solved in the quasi-neutral region of the solar cell, adjacent to the top contact, is found by combining the two equations in (21.73):

$$\frac{\mathrm{d}^2(n-n_{p0})}{\mathrm{d}x^2} = \frac{n-n_{p0}}{L_n^2} - \sum_i \frac{g_i}{D_n}\exp(-k_i\,x)\,, \qquad L_n^2 = \tau_n\,D_n\,, \tag{21.74}$$

whose solution is sought as the sum of the solution of the homogenous equation and of a tentative solution of the nonhomogeneous one:

$$n-n_{p0} = A_n\,\exp(x/L_n) + B_n\,\exp(-x/L_n) + \sum_i C_i\,\exp(-k_i\,x)\,, \tag{21.75}$$

with A_n, B_n, C_i undetermined coefficients. Replacing (21.75) into (21.74) provides, after eliminating the terms deriving from the solution of the homogeneous equation, the relation

$$\sum_i \left[\left(\frac{1}{L_n^2} - k_i^2\right) C_i - \frac{g_i}{D_n}\right]\exp(-k_i\,x) = 0\,, \tag{21.76}$$

which must be fulfilled for any x. Observing that the exponentials are linearly independent from each other, the left-hand side of (21.76) vanishes only if each term in brackets vanishes. It could happen that $k_i = 1/L_n$ for some value of i; in this case, however, the corresponding summand $C_i\,\exp(-k_i\,x) = C_i\,\exp(-x/L_n)$ would be incorporated into the second term at the right-hand side of (21.75), thus becoming a part of the homogeneous solution. In conclusion, when treating the nonhomogeneous part of (21.75), one may always assume that $k_i \neq 1/L_n$; this eventually yields the expression of the coefficients C_i,

$$C_i = \frac{\tau_n\,g_i}{1-k_i^2\,L_n^2}\,. \tag{21.77}$$

The remaining two constants, A_n and B_n, are found as usual from the boundary conditions of the electron concentration in the quasi-neutral region. To this purpose one notes, from Fig. 21.17, that the position of the boundary between the quasi-neutral and space-charge regions is $s = x_j - l_p$ so that, using the Shockley boundary condition (Sect. 21.9.2), the following holds:

$$A_n \exp(s/L_n) + B_n \exp(-s/L_n) + \sum_i C_i \exp(-k_i s) = n_{p0} F , \tag{21.78}$$

with $F = \exp[q V/(k_B T)] - 1$. As for the boundary condition at $x = 0$, that is, at the interface between the top contact and the semiconductor, the minimization of x_j, necessary for the efficient operation of the device, prevents one from using the asymptotic conditions (21.1) typical of the p-n junction's theory; this had already been noted in Sects. 21.2.1 and 21.6.1. On the other hand, as discussed in Sect. 19.5.6, a contact is able to supply the amount of charge necessary to keep the equilibrium and charge-neutrality conditions in the semiconductor layer adjacent to it; as a consequence, here the boundary condition at $x = 0$ reads

$$A_n + B_n + \sum_i C_i = 0 . \tag{21.79}$$

Solving the algebraic system (21.78, 21.79) provides the two coefficients A_n, B_n. The solution is somewhat simplified if $s \ll L_n$ and $s \ll 1/k_i$; comparing with the estimates of the diffusion lengths given in (21.129), the first condition is easily fulfilled. As for the second one, remembering the figures given in Sect. 21.6.1, for very shallow junctions the constraint is fulfilled in the near-infrared and part of the visible range. Then, linearizing the exponentials reduces (21.78) to[21]

$$[(A_n - B_n)/L_n - \sum_i C_i k_i] s = n_{p0} F - (A_n + B_n + \sum_i C_i) , \tag{21.80}$$

where the term in parentheses at the right-hand side vanishes due to (21.79). The ratio $J_n(s)/(q D_n)$ is now determined by combining (21.75) with the second relation in (21.73) and letting $x = s$; using (21.80) one obtains

$$\frac{J_n(s)}{q D_n} = \frac{n_{p0} F}{s} + \left(\frac{A_n + B_n}{L_n^2} + \sum_i C_i k_i^2 \right) s = \frac{n_{p0} F}{s} . \tag{21.81}$$

It is easily found from (21.73), (21.74), and (21.75) that the term in parentheses of (21.81) equals the derivative of $J_n/(q D_n)$ calculated at $x = 0$; such a derivative vanishes due to the equilibrium condition imposed by the contact (namely, $U_{\mathrm{SRH}} - \sum_i G_i = 0$ there).

[21] The solution without the linearization of the exponentials is left to Prob. 21.12.

The calculation of the current density on the other edge of the space-charge region, $x = x_j + l_n = s + l$, is made easier by the fact that, as discussed in Sect. 21.6, the depth l of the space-charge region is made as large as possible, so that the majority of the electron-hole pairs are generated there; for this reason, the optical-generation term can be neglected in the n-type, quasi-neutral region, to yield

$$\frac{J_p(s+l)}{q\,D_p} = \frac{p_{n0}\,F}{L_p}\,. \tag{21.82}$$

Finally, integrating the continuity equation $\mathrm{d}J_n/\mathrm{d}x = q\,(U_{\mathrm{SRH}} - \sum_i G_i)$ over the space-charge region yields

$$J_n(s+l) - J_n(s) = -J_G + J_U\,, \tag{21.83}$$

with J_U the recombination current density (21.23) and

$$J_G = q\sum_i \int_s^{s+l} g_i \exp(-k_i\,x)\,\mathrm{d}x = q\sum_i \frac{g_i}{k_i}\,\exp(-k_i\,s)\,[1-\exp(-k_i\,l)] > 0\,. \tag{21.84}$$

Using $k_i\,s \ll 1$, $k_i\,l \gg 1$ and remembering the definition of g_i (compare with the second relation in (21.72)) yield for the optically generated current density $J_G = q\sum_i \eta_i\,\Phi_i$. In conclusion, the total current density is found from $J = J_p(s+l) + J_n(s+l) = J_p(s+l) - J_G + J_U + J_n(s)$; replacing from (21.81), (21.82), and (21.83), and multiplying both sides by the cross-sectional area A_e yield for the solar cell's current

$$I = I_s'\,F + I_U - I_G\,, \qquad I_s' = q\,A_e\left(\frac{D_n\,n_{p0}}{s} + \frac{D_p\,p_{n0}}{L_p}\right), \tag{21.85}$$

with $I_U = A_e\,J_U$ and $I_G = A_e\,J_G$. As anticipated in Sect. 21.6, the expression of the saturation current I_s' differs from (21.33) due to the modified boundary condition of the shallow junction. The first two terms $I_s'\,F + I_U$ in the expression (21.85) of the solar cell's current have the standard form of the p-n junction's current; the extra term $-I_G$, due to the radiation, does not depend on the voltage drop across the device: therefore, it is represented by a current generator in the equivalent circuit of the cell, shown in Fig. 21.18. When the radiation is absent, the device is at equilibrium and $I_G = I = 0$, $V = 0$.

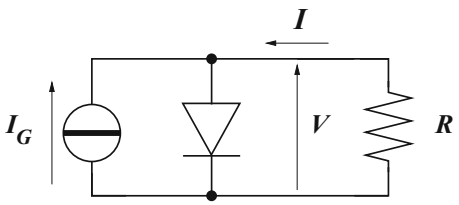


Fig. 21.18 Equivalent circuit of the solar cell

21.8 Bipolar Junction Transistor

The bipolar junction transistor[22] (BJT) was proposed in 1948 [9, 121]. The first prototype of the device, constructed in December 1947 using germanium, was a *point-contact transistor*, namely, a structure where the tip of a triangle-shaped block of material produces a point-contact diode. The structure was improved in 1948 by replacing the point contact with a *p-n* junction.[23] After being used for about two decades for fabricating BJTs, germanium was almost totally replaced by silicon in the mid-sixties of the last century (Sect. 18.7.1). The research that led to the invention of the transistor aimed at replacing vacuum tubes with smaller and less power-consuming devices; the initial applications to low-frequency amplifiers quickly extended to include high-power, high-frequency, and switching operations: thus, the introduction of the BJT had a formidable impact on the semiconductor industry and the related research activities. Until the introduction of the MOS transistor, the BJT has been the basic ingredient in the fabrication of the electronic circuits.

The electric current in a BJT is transported by both type of carriers, electrons and holes; for this reason the device is called *bipolar*. There exist two possible architectures of the devices, the *n-p-n* type and the *p-n-p* type; the schematic cross-section of an *n-p-n* BJT is shown in Fig. 21.19. The starting point is a *p*-type silicon substrate, with an $N_A = \text{const}$ dopant concentration.[24] At the beginning of the process, the upper surface of the semiconductor wafer is in the position marked by the dash-dotted arrows; the superimposed structure does not exist yet. In the first step of the fabrication, a limited area is defined by a suitable mask, and a heavy *n*-type thermal diffusion is carried out in it, to create a region called *buried layer* or *subcollector*; this region is shown as the dark area marked with n^+ in the figure. Purpose of the buried layer is to provide a low-resistance path between the active part of the transistor and the collector contact. Next, a lightly doped, *n*-type epitaxial layer is grown over the entire wafer; this layer, shown as the area marked with n^- in the figure, replicates the crystalline structure of the wafer underneath. As epitaxy

[22] The term is a contraction of *transfer resistor*.

[23] This explains the term "junction" in the name.

[24] The terminology of the process steps is better explained in the chapters devoted to the technological issues.

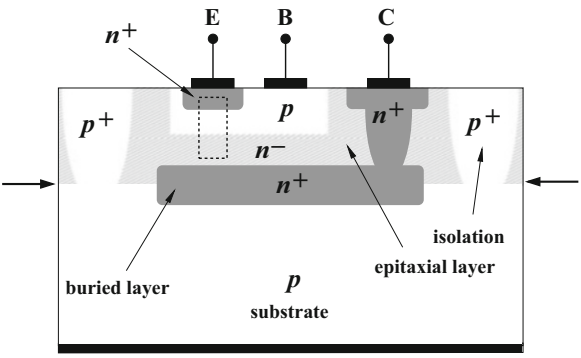


Fig. 21.19 Structure of the bipolar junction transistor of the *n-p-n* type

is a high-temperature process, during the epitaxial growth the buried layer diffuses in all directions; this explains why the upper edge of the buried layer is higher than the interface between the substrate and the epitaxial layer.[25]

As in the normal operation of the device one of its *p-n* junctions is forward biased, it might happen that a part of the junction's current flows into an adjacent device, thus creating a parasitic effect that hampers the circuit's performance. In this respect, BJTs are not *self-isolated*.[26] The problem is fixed by means of *isolation diffusions*, that consist of heavily doped *p* regions diffused across the epitaxial layer down to the substrate. In this way a volume of material belonging to the epitaxial layer, forming a well that will later become the *collector* region of the BJT, is completely surrounded by a *p*-type region and is therefore isolated from the adjacent devices.[27]

Next, the *p*-type *base* region is diffused within the collector region, and the *n*-type *emitter* region is diffused within the base region. In summary, two *p-n* junctions are formed, the base-emitter (B-E) junction and the base-collector (B-C) junction. For reasons dictated by the need of optimizing the device performance, the dopant concentration in the base must be larger than that of the collector, and the dopant concentration in the emitter must be larger than that of the base. Finally, the emitter (E), base (B), and collector (C) contacts are deposited. As indicated in Sect. 21.2.2, to prevent the contact-semiconductor junction from behaving like a rectifying device, a heavy dose of dopant is preliminarily introduced into the semiconductor region onto which the contact is to be deposited. Here, this is necessary for the collector contact, because the epitaxial layer is lightly doped; in general, besides introducing a large dopant concentration in the area near the contact, a heavily doped

[25]The details of the epitaxial process are given in Sects. 24.6 and 24.7; those of thermal diffusion are in Chap. 23.

[26]In contrast, MOS transistors are self-isolated, because their junctions are never forward biased (Sects. 22.7, 22.8). However, in some MOS architectures parasitic currents are possible, and precautions must be taken against them using, e.g., channel stops (Sect. 22.10).

[27]The *p-n* junction formed by the substrate and the epitaxial layer is kept reverse biased by the application of a suitable voltage between the substrate and collector contacts.

Table 21.1 BJT: nomenclature of the functioning regimes

B-E j.	B-C j.	Condition	Comment
F	F	Saturation	Avoided in analog applications
F	R	Forward	High gain
R	R	Cut-off	Negligible currents, practically an open circuit
R	F	Reverse	Poor gain, seldom used

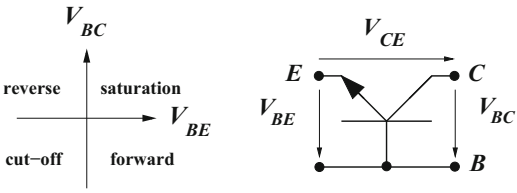


Fig. 21.20 Functioning regimes and symbol of the *n-p-n* type transistor

region (here of the *n* type) is diffused from the collector contact down to the buried layer; purpose of this is again lowering the series resistance.

To distinguish the applied voltages from one another, two letters are used; considering the base metallization as the reference contact, in an *n-p-n* BJT the typical choice of the two independent voltages is $V_{BE} = V_B - V_E$ and $V_{BC} = V_B - V_C$, with $V_{CE} = V_{BE} - V_{BC}$. The situation where the base contact is chosen as reference is called *common-base configuration*. It is clear from the description above that the BJT architecture is structurally asymmetric; this reflects into the different behaviors of the device, depending on the working conditions of the two junctions. The four combinations of the applied biases are termed as shown in Table 21.1, along with the combinations' names and some comments; the indications "F" and "R" stand for "forward bias" and "reverse bias," respectively.[28]

A diagram describing the functioning regimes of the *n-p-n* transistor is shown in Fig. 21.20 along with the device symbol. Note that the arrows associated with the voltages point toward the *p* regions, so that the junctions are forward biased when $V_{BE} > 0$ and $V_{BC} > 0$.

The structure of the BJT of the *p-n-p* type is dual to that of the *n-p-n* type; with reference to Fig. 21.19, the cross-section is obtained by exchanging the "*p*" and "*n*" letters. The voltages to be used in the common-base configuration are $V_{EB} = V_E - V_B$ and $V_{CB} = V_C - V_B$, with $V_{CE} = V_{CB} - V_{EB}$. A diagram describing the functioning regimes of the *p-n-p* transistor is shown in Fig. 21.21 along with the device symbol. Again, the arrows associated with the voltages point toward the *p* regions, so that the junctions are forward biased when $V_{EB} > 0$ and $V_{CB} > 0$.

[28]The nomenclature of Table 21.1 is fully general and is applicable to both the *n-p-n* and *p-n-p* transistors. The forward condition is also called *normal mode* or *active mode*.

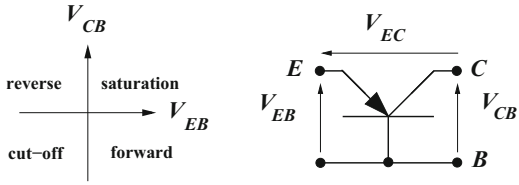


Fig. 21.21 Functioning regimes and symbol of the *p-n-p* type transistor

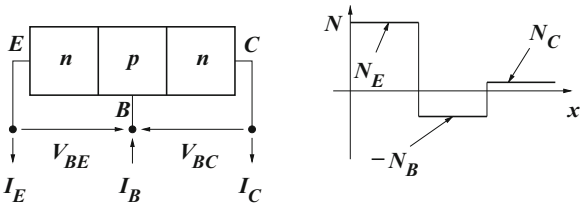


Fig. 21.22 Structure and effective dopant concentration in the intrinsic transistor of the *n-p-n* type

21.8.1 Current-Voltage Characteristics of the P-N-P *BJT*

The structure of the BJT, shown in Fig. 21.19, is rather complicate; for this reason, an analytical derivation of the current-voltage characteristics is not viable for the device as is, and some simplifications are necessary. The standard approach consists in deriving the characteristics of the *intrinsic transistor*, namely, the one-dimensional structure marked by the dashed contour visible in the figure; the dopant concentration is considered constant in the emitter, base, and collector regions: in this way, the structure to be analyzed is reduced to the one-dimensional juxtaposition of two abrupt *p-n* junctions.

The intrinsic transistor, rotated by 90 degrees for convenience, is shown in Fig. 21.22 for the *n-p-n* type, along with the diagram of the net concentration of ionized impurities, $N = N_D^+ - N_A^-$. In the following, the complete-ionization condition is assumed to hold; it follows that, indicating with N_E, N_C the constant donor concentrations of the emitter and collector regions, respectively, and with N_B the constant acceptor concentration of the base region, the diagram of N takes the piecewise-constant shape shown in the right part of the figure. The emitter (I_E), base (I_B), and collector (I_C) currents are also marked in the same figure; note that the arrows associated with the currents point toward the p region, while they point away from the n regions; thus, the relation $I_B = I_E + I_C$ holds. Remembering that V_{BE} and V_{BC} are being used as independent voltages, it is convenient to select I_E and I_C as the independent currents.

By comparing the intrinsic transistor of Fig. 21.22 with the transistor's symbol of Fig. 21.20 one notes that in the former, V_{BE} and V_{BC} are the voltage drops across the junctions of the intrinsic transistor while, in the latter, the same symbols (and V_{CE} as well) indicate the voltages between the external contacts. In principle, different symbols should be used, because the voltage drops of the intrinsic device do not account for other effects due to the rest of the structure; among such effects are the

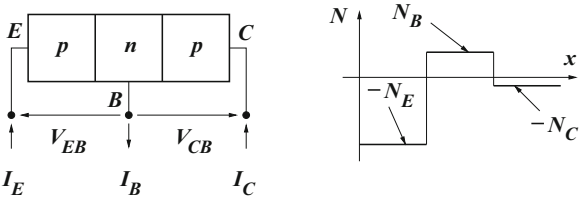


Fig. 21.23 Structure and effective dopant concentration in the intrinsic transistor of the *p-n-p* type

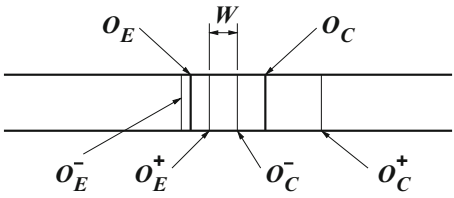


Fig. 21.24 *Symbols* used to indicate the positions of the two junctions, the edges of the space-charge regions, and the width of the quasi-neutral base region. The same nomenclature applies to both the *n-p-n* and *p-n-p* types

extra drop due to the collector current flowing through the buried layer, and that due to the base current that flows laterally (that is, normal to the x direction) across the base region. In the calculations carried out below, the same notation is used for simplicity; distinct symbols are to be used when the extra effects are accounted for.[29]

The intrinsic transistor and net impurity concentration of the *p-n-p* type are those shown in Fig. 21.23, where N_E, N_C are acceptor concentrations and N_B a donor concentration.

The hypotheses upon which the analytical theory of the intrinsic BJT is based are the same as those used for the *p-n* junction (Sect. 21.3). It is assumed that the working conditions are such that the device's partitioning into quasi-neutral and space-charge regions holds: as shown in Fig. 21.24, the position of the emitter junction is indicated with O_E, and the left and right edges of the space-charge region of the same junction are indicated with O_E^-, O_E^+, respectively; a similar nomenclature applies to the collector region. The device has two space-charge regions and three quasi-neutral regions; the length of the quasi-neutral base region is marked with W and is particularly important in determining the electric characteristics of the BJT. The analytical theory carried out below refers to the *p-n-p* type;[30] to distinguish the

[29] A typical way for considering the effect of the lateral flow is to introduce a series resistance (called *intrinsic base resistance*) between the base contact of the intrinsic device and the external base contact.

[30] The choice of the *p-n-p* type is dictated only by a convenience in notation, due to the fact that the current density of the minority carriers in the base (holes) has the same orientation as the x axis.

parameters of the different regions one introduces the symbols for the equilibrium concentrations of the minority carriers,[31]

$$n_E = \frac{n_i^2}{N_E}\,, \qquad p_B = \frac{n_i^2}{N_B}\,, \qquad n_C = \frac{n_i^2}{N_C}\,. \tag{21.86}$$

The diffusion coefficients of the minority carriers in the three quasi-neutral regions are indicated with D_E, D_B, and D_C, respectively; the corresponding lifetimes with τ_E, τ_B, and τ_C, and the diffusion lengths with

$$L_E = \sqrt{\tau_E\, D_E}\,, \qquad L_B = \sqrt{\tau_B\, D_B}\,, \qquad L_C = \sqrt{\tau_C\, D_C}\,. \tag{21.87}$$

The emitter current density can be written $J_E = J_n(O_E^-) + J_p(O_E^-)$ or, expressing the last term at the right-hand side as the sum of a minority-carrier current density and a recombination current density,

$$J_E = J_n(O_E^-) + J_p(O_E^+) + \int_{O_E^-}^{O_E^+} q\, U_{\mathrm{SRH}}\, \mathrm{d}x\,; \tag{21.88}$$

in the same manner, the collector current density can be written $J_C = J_n(O_C^+) + J_p(O_C^+)$, or

$$J_C = J_n(O_C^+) + J_p(O_C^-) - \int_{O_C^-}^{O_C^+} q\, U_{\mathrm{SRH}}\, \mathrm{d}x\,. \tag{21.89}$$

Considering the quasi-neutral emitter region first and following the same reasoning as for the *p*-*n* junction (Sect. 21.3.1) yield

$$J_n(O_E^-) = q\,\frac{D_E\, n_E}{L_E}\, F_E\,, \qquad F_E = \exp\left[q\, V_{EB}/(k_B\, T)\right] - 1\,, \tag{21.90}$$

to be compared with (21.30). Similarly, for the minority-carrier current density in the quasi-neutral collector region one finds[32]

$$J_n(O_C^+) = -q\,\frac{D_C\, n_C}{L_C}\, F_C\,, \qquad F_C = \exp\left[q\, V_{CB}/(k_B\, T)\right] - 1\,. \tag{21.91}$$

To complete the analysis one must determine the minority-carrier current densities at the edges of the quasi-neutral base region, $J_p(O_E^+)$ and $J_p(O_C^-)$ respectively. This

[31] Expressions (21.86) for the equilibrium concentrations imply the assumption of nondegeneracy and complete ionization (Sect. 21.2.1).

[32] As the quasi-neutral emitter region extends from O_C^+ to infinity, after solving $\mathrm{d}^2(n - n_C)/\mathrm{d}x^2 = (n - n_C)/L_C^2$ one must discard the solution $\exp(x/L_C)$ and keep the solution $\exp(-x/L_C)$. Then, the current density is found from $J_n = q\, D_C\, \mathrm{d}n/\mathrm{d}x$; this explains the minus sign in (21.91).

calculation is slightly more involved because the length W of the region is finite, hence the asymptotic conditions for the concentration do not apply. One starts from the differential equation for the minority carriers' concentration within the region, $\mathrm{d}^2(p-p_B)/\mathrm{d}x^2 = (p-p_B)/L_B^2$, whose solution reads

$$p(x) - p_B = A\,\exp(-x/L_B) + B\,\exp(x/L_B) \tag{21.92}$$

with A, B undetermined constants. To proceed, one places the origin of x at O_E^+, so that the position of O_C^- in this reference is W. The minority-carrier concentration at the edges of the quasi-neutral base region fulfills the Shockley boundary conditions (Sect. 21.9.2), that in this case read

$$p(0) = p_B\,\exp\left(\frac{q\,V_{EB}}{k_B\,T}\right), \qquad p(W) = p_B\,\exp\left(\frac{q\,V_{CB}}{k_B\,T}\right); \tag{21.93}$$

replacing (21.93) into (21.92) yields the algebraic system

$$p_B\,F_E = A + B\,, \qquad p_B\,F_C = A\,\exp(-W/L_B) + B\,\exp(W/L_B) \tag{21.94}$$

in the unknowns A, B. Once the two constants are determined, the current density is calculated from

$$J_p = -q\,D_B\,\frac{\mathrm{d}p}{\mathrm{d}x} = q\,\frac{D_B}{L_B}\,\left[A\,\exp(-x/L_B) - B\,\exp(x/L_B)\right]\,. \tag{21.95}$$

Constant A is determined by multiplying by $\exp(W/L_B)$ the first equation in (21.94) and subtracting the second equation from the result; similarly, constant B is determined by multiplying by $\exp(-W/L_B)$ the first equation in (21.94) and subtracting the second equation from the result. Letting $x = 0$ or, alternatively, $x = W$ in (21.95) and using the constants thus found yield the current densities sought:

$$J_p(O_E^+) = J_p(0) = q\,\frac{D_B\,p_B}{L_B}\,\left[F_E\,\coth\left(\frac{W}{L_B}\right) - \frac{F_C}{\sinh(W/L_B)}\right], \tag{21.96}$$

$$J_p(O_C^-) = J_p(W) = q\,\frac{D_B\,p_B}{L_B}\,\left[\frac{F_E}{\sinh(W/L_B)} - F_C\,\coth\left(\frac{W}{L_B}\right)\right]. \tag{21.97}$$

The expressions of the emitter and collector currents are now found by multiplying (21.88), (21.89) by the cross-sectional area A_e of the device and letting

$$I_{UE} = A_e\,\int_{O_E^-}^{O_E^+} q\,U_{\mathrm{SRH}}\,\mathrm{d}x\,, \qquad I_{UC} = A_e\,\int_{O_C^-}^{O_C^+} q\,U_{\mathrm{SRH}}\,\mathrm{d}x\,; \tag{21.98}$$

in turn, $J_n(O_E^-)$, $J_n(O_C^+)$, $J_p(O_E^+)$, and $J_p(O_C^-)$ are taken from (21.90), (21.91), (21.96), and (21.97), respectively. It must be remarked that the relation between

current and current density in the emitter reads $I_E = A_e\,[J_n(O_E^-) + J_p(O_E^-)]$, while that in the collector reads $I_C = -A_e\,[J_n(O_C^+) + J_p(O_C^+)]$; the minus sign in the latter is due to the fact that the orientation of I_C is opposite to that of the x axis (Fig. 21.23). To make the resulting expressions more compact one defines the coefficients

$$a_{11} = q\,A_e \left[\frac{D_E\,n_E}{L_E} + \frac{D_B\,p_B}{L_B}\,\coth\left(\frac{W}{L_B}\right)\right], \quad a_{12} = -q\,A_e\,\frac{D_B\,p_B/L_B}{\sinh(W/L_B)}, \tag{21.99}$$

$$a_{21} = a_{12}\,, \qquad a_{22} = q\,A_e \left[\frac{D_C\,n_C}{L_C} + \frac{D_B\,p_B}{L_B}\,\coth\left(\frac{W}{L_B}\right)\right]; \tag{21.100}$$

such coefficients, whose units are those of a current, form a 2×2 symmetric matrix in which the diagonal entries are positive while the non-diagonal entries are negative. The coefficients depend on the voltages V_{EB}, V_{CB} through the width W of the quasi-neutral base region; in fact,

$$W = (O_C - O_E) - l_{nE}(V_{EB}) - l_{nC}(V_{CB}) = W(V_{EB}, V_{CB})\,, \tag{21.101}$$

where $l_{nE} = O_E^+ - O_E$ ($l_{nC} = O_C - O_C^-$) is the length of the portion of the space-charge region of the emitter (collector) junction belonging to the base region. Thanks to (21.99), (21.100) the emitter and collector currents are written in compact form as

$$I_E = (a_{11}\,F_E + I_{UE}) + a_{12}\,F_C\,, \qquad I_C = a_{21}\,F_E + (a_{22}\,F_C + I_{UC})\,. \tag{21.102}$$

21.8.2 *Equivalent Circuit of the* P-N-P *BJT*

The right-hand sides of (21.102) depend on the voltages V_{EB}, V_{CB} through the coefficients $a_{11}, \ldots, a_{22}$ and the dimensionless factors F_E, F_C defined in (21.90), (21.91). Thus, (21.102) have the form $I_E = I_E(V_{EB}, V_{CB})$, $I_C = I_C(V_{EB}, V_{CB})$ which, from the standpoint of circuit theory, describe a voltage-driven two-port circuit. The first equation in (21.102) shows that the current flowing into the emitter may be thought of as the sum of two currents, $a_{11}\,F_E + I_{UE}$ and $a_{12}\,F_C$, respectively; as a consequence, the two contributions must be represented as two elements in parallel, subjected to the voltage V_{EB}; a similar representation applies to the collector port. To better determine the elements of the circuit it is convenient to neglect the relatively weak dependence of $a_{11}, \ldots, a_{22}$ on the voltages[33] with respect to the exponential

[33]The voltage-dependence of the space-charge region's width in the reverse-bias condition is worked out in Sect. 21.4 for an abrupt junction and in Sect. 21.9.3 for a diffused junction. The dependence is the strongest in the reverse-bias condition. When a forward bias is applied, the width of the space-charge region is small; also, its variation is negligible because the bias itself varies little.

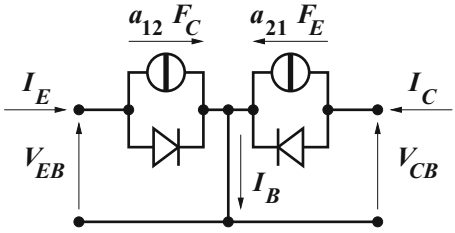


Fig. 21.25 The Ebers and Moll equivalent circuit for the BJT of the *p-n-p* type

dependence of F_E, F_C. It follows that the first element describing I_E becomes identical to the equation of a *p-n* junction subjected to V_{EB}, with a_{11} the saturation current (compare with (21.34)). In turn, the term $a_{12}\,F_C$, which depends on V_{CB} only, is a current driven by the voltage present at the other port; as a consequence, it is represented in the circuit by a voltage-driven current generator. The analysis of the collector port is similar, thus eventually leading to the *Ebers and Moll model* of the BJT, shown in Fig. 21.25. The importance of the width of the quasi-neutral base region can now be appreciated by observing that the coefficients $a_{11}, \ldots, a_{22}$ depend on W through the ratio W/L_B. In the limit $W/L_B \gg 1$, (21.99), (21.100) yield the limiting cases $a_{12}\,, a_{21} \longrightarrow 0$ and

$$a_{11} \longrightarrow q A_e \left(\frac{D_E\, n_E}{L_E} + \frac{D_B\, n_B}{L_B} \right), \quad a_{22} \longrightarrow q A_e \left(\frac{D_C\, n_C}{L_C} + \frac{D_B\, n_B}{L_B} \right). \tag{21.103}$$

As a consequence, the current generators of Fig. 21.25 become open circuits, and the circuit reduces to the assembly of two independent *p-n* junctions in which the current of each port is controlled only by the voltage of the same port. In fact, the limit $W/L_B \gg 1$ is not realistic: the order of magnitude of W is about 1 μm, while the typical diffusion lengths are much larger (compare with the estimate (21.129)); due to this, the full expressions of $a_{11}, \ldots, a_{22}$ must be kept. From the physical point of view, this means that the closeness of the two junctions is such that the current of each port is influenced also by the voltage present at the other port; thus, the behavior of the BJT is substantially different from that of two independent *p-n* junctions.[34]

The typical values of W and L_B are actually such that the other limit $W/L_B \ll 1$ holds; in this case it is $\coth(W/L_B) \longrightarrow 1/\sinh(W/L_B) \longrightarrow L_B/W$, and the coefficients (21.99), (21.100) simplify to

[34] From the technological point of view, the scaling down of W is more difficult than that of the channel length L of the MOS transistor. In fact, W is limited by the positions of the emitter-base and collector-base junctions, which are in turn determined by two diffusion processes (Sect. 21.8). In contrast, L is obtained from a lithographic process, whose control is much finer than that of thermal diffusions. This is the main reason why the evolution of the BJT technology, in terms of scaling, has ceased progressing when W has reached the value of about 0.7 μm.

$$a_{11} \simeq q A_e \left(\frac{D_E n_E}{L_E} + \frac{D_B p_B}{W} \right) , \quad a_{12} \simeq -q A_e \frac{D_B p_B}{W} , \tag{21.104}$$

$$a_{21} = a_{12} , \qquad a_{22} \simeq q A_e \left(\frac{D_C n_C}{L_C} + \frac{D_B p_B}{W} \right) . \tag{21.105}$$

Beside the form (21.102) of the device equations, it is often useful to consider the *mixed form* $I_E = I_E(V_{EB}, V_{CB})$, $I_C = I_C(V_{CB}, I_E)$. Remembering that the dependence of $a_{11}, \ldots, a_{22}$ on the voltages is neglected, this is easily accomplished by eliminating F_E from (21.102). Letting

$$h_{FB} = -\frac{a_{21}}{a_{11}} > 0 , \qquad I_{C0} = \frac{a_{11} a_{22} - a_{12} a_{21}}{a_{11}} \tag{21.106}$$

one finds for the mixed form

$$I_E = (a_{11} F_E + I_{UE}) + a_{12} F_C , \qquad I_C - I_{UC} = -h_{FB} \, (I_E - I_{UE}) + I_{C0} F_C . \tag{21.107}$$

In the forward regime (Table 21.1) the emitter-base junction is forward biased, whence $I_E \gg I_{UE}$; at the same time, the collector-base junction is reverse biased, whence $F_C \simeq -1$: as h_{FB} is of order unity (see below) and I_{C0} is small, the collector current is of the same order of magnitude as the emitter current, whence $I_C \gg I_{UC}$. In summary, in the forward regime the second equation in (21.107) simplifies to $-I_C \simeq h_{FB} I_E + I_{C0}$. Using the expressions (21.104), (21.105) for $a_{11}, \ldots, a_{22}$ one finds

$$h_{FB} \simeq \left(1 + \frac{D_E n_E W}{D_B p_B L_E} \right)^{-1} < 1 , \quad I_{C0} \simeq q A_e \left(\frac{D_C n_C}{L_C} + h_{FB} \frac{D_E n_E}{L_E} \right) . \tag{21.108}$$

From $-I_C \simeq h_{FB} I_E + I_{C0}$ it follows that I_{C0} is the negative collector current when the emitter port is open; such a current is necessarily small because the collector-base junction is reverse biased.[35] As a further approximation, applicable when $I_E \neq 0$, one neglects I_{C0} to find $-I_C \simeq h_{FB} I_E$; this justifies the designation[36] *current gain* ascribed to h_{FB} (suffixes "F" and "B" remind one that $-I_C \simeq h_{FB} I_E$ applies in the forward regime and in the common-base configuration). From the first equation in (21.108) it follows that h_{FB} approaches unity from below when the ratio $D_E n_E W/(D_B p_B L_E)$ is made smaller. Singling out the three factors, and remembering the estimate of W/L_B carried out above, one finds that the ratio W/L_E is smaller than unity; as for the ratio of the diffusion coefficients, the Einstein relation (19.102) yields $D_E/D_B = \mu_{nE}/\mu_{pB}$, where μ_{nE} is the mobility of the

[35] The same result stems from the observation that I_{C0} has the same structure as a saturation current; compare, e.g., with the estimate (21.130).

[36] The term "gain" is used in a broader meaning. In fact, it is $h_{FB} < 1$.

minority carriers (electrons) in the emitter, and μ_{pB} is the mobility of the minority carriers (holes) in the base. Their ratio is larger than one, but still of order unity.[37] Finally, using (21.86) one finds $n_E/p_B = N_B/N_E$; as the two dopant concentrations can be made to differ by orders of magnitude, the ratio N_B/N_E provides an easy way to achieve a current gain close to unity. This explains why the dopant concentration in the emitter must be larger than that in the base, as anticipated in Sect. 21.8.

Still in the forward regime, the contribution $l_{nE}(V_{EB})$ to the width of the quasi-neutral base region (21.101) varies little because in the typical operating conditions the forward bias V_{EB} is practically constant; the variations of the other contribution $l_{nC}(V_{CB})$, instead, may be large because the collector-base junction is reverse biased. To limit the variations, the dopant concentration in the collector is made much smaller than that of the base; in this way, the space-charge region of the collector-base junction is strongly asymmetric, and extends mainly in the collector (compare with the second relation in (21.39)): thus, the variation of $l_{nC}(V_{CB})$ is negligible and W may be considered constant. This explains why the dopant concentration in the collector must be smaller than that in the base, as anticipated in Sect. 21.8. A light dopant concentration in the collector is also useful to prevent the junction's avalanche breakdown (Sect. 21.5); in fact, a larger width of the space-charge region corresponds to a smaller electric field.

The *output characteristics* of the BJT of the *p-n-p* type, in the common-base configuration, represent the $I_C = I_C(V_{CB}, I_E)$ relation, with V_{CB}, I_E used as independent variable and parameter, respectively. The characteristics are drawn in Fig. 21.26 using the second equation in (21.107) with $h_{FB} = 0.95$, after neglecting I_{UC} and I_{UE}. The portion of the characteristics on the right of the vertical line describes the forward regime, while the portion on the left of it describes the saturation regime (Table 21.1). Limiting the discussion to the forward regime, one notes that $-I_C$ is independent of V_{CB} and practically equal to I_E; in other terms, the current at the output port (collector) is dictated by the voltage of the input port (emitter). From the physical point of view, this means that the closeness of the two junctions is such that the majority of the carriers (holes) injected by the emitter into the base are able to diffuse to the space-charge region of the collector without recombining. Once the holes reach the space-charge region, which is reverse biased, the strong electric field present in it pulls them into the collector. The small fraction of holes that recombine within the base region give rise to the base current $I_B = I_E + I_C$; combining the latter with the second equation in (21.107), still neglecting I_{UC} and I_{UE}, and eliminating I_E yield

[37] Moreover, the degradation of μ_{nE} is more severe than that of μ_{pB} because the dopant concentration in the emitter is larger (Sect. 20.5.2). In the BJT of the *n-p-n* type, the ratio of the diffusion coefficients becomes $D_E/D_B = \mu_{pE}/\mu_{nB} < 1$, which is more advantageous; all other parameters being the same, the *n-p-n* BJT is thus preferable. This is true also in time-dependent conditions, because the dynamics of the device is essentially controlled by the mobility of the minority carriers in the base.

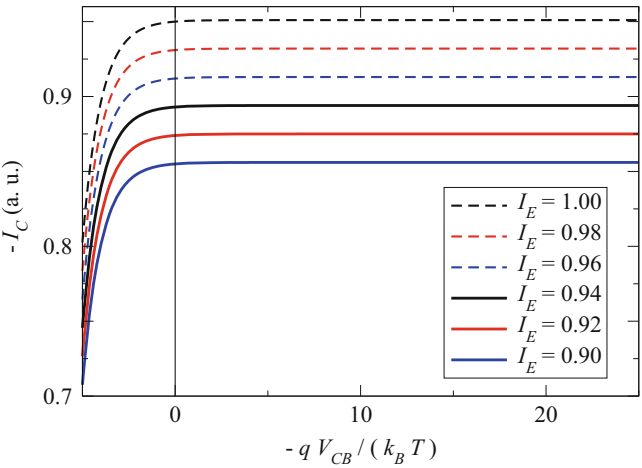


Fig. 21.26 Output characteristics of a BJT of the *p-n-p* type, drawn with $h_{FB} = 0.95$. The units of the voltage axis are normalized to $k_B\,T/q$, those of the current axis are arbitrary

$$-I_C = h_{FE}\,I_B - \frac{I_{C0}\,F_C}{1-h_{FB}}\,, \qquad h_{FE} = \frac{h_{FB}}{1-h_{FB}}\,, \tag{21.109}$$

with h_{FE} the current gain in the *common-emitter* configuration. Using the simplified expression (21.108) for h_{FB} one finds $h_{FE} \simeq D_B\,p_B\,L_E/(D_E\,n_E\,W)$.

The *input characteristics* of the BJT of the *p-n-p* type, in the common-base configuration, represent the $I_E = I_E(V_{EB}, V_{CB})$ relation, with V_{EB}, V_{CB} used as independent variable and parameter, respectively. They are derived from the first equation in (21.107) after neglecting I_{UC} and I_{UE}, namely, $I_E \simeq a_{11}\,F_E + a_{12}\,F_C$. For a qualitative analysis one rewrites the above as $I_E = (a_{11} - |a_{12}|)\,F_E + |a_{12}|\,(F_E - F_C)$; then, using (21.104) and making the same estimate as that used for h_{FB}, one finds $I_E \simeq |a_{12}|\,(F_E - F_C)$. As in the forward regime it is $F_C \simeq -1$, it follows that the curves corresponding to different values of V_{CB} are barely distinguishable from one another, and the input characteristics practically reduce to the single $I(V)$ curve of a *p-n* junction.

21.9 Complements

21.9.1 Weak-Injection Limit of the Drift-Diffusion Equations

In the calculation of the $I(V)$ characteristic of the *p-n* junction carried out in Sect. 21.3.1 it has been stated that in a quasi-neutral region, when the weak-injection condition holds, the diffusion term of the minority carriers dominates over the drift term. To better discuss this issue, the case of a *p*-doped region is considered, so

that the majority-carrier concentration is $c^{\mathrm{eq}} = p^{\mathrm{eq}} = p_{p0}$ and (20.35) become $|p - p_{p0}| \ll p_{p0}$, $|n - n_{p0}| \ll p_{p0}$. The latter may be recast as

$$|p - p_{p0}| \leq \alpha\, p_{p0}\,, \qquad |n - n_{p0}| \leq \alpha\, p_{p0}\,, \tag{21.110}$$

with $\alpha \ll 1$. Indicating with p_m, p_M the minimum and maximum values of p imposed by (21.110), one finds $p_M - p_{p0} = \alpha\, p_{p0}$, $p_{p0} - p_m = \alpha\, p_{p0}$, whence

$$p_M = (1 + \alpha)\, p_{p0}\,, \qquad p_m = (1 - \alpha)\, p_{p0}\,. \tag{21.111}$$

Similarly, for the minority-carrier concentration one finds $n_M - n_{p0} = \alpha\, p_{p0}$, $n_{p0} - n_m = \alpha\, p_{p0}$, whence

$$n_M = n_{p0} + \alpha\, p_{p0}\,, \qquad n_m = n_{p0} - \alpha\, p_{p0}\,. \tag{21.112}$$

The maximum absolute variation of p turns out to be:

$$p_M - p_m = 2\,\alpha\, p_{p0}\,. \tag{21.113}$$

Instead, the maximum variation of n must be treated with some care. In fact, using the nondegenerate case one finds

$$n_M = \frac{n_i^2}{p_{p0}} + \alpha\, p_{p0} = p_{p0} \left(\frac{n_i^2}{p_{p0}^2} + \alpha \right), \tag{21.114}$$

$$n_m = \frac{n_i^2}{p_{p0}} - \alpha\, p_{p0} = p_{p0} \left(\frac{n_i^2}{p_{p0}^2} - \alpha \right). \tag{21.115}$$

Even for a relatively low dopant concentration, say, $N_A \simeq p_{p0} = 10^{16}$ cm^{-3}, at room temperature one has $n_i^2/p_{p0}^2 \simeq 10^{-12}$, which is much smaller than the reasonable values of α. It follows $n_M \simeq \alpha\, p_{p0}$, $n_m \simeq 0$, where the limit of n_m must be chosen as such because n is positive definite. In conclusion, the maximum relative variations of p and n with respect to the equilibrium values are given by

$$\frac{p_M - p_m}{p_{p0}} = 2\,\alpha\,, \qquad \frac{n_M - n_m}{n_{p0}} \simeq \alpha\, \frac{p_{p0}}{n_{p0}} = \alpha\, \frac{p_{p0}^2}{n_i^2} \gg 2\,\alpha\,. \tag{21.116}$$

By way of example, one may let $\alpha = 10^{-3}$, still with $N_A = 10^{16}$ cm^{-3}. While the maximum relative variation of p is 2×10^{-3}, that of n is 10^9; it follows that the constraint imposed onto the derivative is strong in the case of p, much weaker for n. Within the same example, the maximum absolute variation is 2×10^{13} cm^{-3} for p and 10^{13} cm^{-3} for n, in both cases much smaller than the majority-carrier concentration (10^{16} cm^{-3}). The conclusion is that in a quasi-neutral region,

under the weak-injection conditions, the diffusive transport prevails for the minority carriers, whereas the transport of the majority carriers is dominated by drift. With reference to the p-doped region considered here, one has $J_p \simeq q\,\mu_p\,p\,E$ and $J_n \simeq q\,D_n\,\mathrm{grad}n$, respectively.

21.9.2 Shockley's Boundary Conditions

The derivation of the analytical model of the p-n junction's $I(V)$ characteristic, worked out in Sect. 21.3.1, requires the boundary conditions for the minority-carrier concentrations at the boundaries of the space-charge region; specifically, one needs to determine $n(-l_p)$ and $p(l_n)$. The derivation is based on calculating approximate expressions for the ratios $n(-l_p)/n(l_n)$, $p(l_n)/p(-l_p)$, where the denominators are majority-carrier concentrations that are in turn approximated with $n(l_n) \simeq n_{n0} \simeq N_D$ and $p(-l_p) \simeq p_{p0} \simeq N_A$.

To proceed, one considers the electron drift-diffusion equation $J_n = q\,\mu_n\,n\,E + q\,D_n\,\mathrm{d}n/\mathrm{d}x$, and observes that in the space-charge region the drift and diffusion terms have opposite signs; also, their moduli are much larger than that of the current density. In fact, the latter is small due to the weak-injection condition, whereas the terms at the right-hand side of the equation are large because the electric potential and the electron concentration have non-negligible variations over the space-charge region. It follows that the moduli of the drift and diffusion terms are comparable to each other: $-q\,\mu_n\,n\,E \simeq q\,D_n\,\mathrm{d}n/\mathrm{d}x \gg |J_n|$ and, similarly, $-q\,\mu_p\,p\,E \simeq -q\,D_p\,\mathrm{d}p/\mathrm{d}x \gg |J_p|$ for holes. Now, the approximation is introduced, that consists in neglecting J_n and J_p; this yields equilibrium-like expressions for the concentrations, $n \simeq n^{(0)}\,\exp[q\,\varphi/(k_B\,T)]$, $p \simeq n^{(0)}\,\exp[-q\,\varphi/(k_B\,T)]$, which are used to calculate the ratios sought:

$$\frac{n(-l_p)}{n(l_n)} \simeq \exp\left[\frac{q(V-\psi_0)}{k_B\,T}\right] = \frac{n_i^2}{N_A\,N_D}\,\exp\left(\frac{q\,V}{k_B\,T}\right), \tag{21.117}$$

$$\frac{p(l_n)}{p(-l_p)} \simeq \exp\left[\frac{q(V-\psi_0)}{k_B\,T}\right] = \frac{n_i^2}{N_A\,N_D}\,\exp\left(\frac{q\,V}{k_B\,T}\right). \tag{21.118}$$

The last form of (21.117), (21.118) is obtained from the definition (21.5) of the built-in potential. Using $n(l_n) \simeq n_{n0} \simeq N_D$ and $p(-l_p) \simeq p_{p0} \simeq N_A$ along with $n_{p0} = n_i^2/N_A$ and $p_{n0} = n_i^2/N_D$ (compare with (21.1, 21.2)), finally yields the Shockley boundary conditions

$$n(-l_p) \simeq n_{p0}\,\exp\left(\frac{q\,V}{k_B\,T}\right), \qquad p(l_n) \simeq p_{n0}\,\exp\left(\frac{q\,V}{k_B\,T}\right). \tag{21.119}$$

21.9.3 Depletion Capacitance—Arbitrary Doping Profile

The expression of the depletion capacitance worked out in Sect. 21.4 for an abrupt p-n junction is extended here to an arbitrary doping profile, still in one dimension. Let $a < x < b$ be the region where the charge density ρ differs from zero, and assume for the electric potential that $\varphi = \varphi(a) = \text{const}$ for $x < a$ and $\varphi = \varphi(b) = \text{const}$ for $x > b$. Also, it is assumed that there are no single layers or double layers of charge, whence the electric field E and φ are continuous. The constancy of φ in the outside regions implies the global charge neutrality, as is found by integrating $\varepsilon_{\text{sc}}\, \mathrm{d}E/\mathrm{d}x = \rho$ from a to b and using the continuity of E:

$$\int_a^b \rho \, \mathrm{d}x = 0 \,. \tag{21.120}$$

Thanks to (21.120) one finds, for any x,

$$\int_a^x \rho \, \mathrm{d}x + \int_x^b \rho \, \mathrm{d}x = 0 \,, \qquad Q = -\int_a^x \rho \, \mathrm{d}x = \int_x^b \rho \, \mathrm{d}x \,, \tag{21.121}$$

which provides the definition of the charge per unit area Q. The definition holds also if x is outside the interval $[a, b]$; in this case, however, one finds $Q = 0$. In the following it is assumed that x is internal to the space-charge region. The solution of the Poisson equation is taken from Prob. 4.2; using $E(a) = 0$ and the global charge-neutrality condition after letting $\psi = \varphi(b) - \varphi(a)$ yields

$$\varepsilon_{\text{sc}}\, \psi = \int_a^b x \, \rho \, \mathrm{d}x \,. \tag{21.122}$$

If the voltage drop changes by a small amount, $\psi \leftarrow \psi + \mathrm{d}\psi$, the space-charge boundaries are modified, $a \leftarrow a + \mathrm{d}a$, $b \leftarrow b + \mathrm{d}b$, whence Q changes as well:[38]

$$\mathrm{d}Q = \int_b^{b+\mathrm{d}b} \rho \, \mathrm{d}x = \rho(b)\, \mathrm{d}b = \int_a^{a+\mathrm{d}a} \rho \, \mathrm{d}x = \rho(a)\, \mathrm{d}a \,. \tag{21.123}$$

On the other hand, from (21.122) it follows

$$\varepsilon_{\text{sc}}\, \mathrm{d}\psi = \int_{a+\mathrm{d}a}^{b+\mathrm{d}b} x \, \rho \, \mathrm{d}x - \int_a^b x \, \rho \, \mathrm{d}x = b\, \rho(b)\, \mathrm{d}b - a\, \rho(a)\, \mathrm{d}a = (b - a)\, \mathrm{d}Q \,. \tag{21.124}$$

[38] After the change in the boundaries' positions, x in (21.121) is still internal to the space-charge region.

Thus, the capacitance per unit area of the space-charge region is

$$C = \frac{\mathrm{d}Q}{\mathrm{d}\psi} = \frac{\varepsilon_{\mathrm{sc}}}{b-a}\,, \tag{21.125}$$

which is the expected generalization of (21.48). Note that the absence of charge layers makes the variation $\mathrm{d}Q$ to depend on the variations in a and b only. As a consequence it is $a = a(\psi)$, $b = b(\psi)$, whence $C = C(\psi)$. If ψ and $\rho(x)$ are prescribed, the values of a, b are determined by the system of equations (21.120) and (21.122).

It is interesting to note that if the charge density has a power form, $\rho \sim x^n$, then ψ depends on the $(n+2)$th power of $b-a$. Consider by way of example a *diffused junction*, namely, a junction obtained, e.g., by diffusing a dopant of the p type into a uniform n-type substrate. Expanding ρ to first order around the metallurgical junction and using the full-depletion and ASCE approximations yield $\rho \simeq k\,x$ for $-l/2 < x < l/2$ and $\rho = 0$ elsewhere. Using (21.122) and (21.124) then yields

$$\varepsilon_{\mathrm{sc}}\,\psi = \frac{1}{12}\,k\,l^3\,, \qquad C = C_0\left(1-\frac{V}{\psi_0}\right)^{-1/3}\,, \qquad C_0 = \left(\frac{k\,\varepsilon_{\mathrm{sc}}^2}{12\,\psi_0}\right)^{1/3}\,. \tag{21.126}$$

The general expression (21.125) of the capacitance per unit area of the space-charge region finds a useful application in a measuring technique for the doping profile (Sect. 25.6).

21.9.4 Order of Magnitude of Junction's Parameters

Still considering an abrupt, p-n silicon junction with $N_A = 10^{16}$ cm^{-3}, $N_D = 10^{15}$ cm^{-3}, the built-in potential at room temperature is

$$\psi_0 = \frac{k_B\,T}{q}\,\log\left(\frac{N_A\,N_D}{n_i^2}\right) \simeq 0.65\ \mathrm{V} \tag{21.127}$$

(compare with (21.5)). The carrier mobilities have been estimated in Sect. 19.6.6; in fact, hole mobility is smaller than electron mobility and, as outlined in Sect. 20.5.3, the mobility degradation due to impurity scattering is expected to vary from one side of the junction to the other because the dopant concentrations are different. The experimental minority-carrier mobilities for the dopant concentrations and temperature considered here are $\mu_n \simeq 1000$ cm^2 V^{-1} s^{-1} in the p region and $\mu_p \simeq 500$ cm^2 V^{-1} s^{-1} in the n region [128, Sect. 1-5], whence

$$D_n = \frac{k_B\,T}{q}\,\mu_n \simeq 26\ \mathrm{cm}^2\,\mathrm{s}^{-1}\,, \qquad D_p = \frac{k_B\,T}{q}\,\mu_p \simeq 13\ \mathrm{cm}^2\,\mathrm{s}^{-1}\,. \tag{21.128}$$

The experimental values of the minority-carrier lifetimes are $\tau_n \simeq 5 \times 10^{-5}$ s and $\tau_p \simeq 2 \times 10^{-5}$ s. The corresponding diffusion lengths (21.27), (21.31) are

$$L_n = \sqrt{\tau_n D_n} \simeq 360 \ \mu\text{m}\,, \qquad L_p = \sqrt{\tau_p D_p} \simeq 160 \ \mu\text{m}\,. \tag{21.129}$$

The above values provide for the saturation current density (21.33)

$$J_s = q \left(\frac{D_p \, p_{n0}}{L_p} + \frac{D_n \, n_{p0}}{L_n} \right) \simeq 14 \ \text{pA cm}^{-2}\,. \tag{21.130}$$

From (21.44), the width of the depletion region at zero bias is found to be

$$l(V=0) = l_n + l_p = \left[\frac{2\,\varepsilon_{\text{sc}}}{q} \left(\frac{1}{N_D} + \frac{1}{N_A} \right) \psi_0 \right]^{1/2} \simeq 1 \ \mu\text{m}\,, \tag{21.131}$$

with $l_n/l_p = N_A/N_D = 10$. The permittivity of silicon $\varepsilon_{\text{sc}} = 11.7 \times \varepsilon_0$ has been used, with $\varepsilon_0 \simeq 8.854 \times 10^{-14}$ F cm^{-1} the vacuum permittivity. Finally, the value of the differential capacitance per unit area at zero bias (21.47) is

$$C_0 = \left[\frac{q\,\varepsilon_{\text{sc}}/(2\,\psi_0)}{1/N_D + 1/N_A} \right]^{1/2} = \frac{\varepsilon_{\text{sc}}}{l(V=0)} \simeq 11 \ \text{nF cm}^{-2}\,. \tag{21.132}$$

21.9.5 Solar Cell's Parameters

As apparent from (21.85), the $I(V)$ characteristic of the solar cell is obtained by shifting that of a standard *p-n* junction by an amount equal to I_G. The curve is drawn in Fig. 21.27 in arbitrary units; the intercept with the V axis is the *open-circuit voltage* $V_{OC} = V(I=0) = (k_B\,T/q)\,\log[(I_G - I_U)/I'_s + 1] > 0$, that with the I axis is the *short-circuit current* $I_{SC} = I(V=0) = I_U - I_G < 0$. As the intercept of the $I(V)$ curve with the load characteristic $I = -V/R$ belongs to the fourth quadrant, the working voltage is positive while the corresponding current is negative. It follows that the power $P = V\,I$ is negative, namely, it is delivered by the cell to the load resistance; as this reasoning applies to any value of R, the latter can be adapted in order to maximize the transferred power. The voltage that yields the maximum power is in fact determined by letting $\mathrm{d}P/\mathrm{d}V = 0$, namely, $I + V\,(\mathrm{d}I/\mathrm{d}V) = 0$, with I given by (21.85). The solution provides the maximum-power voltage V_{MP} which, replaced into (21.85), renders in turn the maximum-power current I_{MP}. The two values thus found eventually yield the maximum-power resistance $R_{MP} = -V_{MP}/I_{MP}$.

A number of parameters are used to quantify the solar cell's performance; among them, the *fill factor* has two alternative definitions, namely,

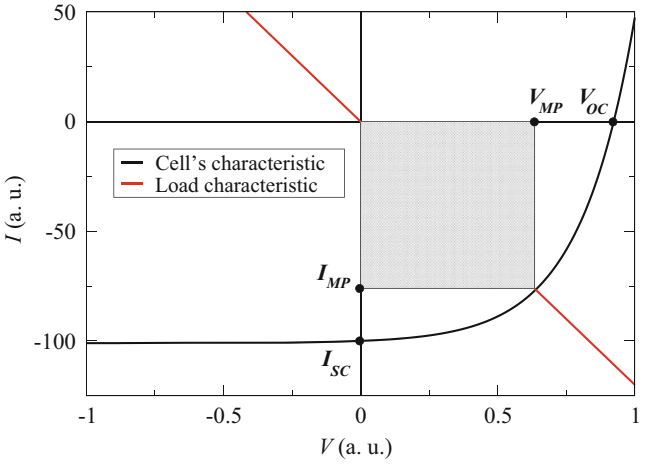


Fig. 21.27 $I(V)$ characteristic of the solar cell, in arbitrary units (*black line*), and maximum-power load characteristic $I = -V/R_{MP}$ (*red line*). The absolute value of the *shaded area* shows the power delivered by the cell

$$f = \frac{V_{MP}\, I_{MP}}{V_{OC}\, I_{SC}}\,, \qquad f = \frac{V_{MP}\, I_{MP}}{\int_0^{V_{OC}} I\, \mathrm{d}V}\,. \tag{21.133}$$

The *efficiency* η_c of the cell is defined as the ratio between the maximum power per unit area delivered to the load and the power per unit area impinging on the cell:

$$\eta_c = \frac{|V_{MP}\, I_{MP}|/A_e}{P_{\mathrm{in}}/A_e}\,. \tag{21.134}$$

Due to the atmospheric absorption and the angle of incidence of sunlight, the denominator of (21.134) is substantially lower than the solar constant[39] $\mathrm{d}P_E/\mathrm{d}A \simeq$ 136 mW cm^{-2}. The maximum theoretical efficiency of a solar cell made of a single *p-n* junction with a 1.34 eV band gap has been calculated in [123] using a 1.5 air mass;[40] the resulting figure is about 33.7%. Silicon, whose band gap at room temperature is about 1.12 eV, has a slightly smaller figure: 32%. The efficiency of actual devices is lower than the theoretical one due to different types of losses: e.g., reflection at the surface, shielding due to the wires deposited over the surface, electron-hole recombinations, lateral resistance in the semiconductor, series

[39] The solar constant is calculated in Prob. 7.1.

[40] In this matter, *air mass* (AM) indicates the length of the optical path, through earth's atmosphere, of light rays coming from an extraterrestrial source. The parameter is used to quantify the attenuation due to absorption and scattering. The AM value at sea level, when the source is at the zenith, is set equal to unity and used as normalization parameter.

resistance of the wires.[41] To date (2017), the maximum efficiency of commercial solar cells made of crystalline materials is about 24%.

It is worth pointing out that the figures indicated so far refer to *single-junction cells*, in which a single band gap exists: such cells cannot absorb the part of the spectrum whose energy is below the band gap itself, and have a limited efficiency in absorbing photons whose energy is much higher than the band gap. For instance, the maximum wavelength that can be absorbed by silicon at room temperature is

$$\lambda_M = \frac{h\,c}{E_G} = \frac{6.63 \times 10^{-34}\ \mathrm{J\ s} \times 3 \times 10^{10}\ \mathrm{cm\ s^{-1}}}{1.12 \times 1.602 \times 10^{-19}\ \mathrm{J}} = 1.11\ \mu\mathrm{m}\,, \tag{21.135}$$

that belongs to the near infrared. A *multi-junction cell* consists of a stack of two or more single-junction cells, each with a different band gap; this arrangement is able to absorb a larger fraction of the spectrum, making the maximum theoretical efficiency to reach 42% or 49% in a two-cell or three-cell stack, respectively.

Another parameter that influences the efficiency of the solar cell is temperature. In the normal functioning of a cell, temperature reaches about 350 K due to the combined effect of the exposure to sunlight and of the internal losses; even higher temperatures are reached when light concentrators are used. The increase in temperature degrades the cell's efficiency; from a thermodynamic viewpoint this was expected, because the cell can be thought of as an engine that delivers energy by working between two temperatures: the larger one, T_H, is that of the sun, the lower one, T_L, is that of cell itself. In turn, the efficiency of a real engine cannot exceed that of a reversible engine, $1 - T_L/T_H$ [47, Sect. III-9].

A more detailed analysis shows that an increase in temperature makes V_{OC} and, consequently, V_{MP}, to decrease. This shows, in passing, that the fill factor (21.133) is not an efficiency: the $I(V)$ curve may deform in a way that makes the product $V_{MP}\,I_{MP}$, hence the efficiency, to decrease, without altering the fill factor.

21.9.6 *Equivalent Circuit of the* P-N-P *BJT (Single Generator)*

The equivalent circuit of Fig. 21.25 is easily transformed into a circuit in which only one current generator is present. To this purpose, the original equations (21.102) are recast as

$$I_E = [(a_{11} + a_{21})\ F_E + I_{UE}] + a_{12}\ (F_C - F_E)\ , \tag{21.136}$$

$$I_C = [(a_{22} + a_{12})\,F_C + I_{UC}] + a_{21}\ (F_E - F_C)\ , \tag{21.137}$$

[41] Due to the need of minimizing at the same time the shielding effect and the series resistance, the total area and layout of the wiring over the top surface of a solar cell must be optimized carefully.

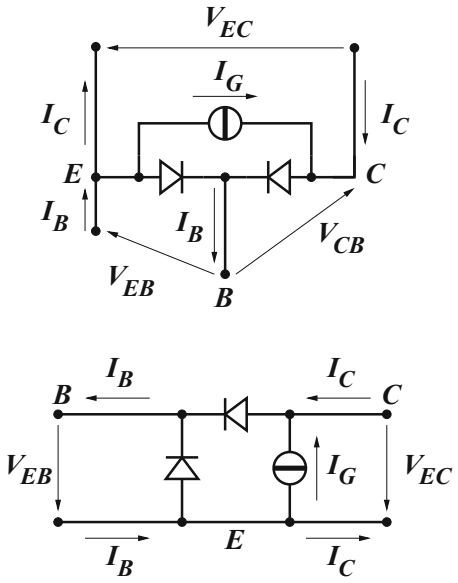


Fig. 21.28 The *upper part of the figure* shows the transformation from the two-generator equivalent circuit of Fig. 21.25 to the single-generator circuit, still in the common-base configuration. The *lower part* shows the transformation of the latter into the common-emitter configuration

where the equality $a_{21} = a_{12}$ has been exploited. The term in brackets in (21.136) is identical to the equation of a *p-n* junction, subjected to V_{EB}, whose saturation current is $a_{11} + a_{21}$; the remaining term $I_G = a_{12}\ (F_C - F_E) = |a_{12}|\ (F_E - F_C)$ is another contribution to I_E, controlled by V_{EB} and V_{CB}. Similarly, (21.137) expresses I_C as the sum of $-I_G$ and the current of a *p-n* junction subjected to V_{CB}, having $a_{22} + a_{12}$ as saturation current. The equivalent circuit corresponding to (21.136, 21.137) is shown in the upper part of Fig. 21.28 where, to the purpose of carrying out a subsequent transformation, and remembering that $I_B = I_E + I_C$, the emitter contact is split into two branches: the upper branch belongs to the loop that closes into the collector, while the lower branch belongs to the loop that closes into the base. A replica of the single-generator equivalent circuit is shown in the lower part of Fig. 21.28, where the two-port circuit is drawn in the *common-emitter configuration* $I_B = I_B(V_{EB}, V_{EC})$, $I_C = I_C(V_{EB}, V_{EC})$. The current gain of such a configuration is given by (21.109). In turn, using the expressions (21.104), (21.105) for $a_{11}, \ldots, a_{22}$ one finds $a_{11} + a_{21} \simeq q A_e D_E n_E / L_E$ and $a_{22} + a_{12} \simeq q A_e D_C n_C / L_C$.

Still with reference to the lower part of Fig 21.28, in the forward regime ($V_{EB} > 0$, $V_{CB} < 0$, $F_C \simeq -1$) one simplifies the circuit by neglecting the current flowing between B and C; it follows $I_G \simeq -I_C$. Thus, from (21.109), $I_G \simeq h_{FE} I_B + I_{C0}/(1 - h_{FB}) \simeq h_{FE} I_B$. In this way, the equivalent circuit reduces to the simplest form, made of two elements only. The corresponding low-frequency, small-signal circuit, also known as *Giacoletto circuit* [53], is found by linearizing the emitter-base junction around the operating point $V_{EB}^{(0)}$, $I_B^{(0)} = I_B(V_{EB}^{(0)})$. To this purpose, one obtains first the base current $I_B = I_E + I_C$ by adding up (21.136) and (21.137); in the forward regime the result reduces to

$$I_B \simeq q A_e \frac{D_E n_E}{L_E} \exp\left(\frac{q V_{EB}}{k_B T}\right) . \tag{21.138}$$

Letting $v_{EB} = V_{EB} - V_{EB}^{(0)}$ and $i_B = I_B - I_B^{(0)}$, the linearized relation for the base-emitter port reads

$$i_B = \frac{v_{EB}}{r_\pi} , \qquad r_\pi = \frac{k_B T/q}{I_B^{(0)}} , \tag{21.139}$$

with r_π the *small-signal input resistance* in the common-emitter configuration. In turn, the relation for the collector-emitter port is

$$i_G = h_{FE}\, i_B = g_m\, v_{EB} , \qquad g_m = \frac{h_{FE}}{r_\pi} = \frac{h_{FE}\, I_B^{(0)}}{k_B T/q} = \frac{I_C^{(0)}}{k_B T/q} , \tag{21.140}$$

with g_m the *transconductance*.

21.9.7 Comment on the Diffusion Length

The analysis of the *p-n* junction and of the BJT carried out in this chapter has shown that the diffusion length of the minority carriers plays an important role in determining the spatial behavior of the carrier concentration. Refer for instance to the *n*-type quasi-neutral region of a *p-n* junction, whose extent is assumed to be large enough to make the asymptotic condition applicable (Sect. 21.2.1); letting the origin of x coincide with the boundary between the quasi-neutral and space-charge regions (Fig. 21.6), the hole concentration reads

$$p = p_{n0} + A_p \exp(-x/L_p) , \qquad L_p = \sqrt{\tau_p D_p} . \tag{21.141}$$

Thus, L_p determines the rate of decay of the minority-carrier concentration toward the asymptotic value. The expression of L_p shows that the diffusion length is in fact the result of two competing mechanisms, thermal recombination and diffusion. In the weak-injection approximation (21.25), the net thermal recombination rate is dictated by τ_p only; in the $\tau_p \to 0$ limit it must be $p \to p_{n0}$ to prevent the left-hand side of (21.25) from diverging. Without the limit, when the lifetime is small the minority-carrier concentration is kept close to the equilibrium value; if, at the same time, the diffusion coefficient D_p is small, the hole current density $J_p = -q D_p \, \mathrm{d}p/\mathrm{d}x$ and average velocity, related to the former by $J_p = q p v_p$, are also small. In these conditions, the holes entering the quasi-neutral region at $x = 0$ move slowly and recombine quickly, thus providing a spatially rapid decay to the asymptotic value; this is rendered by (21.141), where L_p is small. The reasoning can be repeated using other combinations of the values of τ_p and D_p. Also, it can easily

be proven from (21.141) that L_p is the average value of the inverse function $x = x(p)$ (Prob. 21.13); thus, L_p can be thought of as the average distance traveled by a hole within the quasi-neutral region before recombining.

Problems

21.1 The built-in potential ψ_0 of an abrupt p-n junction is calculated in Sect. 21.2.1 assuming that the nondegeneracy and complete-ionization conditions hold. Find a method to determine ψ_0 when the two hypotheses above are disposed of (the outcome of the calculation is not necessarily a closed-form equation).

21.2 Show that avalanche due to impact ionization is possible only if both coefficients k_n and k_p are different from zero.

21.3 Evaluate the widths l_n, l_p of the space-charge region in an abrupt, silicon p-n junction at room temperature, with $N_A = 10^{17}$ cm^3 and $N_D = 10^{14}$ cm^3, subjected to a 2.1 V reverse bias.

21.4 Consider the current of a silicon photodiode operating in the continuous mode, given by (21.66). Using for l the value calculated in Prob. 21.3, and letting $A_e = 1$ cm^2, $\tau_g = 100$ μs (the definition of τ_g is in Sect. 20.2.3, compare also with the values of the minority-carrier lifetimes given in Sect. 21.9.4), estimate the current due to the second term in parentheses in (21.66) at room temperature.

21.5 Calculate the time constant τ_0 of (21.68) letting $A_e = 1$ cm^2 and using for I_G a value five times larger than that of the recombination current found in Prob. 21.4. For ψ_0 and C_0 take the values from (21.127) and (21.132), respectively.

21.6 If the integration time T_i is much shorter than τ_0, one can linearize (21.69). Find by this method the value of V_i corresponding to $T_i = 15$ ms. Let $V_0 = -2.6$ V, $R = 100$ kΩ, $m = 2$, and use for τ_0 and I_G the same values as in Prob. 21.5.

21.7 Using the same values as in Prob. 21.6, find the current gain of the storage-mode operation.

21.8 Solve (21.71) for $m = 2$, using the initial condition $V_D = V_i$ and assuming that I_G is independent of time.

21.9 Solve (21.71) for $m = 3$, using the initial condition $V_D = V_i$ and assuming that I_G is independent of time.

21.10 Draw the graph of $V_D(t)$ within the integration time, with $m = 2$ and $m = 3$. For the parameters use the values $V_0 + R I_G \simeq -2.6$ V, $\psi_0 = 0.65$ V, $\tau_0 = 143$ ms, $T_i = 15$ ms.

21.11 Draw the graph of $V_D(t)$ within the sampling time, with $m = 2$ and $m = 3$. Use the same parameters as in Prob. 21.10, and $\tau = 0.98$ ms. For the initial condition use the value of V_D at the end of the integration time, as found in Prob. 21.10.

21.12 Solve the algebraic system (21.78, 21.79) without linearizing the exponentials.

21.13 Use (21.141) to show that the minority-carrier diffusion length in a quasi-neutral region (L_p in this case) is the average value of the inverse function $x = x(p)$.

Chapter 22
MOS Devices

22.1 Introduction

The mathematical model of semiconductor devices, derived in Chap. 19, is applied here to the description of two fundamental devices of the insulated-gate type: the MIS capacitor, whose most important implementation is the MOS capacitor, and the IGFET, whose most important implementation is the MOSFET. Both devices can be realized starting from either a *p*-doped or an *n*-doped substrate. The analysis of the MOS capacitor is carried out using the simple example of a one-dimensional device in steady state, with the hypotheses of nondegeneracy and complete ionization, that lend themselves to an analytical treatment. Observing that in a steady-state condition the device is in equilibrium, the theory needs the solution of Poisson's equation only. From the solution of the latter, the device's capacitance is calculated, followed by a number of other important relations, that are useful in the subsequent treatment of the MOSFET. The functioning of the photocapacitor is worked out thereafter, by combining the analysis of the MOS capacitor with that of the storage-mode photodiode illustrated in Chap. 21. The theory of the MOSFET is then tackled in two dimensions and in steady-state conditions, first deriving a general expression for the channel current that holds in the case of a well-formed channel. The calculation is then completed by introducing the gradual-channel approximation: the differential conductances are derived first, followed by the expression of the drain current as a function of the applied voltages. A further simplification leads to the linear-parabolic model, which is widely used in the semiqualitative analyses of circuits. Then, still basing upon the linear-parabolic model, the input-output characteristic of the CMOS inverter is derived. The complements address the solution of the Poisson equation in the channel when a nonequilibrium condition holds, to provide a formal proof of the relation between the surface and quasi-Fermi potentials used in the gradual-channel approximation; then, a few phenomena that are not accounted for by the gradual-channel approximation are discussed, and the neglect of the dependence on position

M. Rudan, *Physics of Semiconductor Devices*, DOI 10.1007/978-3-319-63154-7_22

of the average carrier mobility is justified. Finally, the theory of the MOSFET is worked out again after disposing of the hypothesis of a well-formed channel; in this way, the subthreshold operation of the device is described. This chapter is concluded by the illustration of the scaling rules applied to the miniaturization of semiconductor devices, and by a brief account on the design procedures typically adopted in the integrated-circuit production.

22.2 Metal-Insulator-Semiconductor Capacitor

The Metal-Insulator-Semiconductor (MIS) capacitor is a fundamental device that constitutes the basis for the field-effect transistors used in the fabrication of integrated circuits. The device has also extensively been used for studying the properties of semiconductor surfaces [128]. A one-dimensional version of it is sketched in Fig. 22.1: the structure is fabricated by depositing or thermally growing (Chap. 24) an insulator layer over a semiconductor substrate. The fabrication process must obtain an electrically clean interface; in fact, the number of localized electronic states at the interface must be kept to a minimum to avoid carrier trapping-detrapping processes. The contact deposited onto the insulator is called *gate contact*, the other one is called *bulk contact*.

In the standard silicon technology, the insulator is obtained by thermally growing silicon dioxide (Sect. 24.2). For this reason, the thickness of the insulator is indicated in the following with t_{ox} instead of the generic symbol t_i used in Fig. 22.1; by the same token, the insulator's permittivity is indicated with ε_{ox}, and the device is called *MOS capacitor*. In the last years, the progressive scaling down in the size of semiconductor devices has brought the insulator thickness to the range of nanometers. A smaller thickness has the advantage of providing a larger capacitance; however, it may eventually lead to dielectric breakdown and leakage by quantum

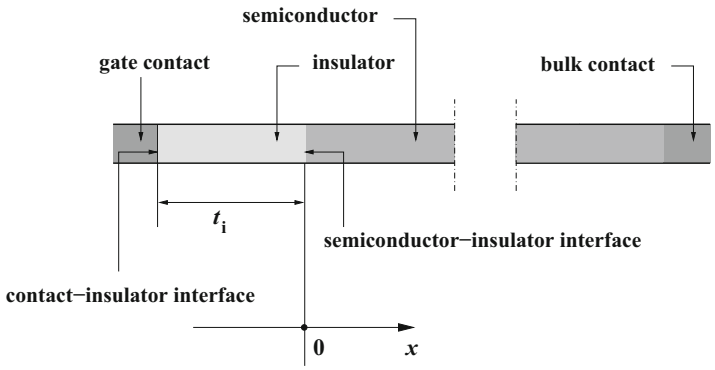


Fig. 22.1 Cross-section of a metal-insulator-semiconductor capacitor. The thickness of the insulator layer is not realistic: in real devices the layer is much thinner than the contacts

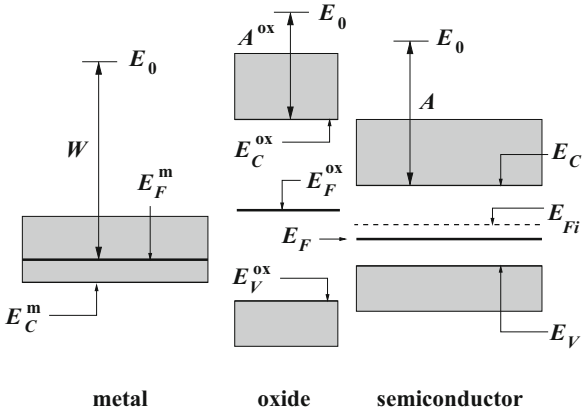


Fig. 22.2 The three materials forming the MOS capacitor shown separately. The symbols' meaning is illustrated in the text

tunneling. Silicon dioxide, which has been used as a gate insulator for decades, is being replaced in advanced devices with insulating layers made of materials having a larger permittivity (*high-k dielectrics*). Such layers are obtained by deposition (Sect. 24.5). Still with reference to the silicon technology, the conductive layers are made of metals, heavily doped polycrystalline silicon, or metal silicides; here they will be indicated with the generic term "metal."

Like in the case of *p-n* junctions, the theory of the MOS capacitor is carried out with reference to a simplified structure, where the device is one dimensional and aligned with the x axis; in this case the semiconductor–insulator interface is a plane normal to x and, as shown in Fig. 22.1, its position is made to coincide with the reference's origin. A constant dopant concentration is present in the semiconductor region; to further simplify the analytical approach one also assumes that the conditions of nondegeneracy and complete ionization hold.

To describe the functioning of the device it is necessary to consider the fact that in a region where the important electric phenomena occur, different materials are brought into an intimate contact. With reference to Fig. 22.2, the three materials (gate metal, oxide, and semiconductor) are initially considered separate from each other, and in the equilibrium condition. The left part of the figure shows the conduction band of the metal, with E_C^{m} the band's lower edge and E_F^{m} the metal's Fermi level. Due to the form of the Fermi-Dirac statistics, the probability that an electron's energy exceeds E_F^{m} is small; remembering the discussion of Sect. 7.2, the minimum energy necessary for an electron to transit from the metal into vacuum is the metal work function $W = E_0 - E_F^m$, with E_0 the vacuum level (left part of Fig. 22.2). For an insulator or a semiconductor, the electrons with maximum energy belong to states in the vicinity of the lower edge of the conduction band, E_C^{ox} (center) or E_C (right); in this case the minimum energy necessary for transiting into vacuum is the *electron affinity* $A^{\mathrm{ox}} = E_0 - E_C^{\mathrm{ox}}$ or $A = E_0 - E_C$, respectively. The semiconductor considered in the figure is uniformly doped of the p type, as shown by the fact that the Fermi level is below the intrinsic Fermi level E_{Fi} (Sect. 18.4.2). However, the analysis carried out here applies also to a semiconductor of the n type.

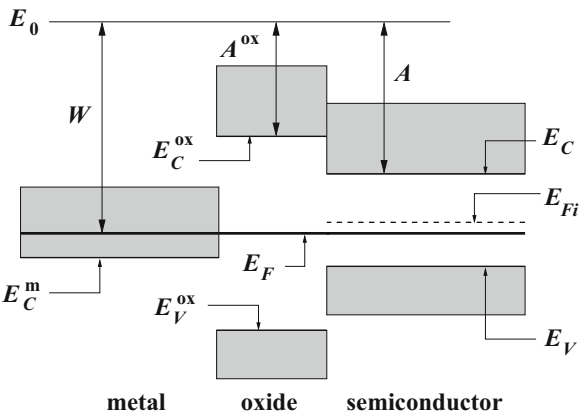


Fig. 22.3 The three materials forming the MOS capacitor after being brought into contact. The symbols' meaning is illustrated in the text

When the materials are brought into contact, they form a single, nonuniform system; as a consequence, in the equilibrium condition the Fermi levels must align with each other. On the other hand, the vacuum levels must align as well, and the bands must adapt to compensate for a possible charge redistribution that occurs when the materials contact each other. The situation is similar to that represented in Fig. 21.3 for the *p*-*n* junction. The values of the parameters $W, A, \ldots$ selected for drawing Fig. 22.2 fulfill the relation

$$W - A = E_C - E_F \,. \tag{22.1}$$

As a consequence, there is no need for the bands to modify their shape; as shown in Fig. 22.3, which represents the situation after the materials have been brought into contact, the bands do not deform. The condition where the semiconductor's bands are everywhere parallel to the Fermi level is indicated with *flat-band condition*.[1] It is important to remark that condition (22.1) seldom occurs in realistic cases; however, as shown below, the case $W - A \neq E_C - E_F$ is easily incorporated into the analysis.

When the bulk contact is considered, Figs. 22.2 and 22.3 must be completed by adding the band structure of the contact's material to the right of that of the semiconductor. Assuming that the gate and bulk contacts are made of the same material, the structure to be added is identical to that already present on the left part of the figures. In the interior of each material, due to spatial uniformity, the electric potential is piecewise constant, thus the electric field is zero.

Consider now the case where a voltage V_G is applied between the gate and bulk contacts; the voltage reference is such that $V_G > 0$ when the electric potential of the gate contact is larger than that of the bulk contact, and vice versa. In steady-state conditions, the insulator prevents a current from flowing through the device; therefore, during the transient consequent to the application of V_G, the electric

[1] The form of 22.1 is general enough to hold for both *p*- and *n*-type semiconductors.

charge adjusts itself to the new boundary conditions. At the end of the transient the device is again in an equilibrium condition, while the form of the bands is different from that of Fig. 22.3. Similarly, the electric potential in the oxide and semiconductor is not constant any longer; its form is found by solving the Poisson equation in each region.

22.2.1 *Surface Potential*—P-*Type Substrate*

The solution of the Poisson equation in the semiconductor region follows the same pattern as for the *p*-*n* junction (Sect. 21.2.1). Here a uniformly *p*-doped region is considered; its extension along the x axis is large, so that, far away from the semiconductor–insulator interface, the semiconductor behaves as if it were isolated. This fixes the carrier-equilibrium concentrations in the bulk; the asymptotic value of the electric potential is set to zero, $\varphi(+\infty) = 0$ whence, remembering that the nondegeneracy and complete-ionization conditions hold, it is

$$p^{(0)} = p(+\infty) = p_{p0} \simeq N_A \,, \qquad n^{(0)} = n(+\infty) = n_{p0} \simeq \frac{n_i^2}{N_A} \,. \tag{22.2}$$

The Poisson equation in the semiconductor then reads

$$u'' = \frac{1}{L_A^2} A(u) \,, \qquad A(u) = \frac{n_i^2}{N_A^2} \left[\exp(u) - 1\right] + 1 - \exp(-u) \,, \tag{22.3}$$

with L_A the Debye length for the holes defined in (21.12). The normalized charge density $A(u)$ has the same sign as u (compare with Sect. 21.2.1). Multiplying by u' both sides of the first equation in (22.3), transforming its left-hand side into $u''\,u' = (1/2)\,[(u')^2]'$, and integrating from $x \geq 0$ to $+\infty$ yield

$$\left(u'\right)^2 = \frac{2}{L_A^2} B(u) \,, \qquad B(u) = \frac{n_i^2}{N_A^2} \left[\exp(u) - 1 - u\right] + u + \exp(-u) - 1 \,. \tag{22.4}$$

Following the same reasoning as for (21.10), one finds that B is nonnegative and u monotonic. However, in contrast with the case of the *p*-*n* junction, where the sign of u' is positive due to the boundary condition at $x \to -\infty$, here u may either increase or decrease monotonically; in fact, the sign of u' is fixed by the boundary condition V_G, which in turn may be either positive or negative. In conclusion, one finds

$$u' = \pm \frac{\sqrt{2}}{L_A} F(u) \,, \qquad F(u) = \sqrt{\frac{n_i^2}{N_A^2} \left[\exp(u) - 1 - u\right] + u + \exp(-u) - 1} \,, \tag{22.5}$$

where the sign must be found on a case-by-case basis. Separating (22.5) finally yields

$$\frac{\mathrm{d}u}{F(u)} = \pm \frac{\sqrt{2}}{L_A}\,\mathrm{d}x\,, \tag{22.6}$$

which must be solved numerically because it has no analytical solution.[2] Much information, however, is gained directly from (22.5), without the need of integrating (22.6). To this purpose, one notes that the electric potential is continuous at the semiconductor–oxide interface; in fact, the normalized charge density in the semiconductor (22.3) has no charge layers in it, hence it cannot contribute to a double charge layer at the interface. As a consequence, one can adopt the same symbol φ_s for the electric potential at $x = 0$, without distinguishing between the two sides of the interface; φ_s is called *surface potential*, whilst $u_s = q\,\varphi_s/(k_B\,T)$ is the *normalized surface potential*. In contrast, the electric field is discontinuous at the same interface; for this reason, one defines

$$u_s' = \lim_{x\to 0^+} \frac{\mathrm{d}u}{\mathrm{d}x}\,, \qquad E_s = -\frac{k_B\,T}{q}\,u_s'\,, \qquad E_{\mathrm{ox}} = -\lim_{x\to 0^-} \frac{\mathrm{d}\varphi}{\mathrm{d}x}\,. \tag{22.7}$$

The relation between E_s and E_{ox} is found by considering a cylinder of thickness c placed across the semiconductor–oxide interface, such that the unit vector $\mathbf{n}_R$ normal to the right face is parallel to the unit vector $\mathbf{i}$ of the x axis, whereas the unit vector $\mathbf{n}_L$ normal to the left face is antiparallel to $\mathbf{i}$ (Fig. 22.4). Letting A_e be the common area of the two faces, the total charge within the cylinder is $A_e\,Q$, with Q the charge per unit area. Integrating $\operatorname{div}\mathbf{D} = \varrho$ over the cylinder's volume and using (A.23) yield

$$A_e\,Q = \int_{A_e} \mathbf{D}\cdot\mathbf{n}\;\mathrm{d}A_e = A_e\;[D_L\,\mathbf{i}\cdot(-\mathbf{i}) + D_R\,\mathbf{i}\cdot\mathbf{i}] = A_e\;(D_R - D_L)\;, \tag{22.8}$$

with D_R (D_L) the electric displacement on the right (left) face. From $\mathbf{D} = \varepsilon\,\mathbf{E}$ one then finds $Q = \varepsilon_{\mathrm{sc}}\,E_R - \varepsilon_{\mathrm{ox}}\,E_L$. It has been shown above that there are no charge layers on the semiconductor's side; as for the oxide layer, in principle it should be free of charges, although some contaminants may be present (Sect. 24.1). Here it is assumed that the oxide is free of charge; in conclusion, letting the cylinder's thickness c go to zero, one obtains $Q \to 0$, whence, using the limits (22.7),

$$\varepsilon_{\mathrm{sc}}\,E_s = \varepsilon_{\mathrm{ox}}\,E_{\mathrm{ox}}\,. \tag{22.9}$$

[2] The numerical evaluation from (22.6) of the inverse relation $x = x(u)$ is straightforward, though. Letting $\xi = \sqrt{2}\,x/L_A$, $\xi_0 = 0$, $u_0 = u_s$, $F_0 = F(u_0)$, it is $u_{k+1} = u_k \mp \delta u$, $F_{k+1} = F(u_{k+1})$, and $\xi_{k+1} = \xi_k + (1/F_k + 1/F_{k+1})\,\delta u/2$, with $k = 0, 1, \ldots$.

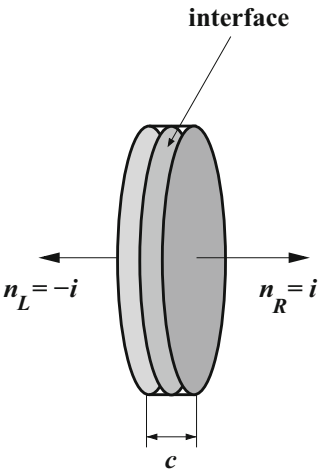


Fig. 22.4 The cylinder used to calculate the relation between electric displacement and charge per unit area across an interface

To find $E_{\rm ox}$ one observes that in a one-dimensional medium free of charge the electric potential is linear, whence $E_{\rm ox}$ is given by the negative potential drop across the oxide divided by the oxide thickness $t_{\rm ox}$. To complete the analysis it is then necessary to consider the interface between oxide and gate metal.

In the interior of the metal the electric potential is uniform; its value with respect to the metal of the bulk contact is V_G. However, in the solution of the Poisson equation within the semiconductor, the asymptotic condition $\varphi(+\infty) = 0$ has been chosen, which holds inside the semiconductor region; the surface potential φ_s is referred to such a zero as well. If (22.1) holds, the distinction between bulk contact and bulk semiconductor does not matter because the electric potential is the same; however, in the realistic cases it is not so, whence V_G and φ_s are referred to two different zeros. Assuming that the interior of the oxide and the semiconductor–oxide interface are free of charge, and remembering the discussion carried out in Sect. 21.2.2, the difference between the external zero (namely, that within the bulk contact) and the internal zero (given by the asymptotic condition) is the built-in potential Φ_{mp} between the bulk contact and the p-type semiconductor; thus, the gate voltage referred to the internal zero is $V'_G = V_G - \Phi_{mp}$. Also, the electric potential is continuous across the interface between the oxide and the gate metal, because no double layer is present there. In contrast, as $E = 0$ within the metal while $E_{\rm ox} \neq 0$, the electric field is generally discontinuous; in fact, a charge layer of density

$$\varrho_{\rm m} = Q_{\rm m}\, \delta(x + t_{\rm ox}^-)\,, \tag{22.10}$$

with $Q_{\rm m}$ the charge per unit area of the metal, builds up at the gate-metal's surface. In conclusion, the electric field within the oxide reads

$$E_{\rm ox} = \frac{V'_G - \varphi_s}{t_{\rm ox}}\,. \tag{22.11}$$

22.2.2 Relation Between Surface Potential and Gate Voltage

Combining (22.5), (22.7), (22.9), and (22.11) one finds

$$C_{\rm ox}\,\left(V'_G-\varphi_s\right)=\mp\varepsilon_{\rm sc}\frac{k_B\,T}{q}\,\frac{\sqrt{2}}{L_A}\,F(\varphi_s)\,,\qquad C_{\rm ox}=\frac{\varepsilon_{\rm ox}}{t_{\rm ox}}\,, \tag{22.12}$$

where $C_{\rm ox}$ is the *oxide capacitance per unit area*, and $F(\varphi_s)$ is obtained by replacing u with $q\,\varphi_s/(k_B\,T)$ in the second relation of (22.5). The left-hand side of (22.12) is the charge per unit area $Q_{\rm m}$ in the gate metal. This is easily found by considering the same cylinder as above, this time placed across the metal-oxide interface. Integrating $\mathrm{div}\,\mathbf{D}=\varrho$ over the cylinder's volume, using (22.10), and observing that $D_L=0$ because the metal's interior is equipotential yield

$$A_e\,Q_{\rm m}=\int_{A_e}\mathbf{D}\cdot\mathbf{n}\;\mathrm{d}A_e=A_e\,D_R=A_e\,C_{\rm ox}\,\left(V'_G-\varphi_s\right)\,. \tag{22.13}$$

Due to the global charge neutrality, the following relation holds between the charge per unit area in the gate metal, $Q_{\rm m}$, and that within the semiconductor, $Q_{\rm sc}$:

$$Q_{\rm m}+Q_{\rm sc}=0\,,\qquad Q_{\rm sc}=\int_0^\infty q\,(p-n-N_A)\;\mathrm{d}x=-C_{\rm ox}\,\left(V'_G-\varphi_s\right)\,. \tag{22.14}$$

In conclusion, (22.12) provides the relation between surface potential and gate voltage. When $\varphi_s=0$, the electric potential vanishes everywhere in the semiconductor, namely, $V'_G=0$ corresponds to the flat-band condition. When $V'_G>0$, the charge in the gate metal is positive; as a consequence, the left-hand side of (22.12) is positive as well, whence $V'_G>\varphi_s$ and the positive sign must be chosen at the right-hand side. The opposite happens when $V'_G<0$. An example of the $\varphi_s=\varphi_s(V'_G)$ relation is given in Fig. 22.5, showing the normalized surface potential u_s in an MOS capacitor as a function of the normalized gate voltage $u'_G=q\,V'_G/(k_B\,T)$. The semiconductor's doping is of the p type with $N_A=10^{16}\ \mathrm{cm}^{-3}$, corresponding to $2\,u_F=2\,q\,\varphi_F/(k_B\,T)=\log(p_{p0}/n_{p0})\simeq 27.6$.

The $\varphi_s=\varphi_s(V'_G)$ relation lends itself to identifying different functioning regimes of the MOS capacitor. This identification can be carried out more accurately basing upon the values of the electron and hole concentrations at the semiconductor surface, $n_s=n(x=0)$ and $p_s=p(x=0)$. In the nondegenerate conditions considered here, the expressions of the surface concentrations read

$$n_s=n_{p0}\,\exp(u_s)=n_i\,\exp(u_s-u_F)\,,\quad p_s=p_{p0}\,\exp(-u_s)=n_i\,\exp(u_F-u_s)\,, \tag{22.15}$$

where $u_F=(1/2)\,\log(p_{p0}/n_{p0})\simeq\log(N_A/n_i)>0$. Depending on the value of u_s, several functioning regimes are identified, which are listed in Table 22.1. The regimes' designations are given by comparing the carrier concentrations at the

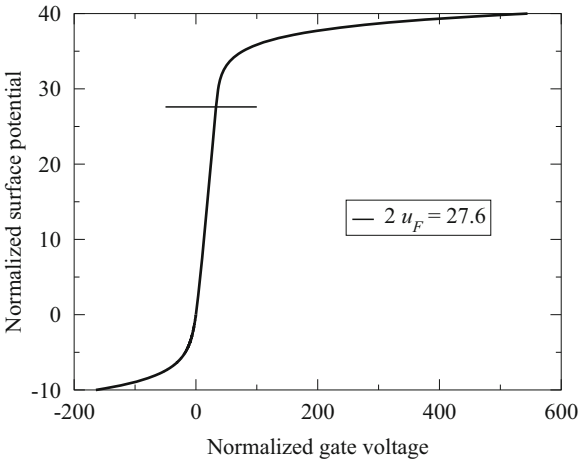


Fig. 22.5 Normalized surface potential u_s in an MOS capacitor with a p-type substrate ($N_A = 10^{16}$ cm^{-3}), as a function of the normalized gate voltage u'_G

Table 22.1 MOS capacitor, p substrate—functioning regimes

Norm. surface potential	Concentrations	Designation
$u_s < 0$	$n_s < n_{p0} < n_i < p_{p0} < p_s$	Accumulation
$u_s = 0$	$n_s = n_{p0} < n_i < p_{p0} = p_s$	Flat band
$0 < u_s < u_F$	$n_{p0} < n_s < n_i < p_s < p_{p0}$	Depletion
$u_s = u_F$	$n_{p0} < n_s = n_i = p_s < p_{p0}$	Mid gap
$u_F < u_s < 2\,u_F$	$n_{p0} < p_s < n_i < n_s < p_{p0}$	Weak inversion
$u_s = 2\,u_F$	$n_{p0} = p_s < n_i < n_s = p_{p0}$	Threshold
$2\,u_F < u_s$	$p_s < n_{p0} < n_i < p_{p0} < n_s$	Strong inversion

surface with the intrinsic and asymptotic ones. When $u_s < 0$ the majority-carrier surface concentration (holes, in the example used here) exceeds the asymptotic one; the regime is called *accumulation*. When $u_s = 0$, both majority- and minority-carrier concentrations equal the corresponding asymptotic concentrations everywhere, and the already-mentioned *flat-band condition* holds. For $0 < u_s < u_F$, the majority-carrier concentration is smaller than the asymptotic one, while the minority-carrier concentration is smaller than the intrinsic one. By continuity, the majority-carrier concentrations is smaller than the asymptotic one not only at the semiconductor's surface, but also in a finite region of width x_d, which is therefore depleted from carriers; for this reason, the condition $0 < u_s < u_F$ is called *depletion regime*, and x_d is called *depletion width*.[3] When $u_s = u_F$, both majority- and minority-carrier concentrations at the surface equal the intrinsic concentration; remembering that in an intrinsic semiconductor the Fermi level practically coincides with the gap's

[3]The depletion width x_d is conceptually the same thing as the extension l_p of the space-charge region on the p side of a metallurgical junction (Sect. 21.2.2). A different symbol is used to avoid confusion in the analysis of the MOSFET (Sect. 22.8).

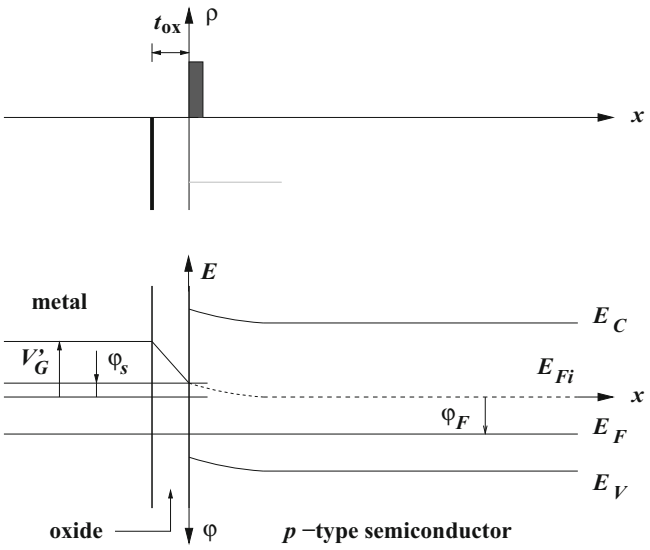


Fig. 22.6 Schematic representation of the charge density and electric potential in a *p*-substrate MOS capacitor in the accumulation regime

midpoint (Sect. 18.3), this regime is called *mid gap*. When $u_F < u_s < 2\,u_F$, the minority-carrier concentration at the surface exceeds that of the majority carriers; however, it is still lower than the asymptotic concentration of the majority carriers: the regime is called *weak inversion*. When $u_s = 2\,u_F$, the surface concentration of the minority carriers equals the asymptotic concentration of the majority carriers, and vice versa; the regime is called *threshold of the strong inversion*, or simply *threshold*. Finally, when $u_s > 2\,u_F$, the minority-carrier concentration at the surface exceeds the asymptotic concentration of the majority carriers, and the regime is called *strong inversion*. In Fig. 22.5 the normalized surface potential at threshold, $2\,u_F$, is marked by the horizontal bar; one notes that in the strong-inversion regime the surface potential rapidly saturates as the gate voltage increases.

The form of the electric potential and charge density is shown in Fig. 22.6 for the accumulation regime. The upper part of the figure shows the charge density, which is schematically represented by a negative charge layer at the metal-oxide interface and by the thicker, positive layer at the semiconductor oxide interface. The lower part of the figure shows the electric potential along with the band structure of the semiconductor; note that two different vertical axes are used, in such a way that energy increases upwards and the electric potential increases downwards. The zero of the electric potential coincides with the horizontal part of the dashed line (in fact, here it is $V'_G < \varphi_s < 0$). The mid-gap condition, $V'_G > 0$, $\varphi_s = \varphi_F$, is illustrated in Fig. 22.7, whose general description is similar to that of Fig. 22.6; here the charge layer on the gate metal is positive and balances the negative charge of the semiconductor. Due to the depletion that occurs in the region adjacent to the

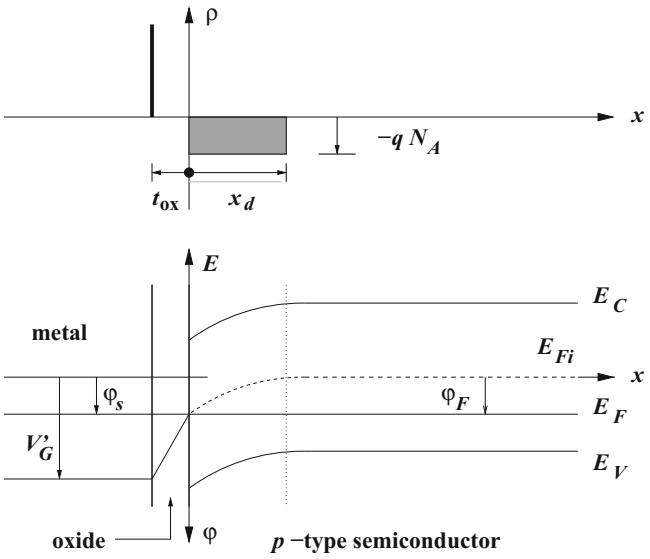


Fig. 22.7 Schematic representation of the charge density and electric potential in a *p*-substrate MOS capacitor in the mid-gap condition

semiconductor–oxide interface, the charge density is dominated by the contribution from the negative acceptor ions, $\varrho \simeq -q\,N_A$. In the figure, the charge density of the semiconductor is schematically indicated by the shaded area, that corresponds to a charge per unit area equal to $-q\,N_A\,x_d$. Finally, Fig. 22.8 shows the form of the electric potential and charge density for the threshold condition, $V'_G > 0$, $\varphi_s = 2\,\varphi_F$. Again, the general description is similar to that of Figs. 22.6 and 22.7; here, there are two contributions to the negative charge of the semiconductor: the first one comes from the contribution of the negative acceptor ions, whose charge density is schematically indicated by the shaded area of width x_d (note that due to the larger value of V_G, the depletion width is larger than that of the mid-gap condition). The second contribution to the semiconductor's charge is due to the electrons, whose concentrations at the interface or near it is sufficiently large to be significant; they form a negative layer, called *inversion layer*, whose width, albeit larger than that of the positive layer located at the metal-oxide interface,[4] is much smaller than x_d.

Numerical solutions of the semiconductor-device model show that, with the exception of the accumulation regime, the semiconductor region of a uniformly doped MOS capacitor can be partitioned into a space-charge and a quasi-neutral region; the quasi-neutral region behaves as an isolated, uniform semiconductor,

[4]The width of the region where the charged layer at the metal-oxide interface is significant is of the order of 1 nm. That of an inversion layer is of the order of 5 nm; an example is given below, with reference to Fig. 22.14.

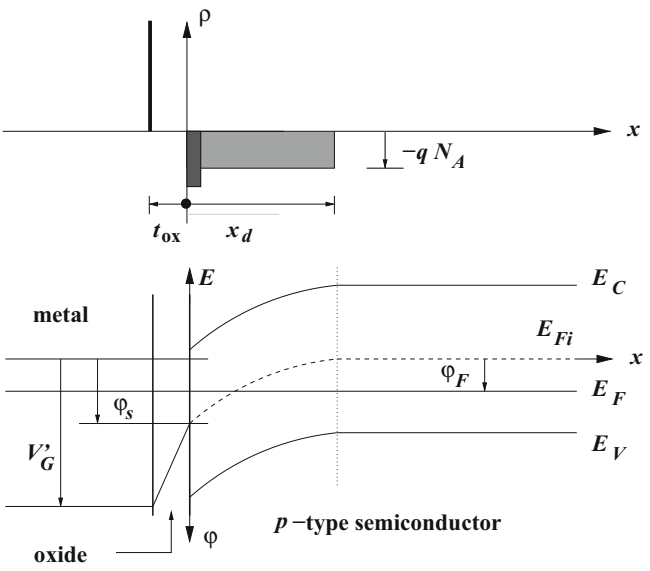


Fig. 22.8 Schematic representation of the charge density and electric potential in a *p*-substrate MOS capacitor at threshold

whereas in the volume of the space-charge[5] region the charge density is essentially dominated by the ionized dopants. In the threshold and strong-inversion regimes, the layer of mobile charges near the semiconductor–oxide interface gives a significant contribution, which must be accounted for; it can be approximated as a charge layer at the interface. Considering the range $\varphi_s > 0$ only, namely, excluding the accumulation regime for the *p*-substrate MOS capacitor, the charge per unit area in the semiconductor is found to be

$$Q_{\rm sc} = \int_0^\infty \varrho \, {\rm d}x \simeq \int_0^{x_d} \varrho \, {\rm d}x \simeq -q \int_0^{x_d} (n + N_A) \, {\rm d}x = Q_i + Q_b \,, \tag{22.16}$$

where the first approximation is due to neglecting the charge of the quasi-neutral region, the second one to neglecting the holes in the space-charge region. Quantities $Q_i, Q_b < 0$ are, respectively, the integral of $-q\,n$ and $-q\,N_A$; they are called *inversion charge per unit area* and *bulk charge per unit area*.

[5] With reference to MOS devices, the space-charge region is also called *depleted region*.

22.3 Capacitance of the MOS Structure—*P*-Type Substrate

The capacitance per unit area of the MOS structure is given by[6]

$$C = \frac{\mathrm{d}Q_\mathrm{m}}{\mathrm{d}V_G} = \frac{\mathrm{d}Q_\mathrm{m}}{\mathrm{d}V'_G}\,. \tag{22.17}$$

Combining (22.17) with (22.12) and (22.14) yields

$$\frac{1}{C} = \frac{\mathrm{d}V'_G}{\mathrm{d}Q_\mathrm{m}} = \frac{\mathrm{d}\left(V'_G - \varphi_s\right) + \mathrm{d}\varphi_s}{\mathrm{d}Q_\mathrm{m}} = \frac{1}{C_\mathrm{ox}} + \frac{\mathrm{d}\varphi_s}{\mathrm{d}(-Q_\mathrm{sc})}\,. \tag{22.18}$$

The above is recast in a more compact form by defining the *semiconductor capacitance per unit area*

$$C_\mathrm{sc} = -\frac{\mathrm{d}Q_\mathrm{sc}}{\mathrm{d}\varphi_s} = -\frac{q}{k_B\,T}\,\frac{\mathrm{d}Q_\mathrm{sc}}{\mathrm{d}u_s} = \pm\frac{\sqrt{2}\,\varepsilon_\mathrm{sc}}{L_A}\,\frac{\mathrm{d}F}{\mathrm{d}u_s} > 0\,, \tag{22.19}$$

where the positive (negative) sign holds for $u_s > 0$ ($u_s < 0$). In conclusion, the capacitance is the series of the oxide and semiconductor capacitances:

$$\frac{1}{C} = \frac{1}{C_\mathrm{ox}} + \frac{1}{C_\mathrm{sc}}\,. \tag{22.20}$$

In (22.20) it is $C_\mathrm{ox} = \text{const}$ while C_sc has a rather complicate dependence on u_s. However, basing on the second equation in (22.5), one may investigate the limiting cases of (22.20). For this, using $\exp(-2\,u_F) = n_i^2/N_A^2$, one finds for the asymptotic behavior of F in a p-substrate device,

$$F \simeq \exp(u_s/2 - u_F)\,, \quad u_s \gg 1\,; \qquad F \simeq \exp(-u_s/2)\,, \quad u_s \ll -1\,. \tag{22.21}$$

When, instead, it is $|u_s| \ll 1$, expanding the exponentials yields $\exp(\pm u_s) \simeq 1 \pm u_s + u_s^2/2$, whence, observing that $\exp(-2\,u_F) \ll 1$,

$$F^2 \simeq \frac{1}{2}\left[1 + \exp(-2\,u_F)\right] u_s^2 \simeq \frac{1}{2}\,u_s^2\,, \qquad F \simeq \pm\frac{u_s}{\sqrt{2}}\,. \tag{22.22}$$

[6]Like in the case of the depletion capacitance of the *p-n* junction (Sect. 21.4), definition $C = \mathrm{d}Q_\mathrm{m}/\mathrm{d}V_G$ is coherent with the choice of the voltage reference described in Sect. 22.2. The units of C are $[C] = \mathrm{F\,m^{-2}}$.

Then, from (22.19) the asymptotic values of $C_{\rm sc}$ are found to be

$$C_{\rm sc} \simeq \frac{1}{\sqrt{2}} \frac{\varepsilon_{\rm sc}}{L_A} \exp(u_s/2 - u_F)\,, \qquad u_s \gg 1\,, \tag{22.23}$$

$$C_{\rm sc} \simeq \frac{1}{\sqrt{2}} \frac{\varepsilon_{\rm sc}}{L_A} \exp(-u_s/2)\,, \qquad u_s \ll -1\,. \tag{22.24}$$

Both limits correspond to the same asymptotic value of the capacitance per unit area,

$$C = \frac{C_{\rm ox}\, C_{\rm sc}}{C_{\rm ox} + C_{\rm sc}} \simeq C_{\rm ox}\,, \qquad |u_s| \gg 1\,. \tag{22.25}$$

Near the origin, instead, one finds

$$C_{\rm sc} \simeq \frac{\varepsilon_{\rm sc}}{L_A}\,, \qquad |u_s| \ll 1\,. \tag{22.26}$$

The limit of C for $u_s \to 0$ is called *flat-band capacitance per unit area*; from (22.26) one finds

$$C \simeq C_{FB} = \frac{C_{\rm ox}}{1 + C_{\rm ox}\, L_A/\varepsilon_{\rm sc}} < C_{\rm ox}\,, \qquad |u_s| \ll 1\,. \tag{22.27}$$

Examples of capacitance's calculations are shown in Fig. 22.9.

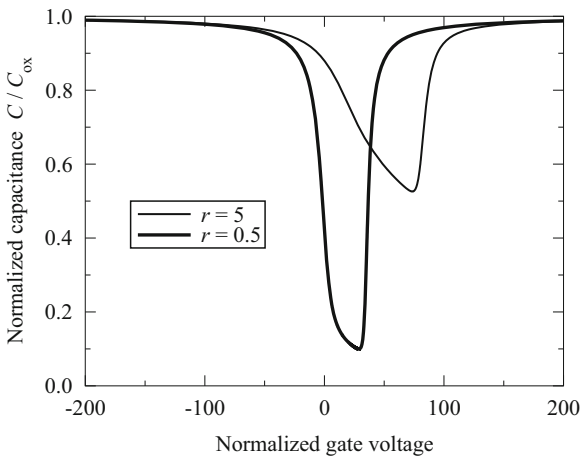


Fig. 22.9 Normalized capacitance $C/C_{\rm ox}$ as a function of the normalized gate voltage u'_G, in a p-substrate MOS capacitor with $N_A = 10^{16}$ cm^{-3}, for different values of $r = \varepsilon_{\rm sc}\, t_{\rm ox}/(\varepsilon_{\rm ox}\, \sqrt{2}\, L_A)$. The details of the calculations are in Prob. 22.1

22.4 Simplified Expression of the Inversion Charge

To the purpose of applying some results of the MOS capacitor's theory to the analysis of MOSFETs, it is convenient to determine a simplified form of the inversion layer's charge, that holds in all the functioning regimes with the exception of accumulation. For this, one starts from the expression of the semiconductor charge per unit area which, combining (22.12) and (22.13), reads

$$Q_{\rm sc} = \pm Q_{\rm sc}^{(1)}\, F(u_s)\,, \qquad Q_{\rm sc}^{(1)} = \varepsilon_{\rm sc}\frac{k_B\,T}{q}\,\frac{\sqrt{2}}{L_A}\,, \tag{22.28}$$

where the negative (positive) sign must be chosen when $u_s > 0$ ($u_s < 0$), and $Q_{\rm sc}^{(1)}$ is the value of $Q_{\rm sc}$ corresponding to $F = 1$. The relation $Q_{\rm sc} = Q_{\rm sc}(u_s)$ is shown in normalized units in Fig. 22.10. In turn, Fig. 22.11 shows, still in normalized form, the individual contributions of electrons, holes, and bulk charge to $F^2 = [Q_{\rm sc}/Q_{\rm sc}^{(1)}]^2$; such contributions are, respectively, $\exp(-2\,u_F)\,[\exp(u_s) - 1]$, $\exp(-u_s) - 1$, and $[1 - \exp(-2\,u_F)]\,u_s$. The contribution of holes dominates for $u_s < 0$, that of the bulk charge dominates for $0 < u_s < 2\,u_F$ and, finally, that of the electrons dominates for $u_s > 2\,u_F$.

When accumulation is excluded, in a p-substrate capacitor one must take $u_s > 0$. The approximate dependence of F on the normalized potential is easily worked out from (22.5), whose limiting case in the depletion and weak-inversion regimes is

$$F \simeq \sqrt{u_s}\,, \qquad 0 < u_s < 2\,u_F\,. \tag{22.29}$$

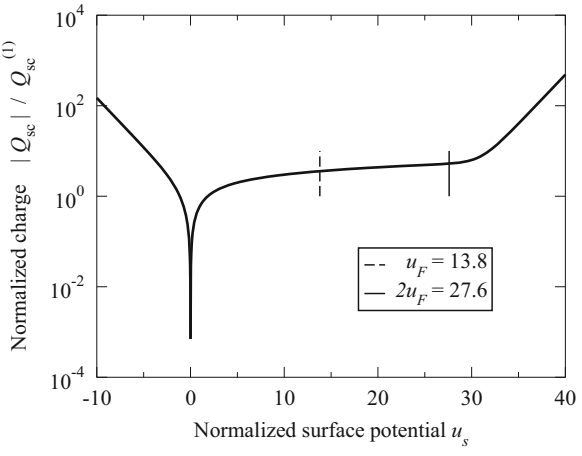


Fig. 22.10 Normalized charge per unit area as a function of the normalized surface potential, in a p-substrate MOS capacitor with $N_A = 10^{16}$ cm^{-3}

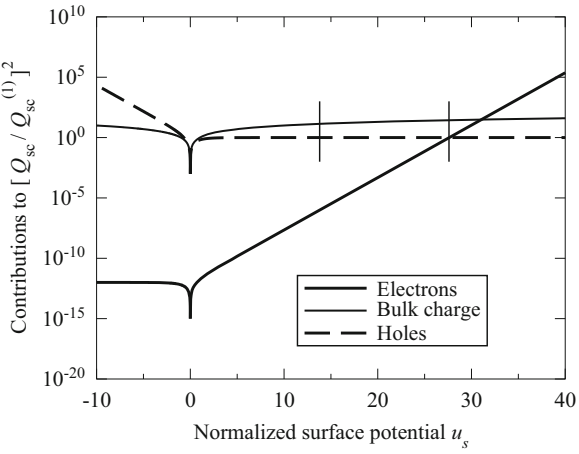


Fig. 22.11 Individual contributions of electrons, holes, and bulk charge to $F^2 = [Q_{sc}/Q_{sc}^{(1)}]^2$, as a function of the normalized surface potential u_s, in a p-substrate MOS capacitor with $N_A = 10^{16}$ cm^{-3}

Introducing (22.29) into (22.28) yields, for $0 < u_s < 2\,u_F$,

$$Q_{sc} \simeq -\frac{\sqrt{2}\,\varepsilon_{sc}\,k_B\,T}{q\,L_A}\,\sqrt{\frac{q\,\varphi_s}{k_B\,T}} = -C_{ox}\,\gamma\,\sqrt{\varphi_s}\,, \qquad \gamma = \frac{\sqrt{2\,\varepsilon_{sc}\,q\,p_{p0}}}{C_{ox}}\,, \tag{22.30}$$

where the expression (21.12) of the Debye length L_A has been used.[7] It is interesting to note that a relation identical to (22.30) is obtained using the full-depletion and ASCE approximations (Sect. 21.4); in fact, letting $\varrho = -q\,N_A$ for $0 < x < x_d$ and $\varrho = 0$ for $x > x_d$ (compare, e.g., with Figs. 22.7 and 22.8) yields a simplified form of the Poisson equation,

$$\varphi'' \simeq \frac{q\,N_A}{\varepsilon_{sc}}\,, \qquad 0 < x < x_d\,. \tag{22.31}$$

The boundary conditions of (22.31) are obtained in the same manner as in the p-n junction, and read $\varphi(x_d) = 0$, $\varphi'(x_d) = 0$; the solution of (22.31) fulfilling the boundary conditions is $\varphi(x) = q\,N_A\,(x - x_d)^2/(2\,\varepsilon_{sc})$. Letting $x = 0$ in the above yields a relation between the surface potential and the depletion width; in turn, in the full-depletion and ASCE approximations the bulk charge per unit area is $Q_b = -q\,N_A\,x_d$. In summary,

$$\varphi_s = \frac{q\,N_A}{2\,\varepsilon_{sc}}\,x_d^2\,, \qquad Q_b = -q\,N_A\,x_d = -\sqrt{2\,\varepsilon_{sc}\,q\,N_A\,\varphi_s}\,. \tag{22.32}$$

Observing that for $0 < u_s < 2\,u_F$ it is $Q_{sc} \simeq Q_b$, and that $N_A \simeq p_{p0}$, one finds that the second relation in (22.32) coincides with the first one in (22.30).

[7] The units of γ are $[\gamma] = \mathrm{V}^{1/2}$.

Combining $Q_{\rm sc} \simeq Q_b$ with (22.30) and with the general expression (22.14) yields $V'_G - \varphi_s = \gamma \sqrt{\varphi_s}$; from it, one finds a simplified $\varphi_s = \varphi_s(V'_G)$ relation, that holds in the depletion and weak-inversion conditions:[8]

$$\sqrt{\varphi_s} = \sqrt{V'_G + (\gamma/2)^2} - \gamma/2\,. \tag{22.33}$$

The contribution of electrons to the semiconductor charge per unit area, $Q_i = -q \int_0^{x_d} n \, {\rm d}x$, becomes relevant from the threshold condition on. Remembering the discussion of Sect. 22.2.2, the electron charge is approximated as a charge layer at the interface, $-q\,n \simeq Q_i\,\delta(x^+)$; as a consequence, the space charge can be considered as entirely due to the ionized dopant atoms also when $\varphi_s > 2\,\varphi_F$, so that (22.31) and (22.32) still hold. From (22.16) one then finds the result sought, that is, a simplified form of the inversion layer's charge of a p-substrate MOS capacitor, that holds in the depletion, weak-inversion, and strong-inversion regimes:

$$Q_i = Q_{\rm sc} - Q_b = -C_{\rm ox}\left[\left(V'_G - \varphi_s\right) - \gamma\sqrt{\varphi_s}\right] < 0\,. \tag{22.34}$$

22.4.1 Flat-Band Voltage

The theory worked out so far was initially based on the assumption that relation (22.1) holds between the gate metal's work function, the semiconductor's affinity, and the position of the semiconductor's Fermi level with respect to the band edges. The assumption was then removed and, to account for the possibility that $W - A \neq E_C - E_F$, the difference $V'_G = V_G - \Phi_{mp}$ has been used to replace V_G when the gate voltage referred to the bulk semiconductor is to be considered, with Φ_{mp} the built-in potential between the bulk contact and the p-type semiconductor (Sect. 22.2.1). This outcome is better understood with the aid of Fig. 22.3, where the electron affinity A of the semiconductor is larger than that used in Fig. 22.3, so that $W - A < E_C - E_F$. As shown in the figure, the distance between the Fermi level of the metal and the vacuum level is prescribed by W; such a distance is the same everywhere in the metal, because the interior of the latter is equipotential. In turn, the distance between the vacuum level and the edge of the semiconductor's conduction band, at the semiconductor–oxide interface, is prescribed by A. This forces the conduction band's edge to come nearer to the Fermi level; in this way, the band's edge shift downwards with respect to its position in the semiconductor bulk, whose distance from the Fermi level (marked with $E_C^\infty - E_F$

[8] The negative sign in front of the square root in (22.33) must be discarded.

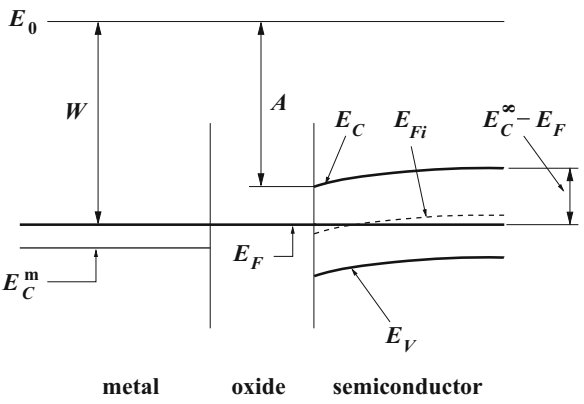


Fig. 22.12 The same materials as in Fig. 22.3, with $W - A < E_C - E_F$. The semiconductor's bands curve downwards near the semiconductor–oxide interface

in the figure) is prescribed by charge neutrality. The equilibrium concentration of electrons, which in the nondegenerate and complete-ionization condition reads $n = N_C \exp[-(E_C - E_F)/(k_B T)]$ (Sect. 18.4), is larger at the semiconductor–oxide interface than in the bulk; in fact, when the materials are connected, a number of electrons flow from the metal to the semiconductor, leaving behind a layer of positive charge at the metal–oxide interface (Fig. 22.12).[9]

The analysis is completed by supposing that a fixed charge density $\varrho_{\text{ox}}(x)$ is present in the oxide. If this is the case, it is not correct to assume that the electric potential is linear within the oxide; however, (22.11) still holds by giving E_{ox} the meaning of average oxide field. The Poisson equation in the oxide is readily solved using the procedure shown in Prob. 4.1; from $-\varepsilon_{\text{ox}}\,\varphi'' = \varrho_{\text{ox}}$ one finds

$$\varepsilon_{\text{ox}}\,\varphi' = \varepsilon_{\text{ox}}\,\varphi'(-t_{\text{ox}}^+) - H\,, \qquad H(x) = \int_{-t_{\text{ox}}}^{x} \varrho_{\text{ox}}(\xi)\,\mathrm{d}\xi\,, \tag{22.35}$$

where the origin has been placed at the semiconductor–oxide interface. Integrating (22.35) by parts, letting $x = 0$, and dividing by t_{ox} yield

$$C_{\text{ox}}\,(V_G' - \varphi_s) = -\varepsilon_{\text{ox}}\,\varphi'(-t_{\text{ox}}^+) - Q_{\text{ox}}\,, \qquad Q_{\text{ox}} = \int_{-t_{\text{ox}}}^{0} \frac{\xi}{t_{\text{ox}}}\,\varrho_{\text{ox}}(\xi)\,\mathrm{d}\xi\,, \tag{22.36}$$

[9] When the materials are put together, the flow of electrons from the metal to the semiconductor, or vice versa, occurs thanks to the electric connection between the gate and bulk contacts provided by the external voltage generator. This connection makes it possible to establish the condition of thermal equilibrium of the whole system. If the contacts were left open, no flow of charge would be possible (apart from the occurrence of tunneling events across the oxide which, given enough time, would be able to restore equilibrium), and the Fermi level of each material would keep the value it possessed when the materials were still separate; this situation would be similar to that of two thermal reservoirs at different temperatures, separated by a thermal insulator.

where $\varphi(0) = \varphi_s$ and $\varphi(-t_{\rm ox}) = V'_G$ have been used. On the other hand, applying (22.8) to the metal–insulator interface and using (22.10), one finds $-\varepsilon_{\rm ox}\,\varphi'(-t_{\rm ox}^+) = Q_{\rm m}$ whence $Q_{\rm m} = C_{\rm ox}\,(V'_G + Q_{\rm ox}/C_{\rm ox} - \varphi_s)$. From $V'_G = V_G - \Phi_{mp}$ it follows

$$V'_G + \frac{Q_{\rm ox}}{C_{\rm ox}} = V_G - V_{FB}\,, \qquad V_{FB} = \Phi_{mp} - \frac{Q_{\rm ox}}{C_{\rm ox}}\,. \tag{22.37}$$

The constant voltage V_{FB} is called *flat-band voltage*. The designation is due to the fact that letting $V_G = V_{FB}$ establishes the flat-band condition in the semiconductor. It is possible that a fixed layer of charge is also present in the oxide, in addition to $\varrho_{\rm ox}$; its effect on V_{FB} is readily incorporated into the analysis leading to (22.37) by adding to $\varrho_{\rm ox}$ a term of the form $Q_f\,\delta(x-h)$, with Q_f the charge per unit area of the layer (compare with Prob. 4.5).

Definition (22.37) of V_{FB} does not account for the fact that, besides the fixed charge, mobile charges may be present within the oxide or at the semiconductor–oxide interface. The former type of mobile charges is typically made of contaminants (Sect. 24.1); they may become mobile at temperatures slightly larger than room temperature, thus making V_{FB} unstable. Consequently, the threshold voltage (22.85) or (22.113) of the MOS transistor becomes unstable as well. The charge at the semiconductor–oxide interface is due in turn to a distribution of interfacial electronic states, called *surface states*, whose energy belongs to the band gap. Such states are due to the disruption of periodicity caused by the finiteness of the crystal's size. In a bare silicon crystal the surface concentration of surface states turns out to be equal to that of the atoms themselves [4]; in silicon, the latter is about $(5.0\times10^{22})^{2/3} \simeq 1.36\times10^{15}$ cm^{-2} (Sect. 24.2). When the semiconductor–oxide interface is considered, instead, the concentration of surface states is much smaller; the typical figures for the silicon–oxide interface range from 10^9 to 10^{11} cm^{-2}, depending on the quality of the oxide growth's process and on the subsequent annealing steps.

In the following it is assumed that the effect of the mobile charges described here is negligible. Also, the analysis carried out in this section shows that the presence of a fixed charge density $\varrho_{\rm ox}$ merely shift V'_G by a constant amount. For this reason, from now on symbol V'_G will be given the more general meaning $V'_G = V_G - V_{FB}$.

22.4.2 Quantitative Relations in the MOS Capacitor

In the p-type silicon substrate considered so far it is $p_{p0} \simeq N_A = 10^{16}$ cm^{-3} and, at room temperature, $n_i \simeq 10^{10}$ cm^{-3}. In turn, the asymptotic minority-carrier concentration is $n_{p0} = n_i^2/N_A \simeq 10^4$ cm^{-3}; it follows $\exp(-2\,u_F) = n_{p0}/p_{p0} \simeq 10^{-12}$ and, as shown, e.g., in Fig. 22.5, $2\,u_F \simeq 27.6$. Using $k_B\,T/q \simeq 26$ mV then yields $2\,\varphi_F \simeq 0.72$ V. As $\varepsilon_{\rm sc} \simeq 11.7\times8.854\times10^{-14} = 1.036\times10^{-12}$ F cm^{-1}, one finds

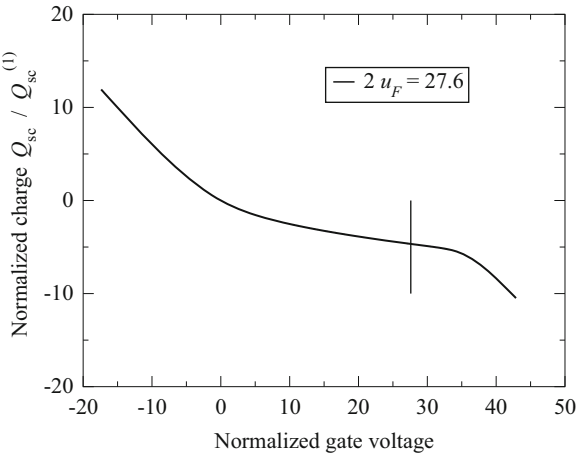


Fig. 22.13 Normalized semiconductor charge $Q_{sc}/Q_{sc}^{(1)}$ as a function of the normalized gate voltage u'_G, for a p-substrate MOS capacitor with $N_A = 10^{16}$ cm^{-3}

$$L_A = \sqrt{\frac{2\,\varepsilon_{sc}\,k_B\,T}{q^2\,p_{p0}}} \simeq 5.8 \times 10^{-2}\ \ \mu\text{m}\,, \tag{22.38}$$

$$Q_{sc}^{(1)} = \frac{2\,\varepsilon_{sc}\,k_B\,T}{q\,L_A} = \sqrt{2\,\varepsilon_{sc}\,k_B\,T\,p_{p0}} \simeq 9.3 \times 10^{-9}\ \ \text{C cm}^{-2}\,. \tag{22.39}$$

The relation $Q_{sc}(V'_G)$ is found from $Q_{sc} = -C_{ox}\,[V'_G - \varphi(V'_G)]$, where $\varphi(V'_G)$ is obtained from (22.12). The result is shown in normalized form in Fig. 22.13.

Note that the results illustrated so far have been obtained without the need of integrating (22.6). The result of a numerical integration of (22.6) is shown in Fig. 22.14, where the dependence on position of n and $N_A - p$ is drawn for a p-substrate MOS capacitor with $N_A = 10^{16}$ cm^{-3} in the strong-inversion regime ($u_s = 2.5\ u_F$). The term $(N_A - p)/p_{p0}$ is significant in a surface region of the semiconductor, whose thickness is several units of x/L_A. The term n/p_{p0} is much larger, but only in a much thinner region near the surface. If the width of the inversion layer is conventionally taken at the intersection between the two curves of Fig. 22.14, that occurs at $x/L_A \simeq 0.1$, one finds from (22.38) a width of about 5 nm.

With reference to Fig. 22.2, and using aluminum as metal and silicon as semiconductor, one has [97, Sect. 8.1] $W = 4.1$, $A = 4.05$, and $A^{ox} = 0.95$ eV. The band gap of thermally grown silicon dioxide is about 9 eV.

22.5 MOS Photocapacitor

The functioning of the *MOS photocapacitor* is easily understood by combining the analysis of the MOS capacitor, carried out in Sect. 22.2, with that of the storage-mode photodiode, carried out in Sect. 21.6.2. A one-dimensional sketch

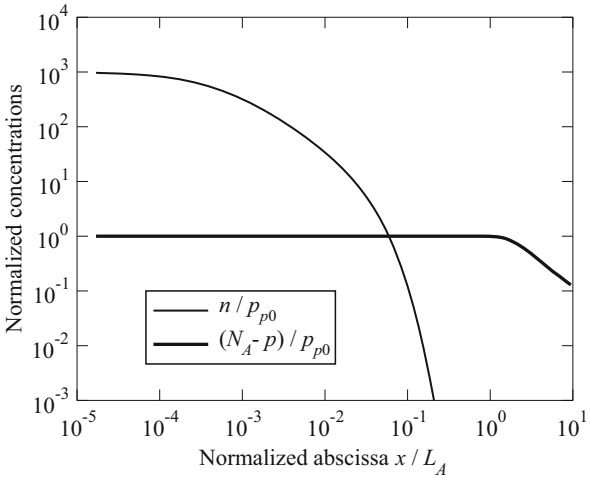


Fig. 22.14 Normalized concentrations n/p_{p0} and $(N_A - p)/p_{p0}$ as a function of position x/L_A, for a p-substrate MOS capacitor with $N_A = 10^{16}$ cm^{-3} in strong inversion ($u_s = 2.5\, u_F$)

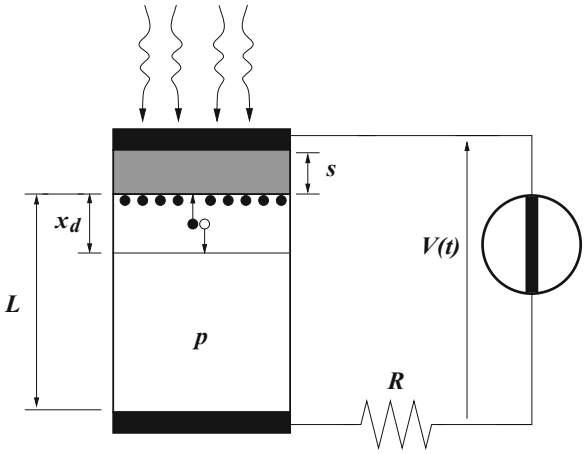


Fig. 22.15 One-dimensional sketch of the MOS photocapacitor with a uniform, p-type substrate. The relative thicknesses of the different layers shown in the figure are not realistic

of the device is shown in Fig. 22.15, where the case of a uniform, p-substrate capacitor is considered. The impinging photons are sketched by the wavy arrows, while the black dots indicate the electrons and the white dot indicates a hole. The considerations about the use of a one-dimensional structure and of a partially transparent contact are the same as in Sect. 21.6. The resistor shown in the figure mimics the input resistance of the circuit that measures the photocapacitor's current.

The applied voltage $V(t)$ is periodically switched between the flat-band voltage V_{FB} and a value V_{GI} that, in the steady-state condition, corresponds to strong inversion. The time extent T_s during which the applied voltage equals V_{FB} is called *sampling time*, while the time extent T_i during which the applied voltage equals V_{GI} is called *integration time*. It is assumed that prior to the switching from V_{FB} to

V_{GI} the earlier transients have vanished, so that the net local charge is negligible.[10] As the applied voltage has a step-like form, the majority carries (holes) are pushed away from the semiconductor–insulator interface and a depleted region of width x_d is formed in a negligible time after the application of the gate voltage; observing that the inversion layer is not present yet, the value of x_d is found by combining the first relation in (22.32) with (22.33), namely,

$$x_d = \sqrt{\frac{2\,\varepsilon_{\rm sc}}{q\,N_A}\,\varphi_s}\,, \qquad \sqrt{\varphi_s} = \sqrt{V_{GI} - V_{FB} + (\gamma/2)^2} - \gamma/2\,, \tag{22.40}$$

with N_A the constant concentration of the acceptor dopant. Due to the absence of the inversion layer, the value of x_d rendered by (22.40) is considerably larger than the one that would correspond to the equilibrium condition. During the integration time, an inversion layer is built up by the combined contributions of the thermal and optical generations in the depleted region; thus, x_d gradually decreases while the modulus of the inverted charge increases. When the gate voltage is switched back to V_{FB}, the majority carriers flood the depleted region, which disappears in a time of the order of that required by a hole to reach the interface; as this time is negligible, the form of the inversion layer is not initially affected by the process. On the other hand, as the disappearing of the depleted region makes the electric field to vanish, the electrons of the inversion layer diffuse toward the substrate and, at the same time, recombine with the holes; this, in turn, produces a current in the external circuit. Under the typical operating conditions of the device, the contribution of optical generation to the inversion layer is dominant with respect to that of thermal generation; it follows that the measured current is essentially due to the illumination, whereas the effect of thermal generation superimposed to it can be considered as a noise signal.

Like in the case of the storage-mode photodiode (Sect. 21.6.2), it is important to identify the order of magnitude of the integration and sampling times. While the former can be made relatively small, the latter cannot be reduced below the limit necessary for the inversion layer to disappear completely; if, in fact, at the moment of the gate voltage's switching from V_{FB} to V_{GI}, a fraction of the electron charge was still present in the semiconductor, it would add to the one generated in the next integration cycle, thus introducing an error in the measured signal.

To analyze the semiconductor's behavior during the sampling time one must solve the minority-carrier continuity equation, whose one-dimensional form reads

$$\frac{\partial n}{\partial t} + \frac{n - n_{p0}}{\tau_n} - \frac{1}{q}\,\frac{\partial J_n}{\partial x} = G_s\,\exp(-k\,x)\,. \tag{22.41}$$

[10] This charge would be exactly zero if the illumination was missing.

In the above it is assumed for simplicity that the weak-injection condition holds (Sect. 20.2.3), and that the absorption coefficient k (Sect. 20.4) is constant with respect to both position and radiation frequency. Remembering (20.67), it is $G_s = \eta\, \Phi_B\, k\, \exp(-k\, s)$, where Φ_B is the flux density of the photons on the inside edge of the contact–oxide interface; the exponential $\exp(-k\, s)$ accounts for the absorption within the oxide, whose thickness is s (the origin is placed at the semiconductor–oxide interface). The integral of the minority-carrier excess concentration $n - n_{p0}$ over the semiconductor thickness L provides the excess charge Q per unit area at time t; observing that $J_n(x = 0) = 0$ due to the presence of the oxide, and that L is large enough to yield $J_n(L) \simeq J_n(\infty) = 0$, the integration of (22.41) over L renders

$$\frac{\mathrm{d}Q}{\mathrm{d}t} + \frac{Q}{\tau_n} = q\,\frac{G_s}{k}\,, \qquad Q = \int_0^L q\,\left(n - n_{p0}\right)\,\mathrm{d}x\,. \tag{22.42}$$

Let $t = 0$ be the instant at which the gate voltage V_G is brought from $V_G(0^-) = V_{GI}$ to $V_G(0^+) = V_{FB}$; from (22.42) one finds[11]

$$Q(t) - q\,\tau_n\,\frac{G_s}{k} = \left[Q_0 - q\,\tau_n\,\frac{G_s}{k}\right]\exp(-t/\tau_n)\,, \tag{22.43}$$

with Q_0 the excess minority charge per unit area at $t = 0$.

The above result shows that the approach to the asymptotic condition[12] is dominated by the minority-carrier lifetime τ_n. The process can be made considerably shorter by building the photocapacitor in a p-type epitaxial layer[13] of thickness a grown over an n-type substrate (Fig. 22.16). From a qualitative standpoint, the advantage of this implementation is that the heavily doped n^+ substrate acts as a contact, so that the excess charge is expected to vanish in the short time that is necessary to diffuse across a. On the other hand, in this case the asymptotic condition $J_n(L) = 0$ does not hold anymore; to determine the time decay of Q one must then solve the continuity equation

$$\frac{\partial n}{\partial t} + \frac{n - n_{p0}}{\tau_n} - D_n\,\frac{\partial^2 n}{\partial x^2} = G_s\,\exp(-k\,x)\,, \tag{22.44}$$

where the drift term of the current has been neglected due to the vanishing of the electric field discussed above. At $t = 0$, the excess minority charge forms a layer at the semiconductor–oxide interface, so that the initial condition of (22.44) reads $n(x, 0) - n_{p0} = Q_0\,\delta(x^+)$. The boundary condition at $x = 0$ is $\partial n/\partial x = 0$ at all

[11] It is assumed that the minimum period of variation of the photon flux density is small with respect to τ_n, so that in the integration of (22.42) G_s can be considered a constant.

[12] Note that in the asymptotic limit the excess charge does not vanish; in fact it is $Q(\infty) = q\,\tau_n\,G_s/k$.

[13] Epitaxy is described in Sect. 24.6.

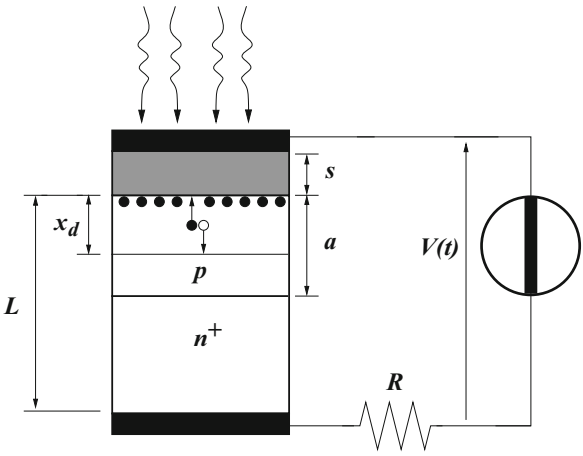


Fig. 22.16 One-dimensional sketch of the MOS photocapacitor built in a *p*-type epitaxial layer. The relative thicknesses of the different layers shown in the figure are not realistic

times, equivalent to $J_n(0) = 0$, while the boundary condition at $x = a$ is $n(a,t) = n_{p0}$ because the heavily doped substrate acts as a contact. Introducing the auxiliary unknown $f = (n - n_{p0})\,\exp(t/\tau_n)$ transforms (22.44) into

$$\frac{\partial f}{\partial t} = D_n \frac{\partial^2 f}{\partial x^2} + G_s\,\exp(-k\,x + t/\tau_n)\,, \tag{22.45}$$

whose initial condition is $f(x,0) = Q_0\,\delta(x^+)$ while the boundary conditions are $\partial f/\partial x = 0$ and $f(a,t) = 0$. To solve the above equation one may resort to a Fourier expansion; to this purpose, let ω be a function defined in $-2\,a \le x \le 2\,a$ and even in such an interval: $\omega(-x,t) = \omega(x,t)$; moreover, let $\omega(x,t) = f(x,t)$ in the interval of interest, $0 \le x \le a$. As ω is even, its Fourier expansion reads

$$\omega = \sum_{k=0}^{\infty} \chi_k(t)\,\gamma_k(x)\,, \quad \chi_k = \frac{1}{2\,a}\int_{-2\,a}^{+2\,a} \gamma_k\,\omega\;\mathrm{d}x\,, \quad \gamma_k = \cos\left(k\,\pi\,\frac{x}{2\,a}\right)\,. \tag{22.46}$$

Clearly ω fulfills the boundary condition $\partial\omega/\partial x = 0$ at $x = 0$, but not the boundary condition $\omega(a) = 0$. To fulfill the latter, one must restrict the choice to those ω such that $\omega(2\,a - x) = -\omega(x)$ for $0 \le x \le a$, namely, that are odd in the interval of length $2\,a$ centered at $x = a$. Letting $\gamma_k(2\,a - x) = -\gamma_k(x)$ it is found $\cos(k\,\pi) = -1$, that is, $k = 2\,i + 1$ with $i = 0, 1, 2, \ldots$; defining $g_i = \cos[(i + 1/2)\,\pi\,x/a]$ and using the properties of ω and g_i then yields

$$c_i = \frac{1}{2\,a}\int_{-2\,a}^{+2\,a} g_i\,\omega\,\mathrm{d}x = \frac{1}{a}\int_0^{2\,a} g_i\,\omega\,\mathrm{d}x = \frac{2}{a}\int_0^{a} g_i\,\omega\,\mathrm{d}x\,. \tag{22.47}$$

As $f = \omega$ in the interval $0 \le x \le a$, the Fourier expansion of f reads

$$f = \sum_{i=0}^{\infty} c_i(t)\, g_i(x)\,, \qquad c_i = \frac{2}{a}\int_0^a f\, \cos\left(\frac{x}{L_i}\right) \mathrm{d}x\,, \tag{22.48}$$

with $1/L_i = (i + 1/2)\,\pi/a$. From $g_i = \cos[(i + 1/2)\,\pi\, x/a]$ it follows $g_i'(0) = 0$, $g(a) = 0$, and $g_i'' = -g_i/L_i^2$; also,

$$\int_0^a g_i \frac{\partial f}{\partial t}\,\mathrm{d}x = \frac{\mathrm{d}}{\mathrm{d}t}\int_0^a g_i f\,\mathrm{d}x = \frac{a}{2}\frac{\mathrm{d}c_i}{\mathrm{d}t} \tag{22.49}$$

and, integrating by parts and using the boundary conditions,

$$\int_0^a g_i \frac{\partial^2 f}{\partial x^2}\,\mathrm{d}x = \int_0^a \frac{\mathrm{d}^2 g_i}{\mathrm{d}x^2} f\,\mathrm{d}x = -\frac{1}{L_i^2}\int_0^a g_i f\,\mathrm{d}x = -\frac{a}{2}\frac{c_i}{L_i^2}\,. \tag{22.50}$$

Multiplying by g_i both sides of (22.45), integrating from 0 to a, and using (22.49), (22.50) finally provide a first-order differential equation for c_i:

$$\frac{\mathrm{d}c_i}{\mathrm{d}t} + \frac{c_i}{\tau_i} = \alpha_i\, \exp(t/\tau_n)\,, \quad \tau_i = \frac{L_i^2}{D_n}\,, \quad \alpha_i = \frac{2}{a}\int_0^a g_i\, G_s \exp(-k\,x)\,\mathrm{d}x\,, \tag{22.51}$$

with the initial condition

$$c_i(0) = \frac{2}{a}\int_0^a g_i f(x,0)\,\mathrm{d}x = \frac{2}{a}\int_0^a g_i\, Q(0)\,\delta(x^+)\,\mathrm{d}x = \frac{2}{a}\,Q_0\,, \tag{22.52}$$

independent of i. Integrating (22.51) yields

$$c_i(t)\, \exp(t/\tau_i) - c_i(0) = \alpha_i\, \tau_i^\star\, \left[\exp(t/\tau_i^\star) - 1\right]\,, \qquad \frac{1}{\tau_i^\star} = \frac{1}{\tau_n} + \frac{1}{\tau_i} \tag{22.53}$$

whence, turning to the original unknown $n - n_{p0} = f\, \exp(-t/\tau_n)$,

$$n - n_{p0} = \sum_{i=0}^{\infty} g_i\, \left[\alpha_i\, \tau_i^\star + \left(c_i(0) - \alpha_i\, \tau_i^\star\right)\, \exp(-t/\tau_i^\star)\right]\,. \tag{22.54}$$

Like in the case of the uniform substrate, the excess charge per unit area $Q(t) = q\int_0^a (n - n_{p0})\,\mathrm{d}x$ does not vanish for $t \to \infty$ as long as $G_s \neq 0$. Using $\tau_i = L_i^2/D_n$ and the definition of the minority-carrier diffusion length $L_n = \sqrt{\tau_n\, D_n}$ one finds

$$\tau_i = \tau_n \frac{L_i^2}{L_n^2} = \tau_n \frac{a^2}{[\pi\, L_n\, (i + 1/2)]^2}\,. \tag{22.55}$$

The slowest decay rate associated with (22.55) is much shorter than that found in the case of a uniform substrate (Prob. 22.14).

22.6 MOS Capacitor—*N*-Type Substrate

The theory of the *n*-substrate MOS capacitor is similar to that of the *p*-substrate one, and is briefly illustrated in this section. A constant concentration N_D of donor atoms is assumed, such that the nondegeneracy and complete-ionization conditions hold. The value of the electric potential is set to zero in the bulk of the semiconductor, $\varphi(\infty) = 0$; the asymptotic conditions for the electron and hole concentrations then read

$$n^{(0)} = n(+\infty) = n_{n0} \simeq N_D\,, \qquad p^{(0)} = p(+\infty) = p_{n0} \simeq \frac{n_i^2}{N_D}\,. \tag{22.56}$$

The Poisson equation in the semiconductor takes the form (compare with (22.3))

$$u'' = \frac{1}{L_D^2}\,A(u)\,, \qquad A(u) = \exp(u) - 1 + \frac{n_i^2}{N_D}\,[1 - \exp(-u)]\,, \tag{22.57}$$

with $u = q\,\varphi/(k_B\,T)$ the normalized electric potential and L_D the Debye length for the electrons defined in (21.9). The equation is solved following the same procedure as in Sect. 22.2.1 to yield

$$u' = \pm\frac{\sqrt{2}}{L_D}\,F(u)\,, \qquad F = \sqrt{\exp(u) - 1 - u + \frac{n_i^2}{N_D^2}\,[u + \exp(-u) - 1]}\,, \tag{22.58}$$

where the sign must be found on a case-by-case basis. Observing that (22.9) and (22.11) hold irrespective of the type of substrate[14] one finds, with the usual meaning of symbols, the relation between surface potential and gate voltage in the *n*-substrate MOS capacitor:

$$C_{\rm ox}\,(V_G' - \varphi_s) = \mp\varepsilon_{\rm sc}\,\frac{k_B\,T}{q}\,\frac{\sqrt{2}}{L_D}\,F(\varphi_s)\,. \tag{22.59}$$

When $\varphi_s = 0$, the electric potential vanishes everywhere in the semiconductor, namely, $V_G' = 0$ corresponds to the flat-band condition. When $V_G' > 0$, the charge in the gate metal is positive; as a consequence, the left-hand side of (22.59) is positive

[14] Here, however, $Q_{\rm sc}$ is the integral of $q\,(p - n + N_D)$, instead of $q\,(p - n - N_A)$, over the semiconductor domain.

Table 22.2 MOS capacitor, *n* substrate—functioning regimes

Norm. surface potential	Concentrations	Designation
$u_s > 0$	$p_s < p_{n0} < n_i < n_{n0} < n_s$	Accumulation
$u_s = 0$	$p_s = p_{n0} < n_i < n_{n0} = n_s$	Flat band
$u_F < u_s < 0$	$p_{n0} < p_s < n_i < n_s < n_{n0}$	Depletion
$u_s = u_F$	$p_{n0} < p_s = n_i = n_s < n_{n0}$	Mid gap
$2\,u_F < u_s < u_F$	$p_{n0} < n_s < n_i < p_s < n_{n0}$	Weak inversion
$u_s = 2\,u_F$	$p_{n0} = n_s < n_i < p_s = n_{n0}$	Threshold
$u_s < 2\,u_F$	$n_s < p_{n0} < n_i < n_{n0} < p_s$	Strong inversion

as well, whence $V'_G > \varphi_s$ and the positive sign must be chosen at the right-hand side. The opposite happens when $V'_G < 0$.

Like in the *p*-substrate case, the $\varphi_s = \varphi_s(V'_G)$ relation lends itself to identifying different functioning regimes of the MOS capacitor. The identification is carried out basing upon the values of the electron and hole concentrations at the semiconductor surface, $n_s = n(x = 0)$ and $p_s = p(x = 0)$. In the nondegenerate conditions considered here, the expressions of the surface concentrations are given by (22.15), where $u_F = (1/2)\,\log(p_{n0}/n_{n0}) \simeq \log(n_i/N_D) < 0$. The functioning regimes are listed in Table 22.2; their designations are given by comparing the carrier concentrations at the surface with the intrinsic and asymptotic ones. The discussion about functioning regimes, depletion width, form of the electric potential and charge density, and differential capacitance is similar to that of the *p*-substrate case.

To obtain a simplified form of the inversion layer's charge that holds in all the functioning regimes with the exception of accumulation, one recasts (22.59) as

$$Q_{\rm sc} = \pm Q_{\rm sc}^{(1)}\,F(u_s)\,, \qquad Q_{\rm sc}^{(1)} = \varepsilon_{\rm sc}\,\frac{k_B\,T}{q}\,\frac{\sqrt{2}}{L_D}\,, \tag{22.60}$$

where the negative (positive) sign must be chosen when $u_s > 0$ ($u_s < 0$), and $Q_{\rm sc}^{(1)}$ is the value of $Q_{\rm sc}$ corresponding to $F = 1$. The relation $Q_{\rm sc} = Q_{\rm sc}(u_s)$ is shown in normalized units in Fig. 22.17. When accumulation is excluded, in an *n*-substrate capacitor one must take $u_s < 0$. The approximate dependence of F on the normalized potential is easily worked out from (22.58), whose limiting case in the depletion and weak-inversion regimes is

$$F \simeq \sqrt{-u_s}\,, \qquad 2\,u_F < u_s < 0\,. \tag{22.61}$$

Introducing (22.61) into (22.60) yields, for $2\,u_F < u_s < 0$ (compare with (22.30)),

$$Q_{\rm sc} \simeq \frac{\sqrt{2}\,\varepsilon_{\rm sc}\,k_B\,T}{q\,L_D}\,\sqrt{-\frac{q\,\varphi_s}{k_B\,T}} = C_{\rm ox}\,\gamma\,\sqrt{-\varphi_s}\,, \qquad \gamma = \frac{\sqrt{2\,\varepsilon_{\rm sc}\,q\,n_{n0}}}{C_{\rm ox}}\,. \tag{22.62}$$

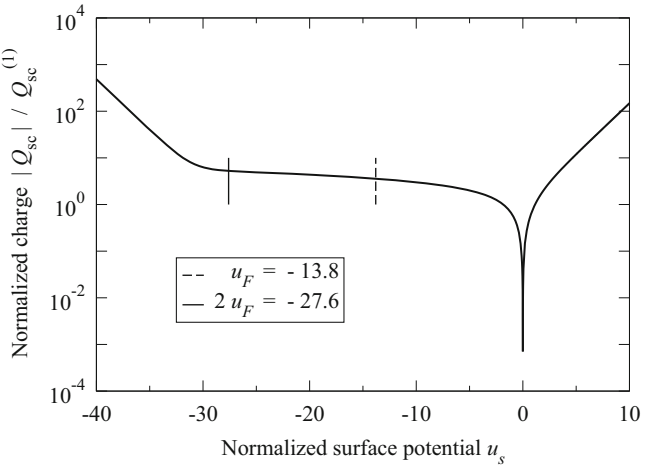


Fig. 22.17 Normalized charge per unit area as a function of the normalized surface potential, in a n-substrate MOS capacitor with $N_D = 10^{16}$ cm^{-3}

A relation identical to (22.62) is obtained using the full-depletion and ASCE approximations (Sect. 21.4); in fact, letting $\varrho = q N_D$ for $0 < x < x_d$ and $\varrho = 0$ for $x > x_d$ yields a simplified form of the Poisson equation in $0 < x < x_d$, namely, $\varphi'' \simeq -q N_D/\varepsilon_{sc}$. Solving the latter using $\varphi(x_d) = 0$, $\varphi'(x_d) = 0$ as boundary conditions yields $\varphi(x) = -q N_D (x - x_d)^2/(2\,\varepsilon_{sc})$. Letting $x = 0$ in the above provides a relation between the surface potential and the depletion width; in turn, in the full-depletion and ASCE approximations the bulk charge per unit area is $Q_b = q N_D x_d$. In summary,

$$-\varphi_s = \frac{q N_D}{2\,\varepsilon_{sc}} x_d^2\,, \qquad Q_b = q N_D x_d = \sqrt{-2\,\varepsilon_{sc}\, q N_D\, \varphi_s}\,. \tag{22.63}$$

Observing that for $2\,u_F < u_s < 0$ it is $Q_{sc} \simeq Q_b$, and that $N_D \simeq n_{n0}$, one finds that the second relation in (22.63) coincides with the first one in (22.62). Combining $Q_{sc} \simeq Q_b$ with (22.62) and with the general expression $Q_{sc} = -C_{ox}\,(V'_G - \varphi_s)$ yields $V'_G - \varphi_s = -\gamma\,\sqrt{-\varphi_s}$; from it, one finds a simplified $\varphi_s = \varphi_s(V'_G)$ relation that holds in the depletion and weak-inversion conditions:[15]

$$\sqrt{-\varphi_s} = \sqrt{-V'_G + (\gamma/2)^2} - \gamma/2\,, \qquad V'_G < 0\,. \tag{22.64}$$

The contribution of holes to the semiconductor charge per unit area, $Q_i = q \int_0^{x_d} p\ \mathrm{d}x$, becomes relevant from the threshold condition on. In the same manner as for the electrons (Sect. 22.2.2), the hole charge is approximated as a charge layer at

[15] The negative sign in front of the square root in (22.64) must be discarded.

the interface, $q\,p \simeq Q_i\,\delta(x^+)$; as a consequence, the space charge can be considered as entirely due to the ionized dopant atoms also when $\varphi_s < 2\,\varphi_F < 0$, so that (22.63) still hold. From $Q_{\rm sc} = Q_i + Q_b$ one then finds the result sought, that is, a simplified form of the inversion layer's charge of an n-substrate MOS capacitor, that holds in the depletion, weak-inversion, and strong-inversion regimes:

$$Q_i = Q_{\rm sc} - Q_b = C_{\rm ox}\left[\left(-V'_G + \varphi_s\right) - \gamma\,\sqrt{-\varphi_s}\right] > 0\,. \tag{22.65}$$

22.7 Insulated-Gate Field-Effect Transistor—MOSFET

The principle of the insulated-gate, field-effect transistor (IGFET) was demonstrated in the early 1930s [128, Chap. 10]. The first structures using a thermally oxidized silicon layer were fabricated in 1960. The IGFET architecture using silicon dioxide as gate dielectric is more commonly called MOSFET. This device architecture, jointly with the continuous improvements in the silicon technology, made it possible the tremendous progress in the fabrication of integrated circuits during the last decades, and is the most common component in digital and analog circuits.

The electric current in a MOSFET is transported by one type of carriers only, electrons or holes; for this reason the device is called *unipolar*. In a p-substrate device, the carriers are the electrons that form the charge layer at the semiconductor–insulator interface; therefore, this type of transistor is called *n-channel* MOSFET. Conversely, in an n-substrate device the carriers are holes, and the transistor is called *p-channel* MOSFET. The schematic cross-section of an *n-channel* MOSFET is shown in Fig. 22.18. The starting point is a p-type silicon substrate, with an $N_A = $ const dopant concentration, onto which a layer of silicon dioxide is thermally grown and patterned; then, the *gate contact* (G) is deposited. The extension L_g of the gate metal in the horizontal direction is called *geometrical length* of the gate. The next step is the introduction of a heavy dose of an n-type dopant on the two sides of the gate. As shown in the figure, lateral diffusion (Sect. 23.8.3) makes the gate oxide to partially overlap the n-doped regions.[16]

The metallizations of the n^+ regions provide two more contacts, called *source* (S) and *drain* (D); the bottom metal layer contacting the p-type substrate is indicated with *bulk* (B), and the term *channel* denotes the interfacial semiconductor region between the two junctions. To distinguish the applied voltages from one another, two letters are used; considering the bulk metallization as the reference contact, in an n-channel MOSFET a typical choice of the three independent voltages is $V_{GB} = V_G - V_B$, $V_{SB} = V_S - V_B$, and $V_{DB} = V_D - V_B$. As the standard MOSFET architecture is structurally symmetric, it is not possible to distinguish the source contact from

[16]The n-type regions are typically obtained by ion implantation, whose lateral penetration is limited. However, ion implantation is followed by a thermal process (*annealing*), during which thermal diffusion takes place.

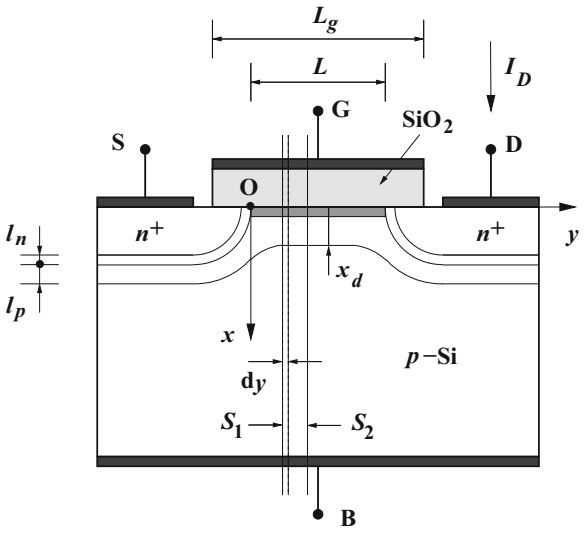


Fig. 22.18 Cross-section of an *n*-channel MOSFET. The *black areas* are the metal contacts

the drain contact basing on geometry or dopant distribution: the distinction is to be based on the applied voltages; in fact, in the typical operating regime of the device the source-bulk and drain-bulk junctions are never forward biased, whence in an *n*-channel MOSFET it is $V_{SB} \geq 0$ and $V_{DB} \geq 0$. The drain contact is identified[17] by the condition $V_{DB} - V_{SB} = V_{DS} > 0$.

22.8 *N*-Channel MOSFET—Current-Voltage Characteristics

To work out the theory of the MOSFET, one introduces a reference whose x axis is normal to the semiconductor–insulator interface, while the y axis is parallel to it. The origin (O) is placed at the intersection of the source *p-n* junction and the interface (Fig. 22.18). The y coordinate corresponding to the intersection of the drain *p-n* junction and the interface is indicated with L; the latter is called *electric length* of the gate, or *channel length*. The device is considered uniform in the z direction; its width along such a direction is indicated with W.

Purpose of the analysis is to derive the steady-state characteristics, namely, the relations between the currents at the contacts and the applied voltages. To proceed, one assumes that the gate voltage V_{GB} is such that at all positions y along the channel the strong-inversion condition holds. Thus, a layer of electrons is present, indicated in Fig. 22.18 with the shaded area underneath the gate oxide; the term

[17] In fact, in some types of logic circuits the source and drain contact may exchange their roles depending on the applied voltages.

well-formed channel is used to denote this situation. The minimum gate voltage necessary for obtaining a well-formed channel will be identified later. In a steady-state condition there is no current through the gate contact because the gate insulator is in series to it; also, the current flowing through the bulk contact is negligibly small because the two junctions are never forward biased. Since the channel layer connects two heavily doped regions of the n type, the application of a drain-source voltage $V_{DS} > 0$ gives rise to a current I_D that flows from the drain to the source contact (its reference is shown in Fig. 22.18); for a given V_{DS}, the drain current I_D is controlled by the amount of charge available in the channel, which is in turn controlled by the gate voltage V_{GB}. In other terms, the device is an electronic valve in which the gate-bulk port controls the current flowing into the drain-source port; moreover, in the steady-state condition the control port does not expend energy, because the gate current is zero. This, among other things, explains the success of the MOSFET concept.

Due to the uniformity in the z direction, the electron and hole current densities have the form

$$\mathbf{J}_n = J_{nx}\,\mathbf{i} + J_{ny}\,\mathbf{j}\,, \qquad \mathbf{J}_p = J_{px}\,\mathbf{i} + J_{py}\,\mathbf{j}\,, \tag{22.66}$$

with $\mathbf{i}$, $\mathbf{j}$ the unit vectors of the x and y axes, respectively. On the other hand, it is $J_{nx} = J_{px} = 0$ because no current can flow through the insulator, so only the y components J_{ny}, J_{py} are left in (22.66). In turn, the condition of a well-formed channel implies that the concentration of holes is negligibly small with respect to that of the electrons,[18] whence $|J_{ny}| \gg |J_{py}|$. It follows $\mathbf{J} = \mathbf{J}_n + \mathbf{J}_p \simeq \mathbf{J}_n = J_{ny}(x, y)\,\mathbf{j}$. The equality $\mathbf{J} = \mathbf{J}_n$ is the mathematical form of the MOSFET's property of being unipolar. It also entails $\operatorname{div}\mathbf{J} = \operatorname{div}\mathbf{J}_n$; therefore, remembering that in a steady-state condition it is $\operatorname{div}\mathbf{J} = 0$, it follows that $\operatorname{div}\mathbf{J}_n = 0$ as well.

Consider now two planes parallel to the x, z plane, placed at different positions y_1 and y_2 in the channel; their intersections with the x, y plane are respectively marked with S_1 and S_2 in Fig. 22.18. From the divergence theorem (A.23) and the property $\operatorname{div}\mathbf{J}_n = 0$ it follows[19]

$$\iint_2 \mathbf{J}_n \cdot \mathbf{j}\; \mathrm{d}x\; \mathrm{d}z - \iint_1 \mathbf{J}_n \cdot \mathbf{j}\; \mathrm{d}x\; \mathrm{d}z = 0\,; \tag{22.67}$$

as a consequence, the channel current

$$I = \int_{z=0}^{W} \int_{x=0}^{x_d} \mathbf{J}_n \cdot \mathbf{j}\; \mathrm{d}x\; \mathrm{d}z = W \int_0^{x_d} J_{ny}\; \mathrm{d}x \tag{22.68}$$

[18] Compare with Fig. 22.14; the latter describes an equilibrium case; however, the situation is similar to the one depicted here.

[19] In the integrals of (22.67) the upper limit of x is given by the depletion width $x_d(y)$ shown in Fig. 22.18. Compare also with Sect. 22.11.1.

is independent of y. The last form of (22.68) derives from the uniformity in z. To express J_{ny} in (22.68) it is convenient to adopt the monomial form (19.141) of the electron current density, $\mathbf{J}_n = -q\,\mu_n\,n\,\operatorname{grad}\varphi_n$, with φ_n the electron quasi-Fermi potential. The components of $\mathbf{J}_n$ in monomial form read

$$J_{nx} = -q\,\mu_n\,n\,\frac{\partial\varphi_n}{\partial x}\,, \qquad J_{ny} = -q\,\mu_n\,n\,\frac{\partial\varphi_n}{\partial y}\,, \tag{22.69}$$

where $J_{nx} = 0$ as found above. In the channel it is $n \neq 0$, whence for J_{nx} to vanish it must be $\partial\varphi_n/\partial x = 0$; as a consequence, φ_n in the channel[20] depends on y only. In conclusion,

$$I = W\,\frac{\mathrm{d}\varphi_n}{\mathrm{d}y}\int_0^{x_d} -q\,\mu_n\,n\,\mathrm{d}x\,. \tag{22.70}$$

In the integral of (22.70) it is $n = n(x, y)$ and $\mu_n = \mu_n(x, y)$; defining the *effective electron mobility* as the average

$$\mu_e(y) = \frac{\int_0^{x_d} -q\,\mu_n\,n\;\mathrm{d}x}{\int_0^{x_d} -q\,n\;\mathrm{d}x} > 0\,, \tag{22.71}$$

yields

$$I = W\,\frac{\mathrm{d}\varphi_n}{\mathrm{d}y}\,\mu_e(y)\,Q_i(y)\,, \qquad Q_i = \int_0^{x_d} -q\,n\,\mathrm{d}x < 0\,, \tag{22.72}$$

where Q_i is the inversion-layer charge per unit area at position y in the channel. In (22.72), μ_e and Q_i are positive- and negative-definite, respectively, and I, W are constant; it follows that $\mathrm{d}\varphi_n/\mathrm{d}y$ has always the same sign and, as a consequence, $\varphi_n(y)$ is invertible. Using the inverse function $y = y(\varphi_n)$ within μ_e and Q_i makes (22.72) separable; integrating the latter over the channel yields

$$\int_0^L I\;\mathrm{d}y = L\,I = W\int_{\varphi_n(0)}^{\varphi_n(L)} \mu_e(\varphi_n)\,Q_i(\varphi_n)\;\mathrm{d}\varphi_n\,. \tag{22.73}$$

In turn, the dependence of μ_e on y is weak,[21] whence

$$I = \frac{W}{L}\,\mu_e\int_{\varphi_n(0)}^{\varphi_n(L)} Q_i(\varphi_n)\,\mathrm{d}\varphi_n\,. \tag{22.74}$$

[20] Far from the channel the semiconductor is practically in the equilibrium condition, whence $\varphi_n \to \varphi_F$ as x increases. However, in the bulk region where the dependence of φ_n on x is significant, the electron concentration is negligible; as a consequence, the integral in (22.68) is not affected.

[21] This issue is discussed in Sect. 22.11.1.

22.8.1 Gradual-Channel Approximation

In the derivation of (22.74) the condition of a well-formed channel has not been exploited yet; this condition makes a number of approximations possible, which are collectively indicated with the term *gradual-channel approximation*;[22] they lead to an expression of (22.74) in closed form. First, one uses the definition of surface potential which, in the two-dimensional analysis considered here, is given by $\varphi_s(y) = \varphi(x = 0, y)$; it is shown in Sect. 22.11.1 that the condition of a well-formed channel entails the relation

$$\varphi_s = \varphi_n + \varphi_F \,. \tag{22.75}$$

It follows that $\mathrm{d}\varphi_n/\mathrm{d}y$ in (22.72) can be replaced with $\mathrm{d}\varphi_s/\mathrm{d}y$, thus showing that the transport in a well-formed channel is dominated by the drift term, $J_{ny} = -q\,\mu_n\,n\,\mathrm{d}\varphi_s/\mathrm{d}y = q\,\mu_n\,n\,E_{sy}$, with E_{sy} the y-component of the electric field at $x = 0$; using (22.75) one changes the variable from φ_n to φ_s in the integral of (22.74). The integration limits in terms of φ_s are found by the same reasoning leading to (22.75), and read (Sect. 22.11.1)

$$\varphi_s(0) = V_{SB} + 2\,\varphi_F \,, \qquad \varphi_s(L) = V_{DB} + 2\,\varphi_F \,. \tag{22.76}$$

In conclusion, (22.74) becomes

$$I = \frac{W}{L}\,\mu_e \int_{V_{SB}+2\,\varphi_F}^{V_{DB}+2\,\varphi_F} Q_i(\varphi_s)\;\mathrm{d}\varphi_s \,. \tag{22.77}$$

The next step of the gradual-channel approximation consists in determining the relation $Q_i(\varphi_s)$, for which the solution of the Poisson equation in two dimensions is necessary. As shown in Sect. 22.11.1, one can exploit the strong difference between the strengths of the electric-field components in the x and y directions, to give the equation a one-dimensional form in which the y coordinate acts as a parameter. This is equivalent to assimilating each elementary portion of the channel, like that marked with dy in Fig. 22.18, to a one-dimensional MOS capacitor whose surface potential is the local value $\varphi_s(y)$. The final step of the gradual-channel approximation is the adoption of the full-depletion and ASCE approximations (Sect. 21.4), so that the inversion-layer charge per unit area at position y in the channel is given by (22.34),

[22] The gradual-channel approximation is not limited to the analysis of the MOSFET shown here. Indeed, it is a widely used method to treat the Poisson equation in devices in which the geometrical configuration and applied voltages are such that the variation of the electric field in one direction is much weaker than those in the other two directions. Typically, the former direction is the longitudinal one (that is, along the channel), the other two the transversal ones. From the mathematical standpoint, the approximation amounts to eliminating a part of the Laplacian operator, so that the dependence on all variables but one becomes purely algebraic.

namely, $Q_i = -C_{\rm ox}\left[\left(V'_{GB}-\varphi_s\right)-\gamma\sqrt{\varphi_s}\right] < 0$. Observing that $V_{DB}+2\varphi_F > V_{SB}+2\varphi_F$ while the integrand in (22.77) is negative, it follows that $I<0$, whence $I_D=-I$ due to the reference chosen for the drain current (Fig. 22.18). In conclusion, (22.77) transforms into[23]

$$I_D=\beta\int_{V_{SB}+2\varphi_F}^{V_{DB}+2\varphi_F}\left[\left(V'_{GB}-\varphi_s\right)-\gamma\sqrt{\varphi_s}\right]\mathrm{d}\varphi_s\,,\qquad \beta=\frac{W}{L}\,\mu_e\,C_{\rm ox}\,. \tag{22.78}$$

22.8.2 *Differential Conductances and Drain Current,* N-*Channel*

The drain current's expression (22.78) of the *n*-channel MOSFET provides a relation of the form $I_D=I_D(V_{GB},V_{DB},V_{SB})$, where $V_{GB}=V'_{GB}+V_{FB}$. In the integrated-circuit operation an important role is played by the *differential conductances* of the device, each of them defined as the partial derivative of I_D with respect to one of the applied voltages. In some cases the differential conductances can be found without the need of actually calculating the integral in (22.78); for this reason, here such conductances are calculated first. Prior to that, it is worth noting that in circuit applications it is often preferred to use the source contact, instead of the bulk contact, as a voltage reference. The transformation from one reference to the other is easily obtained from

$$V_{DS}=V_{DB}-V_{SB}>0\,,\qquad V_{GS}=V_{GB}-V_{SB}\,,\qquad V_{BS}=-V_{SB}\le 0\,. \tag{22.79}$$

Then, the *drain conductance*[24] is defined as the derivative of I_D with respect to V_{DB}, at constant V_{SB} and V_{GB}; or, equivalently, as the derivative with respect to V_{DS}, at constant V_{BS} and V_{GS}:

$$g_D=\left(\frac{\partial I_D}{\partial V_{DB}}\right)_{V_{SB},V_{GB}}=\left(\frac{\partial I_D}{\partial V_{DS}}\right)_{V_{BS},V_{GS}}\,. \tag{22.80}$$

Remembering that the derivative of an integral with respect to the upper limit is the integrand calculated at such limit, from (22.78) one finds

$$g_D=\beta\left[\left(V'_{GB}-V_{DB}-2\varphi_F\right)-\gamma\sqrt{V_{DB}+2\varphi_F}\right]\,. \tag{22.81}$$

[23] The units of β are $[\beta]=\mathrm{A\,V^{-2}}$.

[24] The drain conductance is also called *output conductance*; in this case it is indicated with g_o.

Using (22.76) and (22.34) yields

$$g_D = \beta \left[\left(V'_{GB} - \varphi_s(L) \right) - \gamma \sqrt{\varphi_s(L)} \right] = \frac{W}{L} \mu_e \left[-Q_i(L) \right] , \tag{22.82}$$

namely, the drain conductance is proportional to the inversion charge per unit area at the drain end of the channel. The quantity in brackets in (22.82) is nonnegative by construction; its zero corresponds to the value of $\varphi_s(L)$ obtained from (22.33). Such a zero is indicated with φ_s^{sat} and is termed *saturation surface potential*. From (22.79), the *saturation voltage* in the bulk and source references is found to be

$$V_{DB}^{\text{sat}} = \varphi_s^{\text{sat}} - 2\,\varphi_F , \qquad V_{DS}^{\text{sat}} = V_{DB}^{\text{sat}} - V_{SB} , \tag{22.83}$$

respectively. Similarly, the current $I_D^{\text{sat}} = I_D(V_{DS}^{\text{sat}})$ (which depends on V_{GS}) is called *saturation current*. If a value of V_{DS} larger than V_{DS}^{sat} is used, the right-hand side of (22.81) becomes negative; this result is not physically sound and indicates that the gradual-channel approximation is not applicable in that voltage range.[25]

Still considering the drain conductance, it is also important to determine its limit for $V_{DB} \to V_{SB}$, or $V_{DS} \to 0$. Again, there is no need to calculate the integral in (22.78) which, in this limiting case, is the product of the integration interval V_{DS} times the integrand calculated in the lower integration limit; in turn, the derivative eliminates V_{DS}, whence

$$g_D(V_{DS} \to 0) = \beta \left[\left(V'_{GB} - V_{SB} - 2\,\varphi_F \right) - \gamma \sqrt{V_{SB} + 2\,\varphi_F} \right] . \tag{22.84}$$

Replacing V'_{GB} with $V_{GB} - V_{FB}$ and using (22.79) yield

$$g_D(V_{DS} \to 0) = \beta \, (V_{GS} - V_T) , \qquad V_T = V_{FB} + 2\,\varphi_F + \gamma \sqrt{2\,\varphi_F - V_{BS}} \tag{22.85}$$

where, remembering that the junctions are never forward biased, it is $V_{BS} \leq 0$. The $V_T = V_T(V_{BS})$ voltage defined in (22.85) is called *threshold voltage*, and its dependence on V_{BS} is called *body effect*. Near $V_{DS} = 0$ the relation between I_D and V_{DS} is $I_D = \beta \, (V_{GS} - V_T) \, V_{DS}$: there, the current-voltage characteristics are approximated by straight lines whose slope, for a given V_{BS}, is prescribed by V_{GS}. At larger values of V_{DS} the limiting case (22.85) does not hold any longer: the slope of the $I_D = I_D(V_{DS})$ curves decreases, to eventually vanish when V_{DS} reaches V_{DS}^{sat}. Considering that I_D is nonnegative, the theory depicted above is applicable as long as $\beta \, (V_{GS} - V_T) \geq 0$; this observation allows one to better specify the condition of a well-formed channel, used at the beginning: from the formal standpoint the condition of a well-formed channel is $V_{GS} > V_T$.

[25] In fact, beyond the saturation voltage the Poisson equation near the drain end of the channel cannot be reduced anymore to a one-dimensional equation where y is treated as a parameter (compare with Sect. 22.11.1).

The integration of (22.78) is straightforward and yields $I_D = I'_D - I''_D$, where I'_D is obtained by integrating $\beta\,(V'_{GB} - \varphi_s)$. In this calculation many terms cancel out to yield the relatively simple expression

$$I'_D = \beta \left[(V_{GS} - V_{FB} - 2\,\varphi_F)\; V_{DS} - \frac{1}{2}\, V_{DS}^2 \right] . \tag{22.86}$$

In turn, I''_D is obtained by integrating $\beta\,\gamma\,\sqrt{\varphi_s}$ and reads

$$I''_D = \beta\,\frac{2}{3}\,\gamma \left[(V_{DS} + 2\,\varphi_F - V_{BS})^{3/2} - (2\,\varphi_F - V_{BS})^{3/2} \right] . \tag{22.87}$$

Comparisons with experiments show that the model $I_D = I'_D - I''_D$, where the two contributions are given by (22.86) and (22.87), provides a fair description of the drain current up to the saturation voltage. Beyond saturation, the model is not correct any longer: in fact, the terms with a negative sign within the expression of I_D give rise to a negative slope g_D; instead, the experiments show that for $V_{DS} > V_{DS}^{\rm sat}$ the current tends to saturate. For this reason, the analytical model is given a regional form: for a prescribed pair V_{GS}, V_{BS}, the regional model first separates the *on* condition $V_{GS} > V_T$ from the *off* condition $V_{GS} \le V_T$. The *on* condition is further separated into the *linear region*[26] $0 < V_{DS} \le V_{DS}^{\rm sat}$, where the drain current is described by the $I_D = I'_D - I''_D$ model worked out above, and the *saturation region* $V_{DS} > V_{DS}^{\rm sat}$, where the regional model lets $I_D = I_D^{\rm sat}$. Finally, in the *off* condition the model lets $I_D = 0$.

As the threshold voltage defined in (22.85) refers to a p-type substrate, from the definition (18.26) of the Fermi potential and the analysis of Sect. 18.4.2 one finds $2\,\varphi_F = (k_B\,T/q)\,\log(p_{p0}/n_{p0}) > 0$; on the other hand, the flat-band voltage (22.37) may have either sign, depending on the materials used. It follows that the threshold voltage as a whole may have either sign as well. It is then customary to distinguish two types of n-channel transistors: the *enhancement-type* ones have $V_T > 0$, so that a gate voltage $V_{GS} > V_T > 0$ is necessary to achieve the condition of a well-formed channel, whereas the transistor is in the *off* condition when $V_{GS} = 0$; the *depletion-type* transistors have $V_T < 0$, so that the condition of a well-formed channel is already present when $V_{GS} = 0$, whereas a gate voltage $V_{GS} \le V_T < 0$ is necessary to put the device in the *off* condition.

For a given bulk-source voltage V_{BS}, the $I_D = I_D(V_{DS})$ curves corresponding to different values of V_{GS} are called *output characteristics*. Other types of characteristics are also used to enrich the picture of the MOSFET's behavior: for instance, the *transfer characteristics* are the $I_D = I_D(V_{GS})$ curves drawn using V_{BS} as a

[26] The term *linear* originates from the behavior of the curves near the origin, shown by (22.85). The term is ascribed to the region up to $V_{DS}^{\rm sat}$ despite the fact that far away from the origin the curves are blatantly nonlinear.

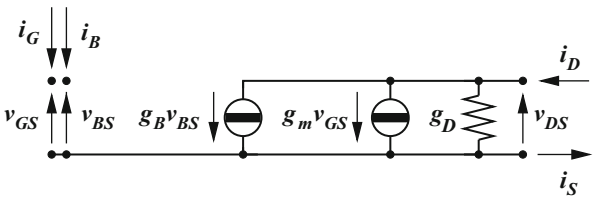


Fig. 22.19 Low frequency, small-signal circuit of an *n*-channel MOSFET

parameter and letting $V_{DS} = \text{const}$, with a value of V_{DS} small enough to let the limiting case (22.85) hold.

Besides the drain conductance (22.80), two more differential conductances are defined in a MOSFET: the first one is the *transconductance* g_m, given by the derivative of I_D with respect to V_{GB}, at constant V_{DB} and V_{SB}; or, equivalently, as the derivative with respect to V_{GS}, at constant V_{DS} and V_{BS}. Observing that V_{GS} appears only in I'_D one finds

$$g_m = \left(\frac{\partial I_D}{\partial V_{GB}}\right)_{V_{SB},V_{DB}} = \left(\frac{\partial I_D}{\partial V_{GS}}\right)_{V_{DS},V_{BS}} = \beta\, V_{DS}\,. \tag{22.88}$$

The second one is the *bulk transconductance* g_B, defined as the derivative of I_D with respect to V_{BS} at constant V_{DS} and V_{GS}:

$$g_B = \left(\frac{\partial I_D}{\partial V_{BS}}\right)_{V_{DS},V_{GS}} = -\left(\frac{\partial I''_D}{\partial V_{BS}}\right)_{V_{DS}}\,. \tag{22.89}$$

The small-signal circuit of an *n*-channel MOSFET is shown in Fig. 22.19. Since the circuit is derived from the steady-state transport model, it holds at low frequencies only. The small-signal voltages are indicated with v_{DS}, v_{GS}, and v_{BS}. The gate and bulk contacts are left open because the corresponding currents are zero; as a consequence, $i_D = i_S$ is the only nonzero small-signal current of the circuit. Observing that

$$i_D = g_D\, v_{DS} + g_m\, v_{GS} + g_B\, v_{BS}\,, \tag{22.90}$$

the drain-source branch of the circuit is made of three parallel branches. One of them is represented as a resistor $1/g_D$ because the current flowing in it is controlled by the voltage v_{DS} applied to the same port; the other two branches are voltage-controlled generators because the current of each branch is controlled by the voltage applied to a different port.

It is worth adding that the body effect mentioned above is actually an inconvenience, because it introduces a complicate dependence on V_{BS} which must be accounted for during the circuit's design. The body effect is suppressed by letting $V_{BS} = 0$: in a circuit's design, this is obtained by shorting the bulk and source

contacts, which amounts to reducing the original four-contact device to a three-contact device.[27] This solution is adopted whenever the circuit's architecture allows for it.

Linear-Parabolic Model, *N*-Channel

In semiqualitative circuit analyses the whole term I_D'' is neglected, thus leading to a simplified model $I_D \simeq I_D'$, called *linear-parabolic model*. As the neglect of I_D'' is equivalent to letting $\gamma \to 0$, it follows $g_B \simeq 0$ and, from the second relation in (22.85), the simplified threshold voltage reads

$$V_T \simeq V_{FB} + 2\,\varphi_F\,. \tag{22.91}$$

In turn, from $I_D \simeq I_D' = \beta\,[(V_{GS} - V_{FB} - 2\,\varphi_F)\,V_{DS} - V_{DS}^2/2]$ one finds for the drain conductance

$$g_D \simeq \beta\;(V_{GS} - V_{FB} - 2\,\varphi_F - V_{DS}) = \beta\;(V_{GS} - V_T - V_{DS})\;, \tag{22.92}$$

whence

$$V_{DS}^{\text{sat}} = V_{GS} - V_T\,. \tag{22.93}$$

The transconductance g_m is the same as in the general case. Note that the linear-parabolic expression of the drain current may be recast as $I_D \simeq \beta\,(V_{DS}^{\text{sat}}\,V_{DS} - V_{DS}^2/2)$, with V_{DS}^{sat} given by (22.93). As a consequence,

$$I_D^{\text{sat}} \simeq \beta\left[\left(V_{DS}^{\text{sat}}\right)^2 - \frac{1}{2}\left(V_{DS}^{\text{sat}}\right)^2\right] = \frac{1}{2}\,\beta\;(V_{GS} - V_T)^2\;. \tag{22.94}$$

The linear-parabolic model then yields for the saturation region

$$I_D = I_D^{\text{sat}}\,, \qquad g_D \simeq 0\,, \qquad g_m \simeq \beta\;(V_{GS} - V_T)\;, \qquad g_B \simeq 0\,. \tag{22.95}$$

An example of the output characteristics of an *n*-channel MOSFET obtained from the linear-parabolic model is given in Fig. 22.20, using $V_T = 1$ V, $\beta = 0.3$ A V^{-2}. The dashed curve represents (22.94). The symbol of the enhancement-type device[28] is shown in Fig. 22.21.

[27] Note that letting $V_{BS} = 0$ also makes g_B to vanish.

[28] Other symbols than that of Fig. 22.21 are used in the literature. For instance, the arrow may be placed over the bulk contact, pointing toward the gate in the *n*-channel device, or away from the gate in the *p*-channel one. The symbol of the depletion-type device is similar; however, the vertical bar representing the channel region is thicker than that of Fig. 22.21 to remind one that the channel is already formed when $V_{GS} = 0$.

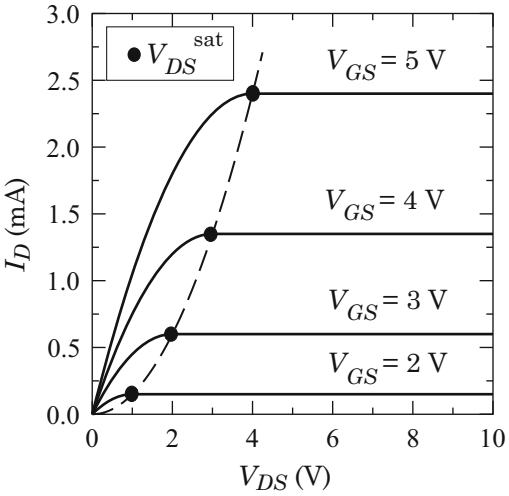


Fig. 22.20 Output characteristics of an *n*-type MOSFET obtained from the linear-parabolic model, with $V_T = 1$ V, $\beta = 0.3$ A V^{-2}. The *dashed curve* represents (22.94)

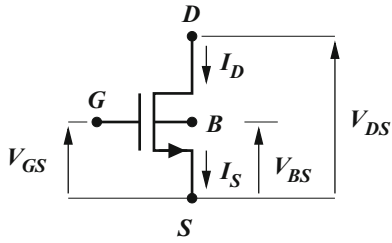


Fig. 22.21 Symbol of the *n*-channel MOSFET of the enhancement type

22.9 *P*-Channel MOSFET—Current-Voltage Characteristics

The structure of the *p-channel* MOSFET is dual to that of the *n*-channel type; with reference to Fig. 22.18, the cross-section is obtained by exchanging the "*p*" and "*n*" letters. Like in the *n*-channel device, the applied voltages are distinguished from one another by means of two letters; considering the bulk metallization as the reference contact, in a *p*-channel MOSFET the standard choice of the three independent voltages is $V_{GB} = V_G - V_B$, $V_{SB} = V_S - V_B$, and $V_{DB} = V_D - V_B$. In the typical operating regime of the *p*-channel MOSFET, the source-bulk and drain-bulk junctions are never forward biased, whence $V_{SB} \leq 0$ and $V_{DB} \leq 0$. The drain contact is identified by the condition $V_{SB} - V_{DB} = V_{SD} > 0$.

To derive the steady-state characteristics of the device, one proceeds in the same manner as for the *n*-channel MOSFET, i.e., by introducing a reference like in Fig. 22.18. Next, one assumes that the gate voltage V_{GB} is such that at all positions y along the channel the strong-inversion condition holds. Thus, a layer of holes is present, providing the condition of a well-formed channel. Since the channel layer connects two heavily doped regions of the p type, the application of a source-drain voltage $V_{SD} > 0$ gives rise to a current I_S that flows from the source to the drain

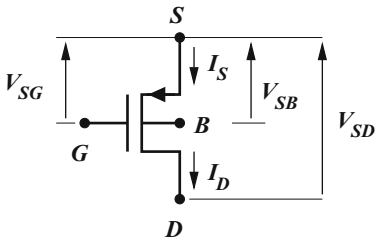


Fig. 22.22 Symbol of the *p*-channel MOSFET of the enhancement type

contact; such a current is indicated in Fig. 22.22, where the symbol of the *p*-channel MOSFET of the enhancement type is depicted. For a given V_{SD}, the source current I_S is controlled by the amount of charge available in the channel, which is in turn controlled by the gate voltage V_{GB}. Like in the *n*-channel device, one obtains an electronic valve in which the gate-bulk port controls the current flowing into the source-drain port; moreover, in the steady-state condition the control port does not expend energy, because the gate current is zero.

Due to the uniformity in the z direction, the electron and hole current densities have the form (22.66), with $\mathbf{i}, \mathbf{j}$ the unit vectors of the x and y axes, respectively. On the other hand, it is $J_{nx} = J_{px} = 0$ because no current can flow through the insulator, so only the y components J_{ny}, J_{py} are left in (22.66). In turn, the condition of a well-formed channel implies that the concentration of electrons is negligibly small with respect to that of the holes, whence $|J_{py}| \gg |J_{ny}|$. It follows $\mathbf{J} = \mathbf{J}_n + \mathbf{J}_p \simeq \mathbf{J}_p = J_{py}(x, y)\,\mathbf{j}$. Following the same reasoning as that leading to (22.68), one finds for the channel current

$$I = \int_{z=0}^{W} \int_{x=0}^{x_d} \mathbf{J}_p \cdot \mathbf{j} \, \mathrm{d}x \, \mathrm{d}z = W \int_0^{x_d} J_{py} \, \mathrm{d}x \,. \tag{22.96}$$

The above is independent of y. To express J_{ny} in (22.96) it is convenient to adopt the monomial form (19.142) of the hole current density, $\mathbf{J}_p = -q\,\mu_p\,p\ \mathrm{grad}\,\varphi_p$, with φ_p the hole quasi-Fermi potential; this leads to (compare with (22.70))

$$I = -W \frac{\mathrm{d}\varphi_p}{\mathrm{d}y} \int_0^{x_d} q\,\mu_p\,p \, \mathrm{d}x \,. \tag{22.97}$$

In the integral of (22.97) it is $p = p(x, y)$ and $\mu_p = \mu_p(x, y)$; defining the *effective hole mobility* as the average

$$\mu_h(y) = \frac{\int_0^{x_d} q\,\mu_p\,p \, \mathrm{d}x}{\int_0^{x_d} q\,p \, \mathrm{d}x} > 0 \,, \tag{22.98}$$

yields

$$I = -W \frac{\mathrm{d}\varphi_p}{\mathrm{d}y} \, \mu_h(y)\, Q_i(y) \,. \qquad Q_i = \int_0^{x_d} q\,p \, \mathrm{d}x > 0 \,, \tag{22.99}$$

where Q_i is the inversion-layer charge per unit area at position y in the channel. In (22.99), both μ_h and Q_i are positive-definite, and I, W are constant; it follows that $\mathrm{d}\varphi_p/\mathrm{d}y$ has always the same sign and, as a consequence, $\varphi_p(y)$ is invertible. Using the inverse function $y = y(\varphi_p)$ within μ_h and Q_i makes (22.99) separable; integrating the latter over the channel yields

$$\int_0^L I \, \mathrm{d}y = L\,I = W \int_{\varphi_p(L)}^{\varphi_p(0)} \mu_h(\varphi_p)\, Q_i(\varphi_p) \, \mathrm{d}\varphi_p \,. \tag{22.100}$$

Due to the choice of the reference, which is the same as for the n-channel device, the positions $y = 0$ and $y = L$ correspond to the source and drain ends of the channel, respectively. As the dependence of μ_h on y is weak,[29] (22.100) becomes (compare with (22.74))

$$I = \frac{W}{L}\,\mu_h \int_{\varphi_p(L)}^{\varphi_p(0)} Q_i(\varphi_p)\,\mathrm{d}\varphi_p \,. \tag{22.101}$$

To complete the calculation of (22.101) one makes use of the gradual-channel approximation, starting from the condition of a well-formed channel. Using the definition $\varphi_s(y) = \varphi(x = 0, y)$ of the surface potential, the latter condition in a p-channel device reads

$$\varphi_s = \varphi_p + \varphi_F \,. \tag{22.102}$$

It follows that $\mathrm{d}\varphi_p/\mathrm{d}y$ in (22.99) can be replaced with $\mathrm{d}\varphi_s/\mathrm{d}y$, that is, transport in a well-formed channel is dominated by the drift term, $J_{py} = -q\,\mu_p\,p\,\mathrm{d}\varphi_s/\mathrm{d}y = q\,\mu_p\,p\,E_{sy}$, with E_{sy} the y-component of the electric field at $x = 0$; using (22.102) one changes the variable from φ_n to φ_s in the integral of (22.101). The integration limits in terms of φ_s are found by the same reasoning leading to (22.102) and read (Sect. 22.11.1)

$$\varphi_s(0) = V_{SB} + 2\,\varphi_F \,, \qquad \varphi_s(L) = V_{DB} + 2\,\varphi_F \,. \tag{22.103}$$

In conclusion, (22.101) becomes (compare with (22.77))

$$I = \frac{W}{L}\,\mu_h \int_{V_{DB}+2\,\varphi_F}^{V_{SB}+2\,\varphi_F} Q_i(\varphi_s) \, \mathrm{d}\varphi_s \,. \tag{22.104}$$

The next step consists in determining the relation $Q_i(\varphi_s)$, for which the solution of the Poisson equation in two dimensions is necessary. Following again the reasoning of Sect. 22.11.1, one gives the equation a one-dimensional form in which the y

[29] This issue is discussed in Sect. 22.11.1.

coordinate acts as a parameter. This is equivalent to assimilating each elementary portion of the channel, of width dy, to a one-dimensional, n-substrate MOS capacitor whose surface potential is the local value $\varphi_s(y)$. Finally, introducing the full-depletion and ASCE approximations (Sect. 21.4), one finds the expression of the inversion-layer charge per unit area at position y in the channel, for a p-channel MOSFET. The result is similar to that already obtained for a p-substrate capacitor and is given by (22.65), namely:

$$Q_i = C_{\rm ox}\left[\left(-V'_{GB}+\varphi_s\right)-\gamma\sqrt{-\varphi_s}\right] > 0\,, \tag{22.105}$$

with

$$V'_{GB} = V_{GB}-V_{FB}\,, \qquad \gamma = \frac{\sqrt{2\,\varepsilon_{\rm sc}\,q\,n_{n0}}}{C_{\rm ox}}\,. \tag{22.106}$$

Observing that $V_{SB}+2\,\varphi_F > V_{DB}+2\,\varphi_F$ and that the integrand in (22.104) is positive, it follows that $I>0$, whence $I_S = I$ due to the reference chosen for the drain current (Fig. 22.22). In conclusion, (22.104) transforms into

$$I_S = \beta\int_{V_{DB}+2\,\varphi_F}^{V_{SB}+2\,\varphi_F}\left[\left(-V'_{GB}+\varphi_s\right)-\gamma\sqrt{-\varphi_s}\right]{\rm d}\varphi_s\,, \qquad \beta = \frac{W}{L}\,\mu_h\,C_{\rm ox}\,. \tag{22.107}$$

22.9.1 *Differential Conductances and Drain Current,* P-*Channel*

The differential conductances are calculated from (22.107) in the same way as for the n-channel device. Prior to that, one remembers that in circuit applications it is often preferred to use the source contact, instead of the bulk contact, as a voltage reference. The transformation from one reference to the other is easily obtained from

$$V_{SD} = V_{SB}-V_{DB} > 0\,, \qquad V_{SG} = V_{SB}-V_{GB}\,, \qquad V_{SB} = -V_{BS} \le 0\,. \tag{22.108}$$

Then, the *source conductance*, or *output conductance*, is defined as the derivative of I_S with respect to V_{BS}, at constant V_{BD} and V_{GB}; or, equivalently, as the derivative with respect to V_{SD}, at constant V_{BD} and V_{GD}:

$$g_S = \left(\frac{\partial I_S}{\partial V_{BS}}\right)_{V_{BD},V_{BG}} = \left(\frac{\partial I_S}{\partial V_{SD}}\right)_{V_{BD},V_{GD}}\,. \tag{22.109}$$

Remembering that the derivative of an integral with respect to the upper limit is the integrand calculated at such limit, from (22.107) one finds

$$g_S = \beta \left[\left(-V'_{GB} + V_{SB} + 2\,\varphi_F \right) - \gamma \sqrt{-V_{SB} - 2\,\varphi_F} \right] = \frac{W}{L}\,\mu_h\,Q_i(0)\,. \tag{22.110}$$

The last form of (22.110) is obtained by combining the first relation in (22.103) with (22.105), and shows that the source conductance is proportional to the inversion charge per unit area at the source end of the channel. The zero of g_S, which is nonnegative by construction, corresponds to a value of $\varphi_s(0)$ equal to the *saturation surface potential*; the latter is given in turn by (compare with (22.33))

$$\sqrt{-\varphi_s^{\text{sat}}} = \sqrt{-V'_{GB} + (\gamma/2)^2} + \gamma/2\,. \tag{22.111}$$

The *saturation voltage* in the bulk and source references is then found to be

$$V_{SB}^{\text{sat}} = \varphi_s^{\text{sat}} - 2\,\varphi_F\,, \qquad V_{SD}^{\text{sat}} = V_{SB}^{\text{sat}} - V_{DB}\,, \tag{22.112}$$

respectively. Similarly, the current $I_S^{\text{sat}} = I_S(V_{SD}^{\text{sat}})$ (which depends on V_{SG}) is called *saturation current*. If a value of V_{SD} larger than V_{SD}^{sat} is used, the right-hand side of (22.110) becomes negative; this result is not physically sound and indicates that the gradual-channel approximation is not applicable in that voltage range.[30] The limit of the source conductance for $V_{SB} \to V_{DB}$, or $V_{SD} \to 0$, is

$$g_S(V_{SD} \to 0) = \beta\ (V_{SG} - V_T)\ , \qquad V_T = -V_{FB} - 2\,\varphi_F + \gamma\ \sqrt{V_{BS} - 2\,\varphi_F}\,, \tag{22.113}$$

where $V'_{GB} = V_{GB} - V_{FB}$ has been used. Remembering that the junctions are never forward biased, it is $V_{BS} \geq 0$. The $V_T = V_T(V_{BS})$ voltage defined in (22.113) is the *threshold voltage* of the *p*-channel MOSFET. Near $V_{SD} = 0$ the relation between I_S and V_{SD} is $I_S = \beta\ (V_{SG} - V_T)\ V_{SD}$: considering that I_S is nonnegative, the theory depicted above is applicable as long as $\beta\ (V_{SG} - V_T) \geq 0$; this observation allows one to better specify the condition of a well-formed channel, used at the beginning: from the formal standpoint the condition of a well-formed channel in a *p*-channel MOSFET is $V_{SG} > V_T$.

The integration of (22.107) is straightforward and yields $I_S = I'_S - I''_S$, where I'_S is obtained by integrating $\beta\,(-V'_{GB} + \varphi_s)$. In this calculation many terms cancel out to yield the relatively simple expression

$$I'_S = \beta \left[(V_{SG} + V_{FB} + 2\,\varphi_F)\ V_{SD} - \frac{1}{2}\,V_{SD}^2 \right]. \tag{22.114}$$

[30] In fact, beyond the saturation voltage the Poisson equation near the drain end of the channel cannot be reduced anymore to a one-dimensional equation where y is treated as a parameter (compare with Sect. 22.11.1).

In turn, I_S'' is obtained by integrating $\beta\,\gamma\,\sqrt{-\varphi_s}$, and reads

$$I_S'' = \beta\,\frac{2}{3}\,\gamma\,\left[(V_{SD} - 2\,\varphi_F + V_{BS})^{3/2} - (-2\,\varphi_F + V_{BS})^{3/2}\right]. \tag{22.115}$$

Like in the case of the *n*-channel device, comparisons with experiments show that the model $I_S = I_S' - I_S''$, where the two contributions are given by (22.114) and (22.115), provides a fair description of the source current up to the saturation voltage. Beyond saturation, the model is not correct any longer: in fact, the terms with a negative sign within the expression of I_S give rise to a negative slope g_S; instead, the experiments show that for $V_{SD} > V_{SD}^{\rm sat}$ the current tends to saturate. For this reason, the analytical model is given a regional form: for a prescribed pair V_{SG}, V_{BS}, the regional model first separates the *on* condition $V_{SG} > V_T$ from the *off* condition $V_{SG} \le V_T$. The *on* condition is further separated into the *linear region* $0 < V_{SD} \le V_{SD}^{\rm sat}$, where the source current is described by the $I_S = I_S' - I_S''$ model worked out above, and the *saturation region* $V_{SD} > V_{SD}^{\rm sat}$, where the regional model lets $I_S = I_S^{\rm sat}$. Finally, in the *off* condition the model lets $I_S = 0$.

As the threshold voltage defined in (22.113) refers to an *n*-type substrate, it is $2\,\varphi_F = (k_B\,T/q)\,\log(p_{n0}/n_{n0}) < 0$; on the other hand, the flat-band and threshold voltages may have either sign. Like in the case of the *n*-channel devices, it is then customary to distinguish two types of *p*-channel transistors: the *enhancement-type* ones have $V_T > 0$, so that a gate voltage $V_{SG} > V_T > 0$ is necessary to achieve the condition of a well-formed channel, whereas the transistor is in the *off* condition when $V_{SG} = 0$; the *depletion-type* transistors have $V_T < 0$, so that the condition of a well-formed channel is already present when $V_{SG} = 0$, whereas a gate voltage $V_{SG} \le V_T < 0$ is necessary to put the device in the *off* condition.

The *output characteristics* of the *p*-channel MOSFET are the $I_S = I_S(V_{SD})$ curves corresponding to different values of V_{SG}, for a given bulk-source voltage V_{BS}; the *transfer characteristics* are the $I_S = I_S(V_{SG})$ curves drawn using V_{BS} as a parameter and letting $V_{SD} = \text{const}$, with a value of V_{SD} small enough to let the limiting case (22.113) hold.

The *transconductance* g_m of the *p*-channel MOSFET is the derivative of I_S with respect to V_{BG}, at constant V_{BS} and V_{BD}; or, equivalently, the derivative with respect to V_{SG}, at constant V_{SD} and V_{SB}. Observing that V_{SG} appears only in I_S' one finds

$$g_m = \left(\frac{\partial I_S}{\partial V_{BG}}\right)_{V_{BS},V_{BD}} = \left(\frac{\partial I_S}{\partial V_{SG}}\right)_{V_{SD},V_{SB}} = \beta\,V_{SD}\,. \tag{22.116}$$

The *bulk transconductance* g_B is the derivative of I_S with respect to V_{SB} at constant V_{SD} and V_{SG}:

$$g_B = \left(\frac{\partial I_S}{\partial V_{SB}}\right)_{V_{SD},V_{SG}} = -\left(\frac{\partial I_S''}{\partial V_{BS}}\right)_{V_{SD}}. \tag{22.117}$$

Like in the n-channel device, the body effect is suppressed by letting $V_{BS} = 0$, that is, by shorting the bulk and source contacts, whenever the circuit's architecture allows for it.

Linear-Parabolic Model, *P*-Channel

The expressions worked out so far may further be simplified by neglecting I_S'', this yielding the *linear-parabolic model* $I_S \simeq I_S'$ of the p-channel MOSFET. As the neglect of I_S'' is equivalent to letting $\gamma \to 0$, it follows $g_B \simeq 0$ and, from the second relation in (22.113), the simplified threshold voltage reads

$$V_T \simeq -V_{FB} - 2\,\varphi_F\,. \tag{22.118}$$

In turn, from $I_S \simeq I_S' = \beta\,[(V_{SG} + V_{FB} + 2\,\varphi_F)\,V_{SD} - V_{SD}^2/2]$ one finds for the source conductance

$$g_S \simeq \beta\,\left(V_{SG} + V_{FB} + 2\,\varphi_F - V_{SD}\right) = \beta\,\left(V_{SG} - V_T - V_{SD}\right)\,, \tag{22.119}$$

whence

$$V_{SD}^{\text{sat}} = V_{SG} - V_T\,. \tag{22.120}$$

The transconductance g_m is the same as in the general case. Note that the linear-parabolic expression of the drain current may be recast as $I_S \simeq \beta\,(V_{SD}^{\text{sat}}\,V_{SD} - V_{SD}^2/2)$, with V_{SD}^{sat} given by (22.120). As a consequence,

$$I_S^{\text{sat}} \simeq \beta\,\left[\left(V_{SD}^{\text{sat}}\right)^2 - \frac{1}{2}\,\left(V_{SD}^{\text{sat}}\right)^2\right] = \frac{1}{2}\,\beta\,\left(V_{SG} - V_T\right)^2\,. \tag{22.121}$$

The linear-parabolic model then yields for the saturation region

$$I_S = I_S^{\text{sat}}\,, \qquad g_S \simeq 0\,, \qquad g_m \simeq \beta\,\left(V_{SG} - V_T\right)\,, \qquad g_B \simeq 0\,. \tag{22.122}$$

22.10 CMOS Inverter

The CMOS architecture (the designation stands for "Complementary Metal-Oxide-Semiconductor") is to date the most successful technology for constructing integrated circuits. The scheme of the CMOS inverter, which is the basic structure of the logic circuits manufactured with this technology, is shown in Fig. 22.23; it is obtained by connecting a p-channel MOSFET with an n-channel MOSFET, in

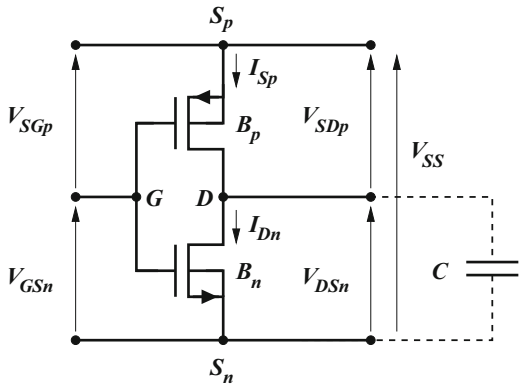


Fig. 22.23 The CMOS inverter

such a way that the two gate contacts are connected to each other, and the two drain contacts are connected to each other. Whenever possible, the bulk contact is shorted with the source contact of the same device, thus eliminating the body effect (Sect. 22.8.2); in this case, each MOSFET is described in terms of two voltages only, instead of three, and the threshold voltages (22.85) and (22.113) reduce to two constants.[31] From the technological standpoint, even the simple arrangement shown in Fig. 22.23 requires the presence of two different substrates in the same wafer: a *p*-type substrate to build the *n*-channel device, and an *n*-type substrate to build the *p*-channel device. This is achieved, starting from a given wafer, by preliminarily diffusing a region (called *well* or *tub*) of the opposite polarity with respect to that of the wafer, in which one of the two transistors is then accommodated. An example of fabrication of the *n*-well structure in silicon is illustrated in the following.

A CMOS structure can be realized using either the *bulk technology*, where the devices are fabricated directly within the silicon wafer, or the *silicon on insulator (SOI) technology*, where the devices are fabricated within an epitaxial layer of crystalline silicon grown on an insulator, e.g., sapphire (Sect. 24.6). The simpler bulk technology is shown here, with reference to Fig. 22.24. Starting from a lightly doped, *p*-type substrate, the area of the well is defined by a suitable mask, and the well is obtained by implanting *n*-type dopants.[32] The implant is followed by a thermal process that drives-in the impurities. As the duration of the thermal process is rather long, the final gradient of the impurity concentration is small; this makes the impurity profile of the well relatively flat. This step is followed by a *p*-type implant, called *channel stop*, in the area that separates the regions where the two transistors

[31] Elimination of the body effect is possible in simple structures like that shown in Fig. 22.23, where the source contact of the *p*-channel MOSFET is connected to the highest voltage of the circuit, and the source contact of the *n*-channel MOSFET is connected to the lowest one. However, when the circuit becomes slightly more complicate like, e.g., in a NAND or a NOR gate, shorting the bulk and source contact of all devices becomes impossible.

[32] The terminology of the process steps is better explained in the chapters devoted to the technological issues.

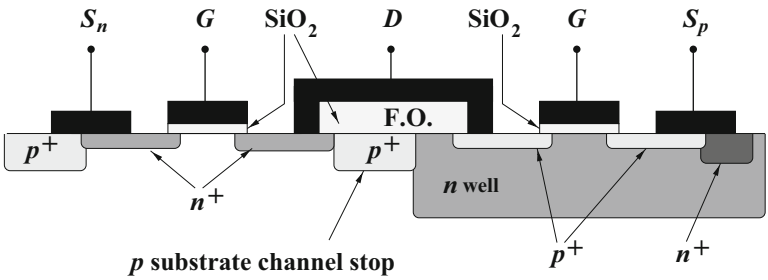


Fig. 22.24 Cross-section of an *n*-well CMOS inverter

will be fabricated; to further separate the *n*-channel from the *p*-channel transistor, a thick oxide layer is grown in the area above the channel stop; this layer is called *field oxide* and is marked with "F.O." in Fig. 22.24. Purpose of the channel stop and of the field oxide is to prevent the formation of an inversion layer of electrons at the silicon surface: such a layer, in fact, would form a parasitic channel connecting the two transistors. As shown in the figure, the drain metallization is right above the field oxide. In some operating conditions of the inverter the voltage drop between drain and substrate would be large enough to form the parasitic channel; this is averted by the combined effect of the field oxide, that lowers the vertical component of the electric field, and of the channel stop, that lowers the surface concentration of the electrons.

In the next step, the gate oxide is thermally grown over the active areas of the two transistors, and suitably patterned; this is followed by the deposition and patterning of the gate contacts, usually made of polysilicon.[33] The source and drain regions of the *n*-channel transistor are obtained by implanting a donor-type dopant while protecting from the implant the active area of the *p*-channel device; a successive implant of acceptor-type dopant creates in the same manner the source and drain regions of the *p*-channel transistor. Both implants use the corresponding gate electrode as a mask; this technique, called *self-aligned gate*, has the advantage that the implantation process, along with the subsequent annealing step, ensure that the gate overlaps the edges of the source and drain regions. Finally, the drain and source contacts are deposited and patterned; as shown in Fig. 22.24, the source contacts of both transistors extend laterally with respect to the source region: in the *n*-channel device, part of the source contact overlaps the substrate region, while in the *p*-channel device part of the source contact overlaps the well region. By this contrivance, the source and bulk contacts of each transistor are shorted, and the body effect eliminated. As indicated in Sect. 21.2.2, to prevent the contact-semiconductor junction from behaving like a rectifying device, a heavy dose of dopant is preliminarily introduced into the semiconductor region onto which the

[33] The gate contacts are connected with each other in a position not visible in the cross-section of Fig. 22.24.

contact is to be deposited; as shown in the figure, in the CMOS inverter two heavily doped regions of this type are formed, of the p and n type respectively. The p-well structure is dual with respect to that of Fig. 22.24; it is obtained from the latter by exchanging "p" and "n" letters, including those of S_p and S_n.

In the standard operating conditions, the input signal of the inverter is applied to the common gate contact, and the output signal is taken from the common drain contact; concerning the latter, in complex digital circuits the output contact is connected to several input contacts belonging to other logic gates. The *fan out* of the inverter is the maximum number of input contacts it can be connected to without altering its functioning. In general, such input contacts are the gate contacts of other MOS transistors: as a consequence, from the circuit viewpoint they can globally be mimicked as a nonlinear capacitor. The analysis that follows is limited to the steady-state functioning of the inverter, whence one can assume that no current is absorbed from the drain contact; the capacitor sketched in Fig. 22.23 is simply meant to remind which type of load should be used in dynamic conditions.

22.10.1 I-O Characteristic of the CMOS Inverter

To describe the functioning of the CMOS inverter, it is necessary to distinguish the quantities pertaining to the p-channel device from those of the n-type one. For this, it suffices to add an extra index; in particular, the voltages and current of the p-channel MOSFET are denoted with V_{SGp}, V_{SDp}, and I_{Sp}, while symbols V_{GSn}, V_{DSn}, and I_{Dn} are reserved for the n-channel device. The threshold voltages (22.85) and (22.113), are indicated with

$$V_{Tn} = V_{FBn} + 2\,\varphi_{Fn} + \gamma_n \sqrt{2\,\varphi_{Fn}}\,, \quad V_{Tp} = -V_{FBp} - 2\,\varphi_{Fp} + \gamma_p \sqrt{-2\,\varphi_{Fp}}\,. \tag{22.123}$$

For the correct functioning of the inverter, both transistors must be of the enhancement type, namely, remembering the definitions of Sects. 22.8.2 and 22.9.1, the choice of the materials and the fabrication process must be such that $V_{Tn} > 0$ and $V_{Tp} > 0$. A fixed bias $V_{SS} > 0$ is applied between the source contact of the p-channel transistor and that of the n-channel one, as shown in Fig. 22.23; it is assumed that the inequality

$$V_{SS} > V_{Tn} + V_{Tp} \tag{22.124}$$

holds, which is easily achieved in the practical cases. The voltages V_{GSn} and V_{DSn} are considered the input (V_{in}) and output (V_{out}) voltages of the inverter, respectively. With these premises, the qualitative functioning of the inverter is readily discussed, basing on the linear-parabolic model for the current-voltage characteristics worked out in Sects. 22.8 and 22.9. Let, for instance, be $0 \le V_{\text{in}} \le V_{Tn}$, so that the n-channel transistor is in the *off* condition and $I_{Dn} = 0$. As a consequence it is also $I_{Sp} = 0$ due

to the steady-state condition. On the other hand, it is $V_{SGp} - V_{Tp} = V_{SS} - V_{\text{in}} - V_{Tp}$ which, after adding and subtracting V_{Tn}, yields $V_{SGp} - V_{Tp} = (V_{SS} - V_{Tn} - V_{Tp}) + (V_{Tn} - V_{\text{in}})$; in the latter, the first term in parentheses is positive due to (22.124), and the second one is nonnegative. In conclusion, $V_{SGp} - V_{Tp} > 0$ and the p-channel device is in the *on* condition. For the latter to be compatible with $I_{Sp} = 0$, it is necessary for the device to work in the linear region with $V_{SDp} = 0$; in fact, the saturation region is ruled out because it implies $I_{Sp} \neq 0$. From the general relation $V_{SS} = V_{DSn} + V_{SDp}$ it then follows $V_{\text{out}} = V_{SS}$. As long as $0 \leq V_{\text{in}} \leq V_{Tn}$, the output voltage is the highest and no power is dissipated by the inverter.

Let now the input voltage be increased by a small amount $\delta V_{\text{in}} > 0$ starting from the threshold value, so that $V_{\text{in}} = V_{Tn} + \delta V_{\text{in}}$. It follows $V_{GSn} - V_{Tn} = \delta V_{\text{in}}$, which makes the n-channel device to turn on, yielding $I_{Sp} = I_{Dn} \neq 0$. In turn, the current flow makes V_{out} to decrease and become $V_{\text{out}} = V_{SS} - \delta V_{\text{out}}$, with $\delta V_{\text{out}} > 0$, whence $V_{SDp} = V_{SS} - V_{\text{out}} = \delta V_{\text{out}}$ and $V_{DSn} = V_{SS} - \delta V_{\text{out}}$. Then,

$$(V_{GSn} - V_{Tn}) - V_{DSn} = \delta V_{\text{in}} - (V_{SS} - \delta V_{\text{out}}) = (\delta V_{\text{in}} + \delta V_{\text{out}}) - V_{SS}\,, \tag{22.125}$$

$$\big(V_{SGp} - V_{Tp}\big) - V_{SDp} = V_{SS} - \big(V_{Tn} + V_{Tp}\big) - (\delta V_{\text{in}} + \delta V_{\text{out}})\ . \tag{22.126}$$

As $\delta V_{\text{in}} + \delta V_{\text{out}}$ can be made as small as we please, from (22.125) it follows $(V_{GSn} - V_{Tn}) - V_{DSn} < 0$, which shows that the n-channel device works in the saturation region; at the same time, thanks to (22.126) it is $(V_{SGp} - V_{Tp}) - V_{SDp} > 0$, namely, the p-channel device is in the linear region. If the input voltage is further increased, the p-channel transistor will eventually turn from the linear to the saturation region. This happens for $V_{SGp} - V_{Tp} = V_{SDp}$, namely, $V_{SS} - V_{\text{in}} - V_{Tp} = V_{SS} - V_{\text{out}}$, equivalent to $V_{\text{out}} = V_{\text{in}} + V_{Tp}$. When this occurs, the operating region of the n-channel device is given by

$$V_{GSn} - V_{Tn} - V_{DSn} = V_{\text{in}} - V_{Tn} - V_{\text{out}} = -\big(V_{Tn} + V_{Tp}\big)\ . \tag{22.127}$$

As $V_{Tn}, V_{Tp} > 0$, the n-channel device is still in the saturation region, namely, both transistors are working in the saturation region. In this regime, equating the currents at the output node, $I_{Sp} = I_{Dn}$, provides an expression where only the input voltage appears. This is so, because in the saturation region the expression of the current provided by the model is independent of the drain-source voltage. In fact, equating the currents yields $\beta_p\,(V_{SS} - V_{\text{in}} - V_{Tp})^2 = \beta_n\,(V_{\text{in}} - V_{Tn})^2$, whence $(V_{SS} - V_{Tn} - V_{Tp})/(V_{\text{in}} - V_{Tn}) = 1 \pm \sqrt{\beta_n/\beta_p}$. The minus sign in the latter must be discarded because the left-hand side is positive definite and finite. In conclusion, the input voltage at which both transistors work in the saturation region is found to be

$$V_{\text{in}}^{\text{sat}} = V_{Tn} + \frac{V_{SS} - V_{Tn} - V_{Tp}}{1 + \sqrt{\beta_n/\beta_p}}\,. \tag{22.128}$$

It is worth observing that this result is not physically sound because the output voltage corresponding to $V_{\text{in}}^{\text{sat}}$ is not defined. This happens because the model

assumes perfect saturation,[34] which never occurs in practice. However, in real CMOS inverters the slope $|\mathrm{d}V_{\mathrm{out}}/\mathrm{d}V_{\mathrm{in}}|$, although finite, is still very large when both transistors operate in the saturation region.

In the frame of the linear-parabolic model, the condition in which the *n*-channel transistor turns from the saturation to the linear region must be found by keeping $V_{\mathrm{in}} = V_{\mathrm{in}}^{\mathrm{sat}}$ and decreasing V_{out} until the condition $V_{GSn} - V_{Tn} = V_{DSn}$ is fulfilled. When $V_{\mathrm{in}} = V_{\mathrm{in}}^{\mathrm{sat}}$ and V_{out} decreases, the *n*-channel transistor turns from the saturation to the linear region as soon as $V_{GSn} - V_{Tn} = V_{DSn}$. This is equivalent to $V_{\mathrm{in}} - V_{Tn} = V_{\mathrm{out}}$, namely

$$V_{\mathrm{out}} = V_{\mathrm{in}}^{\mathrm{sat}} - V_{Tn}\,. \tag{22.129}$$

When V_{in} is increased beyond $V_{\mathrm{in}}^{\mathrm{sat}}$, more current is drawn by the *n*-channel transistor, which makes V_{out} to decrease and the saturation condition of the *p*-channel one to become deeper. The description of this region of the $V_{\mathrm{out}}(V_{\mathrm{in}})$ curve is similar to that of the region $V_{Tn} \le V_{\mathrm{in}} \le V_{\mathrm{in}}^{\mathrm{sat}}$. The *p*-channel transistor reaches the *off* condition as soon as the relation $V_{SGp} = V_{Tp}$ is fulfilled, namely,

$$V_{SS} - V_{\mathrm{in}} = V_{Tp}\,, \qquad V_{\mathrm{in}} = V_{SS} - V_{Tp}\,. \tag{22.130}$$

As long as $V_{SS} - V_{Tp} \le V_{\mathrm{in}} \le V_{SS}$, the *p*-channel transistor remains in the *off* condition, whence $V_{\mathrm{out}} = 0$ and $I_{Sp} = I_{Dn} = 0$. Thus, the output voltage is the lowest and no power is dissipated by the inverter.

The functioning conditions of the CMOS inverter are summarized in Table 22.3, where columns T_n, T_p indicate the operating regions of the *n*-channel and *p*-channel transistor, respectively. Letters *A*, *B*, *C*, or *D* in the first column mark some significant points of the $V_{\mathrm{out}}(V_{\mathrm{in}})$ curve. The full curve is shown by the black, thick line in Fig. 22.25; in the same figure, the continuous, blue line represents the relation $V_{\mathrm{out}} = V_{\mathrm{in}} - V_{Tn}$, which is equivalent to $V_{DSn} = V_{GSn} - V_{Tn}$. Remembering (22.93), the line marks the boundary between the linear and saturation regions of the *n*-channel transistor; such a line intercepts the horizontal axis at V_{Tn}. In turn, the

Table 22.3 CMOS inverter—functioning conditions

Point	V_{in}	V_{out}	T_n	T_p
–	0	V_{SS}	*off*	*lin*
A	V_{Tn}	V_{SS}	*off* → *sat*	*lin*
B	$V_{\mathrm{in}}^{\mathrm{sat}}$	$V_{\mathrm{in}}^{\mathrm{sat}} + V_{Tp}$	*sat*	*lin* → *sat*
C	$V_{\mathrm{in}}^{\mathrm{sat}}$	$V_{\mathrm{in}}^{\mathrm{sat}} - V_{Tn}$	*sat* → *lin*	*sat*
D	$V_{SS} - V_{Tp}$	0	*lin*	*sat* → *off*
–	V_{SS}	0	*lin*	*off*

[34] The conclusion would be the same if the terms I_D'', I_S'', calculated from (22.87), (22.115) with the saturation value of the drain-source voltages, were kept. The expression of $V_{\mathrm{in}}^{\mathrm{sat}}$, instead, would become more complicate than (22.128).

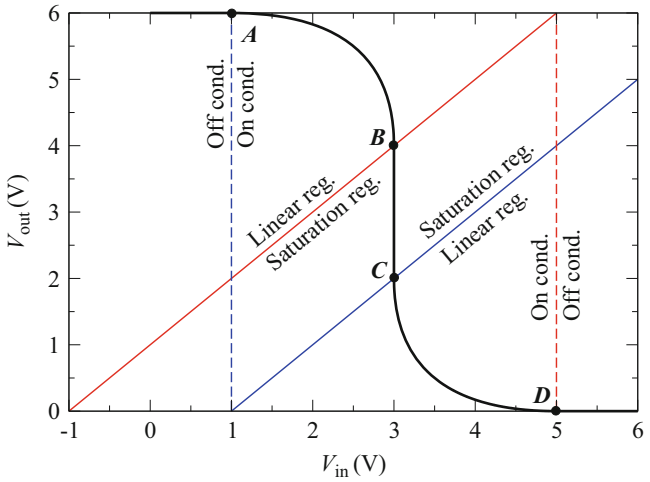


Fig. 22.25 The *black, continuous line* shows the input-output curve of the CMOS inverter. The curve has been drawn using $V_{SS} = 6$ V, $V_{Tp} = V_{Tn} = 1$ V, and $\beta_p = \beta_n$

continuous, red line represents the relation $V_{\text{out}} = V_{\text{in}} + V_{Tp}$, which is equivalent to $V_{SDp} = V_{SGp} - V_{Tp}$. Remembering (22.120), the line marks the boundary between the linear and saturation regions of the p-channel transistor; such a line intercepts the horizontal axis at $-V_{Tp}$. The vertical, dashed lines mark the boundary between the *off* and *on* condition of the n-channel transistor (blue line) and of the p-channel transistor (red line). Other details on the derivation of the curve are given in Prob. 22.6.

22.11 Complements

22.11.1 Poisson's Equation in the MOSFET Channel

The derivation of the MOSFET's current carried out in Sect. 22.8 is based upon two integrals; the first one, (22.68), is calculated over a section of the channel at some position y, the second one, (22.73), is calculated along the channel from $y = 0$ to $y = L$. Apparently this procedure eliminates the need of solving the Poisson equation. In fact, the solution of the latter is deeply rooted in the relation (22.75), which is a fundamental point of the procedure itself, and in the choice of the integration limits (22.76), which are also related to (22.75).

The Poisson equation in a nonequilibrium condition is conveniently tackled by expressing the carrier concentrations in terms of the quasi-Fermi potentials φ_n and φ_p; the device considered here is the same n-channel MOSFET of Sect. 22.8. Using the normalized form $u_n = q\,\varphi_n/(k_B\,T)$ and $u_p = q\,\varphi_p/(k_B\,T)$, the concentrations

read $n = n_i \exp(u - u_n)$ and $p = n_i \exp(u_p - u)$, respectively. Remembering that in the equilibrium limit it is $u_n, u_p \to u_F$, with u_F the normalized Fermi potential, it is useful to introduce the differences

$$\chi_n = u_n - u_F \,, \qquad \chi_p = u_p - u_F \,, \tag{22.131}$$

by which the concentrations take the form

$$n = n_{p0} \exp(u - \chi_n) \,, \qquad p = p_{p0} \exp(\chi_p - u) \,. \tag{22.132}$$

In the equilibrium limit it is $\chi_n, \chi_p \to 0$. Moreover, when a nonequilibrium condition holds, at any position y in the channel it is $\lim_{x\to\infty} \chi_n, \chi_p = 0$; in fact, as observed in Sect. 22.8, far from the channel the semiconductor is practically in the equilibrium condition, whence $\varphi_n \to \varphi_F$ as x increases. The same applies to φ_p. With these provisions, the charge density in the semiconductor reads $\varrho = q \left[p_{p0} \exp(\chi_p - u) - n_{p0} \exp(u - \chi_n) - N_A \right]$, namely,

$$\varrho = -q\, p_{p0}\, A \,, \qquad A = \frac{n_i^2}{N_A^2} \left[\exp(u - \chi_n) - 1 \right] + 1 - \exp(\chi_p - u) \,, \tag{22.133}$$

and the Poisson equation takes the form

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = \frac{1}{L_A^2} A \,, \tag{22.134}$$

with L_A the holes' Debye length defined in (21.12). One notes that (22.133), (22.134) are generalizations of (22.3). In the description of the MOSFET the accumulation condition is not considered, hence the holes' contribution $\exp(\chi_p - u)$ to A is negligible in the channel region; thus,

$$A \simeq \frac{n_i^2}{N_A^2} \left[\exp(u - \chi_n) - 1 \right] + 1 \simeq \frac{n_i^2}{N_A^2} \exp(u - \chi_n) + 1 > 0 \,, \tag{22.135}$$

where the term $n_i^2 / N_A^2 = \exp(-2\, u_F)$ is negligible with respect to unity. Remembering that the quasi-Fermi potential in the channel does not depend on x, in (22.135) it is $u = u(x, y)$, $\chi_n = \chi_n(y)$, with $u(\infty, y) = 0$, $u'(\infty, y) = 0$ due to the charge neutrality of the bulk region; in fact, as shown by numerical solutions, both u and u' practically vanish when x reaches the value of the depletion width $x_d(y)$.

The Poisson equation is to be solved in two dimensions. If the condition of a well-formed channel holds, the components of the electric field along the x and y directions are quite different from each other. The x component at the semiconductor–oxide interface, E_{sx}, which is due to the voltage applied to the gate contact, is large because it maintains the strong-inversion condition of the surface. Moreover, the derivative $\partial E_x / \partial x = -\partial^2 u / \partial x^2$ is also large, because E_x changes

from E_{sx} to zero in the short distance $x_d(y)$. In contrast, the y component of the electric field at the interface, E_{sy}, which is due to the voltage V_{DS} applied between the drain and source contacts, is small; in fact, V_{DS} is small in itself because the linear region only is considered, and the channel length L is larger than the insulator thickness. Moreover, numerical solutions show that for $0 < V_{DS} < V_{DS}^{\text{sat}}$ the dependence of both E_{sx} and E_{sy} on y is weak, which in particular makes $\partial E_y/\partial y = -\partial^2 u/\partial y^2$ also small at the semiconductor–insulator interface. In conclusion, one approximates (22.134) as

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \simeq \frac{\mathrm{d}^2 u}{\mathrm{d}x^2} = \frac{1}{L_A^2} A \,. \tag{22.136}$$

The dependence of A on y remains, and y is treated as a parameter in the solution procedure. Due to the form of (22.136), the solution method is identical to that used in Sect. 22.2.1 to treat the equilibrium case; it yields

$$\left(\frac{q E_{sx}}{k_B T}\right)^2 = \frac{2}{L_A^2} F^2 \,, \qquad F^2 = \exp(-\chi_n - 2\,u_F)\,[\exp(u_s) - 1] + u_s \,. \tag{22.137}$$

Remembering that the accumulation condition is excluded, here it is $u_s(y) \geq 0$; the flat-band condition $u_s = 0$ corresponds to $F = 0$. In the strong-inversion condition the contribution of the electron charge (proportional to $\exp(u_s) - 1$ in (22.137)) is dominant; for this to happen it is necessary that the exponent $u_s - \chi_n - 2\,u_F$ in (22.136) be positive; it follows that the threshold condition is identified by $u_s = \chi_n + 2\,u_F$. Remembering the definition (22.131) of χ_n one then finds

$$u_s = u_n + u_F \,, \qquad 0 \leq y \leq L \,, \tag{22.138}$$

that is, the normalized form of (22.75). Note that $u_s = \chi_n + 2\,u_F$ is coherent with the definition of the threshold condition at equilibrium (Table 22.1), which is obtained by letting $\chi_n = 0$. Also, specifying $u_s = \chi_n + 2\,u_F$ at the source and drain ends of the channel provides the integration limits (22.76) [128, Sec. 10-2].

The neglect of the variation in the y component of the electric field along the channel makes the general relation

$$Q_{\text{sc}} = -C_{\text{ox}}\,(V'_{GB} - \varphi_s) = -C_{\text{ox}}\,\frac{k_B T}{q}\,(u'_{GB} - u_s) \tag{22.139}$$

still valid at each position y along the channel, with $Q_{\text{sc}} < 0$ because $u_s > 0$ ($Q_{\text{sc}} = 0$ in the flat-band condition $u_s = 0$). Also, when the inversion charge is approximated by a charge layer at $x = 0^+$, the volume charge is entirely due to the ionized dopants, whose contribution in the second relation of (22.137) is proportional to u_s like in the equilibrium case. Using the same relation $Q_i = Q_{\text{sc}} - Q_b$ as in Sect. 22.4 one finally finds that the theory worked out in this section makes (22.34) applicable also in a nonequilibrium condition in which the channel is well formed.

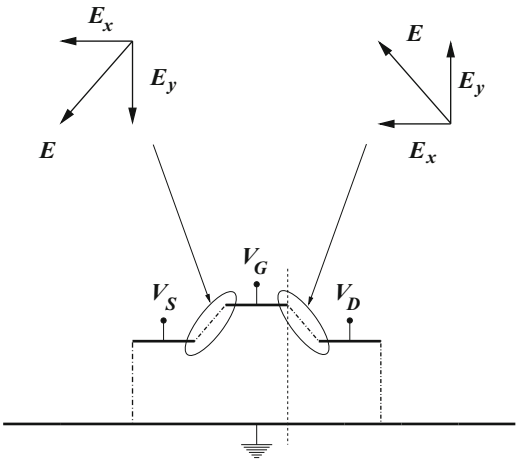


Fig. 22.26 Illustration of the electric field's components at the channel ends in the saturation condition

22.11.2 Inversion-Layer Charge and Mobility Degradation

In the model (22.78) for the drain current worked out in the previous sections, the inversion-layer charge Q_i decreases from source to drain due to the increase in φ_s along the channel (compare with (22.34)). Considering that the MOSFET current is carried by the inversion-layer charge, the vanishing of the latter occurring at the drain end of the channel when $V_{DS} \to V_{DS}^{\text{sat}}$ may seem an oddity. However, it is important to remember that a number of approximations are necessary to reach the result expressed by (22.86) and (22.87); such approximations make the theory applicable only in the linear region and in the condition of a well-formed channel.

When $V_{DS} \to V_{DS}^{\text{sat}}$, the vertical component of the electric field at the interface, E_{sx}, is made weaker by the interplay between the voltages applied to the gate electrode and to the nearby drain electrode; for this reason, at the drain end of the channel the flow lines of $\mathbf{J}_n$ do not keep close to the interface any longer, but spread into the substrate, thus decreasing the carrier density. The phenomenon is better understood with the aid of Fig. 22.26, where the linear-parabolic model is used with, e.g., $V_T = 0.5$, $V_S = 0$, $V_{GS} = 1.5$, $V_{DS} = 2$ V. As $V_{DS}^{\text{sat}} = 1$ V, the saturation condition $V_{DS} > V_{DS}^{\text{sat}} = 1$ holds. Along the dash-dotted line enclosed in the right oval, the direction of the electric field is that shown in the vector diagram in the upper-right part of the figure; in particular, the vertical component of the field at the position marked by the vertical dashed line points upwards. As a consequence, the channel electrons are repelled downwards and the flow lines of the current density detach from the interface. On the source end of the channel, instead, the vertical component of the field points downwards. By continuity, a position within the channel exists where, in saturation, the vertical component of the field vanishes; such a position (not shown in the figure) is called *inversion point*. Also, the large component of the electric field along the y direction, which exists within the space-charge region of the reverse-biased drain junction, makes the carriers'

average velocity to increase at the drain end of the channel; as the total current is constant, a further decrease in the carrier concentration occurs. The two effects briefly illustrated above are not accounted for by the simplified model and require a more elaborate approach in which the two-dimensional structure of the electric field is accounted for.

Another comment is useful, this time related to the *off* condition. When $V_{GS} \to V_T$, the current of a real device does not actually vanish; in fact, a current (called *subthreshold current*), due to the carriers present in the channel in the weak-inversion condition, flows for $V_{GS} < V_T$. Also in this case a more elaborate theory is necessary, showing that the subthreshold current vanishes exponentially as V_{GS} decreases (Sect. 22.11.5).

It is worth concluding this section by commenting the simplification used in (22.74), where the effective electron mobility μ_e is assumed to be independent of the position y along the channel. The factors that affect mobility in a MOSFET are collisions with phonons, ionized impurities, and semiconductor–oxide interface; remembering the features of the macroscopic mobility models (Sect. 20.5), the electron mobility $\mu_n(x, y)$ is made to depend on the lattice temperature T, concentration of ionized impurities N_A, and x component of the electric field $E_x(x, y)$. The first two parameters, T and N_A, do not introduce a dependence on position because they are themselves constant. The x dependence of E_x is absorbed by the integral (22.71) that defines μ_e; it follows that the average mobility depends on y because E_x does. Such a dependence, in turn, is relatively weak in the strong-inversion condition as remarked in Sect. 22.11.1. Therefore, the dependence of μ_e on position is considered negligible.

22.11.3 Comments About the CMOS Inverter

With respect to other technologies, the CMOS inverter has a number of advantages; they are briefly listed below, using for comparison the old-fashioned inverters made of an n-channel MOSFET with a resistive load (Fig. 22.27), or a p-channel MOSFET still with a resistive load (Fig. 22.28). For a given bias V_{SS}, the CMOS inverter has the largest difference between the maximum and minimum output voltages (*dynamic range*). In fact, as shown in Sect. 22.10, in a CMOS inverter described with the linear-parabolic model such voltages are V_{SS} and zero, respectively; a resistive-load inverter (Probs. 22.7 and 22.8) cannot attain both limits.[35] Connected to the above comment is the fact that the transition of V_{out} from V_{SS} to zero is actively forced by the turning on of the n-channel transistor (*pull-down device*); similarly, the transition of V_{out} from zero to V_{SS} is actively forced by the turning on of the

[35] As shown in Sect. 22.11.5, in a real CMOS inverter the dynamic range is slightly decreased, with respect to V_{SS}, due to the subthreshold current; however, the conclusions about the CMOS inverter's superiority do not change.

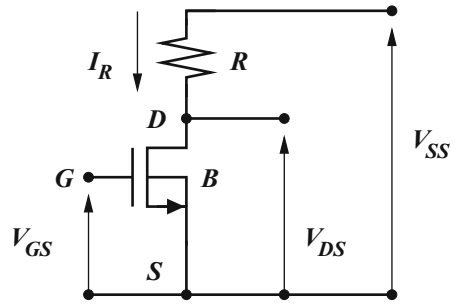


Fig. 22.27 The inverter made of an *n*-channel MOSFET with a resistive load. It is $V_{\text{in}} = V_{GS}$, $V_{\text{out}} = V_{DS}$

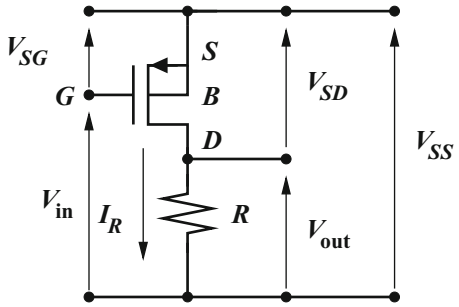


Fig. 22.28 The inverter made of a *p*-channel MOSFET with a resistive load

p-channel transistor (*pull-up device*). In a resistive-load inverter, instead, one of the two transitions is produced by the much less efficient charging (discharging), through the load resistor, of the capacitance associated with the logic gates connected to the inverter's output.

Another advantage of the CMOS inverter is that it has almost zero power consumption in the steady-state condition. In the linear-parabolic model the static power consumption is exactly zero because the current is such; in a real CMOS inverter, some power is dissipated due to subthreshold conduction; however, it is still much smaller than in other technologies.

More energy is consumed during the transients; assuming for simplicity that the capacitor C associated with the output's logic gates is linear, the energy dissipated during either transient is $C\,V_{SS}^2/2$: such an energy is still very small due to the fact that MOS transistors are an easily scalable technology, which is another advantage of the CMOS inverter with respect to earlier implementations.[36] The combination of small power consumption and ease of miniaturization, along with the relatively simple and easy-to-replicate structure of the CMOS gates, makes it possible to manufacture densely integrated and highly performing circuits. It is mainly thanks to these properties that the CMOS architecture in silicon is to date (2017) the most successful technology for manufacturing microprocessors or other integrated circuits.

[36] An integrated resistor occupies a larger area than a MOSFET; this is another reason why resistors are not suited for miniaturization. The scaling procedures are illustrated in some detail in Sect. 22.11.6.

22.11.4 Exact Charge Partitioning in the MOS Capacitor

In view of the application to the theory of MOS transistors, the analysis of the charge distribution within the MOS capacitor, in all functioning regimes with the exclusion of accumulation, has been carried out using the approximation of Sect. 22.4; with reference to the p-substrate capacitor, the procedure consists in partitioning the semiconductor's volume into the space-charge and quasi-neutral regions, and using the expression (22.32) of the bulk charge per unit area Q_b, with x_d the extension of the space-charge region. From this, the simplified form (22.34) is found, that describes the inversion layer's charge of a p-substrate MOS capacitor in the depletion, weak-inversion, and strong-inversion regimes.

One may argue that a more precise analysis could be carried out without an *a priori* partitioning of the semiconductor's volume into regions; this is in fact possible, starting from the integral expression (22.14) of the semiconductor charge per unit area,[37] and considering the individual charge contributions within the integral. In this process, however, care must be taken to avoid diverging terms. To investigate this issue one observes that if the partitioning into regions is not carried out, the integral defining the semiconductor charge per unit area extends over an infinite domain,

$$Q_{\rm sc} = \int_0^\infty q\,(p - n - N_A)\,\mathrm{d}x\,. \tag{22.140}$$

While the integral as a whole does converge because the asymptotic vanishing of the integrand is sufficiently fast (see below), the integral of p or n alone may not converge, and the integral of N_A obviously diverges. This outcome is an artifact due to the consideration of an infinite domain; however, as shown below, a separate calculation of the charges is still possible, keeping the domain as is: in this way, the complicacy of placing the neutrality condition at a finite distance is avoided. To proceed, one remembers that in a p-type substrate the relation between the normalized electric potential and normalized field at any position within the semiconductor is given by (22.5), where the sign is negative because accumulation is excluded. Remembering that the charge density reads $\varrho = -q\,p_{p0}\,A(u)$, with A given by the second relation in (22.3), one changes the integration variable in (22.140) by letting $\mathrm{d}x = \mathrm{d}u/u' = -L_A\,\mathrm{d}u/(\sqrt{2}\,F)$. Using the integration limits $u(0) = u_s$ and $u(\infty) = 0$, one finds

$$Q_{\rm sc} = -Q_{\rm sc}^{(1)} \int_0^{u_s} \frac{A(u)}{2\,F(u)}\,\mathrm{d}u\,, \tag{22.141}$$

[37] The effect of a fixed oxide charge is included in (22.14) by redefining V'_G as indicated at the end of Sect. 22.4.1.

with $Q_{\rm sc}^{(1)}$ given by the second relation in (22.28). In the $u \to 0$ limit, corresponding to the asymptotic condition in the x coordinate, the integrand takes the indeterminate form $0/0$; however, near $u = 0$ it is $A \sim (1 + n_i^2/N_A^2)\, u$ and $F \sim (1 + n_i^2/N_A^2)^{1/2}\, u/\sqrt{2}$, which shows that the limit of A/F is finite, whence the integral converges. On the contrary, if one replaces A with $(n_i^2/N_A^2)\, \exp(u)$, to calculate the electrons' contribution alone, or with $-\exp(-u)$, to calculate that of the holes, the corresponding integrals diverge. If, instead, one uses in the numerator the quantities $(n_i^2/N_A^2)\, [\exp(u) - 1]$ and $1 - \exp(-u)$, the integrals converge; when the normalization is eliminated, such quantities correspond $n - n_{p0}$ and $N_A - p$, respectively. Observing that their sum is A, and that in the depletion, weak-inversion, and strong-inversion regimes it is $n \gg n_{p0}$ and $N_A \gg p$, one lets

$$Q_i = -Q_{\rm sc}^{(1)}\, \exp(-2\, u_F) \int_0^{u_s} \frac{\exp(u) - 1}{2\, F(u)}\, {\rm d}u\,, \qquad \exp(-2\, u_F) = \frac{n_i^2}{N_A^2}\,, \tag{22.142}$$

$$Q_b = -Q_{\rm sc}^{(1)} \int_0^{u_s} \frac{1 - \exp(-u)}{2\, F(u)}\, {\rm d}u\,. \tag{22.143}$$

The above fulfills the relation $Q_{\rm sc} = Q_i + Q_b$ and provide a more exact partitioning of the charge in the MOS capacitor. The analysis of the n-substrate capacitor is similar. One observes that (22.142), (22.143) cannot be integrated analytically as they stand; their sum (22.141), instead, is integrable analytically: remembering (22.28) and combining it with (22.141), one finds in fact

$$F(u_s) = \int_0^{u_s} \frac{A(u)}{2\, F(u)}\, {\rm d}u\,, \tag{22.144}$$

which can also be proven by a direct calculation. Relations (22.142) and (22.143) are useful to check the accuracy of the approximations used in Sect. 22.4; as shown in [5], when accumulation is excluded the approximations are sufficiently accurate and can also be extended to the treatment of the nonequilibrium case (Sect. 22.11.5).

22.11.5 MOSFET Theory Including the Subthreshold Current

The current-voltage relation of the n-channel MOSFET has been derived in Sect. 22.8.1 starting from (22.74) and applying the gradual-channel approximation. Such an approximation entails the use of (22.34) for the inversion-layer charge per unit area at any position y in the channel, and of the condition (22.75) of a well-formed channel. As indicated in Sect. 22.11.4, the expression of the charge is accurate enough, both in the equilibrium and nonequilibrium conditions. In order to improve the theory, one may then dispose of the condition of a well-formed channel; in this case, when changing the variable from φ_n to φ_s in (22.74), one finds

$$I = \frac{W}{L}\,\mu_e \int_{\varphi_s(0)}^{\varphi_s(L)} Q_i(\varphi_s)\,\frac{\mathrm{d}\varphi_n}{\mathrm{d}\varphi_s}\,\mathrm{d}\varphi_s\,, \tag{22.145}$$

to be compared with (22.77). To proceed, one exploits the results of Sect. 22.11.1, specifically, the definitions (22.131) and the relation (22.137), that holds between surface potential, electron quasi-Fermi potential, and gate voltage when the accumulation condition is excluded. Remembering that at each channel position y it is $E_{sx} = -Q_{\rm sc}/\varepsilon_{sc}$, one recasts (22.137) as[38]

$$r^2\,\left(u'_{GB} - u_s\right)^2 = \exp(-\chi_n - 2\,u_F)\,\left[\exp(u_s) - 1\right] + u_s\,, \quad r = \frac{1}{\sqrt{2}}\,\frac{\varepsilon_{\rm ox}}{\varepsilon_{\rm sc}}\,\frac{L_A}{t_{\rm ox}}\,, \tag{22.146}$$

with $\chi_n = u_n - u_F$. For a given $u'_{GB} = q\,V'_{GB}/(k_B\,T)$, the first relation in (22.146) has the form $\chi_n = \chi_n(u_s)$; it is somewhat simplified by neglecting the unity with respect to $\exp(u_s)$, to find[39]

$$\chi_n = u_s - 2\,u_F - \log\left[r^2\,\left(u'_{GB} - u_s\right)^2 - u_s\right]\,. \tag{22.147}$$

The integration limits of (22.145) are found by inverting (22.147) numerically after letting $\chi_n(0) = q\,V_{SB}/(k_B\,T)$ and $\chi_n(L) = q\,V_{DB}/(k_B\,T)$, respectively. The factor $\mathrm{d}\varphi_n/\mathrm{d}\varphi_s = \mathrm{d}\chi_n/\mathrm{d}u_s$ to be inserted into the integral of (22.145) is readily found to be

$$\frac{\mathrm{d}\chi_n}{\mathrm{d}u_s} = 1 + \frac{2\,r^2\,\left(u'_{GB} - u_s\right) + 1}{r^2\,\left(u'_{GB} - u_s\right)^2 - u_s}\,. \tag{22.148}$$

Note that neglecting the term with the logarithm in (22.147) reproduces (22.75). It follows that the use of (22.147) disposes of the condition of a well-formed channel and allows one to incorporate the subthreshold current in the analysis. Also, it is $\mathrm{d}\chi_n/\mathrm{d}u_s \neq 1$, thus indicating that the transport is not necessarily dominated by the drift term; in fact, the diffusion component of the electron current is included through the quasi-Fermi potential. Using the normalized potential, the relation $I_D = -I$ (Sect. 22.8.1), and the definition (22.78) of β transforms (22.145) into

$$I_D = \beta\,\left(\frac{k_B\,T}{q}\right)^2 \int_{u_s(0)}^{u_s(L)} \left[\left(u'_{GB} - u_s\right) - (1/r)\,\sqrt{u_s}\right]\,\frac{\mathrm{d}\chi_n}{\mathrm{d}u_s}\,\mathrm{d}u_s\,. \tag{22.149}$$

[38] From (22.30) one notes that $r\,\gamma = \sqrt{k_B\,T/q}$.

[39] Prior to the simplification, (22.146) yields $u'_{GB} = 0$ for $u_s = 0$, that is, the flat-band condition as expected. When the unity is neglected, the condition $u_s = 0$ yields $r\,u'_{GB} = \exp(-u_F - \chi_n/2) \leq \exp(-u_F) = n_i/N_A \ll 1$. In turn, at room temperature it is $L_A \simeq 41$ nm when $N_A = 10^{16}$ cm^{-3}, whence $(\varepsilon_{\rm ox}/\varepsilon_{\rm sc})\,L_A/\sqrt{2} \simeq 10$ nm in silicon. Thus, for typical oxide thicknesses, r turns out to be of order unity.

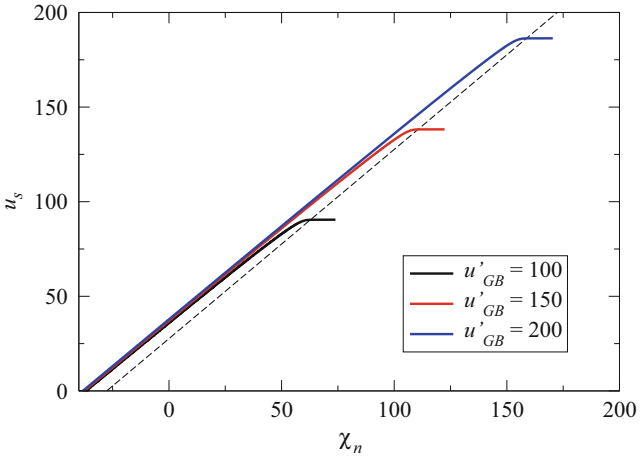


Fig. 22.29 The $u_s = u_s(\chi_n)$ relation calculated from (22.147) for different values of u'_{GB}. The constant acceptor concentration of the substrate is $N_A = 10^{16}$ cm^{-3}, corresponding to $2\,u_F \simeq 27.6$. The *dashed line* shows the $u_s = \chi_n + 2\,u_F$ approximation

It is found by inspection that I_D is made of two parts: the first one derives from the first term (that is, the unity) at the right-hand side of (22.148), and is structurally identical to the current $I'_D - I''_D$ found in Sect. 22.8.2; the only difference here with respect to the case of Sect. 22.8.2 is that one must obtain the integration limits from (22.147) instead of using $V_{SB} + 2\,\varphi_F$ and $V_{DB} + 2\,\varphi_F$ directly. The other part of I_D derives from the second term at the right-hand side of (22.148); its analytical calculation, although a little more elaborate, is feasible (Prob. 22.11).

The form of the $u_s = u_s(\chi_n)$ relation calculated from (22.147) is shown in Fig. 22.29 for different values of u'_{GB}. The dashed line shows the $\chi_n = u_s + 2\,u_F$ approximation; it is found that the approximation is fair when the surface potential is sufficiently smaller than the lower zero of the logarithm's argument in (22.147). Then, the surface potential u_s saturates at the value corresponding to such a zero (the latter is calculated in Prob. 22.11). The observation is important because it shows that (22.149) provides a smooth transition from the linear to the saturation region: when V_{DB} is such that $\chi_n(L) = q\,V_{DB}/(k_B\,T)$ enters the saturation region, the upper limit $u_s(L)$ of the integral (22.149) is pinned to the saturation value; as a consequence, the current does not increase anymore from that value of V_{DB} on.

In contrast to the simpler model of Sect. 22.8, the theory worked out here makes the hypothesis of a strong-inversion regime at each point of the channel unnecessary; with reference to Eq. (22.147) and Fig. 22.29, at some position y, depending on the values of χ_n and u'_{GB} the condition may or may not hold. The latter case is also indicated with *subthreshold condition*. If both V_{DB} and V_{SB} are such that the drain and source ends of the channel are in the subthreshold condition, the whole channel is such, and the expression (22.149) of the drain current becomes (Prob. 22.12):

$$I_D \simeq \beta \, \frac{\exp(u_s^- - 2\,u_F)}{q^2/(k_B\,T)^2} \left[1 + \frac{1/(2\,r^2)}{u'_{GB} - u_s^-}\right] \frac{\exp[-\chi_n(0)] - \exp[-\chi_n(L)]}{(1 + 4\,r^2\,u'_{GB})^{1/2}} \,, \tag{22.150}$$

where $u_s^- = u'_{GB} + [1 - (1 + 4\,r^2\,u'_{GB})^{1/2}]/(2\,r^2)$ (Prob. 22.11). The above result is important because it provides the dependence of the drain current on the gate voltage near threshold, namely, when the channel is making the transition from the *on* to the *off* condition, or vice versa. For a reliable use of the device as a switch, the transitions must occur in a small interval of gate voltages. The strongest dependence of (22.150) on u'_{GB} is the exponential one embedded in u_s^- so that when the other voltages are kept fixed, the subthreshold current reads $I_D \simeq \text{const} \times \exp[q\,V_{GB}/(k_B\,T)]$. In fact, the experimental $\log(I_D)$ vs. V_{GB} relation is to a good approximation a straight line; the *subthreshold slope* (indicated with S, or SS) is defined[40] as

$$S = \left[\frac{\mathrm{d}\log_{10}(I_D)}{\mathrm{d}V_{GB}}\right]^{-1} \simeq \frac{k_B\,T}{q}\,\log(10)\,. \tag{22.151}$$

From the above definition it follows that $(k_B\,T/q)\,\log(10)$ is the variation ΔV_{GB} that makes I_D to vary by a factor 10; at room temperature it is $\Delta V_{GB} \simeq 60$ mV.

22.11.6 Scaling Rules for MOSFETs

An integrated circuit (IC, or *chip*) is a complex electric circuit (e.g., a computer's processor) whose elements are simultaneously fabricated[41] on the same physical support (*wafer*); it may incorporate transducers (e.g., optical, thermal, chemical, or mechanical sensors; solid-state displays). An IC is fabricated by means of a sequence of optical, chemical, and physical processes applied onto a wafer of semiconductor material.[42] The above processes involve only the surface region of the wafer and a limited portion of its thickness; this explains the term *planar technology* by which the fabrication of integrated circuits is typically indicated.

The *integration level* is determined by the chip's area and by the number of elementary devices per unit area. Since their introduction in the 1960s, the speed of the integrated circuits, along with their computing and storage capacity, have increased exponentially; this trend is consistent with the well-known *Moore law*,

[40] The definition of subthreshold slope is not limited to the MOS transistor, but applies in general to the devices where a conductive channel is controlled by a gate. The value of 60 mV per decade calculated in (22.151) is the minimum (that is, the most favorable one); in real devices, S is larger (that is, worse).

[41] In contrast, in a *discrete circuit* the individual components are fabricated independently and assembled afterwards.

[42] Some process steps are illustrated in the chapters devoted to the technological issues.

that is, the observation that the number of transistors in an integrated circuit approximately doubles every two years. This steady improvement in the fabrication of chip makes it possible to identify subsequent "generations" of the integrated-circuit technology, that are referred to as *nodes*, each node being marked by a characteristic length.[43] By way of example, the 1971 node was 10 μm, the 2017 one is 10 nm [68]. Considering that the chip as a whole, and each elementary device in it, are essentially two-dimensional structures, it is sensible to associate a "radius" with each of them; the steady improvement in the resolution and quality of the fabrication processes has made it possible, at each generation of integrated circuits, to decrease by a constant factor the radius of the elementary device and to increase by a constant factor that of the chip: as a consequence, the number of elementary devices of the chip, hence the chip's capability to perform complex functions, have grown exponentially.

In addition to the above, other advantages of miniaturization of the elementary devices are that the electric power absorbed by the individual device decreases, and that of the whole chip decreases as well thanks to clever circuit design made possible by complexity; as a consequence, portable equipments become lighter as the number and size of batteries decrease. Also, as the transistors' switching time decreases along with size, computers of the new generation perform more operations per unit time with respect to those of the old generation.

If the applied voltages are left unchanged, the decrease in size of a device brings about an increase in the electric field. This is not acceptable because it eventually gives rise to, e.g., junction breakdown due to avalanche (Sect. 21.5), or gate-oxide breakdown; it is then necessary to decrease the applied voltages in order to prevent these phenomena from happening. The methodologies by which this goal is achieved are termed *scaling rules*; if, in particular, the latter are such that the electric field within the new-generation device is the same as in the old-generation one, they are referred to as *constant-field scaling rules*.

As the spatial dependence of the electric field is found by solving the Poisson equation $-\varepsilon \nabla^2 \varphi = \varrho$, it is necessary to examine how the equation behaves under scaling. It is assumed that the material does not change from the old-generation to the new-generation device, so that the permittivity is the same in the two cases. The shrinking of the device can be described as a variable transformation, such that the coordinates x_i^{old} of the old-generation device are replaced with $x_i^{\text{new}} = x_i^{\text{old}}/\lambda$, where $\lambda > 1$ is the common scaling factor of the coordinates. At the same time, the electric potential φ^{old} becomes $\varphi^{\text{new}} = \varphi^{\text{old}}/\kappa$, where $\kappa > 1$ is the corresponding scaling factor. For the time being it is assumed that no relation exists between λ and κ; this approach is more general than the constant-field scaling rules, and is referred to as *generalized scaling theory* [6]. It is important to note that the transformation of the electric potential does not involve the interfaces between different media like, for instance, the metal–semiconductor interfaces of the contacts; in fact, the work-

[43] Such a length is that of the *half-pitch*, namely, half the distance between identical features in a DRAM array.

function differences present at the interfaces are not scalable and must be treated separately. With this provision, the Poisson equation in the new-generation device reads

$$-\varepsilon \sum_{i=1}^{3} \frac{\mathrm{d}^2(\varphi^{\mathrm{old}}/\kappa)}{\mathrm{d}(x_i^{\mathrm{old}}/\lambda)^2} = \varrho^{\mathrm{new}}\,, \tag{22.152}$$

that becomes identical to that of the old-generation device if $\varrho^{\mathrm{new}} = \varrho^{\mathrm{old}}\,\lambda^2/\kappa$. At the same time, the components of the electric field transform as

$$E_i^{\mathrm{new}} = -\frac{\partial(\varphi^{\mathrm{old}}/\kappa)}{\partial(x_i^{\mathrm{old}}/\lambda)} = \frac{\lambda}{\kappa}\,E_i^{\mathrm{old}}\,. \tag{22.153}$$

It is readily found from (22.153) that the constant-field scaling rules are derived from the generalized scaling theory by letting $\kappa = \lambda = x_i^{\mathrm{old}}/x_i^{\mathrm{new}}$; in this case, the charge density must scale as $\varrho^{\mathrm{new}} = \lambda\,\varrho^{\mathrm{old}}$. The constant-field scaling relation for ϱ is fulfilled in the insulator regions ($\varrho_{\mathrm{ox}} \simeq 0$); as for the semiconductor regions, where $\varrho = q\,(p - n + N)$, in principle the relation is not fulfilled due to the presence of the electron and hole concentrations. However, one remembers that in most operating conditions the semiconductor domain can be divided into neutral and space-charge regions (Sects. 21.2.2 and 22.2.2). In the space-charge regions, the mobile-carrier concentrations are negligible with respect to the dopant's, $\varrho = q\,(p - n + N) \simeq q\,N$; thus, the scaling relation is fulfilled if

$$N^{\mathrm{new}} = \lambda\,N^{\mathrm{old}}\,. \tag{22.154}$$

In the neutral regions it is $\varrho \simeq 0$, whence the scaling relation is fulfilled anyhow.

The most important application of the scaling rules relates to the MOS technology. Considering the linear-parabolic model for the drain current in the *n*-channel transistor (Sect. 22.8.2), one has

$$I_D = \beta\,[(V_{GS} - V_T)\,V_{DS} - V_{DS}^2/2]\,, \qquad \beta = \frac{W}{L}\,\mu_e\,\frac{\varepsilon_{\mathrm{ox}}}{t_{\mathrm{ox}}}\,. \tag{22.155}$$

The use of (22.155), where the contact voltages appear, seems to contradict the prescription that the interfaces with the contacts should not be considered; one notes, however, that with the exception of the threshold voltage V_T, the voltages appearing in (22.155) are actually voltage differences, from which the non-scalable constants are automatically eliminated.[44] The effect of the remaining non-scalable voltage, V_T, is provisionally neglected by assuming $V_{GS} \gg |V_T|$. In conclusion, the term in brackets of (22.155) can be viewed as the square of a scalable voltage. As far as coefficient β is concerned, the permittivity does not change because, as said

[44] The reasoning implies that all contacts are made of the same material.

before, the material is the same as in the old generation; the geometrical factor, in turn, scales like $W^{\text{new}}/(L^{\text{new}}\, t_{\text{ox}}^{\text{new}}) = \lambda\, W^{\text{old}}/(L^{\text{old}}\, t_{\text{ox}}^{\text{old}})$. Finally, to determine how μ_e behaves under scaling, one refers to the discussion carried out in Sect. 20.5.3 about the mobility dependence on macroscopic parameters in surface-channel devices. Specifically, such parameters are lattice temperature, dopant concentration, and intensity of the electric field's component normal to the semiconductor–insulator interface. Experimental evidence shows that a difference in lattice temperature from one generation to another can be neglected; in turn, the intensity of the electric field's components is obviously invariant. In contrast, one expects a mobility degradation to occur, due to the increase in the dopant concentration as prescribed by (22.154); however, if one limits the analysis to the single step that brings from one generation to the next one, factor λ in (22.154) is limited between 1 and 2, which makes the effect of the increase in the dopant concentration negligible as well.[45] In conclusion, assuming that μ_e is invariant under scaling and using symbol V_X to collectively indicate the voltages in (22.155) yield

$$I_D^{\text{new}} = \text{cost} \times \frac{\left(V_X^{\text{new}}\right)^2}{t_{\text{ox}}^{\text{new}}} = \text{cost} \times \frac{\left(V_X^{\text{old}}/\lambda\right)^2}{t_{\text{ox}}^{\text{old}}/\lambda} = \frac{I_D^{\text{old}}}{\lambda}\,. \tag{22.156}$$

Remembering that in an MOS transistor the electric power is dissipated in the source-drain port (Sect. 22.8), one finds that the power P dissipated within a single transistor scales as

$$P^{\text{new}} = V_{DS}^{\text{new}}\, I_D^{\text{new}} = \frac{V_{DS}^{\text{old}}}{\lambda}\,\frac{I_D^{\text{old}}}{\lambda} = \frac{P^{\text{old}}}{\lambda^2}\,. \tag{22.157}$$

Letting ζ be the number of transistors per unit area, it is $\zeta^{\text{new}} = \lambda^2\, \zeta^{\text{old}}$. Combining the latter with (22.157) shows that, as anticipated above, the power $\tilde{P}$ dissipated per unit area within the chip is invariant under the constant-field scaling:

$$\tilde{P}^{\text{new}} = \zeta^{\text{new}}\, P^{\text{new}} = \lambda^2\, \zeta^{\text{old}}\, \frac{P^{\text{old}}}{\lambda^2} = \zeta^{\text{old}}\, P^{\text{old}} = \tilde{P}^{\text{old}}\,. \tag{22.158}$$

This is a fundamental result: it shows that scaling makes it possible to concentrate a larger number of devices in the same area without producing an excessive heat.[46] Moreover, as anticipated above, smaller transistors exhibit a shorter switching time. To demonstrate this it is convenient to refer to the CMOS architecture (Fig. 22.23) and to examine the charging or discharging processes of the capacitive load that, as indicated in Sect. 22.10, mimics the inverter's fan-out. Letting A be the

[45] In the planar technology, the scaling factor that makes the number of transistors in an integrated circuit to double is $\lambda = \sqrt{2}$.

[46] This also explains why variations in μ_e are negligible as far as the effect of temperature is concerned.

capacitor area, and considering a linear capacitor for simplicity, one finds $C^{\text{new}} = \varepsilon_{\text{ox}} A^{\text{new}}/t_{\text{ox}}^{\text{new}} = \varepsilon_{\text{ox}} (A^{\text{old}}/\lambda^2)/(t_{\text{ox}}^{\text{old}}/\lambda) = C^{\text{old}}/\lambda$. Remembering the description of Sect. 22.10.1, the voltage V across the capacitor coincides with V_{DSn}; it is equal to V_{SS} when the p-channel transistor is in the *on* state and the n-channel transistor is in the *off* state. When the states reverse, the capacitor discharges through the n-channel transistor, which is now in the *on* state; as a consequence, a current flows in the circuit made of the capacitor and the n-channel transistor. In the opposite situation, in which V is initially equal to zero, when the transistors' states reverse the capacitor charges through the p-channel transistor. In both cases, the current I is the drain (source) current of an MOS transistor, whose scaling rule is (22.156); in turn, the equation describing the circuit is $I = C\,\mathrm{d}V/\mathrm{d}t$. Observing that the scaling factors of V and I cancel each other, one finds

$$\mathrm{d}t^{\text{new}} = C^{\text{new}} \frac{\mathrm{d}V^{\text{new}}}{I^{\text{new}}} = \frac{C^{\text{old}}}{\lambda} \frac{\mathrm{d}V^{\text{old}}}{I^{\text{old}}} = \frac{\mathrm{d}t^{\text{old}}}{\lambda}\,, \tag{22.159}$$

namely, within the constant-field scaling rules, the scaling factor of time is the same as that of lengths.

The presence of quantities that are not scalable introduces some complicacies into the plain form of the scaling theory depicted so far; the general conclusions, however, do not change. As mentioned above, non-scalable quantities typically arise at the interfaces among different media; the problem is not necessarily restricted to the device boundaries: for instance, in a reverse-biased abrupt p-n junction (Sect. 21.4), the expression of the electric field at the junction reads

$$E_0 = -\sqrt{\frac{2\,q}{\varepsilon_{\text{sc}}} \frac{\psi_0 - V}{1/N_D + 1/N_A}}\,, \qquad \psi_0 = \frac{k_B\,T}{q} \log\left(\frac{N_D\,N_A}{n_i^2}\right). \tag{22.160}$$

One finds that ψ_0 increases under scaling by the additive term $(2\,k_B\,T/q)\,\log\lambda$; only in the limiting case $|V| \gg \psi_0$ the scaling factors of V and $1/N_D + 1/N_A$ cancel each other, thus making E_0 invariant under scaling. Another non-scalable quantity is the threshold voltage, whose expression for an n-channel transistor is given by (22.85); it embeds the Fermi potential $\varphi_F = (k_B\,T/q)\,\log(N_D/n_i)$, the work-function difference at the interface between semiconductor and bulk contact (through the flat-band voltage, Sect. 22.4.1) and, possibly, the fixed oxide charge and the surface-state charge.[47]

Other problems connected with miniaturization exist, deriving either from fundamental physical limits or practical issues. The insulator thickness belongs to

[47] It is interesting to observe that if $|V_{BS}| \gg 2\,\varphi_F$, the last term of the threshold voltage (22.85) scales by the correct factor λ. Remembering in fact that in the expression (22.30) of γ the factor C_{ox} is a capacitance per unit area, in the approximation $|V_{BS}| \gg 2\,\varphi_F$ it is $\gamma \sqrt{-V_{BS}} = \text{const} \times t_{\text{ox}} \sqrt{-V_{BS}\,N_A}$. As the scaling factors within the square root cancel each other, the whole product $\gamma \sqrt{-V_{BS}}$ scales like t_{ox}.

the first class: as mentioned in Sect. 22.2, the progressive scaling has brought the insulator thickness to the range of nanometers. This may eventually lead to dielectric breakdown and leakage by quantum tunneling; for this reason, silicon dioxide is being replaced in advanced devices with high-k dielectrics. Another problem that is aggravated by miniaturization is the statistical variability produced by the random distribution of the dopant atoms. For instance, implanting a dopant with a 6×10^{18} cm^{-3} concentration into a channel whose size is $40 \times 20 \times 20$ nm^3 yields about 100 implanted atoms; therefore, considering the dopant distribution still a continuous function becomes a rather simplistic approximation. Practical problems, in turn, derive from the increase (of exponential nature) in the cost of investments that are necessary to move from one generation of integrated circuits to the next one, and from the increasing complexity of the integrated circuits' design: the latter issue is in fact tackled by means of advanced techniques for automatic design, and of sophisticated models describing each phase of the design. A brief outline of this aspect is given in Sect. 22.11.7.

22.11.7 A Brief Account on IC Design

A rough scheme of the IC-design tree is depicted in Fig. 22.30. The design of a new IC can be stimulated by a direct request from a customer or by a market analysis carried out by the manufacturer. With reference to the left section of Fig. 22.30, the first steps of the design are a feasibility study including the description of the functional specs, the choice of the technology (e.g., CMOS) along with an estimate of the chip size, followed by chip planning, namely, the placement of the individual blocks forming the chip. The next steps are the design of the circuits able to abide by the functional specs, of the devices able to provide the static and dynamic characteristics required by the circuits, and of the processes able to yield the devices of interest. The design of equipments that make the processes feasible is also important; it is uncommon, however, that the design of equipments is accomplished by the same manufacturer that produces the integrated circuits.

As a whole, the steps listed above are also called *top-down design*, because they start from an overview of the global system and break it successively into smaller blocks. This part of the process is followed by the validation part, in which the functionality of the individual blocks is checked separately and in detail (*bottom-up design*). To complete the design it may be necessary to repeat several times the top-down, bottom-up suite. When the design is complete, the layout of the circuit is represented in a figure called *stick diagram*, useful for planning the relative placement of the devices. The stick diagram for the CMOS inverter of Fig. 22.23 is shown in Fig. 22.31 by way of example; the colors have a meaning: the red (light blue) stick represents the p-channel (n-channel) transistor, the green sticks are the input contact, the black sticks represent the bias or ground lines, and the output contact; finally, the dots indicate the positions where different layers contact each other.

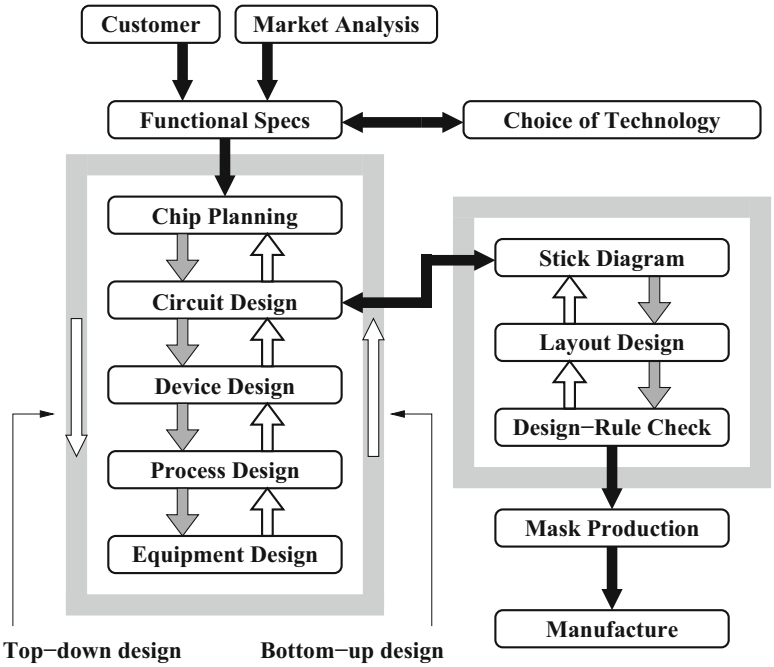
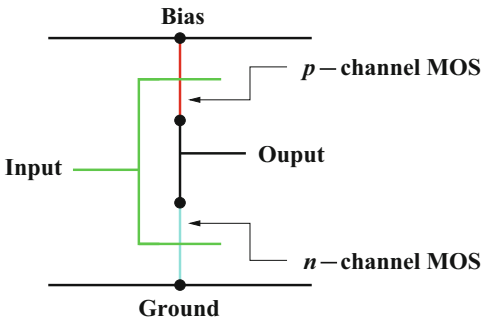


Fig. 22.30 Pictorial scheme of the *top-down* and *bottom-up* design of integrated circuits

Fig. 22.31 Stick diagram of the CMOS inverter of Fig. 22.23



The placement of each device in the layout is accomplished thanks to the stick diagram; after its completion, the circuit's layout is examined by means of an automatic procedure (*design-rule checker*) to verify whether a set of recommended design rules is satisfied; the rules specify restrictions about geometry and connectivity (for instance, minimum width of each shape, minimum distance between two adjacent objects, and many others), able to account for variability in the manufacturing process. The steps involving the production of the stick-diagram, of layout design, and of design-rule checking are depicted in the right section of Fig. 22.30; again, it may be necessary to repeat the suite several times. The final

layout is then used for the production of masks and the rest of the manufacturing process.

It is important to mention that the steps encircled by the left and right frames in Fig. 22.30 are carried out by computer simulation. Considering in particular the design of circuits, devices, processes, and equipment (left frame), two aspects are fundamental for each step: the accurate physical modeling of the problem in hand, and the ability of solving the equations resulting from the physical modeling. An example of numerical methods useful for solving the semiconductor-device equations is given in Sect. A.13.

Problems

22.1 Work out a method for drawing the curves of Fig. 22.9 without approximations.

22.2 Prove that the condition of a well-formed channel in a p-channel MOSFET is given by (22.102).

22.3 Using the linear-parabolic model, find the current of the CMOS inverter when both transistors work in the saturation region.

22.4 Using the linear-parabolic model, determine the analytical form of the $V_{\text{out}}(V_{\text{in}})$ curve between points A and B of Table 22.3. Repeat the calculation for the part of the curve between points C and D.

22.5 Find the conditions that make the $V_{\text{out}}(V_{\text{in}})$ curve antisymmetric around the point $(V_{SS}/2, V_{SS}/2)$.

22.6 Discuss the $V_{\text{out}}(V_{\text{in}})$ curve of the CMOS inverter shown in Fig. 22.25.

22.7 Calculate the $V_{\text{out}}(V_{\text{in}})$ curve of the inverter made of an n-channel MOSFET with a resistive load.

22.8 Draw the $V_{\text{out}}(V_{\text{in}})$ curve of the inverter made of an n-channel, silicon MOSFET with a resistive load, using the expressions found in Prob. 22.7 and the following values for the parameters: $\mu_n = 993\ \text{cm}^2/(\text{V s})$, $t_{\text{ox}} = 5$ nm, $W/L = 1$, $V_{SS} = 5$ V, $R = 3.5\ \text{k}\Omega$, $V_{Tn} = 0.5$ V.

22.9 Calculate the $V_{\text{out}}(V_{\text{in}})$ curve of the inverter made of a p-channel MOSFET with a resistive load.

22.10 Draw the $V_{\text{out}}(V_{\text{in}})$ curve of the inverter made of a p-channel, silicon MOSFET with a resistive load, using the expressions found in Prob. 22.9 and the following values for the parameters: $\mu_p = 662\ \text{cm}^2/(\text{V s})$, $t_{\text{ox}} = 5$ nm, $W/L = 1$, $V_{SS} = 5$ V, $R = 5.25\ \text{k}\Omega$, $V_{Tp} = 0.5$ V.

22.11 Complete the calculation of (22.149).

22.12 Starting from (22.149), calculate the subthreshold expression (22.150) of the drain current.

22.13 Given V'_{GB}, V_{SB}, and V_{DB}, use the concepts of Sects. 22.9 and 22.11.5 to calculate the normalized surface potential u_s at each position y along the channel. Repeat the calculation for χ_n.

22.14 Letting $\tau_n = 10\ \mu$s, $D_n = 10\ \text{cm}^2/\text{s}$, and $a = 10\ \mu$m, find the slowest decay rate associated with (22.55).

Part VIII
Miscellany

Chapter 23
Thermal Diffusion—Ion Implantation

23.1 Introduction

The fabrication of integrated circuits requires the introduction into the semiconductor material of atoms belonging to specifically selected chemical species. Such atoms are called *impurities* or *dopants*. As shown in Chap. 18, the inclusion of dopants into the semiconductor lattice attains the important goals of fixing the concentration of mobile charges in the material and making it practically independent of temperature.

Dopants are divided into two classes, termed *n-type* and *p-type*. With reference to silicon (Si), the typical *n*-type dopants are phosphorus (P), arsenic (As), and antimony (Sb), while the typical *p*-type dopants are boron (B), aluminum (Al), gallium (Ga), and Indium (In). When a dopant atom is introduced into the semiconductor lattice, in order to properly act as a dopant it must replace an atom of the semiconductor, namely, it must occupy a lattice position. When this happens, the dopant atom is also called *substitutional impurity*. An impurity atom that does not occupy a lattice position is called *interstitial*. Interstitials cannot properly act as dopants; in fact, they degrade the conductivity and other electrical properties of the semiconductor.

The concentration of the dopant atoms that are introduced into a semiconductor is smaller by orders of magnitude than the concentration of the semiconductor atoms themselves. As a consequence, the average distance between dopant atoms within the lattice is much larger than that between the semiconductor atoms. Thus, the material resulting from a doping process is not a chemical compound: it is still the semiconductor in which some of the electrical properties are modified by the presence of the dopant atoms. In fact, while the presence and type of dopants are easily revealed by suitable electrical measurements, they may remain undetectable by chemical analyses.

M. Rudan, *Physics of Semiconductor Devices*, DOI 10.1007/978-3-319-63154-7_23

The first part of this chapter illustrates the diffusive transport with reference to the processes that are used for introducing impurities into a semiconductor in a controlled way. First, the expressions of the continuity equation and of the diffusive flux density are derived. These expressions are combined to yield the diffusion equation, whose form is reduced to a one-dimensional model problem. The model problem allows for an analytical solution, based on the Fourier-transform method, that expresses the diffused profile at each instant of time as the convolution of the initial condition and an auxiliary function.

Then, the solution of the model problem is used to calculate the impurity profiles resulting from two important processes of semiconductor technology, namely, the predeposition and the drive-in diffusion. In the next part of this chapter the solution of the model problem is extended to more general situations. Specific data about the parameters governing the diffusion processes in semiconductors are in [61, Chap. 3], [129, Chap. 10], [130, Chap. 7], [88, Chap. 12]. Many carefully drawn illustrations of the diffusion process are found in [93, Sec. 1.5]. The properties of the Fourier transform are illustrated in [89, 146] and Sect. C.2.

The last part of this chapter illustrates the features of ion implantation; it starts from the description of the ion implanter, followed by the analysis of the energy exchange between the ion and the crystal, and the statistical construction of the implanted profile. Some remarks about the annealing techniques are given at the end of this chapter. Specific data about the parameters governing ion implantation may be found in [130, Chap. 8] and [88, Chap. 11]. Many carefully drawn illustrations of the diffusion process are found in [93, Sec. 1.6].

23.2 Continuity Equation

The continuity equation described in this section is a balance relation for the number of particles. Here it is not necessary to specify the type of particles that are being considered: they may be material particles, like molecules or electrons, particles associated with the electromagnetic field (photons), those associated with the vibrational modes of atoms (phonons), and so on.[1] Although the type of particles is not specified, it is assumed that all particles considered in the calculations are of the same type.

The balance relation is obtained by considering the space where the particles belong and selecting an arbitrary volume V in it, whose boundary surface is denoted with S. The position of the volume is fixed. Let $\mathcal{N}(t)$ be the number of particles

[1] The continuity equations for the free carriers in semiconductors, and the corresponding transport equations, are deduced and extensively discussed in Chap. 19. The derivation of the continuity equation for the impurity atoms is repeated here in a more concise manner, to let the reader use the material of this chapter without the need of referring to concepts introduced elsewhere in the book. The transport equation here is simpler than those of Chap. 19 because the drift term is missing due to the atoms' neutrality. Other types of continuity equations are illustrated in Chaps. 4 and 5 with reference to the electromagnetic field, and in Chap. 9 with reference to the wave function.

that are inside S at time t. Due to the motion of the particles, in a given time interval some of them move across S in the outward direction, namely, from the interior to the exterior of S. In the same interval of time, other particles move across S in the inward direction. Let $\mathscr{F}_{\text{out}}(t)$ and $\mathscr{F}_{\text{in}}(t)$ be the number of particles per unit time that cross S in the outward or inward direction, respectively, and let $\mathscr{F} = \mathscr{F}_{\text{out}} - \mathscr{F}_{\text{in}}$. The quantity $\mathscr{F}$, whose units are s^{-1}, is the *flux* of the particles across the surface S. If the only reason that makes $\mathscr{N}$ to change is the crossing of S by some particles, the balance relation takes the form of the first equation in (23.1). The minus sign at the right-hand side is due to the definition of $\mathscr{F}$; in fact, $\mathscr{N}$ decreases with time when $\mathscr{F} > 0$, and vice versa.

Besides the crossing of the boundary S by some particles, there is another mechanism able to contribute to the time variation of $\mathscr{N}$, namely, the generation or destruction of particles inside the volume V. This possibility seems to violate some commonly accepted conservation principle. However, it is not so, as some examples given in Sect. 23.8.1 will show. As a consequence, the description of the particle generation or destruction must be included. This is accomplished by letting $\mathscr{W}_{\text{ge}}(t)$ and $\mathscr{W}_{\text{de}}(t)$ be the number of particles per unit time that are generated or, respectively, destroyed within the volume V. Defining $\mathscr{W} = \mathscr{W}_{\text{ge}} - \mathscr{W}_{\text{de}}$, the balance relation that holds when generation or destruction are present takes the form of the second equation in (23.1):

$$\frac{\mathrm{d}\mathscr{N}}{\mathrm{d}t} = -\mathscr{F}\,, \qquad \frac{\mathrm{d}\mathscr{N}}{\mathrm{d}t} = -\mathscr{F} + \mathscr{W}\,. \tag{23.1}$$

The quantity $\mathscr{W}$, whose units are s^{-1}, is the *net generation rate* within volume V.

It is convenient to recast (23.1) in local form. This is done basing on the second equation of (23.1), which is more general, and is accomplished by describing the motion of the particles as that of a continuous fluid. Such a description is legitimate if V can be partitioned into equal cells of volume $\Delta V_1, \Delta V_2, \ldots$ having the following properties: *i)* the cells can be treated as infinitesimal quantities in the length scale of the problem that is being considered and, *ii)* the number of particles within each cell is large enough to make their average properties significant. If the above conditions are fulfilled one lets $\Delta V_k \to \mathrm{d}V$ and introduces the *concentration* $N(\mathbf{r}, t)$, such that $N\,\mathrm{d}V$ is the number of particles that at time t belong to the volume $\mathrm{d}V$ centered at position $\mathbf{r}$. Similarly, one defines the *net generation rate per unit volume* $W(\mathbf{r}, t)$ such that $W\,\mathrm{d}V$ is the net generation rate at time t within $\mathrm{d}V$. The units of N and W are m^{-3} and $\text{m}^{-3}\text{s}^{-1}$, respectively. From the definitions of N, W it follows that the number $\mathscr{N}(t)$ of particles that are inside S at time t is found by integrating $N(\mathbf{r}, t)$ over V, and that the net generation rate $\mathscr{W}(t)$ is found by integrating W over V.

To recast in local form the part of (23.1) related to the flux $\mathscr{F}$, one associates a velocity $\mathbf{v}(\mathbf{r}, t)$ to the concentration $N(\mathbf{r}, t)$. In general, such a velocity is different from the velocity of each individual particle that contributes to the concentration N. In fact, $\mathbf{v}$ is a suitable average of the particles' velocities, whose definition (6.6) is given in Sect. 6.2. In the elementary time $\mathrm{d}t$ the concentration originally in $\mathbf{r}$ moves over a distance $\mathbf{v}\,\mathrm{d}t$ in the direction of $\mathbf{v}$. As a consequence, if $\mathbf{r}$ belongs to the

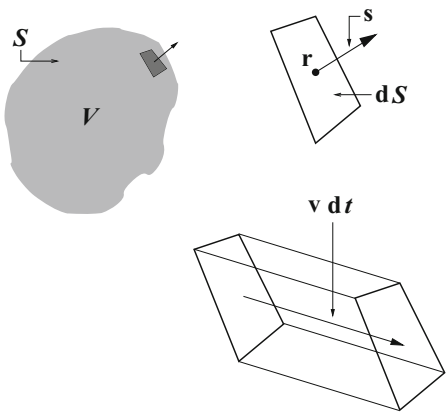


Fig. 23.1 Illustration of the symbols used in the calculation of the flux

boundary surface S, a crossing of S by the particles may occur, thus contributing to the flux. To calculate the contribution to the flux at a point $\mathbf{r}$ belonging to S, one takes the plane tangent to S at $\mathbf{r}$ and considers an elementary area $\mathrm{d}S$ of this plane centered at $\mathbf{r}$ (Fig. 23.1). The construction implies that the surface S is smooth enough to allow for the definition of the tangent plane at each point of it.

Let $\mathbf{s}$ be the unit vector normal to $\mathrm{d}S$, oriented in the outward direction with respect to S. If $\mathbf{v}$ is normal to $\mathbf{s}$, no crossing of S occurs and the contribution to the flux at point $\mathbf{r}$ is zero. If the scalar product $\mathbf{v}\cdot\mathbf{s}$ is positive, the crossing occurs in the outward direction and contributes to $\mathscr{F}_{\mathrm{out}}$. Its contribution is found by observing that the elementary cylinder, whose base area and side are $\mathrm{d}S$ and, respectively, $\mathbf{v}\,\mathrm{d}t$, has a volume equal to $\mathbf{v}\cdot\mathbf{s}\,\mathrm{d}S\,\mathrm{d}t$. Due to the sign of $\mathbf{v}\cdot\mathbf{s}$, the cylinder is outside the surface S. The number of particles in the cylinder is found by multiplying its volume by the concentration $N(\mathbf{r},t)$. Letting $\mathbf{F}=N\mathbf{v}$, such a number reads $\mathbf{F}\cdot\mathbf{s}\,\mathrm{d}S\,\mathrm{d}t$. As the particles that are in the cylinder at time $t+\mathrm{d}t$ were inside the surface S at time t, dividing the above expression by $\mathrm{d}t$ yields the elementary contribution of point $\mathbf{r}$ to the flux, $\mathrm{d}\mathscr{F}=\mathbf{F}\cdot\mathbf{s}\,\mathrm{d}S>0$. The contribution from a point $\mathbf{r}$ where $\mathbf{v}\cdot\mathbf{s}<0$ is calculated in a similar way. The flux $\mathscr{F}$ is then found by integrating $\mathbf{F}\cdot\mathbf{s}$ over the surface S. The quantity $\mathbf{F}\cdot\mathbf{s}=\mathrm{d}\mathscr{F}/\mathrm{d}S$, whose units are $\mathrm{m}^{-2}\mathrm{s}^{-1}$, is the *flux density*.

Introducing the relations found so far into the second form of (23.1) and interchanging the derivative with respect to t with the integral over V yield

$$\int_V\left(\frac{\partial N}{\partial t}-W\right)\mathrm{d}V=-\int_S\mathbf{F}\cdot\mathbf{s}\,\mathrm{d}S=-\int_V\operatorname{div}\mathbf{F}\,\mathrm{d}V\,. \tag{23.2}$$

The last equality in (23.2) is due to the divergence theorem (A.23), whereas the use of the partial-derivative symbol is due to the fact that the N, in contrast with $\mathscr{N}$, depends also on $\mathbf{r}$. The procedure leading to (23.2) does not prescribe any constraint on the choice of the volume V. As a consequence, the two integrals over V that appear in (23.2) are equal to each other for any V. It follows that the corresponding integrands must be equal to each other, thus yielding the *continuity equation*

$$\frac{\partial N}{\partial t} + \operatorname{div} \mathbf{F} = W\,, \qquad \mathbf{F} = N\mathbf{v}\,. \tag{23.3}$$

As mentioned above, (23.3) is the local form of the second equation of (23.1), which in turn is a balance relation for the number of particles. In the steady-state condition the quantities appearing in (23.3) do not depend explicitly on time, hence (23.3) reduces to $\operatorname{div} \mathbf{F} = W$. In the equilibrium condition it is $\mathbf{v} = 0$ and (23.3) reduces to the identity $0 = 0$. It is worth noting that in the equilibrium condition the velocity of each particle may differ from zero; however, the distribution of the individual velocities within any elementary volume $\mathrm{d}V$ is such that the average velocity $\mathbf{v}$ vanishes. Similarly, in the equilibrium condition the generation or destruction of particles still occurs; however, they balance each other within any $\mathrm{d}V$.

To proceed it is assumed that the net generation rate per unit volume W, besides depending explicitly on $\mathbf{r}$ and t, may also depend on N and $\mathbf{F}$, but not on other functions different from them.

23.3 Diffusive Transport

The continuity equation (23.3) provides a relation between the two quantities N and $\mathbf{F}$ (or, equivalently, N and $\mathbf{v}$). If both N and $\mathbf{F}$ are unknown it is impossible, even in the simple case $W = 0$, to calculate them from (23.3) alone. However, depending on the specific problem that is being considered, one can introduce additional relations that eventually provide a closed system of differential equations. The important case of the *diffusive transport* is considered in this section.

It is convenient to specify, first, that the term *transport* indicates the condition where an average motion of the particles exists, namely $\mathbf{F} \neq 0$ for some $\mathbf{r}$ and t. The type of transport in which the condition $\mathbf{F} \neq 0$ is caused only by the spatial nonuniformity of the particles' concentration N is called *diffusive*. Simple examples of diffusive transport are those of a liquid within another liquid, or of a gas within another gas. They show that in the diffusive motion of the particles, the flux is oriented from the regions where the concentration is larger towards the regions where the concentration is smaller.

The analytical description of the diffusion process dates back to 1855 [48]. Here the relation between $\mathbf{F}$ and N in the diffusive case is determined heuristically, basing on the observation that $\operatorname{grad} N$ is a sensible indicator of the spatial nonuniformity of N. Specifically it is assumed, first, that $\mathbf{F}$ depends on N and $\operatorname{grad} N$, but not on higher-order derivatives of N. The dependence on $\operatorname{grad} N$ is taken linear, $\mathbf{F} = \mathbf{F}_0 - D \operatorname{grad} N$, with $\mathbf{F}_0 = 0$ because $\mathbf{F}$ must vanish when the concentration is uniform. Finally, one remembers that the particles' flux is oriented in the direction of a decreasing concentration, namely, opposite to $\operatorname{grad} N$. It follows that $D > 0$, so that the relation takes the form

$$\mathbf{F} = -D \operatorname{grad} N\,, \qquad D > 0\,. \tag{23.4}$$

The above is called *transport equation of the diffusion type*, or *Fick's first law of diffusion*. Parameter D is the *diffusion coefficient*, whose units are m^2s^{-1}. From the derivation leading to (23.4) it follows that if a dependence of $\mathbf{F}$ on N exists, it must be embedded in D. In the case $D = D(N)$ the relation (23.4) is linear with respect to $\operatorname{grad} N$, but not with respect to N. The diffusion coefficient may also depend explicitly on $\mathbf{r}$ and t. For instance, it depends on position when the medium where the diffusion occurs is nonuniform; it depends on time when an external condition that influences D, e.g., temperature, changes with time.

For the typical dopants used in the silicon technology, and in the temperature range of the thermal-diffusion processes, the experimentally determined dependence on temperature of the diffusion coefficient can be approximated by the expression

$$D = D_0 \exp[-E_a/(k_B T)] , \tag{23.5}$$

where k_B (J K^{-1}) is the Boltzmann constant and T (K) the process temperature. In turn, the *activation energy* E_a and D_0 are parameters whose values depend on the material involved in the diffusion process. The form of (23.5) makes it more convenient to draw it as an *Arrhenius plot*, that displays the logarithm of the function using the inverse temperature as a variable: $\log D = \log(D_0) - (E_a/k_B)\,(1/T)$. At the diffusion temperatures, E_a and D_0 can often be considered independent of temperature. In this case the Arrhenius plot is a straight line (examples of Arrhenius plots are given in Chap. 24). At room temperature the diffusion coefficient of dopants in silicon is too small to make diffusion significant. In order to activate the diffusion mechanism a high-temperature process is necessary, typically between 900°C and 1,100°C. Data about the diffusion coefficients of different dopants in silicon are reported, e.g., in [130, Sect. 7.5].

23.4 Diffusion Equation—Model Problem

Inserting (23.4) into (23.3) yields the *diffusion equation*

$$\frac{\partial N}{\partial t} = \operatorname{div}(D \operatorname{grad} N) + W , \tag{23.6}$$

where W depends on $\mathbf{r}$, t, N, and $\operatorname{grad} N$ at most, while D depends on $\mathbf{r}$, t, and N at most. The above is a differential equation in the only unknown N. It must be supplemented with the initial condition $N_0(\mathbf{r}) = N(\mathbf{r}, t = 0)$ and suitable boundary conditions for $t > 0$. If the diffusion coefficient is constant, or depends on t at most, (23.6) becomes

$$\frac{\partial N}{\partial t} = D\,\nabla^2 N + W , \qquad D = D(t) . \tag{23.7}$$

It is convenient to consider a simplified form of (23.7) to be used as a model problem. For this, one takes the one-dimensional case in the x direction and lets $W = 0$, thus yielding

$$\frac{\partial N}{\partial t} = D \frac{\partial^2 N}{\partial x^2} . \tag{23.8}$$

Equation (23.8) is also called *Fick's second law of diffusion*. Thanks to the linearity of (23.8), the solution can be tackled by means of the Fourier-transform method, specifically, by transforming both sides of (23.8) with respect to x. Indicating[2] with $G(k, t) = \mathscr{F}_x N$ the transform of N with respect to x, and using some of the properties of the Fourier transform illustrated in Appendix C.2, one finds

$$\mathscr{F}_x \frac{\partial N}{\partial t} = \frac{\mathrm{d}G}{\mathrm{d}t} , \qquad \mathscr{F}_x D \frac{\partial^2 N}{\partial x^2} = D \mathscr{F}_x \frac{\partial^2 N}{\partial x^2} = -k^2 D G . \tag{23.9}$$

The symbol of total derivative is used at the right-hand side of the first of (23.9) because k is considered as a parameter. The Fourier transform of the initial condition of N provides the initial condition for G, namely, $G_0 = G(k, t = 0) = \mathscr{F}_x N_0$. Equating the right-hand sides of (23.9) and rearranging yield $\mathrm{d}G/G = -k^2 D(t)\,\mathrm{d}t$. Integrating the latter from 0 to t,

$$\log(G/G_0) = -k^2 a(t) , \qquad a(t) = \int_0^t D(t')\,\mathrm{d}t' , \tag{23.10}$$

with a an area. The concentration N is now found by antitransforming the expression of G extracted from the first of (23.10):

$$N(x, t) = \mathscr{F}_k^{-1} G = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} G_0 \, \exp(\mathrm{i}kx - ak^2)\,\mathrm{d}k . \tag{23.11}$$

In turn, G_0 within the integral of (23.11) is expressed as the transform of N_0. After rearranging the integrals one finds

$$N(x, t) = \int_{-\infty}^{+\infty} N_0(\xi) \left\{ \int_{-\infty}^{+\infty} \frac{1}{2\pi} \exp[\mathrm{i}k\,(x - \xi) - ak^2]\,\mathrm{d}k \right\} \mathrm{d}\xi . \tag{23.12}$$

As shown in Appendix C.7, the expression in braces in (23.12) is the integral form of the function $\Delta(x - \xi, t)$ defined by (C.76). As a consequence, the solution of the simplified form (23.8) of the diffusion equation is the convolution between Δ and the initial condition N_0, namely,

[2] Symbol $\mathscr{F}_x$ indicating the Fourier transform should not be confused with the symbol $\mathscr{F}$ used for the particles' flux in Sect. 23.2.

$$N(x,t) = \int_{-\infty}^{+\infty} N_0(\xi)\, \Delta(x-\xi,t)\, \mathrm{d}\xi\,. \tag{23.13}$$

A straightforward calculation shows that Δ fulfills (23.8) for all ξ. As a consequence, (23.13) is a solution as well. In addition, due to (C.80), (23.13) also fulfills the initial condition N_0.

23.5 Predeposition and Drive-in Diffusion

Basing on the model problem worked out in Sect. 23.4 it is possible to describe the thermal diffusion of dopants in silicon. The modification induced in the electrical properties of the silicon lattice by the inclusion of atoms belonging to different chemical species (e.g., phosphorus or boron) is described elsewhere (Sects. 18.4.1 and 18.4.2). Here the analysis deals with the diffusion process in itself.

The formation of a diffused profile in silicon is typically obtained in a two-step process [61, 93, 130]. In the first step, called *predeposition*, a shallow layer of dopants is introduced into the semiconductor. The most common predeposition methods are the diffusion from a chemical source in a vapor form or the diffusion from a solid source (e.g., polycrystalline silicon) having a high concentration of dopants in it. In both methods the silicon wafers are placed in contact with the source of dopant within a furnace kept at a high temperature.

During a predeposition step, new dopant atoms are continuously supplied by the source to the silicon region. As a consequence, the number of dopant atoms in the silicon region increases with time. When the desired amount of atoms is reached, the supply of dopants is blocked, whereas the diffusion process is continued. During this new step, called *drive-in diffusion*, the number of dopant atoms in the silicon region remains constant. The drive-in diffusion is continued until a suitable profile is reached.

Typically, the blocking of the flow of dopant atoms from the source to the silicon region is achieved by introducing oxidizing molecules into the furnace atmosphere, thus resulting in the growth of a silicon-dioxide layer at the silicon surface (the details of the oxidation process are given in Chap. 24).

It is worth anticipating that in some processes the predeposition step is skipped, and the dopant atoms are introduced into the silicon wafers at low temperature by means of an ion-implantation process (Sect. 23.7).

23.5.1 Predeposition

Figure 23.2 provides a schematic picture of the source–wafer structure during a predeposition step. The interface between wafer and source is assumed to coincide

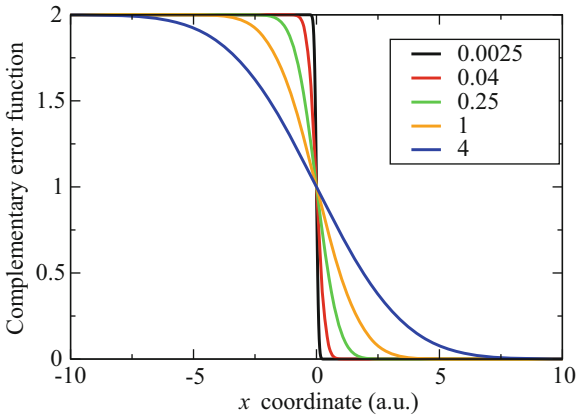


Fig. 23.2 Normalized profiles N/C produced at different instants by a predeposition, using the first of (23.15) as initial condition with arbitrary units for the x coordinate. The outcome is a set of complementary error functions whose expression is the first of (23.16). The legends show the value of $4\,a$ for each curve, also in arbitrary units, with $a = a(t)$ given by the second of (23.10)

with the y, z plane, with the x axis oriented towards the wafer's bulk, and the initial condition N_0 is assumed constant in the source region. The diffusion coefficients in the source and wafer regions are provisionally taken equal to each other. Thanks to these assumptions the problem has no dependencies on the y, z variables, and the one-dimensional form (23.8) of the diffusion equation holds. In the practical cases the extent of the source region in the x direction is large and the concentration of the dopant atoms in it is high. As a consequence, the source is not depleted when the atoms diffuse into the wafer. The spatial form of the concentration N at a given time $t = t'$ is called *diffused profile*. Its integral over the semiconductor region,

$$Q(t') = \int_0^{+\infty} N(x, t = t')\,\mathrm{d}x \qquad (\mathrm{m}^{-2})\,, \tag{23.14}$$

is called *dose*. For convenience the constant value of the initial condition in the source region is indicated with $2\,C$ (m^{-3}). It follows that the initial condition of the predeposition step is given by the first of (23.15). In turn, the general expression (23.13) of the dopant concentration reduces to the second of (23.15):

$$N_0(\xi) = \begin{cases} 2\,C & \xi < 0 \\ 0 & \xi > 0 \end{cases}; \qquad N(x,t) = 2\,C \int_{-\infty}^{0} \Delta(x - \xi, t)\,\mathrm{d}\xi\,. \tag{23.15}$$

Using (C.79) and (C.71) one finds the following expressions for the diffused profile and dose of the predeposition step,

$$N(x,t) = C\,\mathrm{erfc}\left(\frac{x}{\sqrt{4\,a}}\right), \qquad Q(t) = C\sqrt{\frac{4\,a}{\pi}}\,, \tag{23.16}$$

where the dependence on t derives from the second of (23.10). As parameter a increases with time, the dose increases with time as well, consistently with the qualitative description of predeposition given earlier in this section. In most cases the diffusion coefficient is independent of time, $a = D\,t$, thus yielding $Q \propto \sqrt{t}$.

Still from the second of (23.10) one finds $a(0) = 0$. Combining the latter with the properties (C.69) of the complementary error function shows that $\lim_{t\to 0+} N(x,t)$ coincides with the initial condition given by the first of (23.15). Also, the solution (23.16) fulfills the boundary conditions $N(-\infty,t) = 2C$, $N(+\infty,t) = 0$ at any $t > 0$. Finally it is $N(0,t) = C$ at any $t > 0$. This explains the term *constant-source diffusion* that is also used to indicate this type of process. In fact, the concentration at the wafer's surface is constant in time. Figure 23.2 shows the normalized concentration N/C calculated from the first of (23.16) at different values of a.

The analysis of the diffusion process carried out so far was based on the assumption of a position-independent diffusion coefficient D. In the actual cases this assumption is not fulfilled because the dopant source and the wafer are made of different materials. As a consequence, the solution of (23.8) must be reworked. In the case of predeposition this is accomplished with little extra work, which is based on the first of (23.16) as shown below.

One assumes, first, that the diffusion coefficient in either region is independent of time, as is the standard condition of the typical processes. In each region the diffusion coefficient takes a spatially constant value, say, D_S in the source and $D_W \neq D_S$ in the wafer. Now, observe that (23.8) is homogenous and contains the derivatives of N, but not N itself. It follows that if $C\ \mathrm{erfc}[x/(4\,a)^{1/2}]$ is the solution of (23.8) fulfilling some initial and boundary conditions, then $A\ \mathrm{erfc}[x/(4\,a)^{1/2}] + B$ is also a solution of (23.8), fulfilling some other conditions that depend on the constants A and B. One then lets, with $t > 0$,

$$N_S = A_S\ \mathrm{erfc}\left(\frac{x}{\sqrt{4D_S t}}\right) + B_S\,, \qquad x < 0\,, \tag{23.17}$$

$$N_W = A_W\ \mathrm{erfc}\left(\frac{x}{\sqrt{4\,D_W t}}\right) + B_W\,, \qquad x > 0\,, \tag{23.18}$$

and fixes two relations among the constants in order to fulfill the initial conditions (23.15):

$$\lim_{t=0^+} N_S = 2\,A_S + B_S = 2\,C\,, \qquad \lim_{t=0^+} N_W = B_W = 0\,. \tag{23.19}$$

In order to fix the remaining constants one must consider the matching conditions of the two regional solutions (23.17, 23.18) at the source–wafer interface. The concentrations across an interface between two different media are related by the *segregation coefficient* k [130, Sect. 1.3.2]. Also, given that no generation or

destruction of dopant atoms occurs at the interface, the flux density $-D\,\partial N/\partial x$ must be continuous there. In summary, the matching conditions at the source–wafer interface are

$$N_W(0^+,t) = k\,N_S(0^-,t)\,, \qquad D_W\left(\frac{\partial N_W}{\partial x}\right)_{0+} = D_S\left(\frac{\partial N_S}{\partial x}\right)_{0-}\,. \tag{23.20}$$

Using (23.17, 23.18, 23.19) transforms (23.20) into $A_W = k\,(2\,C - A_S)$ and, respectively, $\sqrt{D_W}\,A_W = \sqrt{D_S}\,A_S$ whence, remembering the first of (23.19) and letting $\eta = D_W/D_S$,

$$A_S = \frac{k\,\sqrt{\eta}}{1+k\,\sqrt{\eta}}\,2\,C\,, \qquad B_S = \frac{1-k\,\sqrt{\eta}}{1+k\,\sqrt{\eta}}\,2\,C\,, \qquad A_W = \frac{k}{1+k\,\sqrt{\eta}}\,2\,C\,. \tag{23.21}$$

Thanks to (23.21), the concentration of the dopant atoms in the source region at the source–wafer interface at $t > 0$ turns out to be $N_S(0^-,t) = A_S + B_S = 2\,C/(1 + k\,\sqrt{\eta})$. If, in particular, the source of dopant is in the gaseous phase, it is $\eta \ll 1$. As k is of order unity, one finds for the gaseous source $N_S(0^-,t) \simeq 2\,C$, namely, the interface concentration of the source region is practically equal to the asymptotic one. Figure 23.3 shows the diffused profile N calculated from (23.17, 23.18) at two different instants t_1 and $t_2 = 16\,t_1$, with $D_S = 400\,D_W$. The coefficients are, in arbitrary units, $A_S = 2$, $B_S = 78$, $A_W = 40$, $B_W = 0$. From the first of (23.19) it follows $C = 41$. Letting $(4\,D_W\,t_1)^{1/2} = 1$ (a.u.) one has $(4\,D_S\,t_1)^{1/2} = 20$, $(4\,D_W\,t_2)^{1/2} = 4$, and $(4\,D_W\,t_2)^{1/2} = 80$. These values are used to calculate the four curves shown in the figure.

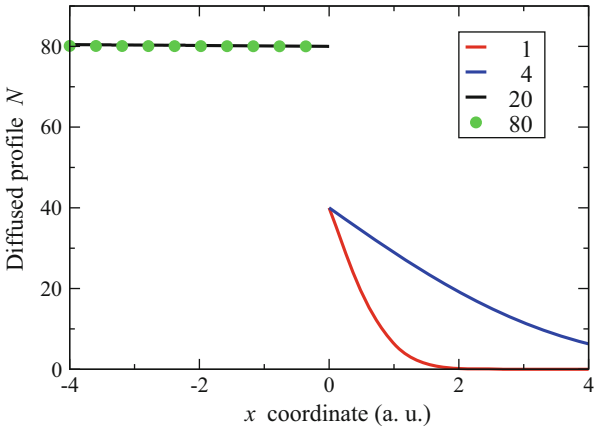


Fig. 23.3 Diffused profiles calculated at t_1 and $t_2 = 16\,t_1$ when two different materials are involved. The calculation is based on (23.17), (23.18) as described at the end of Sect. 23.5.1. The legends show the $(4\,D\,t)^{1/2}$ value for each curve

23.5.2 *Drive-in Diffusion*

As indicated at the beginning of this section, the drive-in diffusion is started when the desired amount of atoms has been introduced into the silicon lattice, and is continued until a suitable profile is reached.

In principle, the profile to be used as initial condition of a drive-in diffusion is not exactly equal to the final profile of the predeposition step. In fact, the boundary condition $(\partial N_W/\partial x)_{0+}$ is different from zero during the predeposition step. Instead, during the growth of the silicon-dioxide layer that blocks the supply of dopant atoms from the source region, the boundary condition becomes equal to zero to adapt to the situation of a vanishing flux density of dopants across the interface.

The calculation of the drive-in diffusion is tackled more easily by assuming that the blocking of the supply of dopants atoms is instantaneous, so that the final profile of the predeposition step is "frozen." Then, one considers the full domain $-\infty < x < +\infty$ instead of the wafer domain $0 \le x < +\infty$, with the same diffusion coefficient $D = D_W$ everywhere. In this way one can still use the model problem (23.8). As for the initial condition N_0, one mirrors the final profile of the predeposition step over the negative axis, thus making the initial condition even with respect to x. Letting $x \leftarrow -x$ in (23.13) one easily proves that $N(-x,t) = N(x,t)$ if $N_0(-\xi) = N_0(\xi)$; namely, if the initial condition is even, then the solution is even at all times. With the provisions above one finds $(\partial N_W/\partial x)_{0+} = -(\partial N_W/\partial x)_{0-}$, which automatically fulfills the condition of a vanishing flux density of dopants across the origin. Then, the application of (23.13) provides the profile of the drive-in diffusion in the wafer region $0 \le x < +\infty$.

The final profile (23.16) of the predeposition step, used as initial condition, does not lend itself to an analytical calculation of the drive-in diffusion. Some examples of calculation are given below, in which profiles of a simpler form than (23.16) are used as approximations. Let Q be the dose present within the wafer region. As a first example one lets

$$N_0(\xi) = 2\,Q\,\delta(\xi)\,, \qquad N(x,t) = 2\,Q\int_{-\infty}^{+\infty} \delta(\xi)\,\Delta(x-\xi,t)\,\mathrm{d}\xi\,. \tag{23.22}$$

From the properties of the Dirac δ (Sect. C.4) it follows

$$N(x,t) = 2\,Q\,\Delta(x,t) = 2\,Q\,\frac{\exp[-x^2/(4\,a)]}{\sqrt{4\pi\,a}}\,, \tag{23.23}$$

showing that when the initial condition is a Dirac δ, the profile resulting from a diffusion process is Gaussian. Only the portion of (23.23) belonging to the wafer region, that is, $x \ge 0$, must in fact be considered. Integrating (23.23) from 0 to $+\infty$ and using (C.78) yield the expected value Q of the dose at all times. Although rather crude, the approximation of using a Dirac δ as initial condition is acceptable, because the profile obtained from a predeposition or an ion-implantation process is typically very thin.

As a second example one takes a Gaussian profile as the initial condition, specifically, the second of (23.23) where, to better distinguish the symbols, a is replaced with a_1. It is assumed that the drive-in diffusion to be calculated is characterized by another value of the parameter, say, a_2. The difference between a_2 and a_1 may be due to the duration of the diffusion process under investigation, to a temperature-induced difference in the diffusion coefficients, or both. As usual the instant $t = 0$ is set as the initial time of the diffusion process. Applying (23.13) yields

$$N(x,t) = 2\,Q \int_{-\infty}^{+\infty} \frac{\exp[-\xi^2/(4\,a_1)]}{\sqrt{4\pi\,a_1}} \frac{\exp[-(x-\xi)^2/(4\,a_2)]}{\sqrt{4\pi\,a_2}}\,\mathrm{d}\xi\,. \tag{23.24}$$

Using the auxiliary variable $\eta = \xi - a_1\,x/(a_1+a_2)$, whence $x-\xi = -\eta + a_2\,x/(a_1+a_2)$, transforms the exponent of (23.24) as

$$-\frac{\xi^2}{4\,a_1} - \frac{(x-\xi)^2}{4\,a_2} = -\frac{x^2}{4\,(a_1+a_2)} - \frac{a_1+a_2}{4\,a_1 a_2}\,\eta^2\,. \tag{23.25}$$

Then, integrating with respect to $\sqrt{(a_1+a_2)/(4a_1a_2)}\,\eta$ and using again (C.78) yield

$$N(x,t) = 2\,Q\,\frac{\exp[-x^2/(4\,a_1+4\,a_2)]}{\sqrt{4\pi\,(a_1+a_2)}}\,. \tag{23.26}$$

As before, the integral of the profile from 0 to $+\infty$ yields the dose Q at all times. The result expressed by (23.26) is important because it shows that a diffusion process whose initial condition is a Gaussian profile yields another Gaussian profile. The parameter of the latter is found by simply adding the parameter $a_2 = \int_0^t D(t')\,\mathrm{d}t'$ of the diffusion process in hand, whose duration is t, to the parameter a_1 of the initial condition. Clearly, the result is also applicable to a sequence of successive diffusion processes. In fact, it is used to calculate the final profiles after the wafers have undergone the several thermal processes that are necessary for the integrated-circuit fabrication.

23.6 Generalization of the Model Problem

The generalization of the model problem (23.8) to three dimensions, that is, equation (23.7) with $W = 0$ and initial condition $N_0(\mathbf{r}) = N(\mathbf{r}, t = 0)$, is still tackled by means of the Fourier transform. For this, it is necessary to define the vectors $\mathbf{r} = (r_1, r_2, r_3)$, $\mathbf{s} = (s_1, s_2, s_3)$, $\mathbf{k} = (k_1, k_2, k_3)$, and the elements $\mathrm{d}^3k = \mathrm{d}k_1\,\mathrm{d}k_2\,\mathrm{d}k_3$, $\mathrm{d}^3s = \mathrm{d}s_1\,\mathrm{d}s_2\,\mathrm{d}s_3$. Using (C.20) and following the procedure of Sect. 23.4, one finds again the relations (23.10). This time, however, it is $k^2 = k_1^2 + k_2^2 + k_3^2$. The solution $N(\mathbf{r}, t)$ is readily found as a generalization of (23.12), namely

$$N(\mathbf{r},t) = \iiint_{-\infty}^{+\infty} N_0(\mathbf{s}) \left\{ \iiint_{-\infty}^{+\infty} \frac{1}{(2\pi)^3} \exp[\mathrm{i}\mathbf{k}\cdot(\mathbf{r}-\mathbf{s}) - ak^2]\, \mathrm{d}^3k \right\} \mathrm{d}^3s\,. \tag{23.27}$$

The expression in braces in (23.27) is the product of three functions of the same form as (C.76). It follows

$$N(\mathbf{r},t) = \iiint_{-\infty}^{+\infty} N_0(\mathbf{s}) \frac{\exp[-|\mathbf{r}-\mathbf{s}|^2/(4a)]}{(4\pi\, a)^{3/2}}\, \mathrm{d}^3s\,. \tag{23.28}$$

When the net generation rate per unit volume, W, is different from zero, it is in general impossible to find an analytical solution of (23.7). An important exception is the case where W is linear with respect to N and has no explicit dependence on $\mathbf{r}$ or t. In this case (23.7) reads

$$\frac{\partial N}{\partial t} = D\,\nabla^2 N - \frac{N - N_a}{\tau}\,, \qquad D = D(t)\,, \tag{23.29}$$

where the two constants N_a (m^{-3}) and τ (s) are positive. This form of W is such that the particles are generated if $N(\mathbf{r},t) < N_a$, while they are destroyed if $N(\mathbf{r},t) > N_a$. Equation (23.29) is easily solved by introducing an auxiliary function N' such that $N = N_a + N'\exp(-t/\tau)$. In fact, N' turns out to be the solution of the three-dimensional model problem, so that using (23.28), the solution of (23.29) reads

$$N(\mathbf{r},t) = N_a + \exp(-t/\tau) \iiint_{-\infty}^{+\infty} N_0(\mathbf{s}) \frac{\exp[-|\mathbf{r}-\mathbf{s}|^2/(4a)]}{(4\pi\, a)^{3/2}}\, \mathrm{d}^3s\,. \tag{23.30}$$

23.7 Ion Implantation

The reproducibility of thermal diffusion, as a method for introducing dopant atoms into a semiconductor material, is limited. The analysis carried out in the previous sections has shown that the predeposition process is controlled by the boundary condition, where a constant concentration of the dopant is kept, and by the diffusion coefficient. Similarly, the drive-in diffusion is controlled by the zero-flux boundary condition and by the diffusion coefficient. Thus, the profile obtained from a thermal diffusion is invariably of the same type; another drawback of the process is the strong dependence on temperature of the diffusion coefficient (23.5), which makes the reproducibility more difficult.

In the sixties of the last century, *ion implantation* has emerged as a method, alternative to thermal diffusion, for introducing dopant atoms in the semiconductor. The *ion implanter* (Fig. 23.4) consists of an *ion source* where the dopant atoms (in solid, liquid, or gaseous phase) are ionized by a confined electric discharge, one

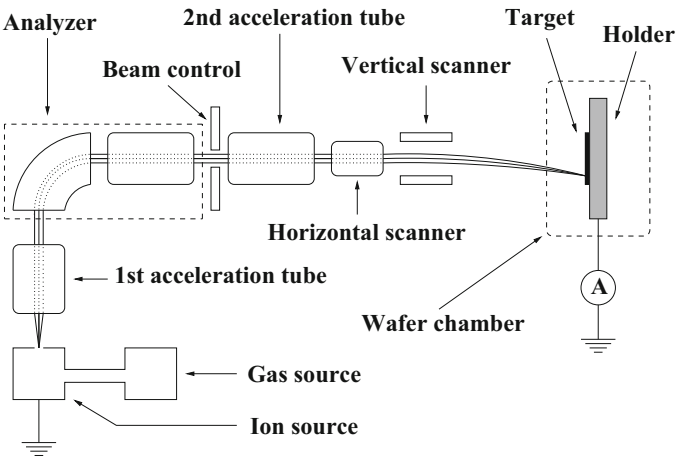


Fig. 23.4 Schematic cross-section of an ion implanter

or more *acceleration systems*, and an *analyzing system* where the ions are filtered according to their mass and kinetic energy. The ion beam thus obtained is uniformly distributed, by means of a *scanning system*, over the wafer (*target*). When the ions enter the wafer, they gradually lose energy due to collisions, until they come to rest. The collision between an ion and a wafer's nucleus may displace the nucleus from its lattice site; this happens if the energy released by the ion during the collision is larger than the *displacement threshold energy* (about 15 eV in silicon). The damage is removed at a later stage by a suitable heat treatment (*annealing*); besides removing the damage, annealing drives the implanted ions into the substitutional positions.

The neutrality of the implanted ions is restored by electrically connecting the ion source with the target holder (Fig. 23.4). This also provides a method for measuring the current flowing through the connection; since all ions are equal to each other, this yields a very precise measurement of the ion current and, as a consequence, of the implanted dose (more details are given below). From the description of the process, it is clear that the dose is controlled independently from the penetration depth of the ions; such a depth is in fact determined by prescribing the kinetic energy of ions through the acceleration systems. It follows that a large variety of profiles can be obtained from the convolution of subsequent implantation steps with different doses and penetrations. In summary, the advantages of ion implantation are:

- Fine control of dose and penetration.
- Flexibility in building-up profiles.
- Reproducibility.
- Small lateral penetration.
- Applicability to compound semiconductors.

- Low temperature.[3]

The disadvantages with respect to thermal diffusion are the cost of the equipments, the higher complexity of the process, and crystal damage.[4]

23.7.1 Ion Implanter

The ion implanter is made of a number of blocks that are illustrated with reference to Fig. 23.4. The whole process is carried out *in vacuo* to prevent the molecules of the atmosphere from interfering with the motion of the ions. The first block of the implanter is the ion source, where a suitable compound containing the dopant is subjected to a high electric field. Typical compounds are *boron trifluoride* (BF_3), *arsine* (AsH_3), or *phosphine* (PH_3). A fraction of the atoms ionize; as a consequence, free electrons and ionized atoms of different chemical species are present in the ion source. While the electrons are collected by a suitable contact, some of the ions escape through a hole in the wall of the ion source and transfer from the latter to the second block, where an electric field is present. If V_A is the potential drop across the second block, the ions are accelerated and their kinetic energy increases by $q\,V_A$ (Fig. 23.5).

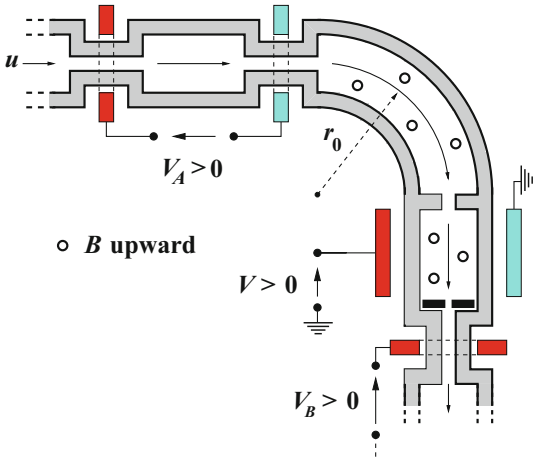


Fig. 23.5 Detail of the analyzing system

[3]This aspect is particularly relevant because each high-temperature process activates the diffusion coefficients and deforms the doping profiles already present in the semiconductor.

[4]In contrast, thermal diffusion has simpler process and equipments, and does not produce a crystal damage; this, however, does not compensate for its drawbacks (high temperature, modest reproducibility, difficult application to compound semiconductors, large lateral penetration).

It is important to note that although the increase in kinetic energy is known, the actual kinetic energy of the ions is not, because there is no information about the distribution of the ion velocities when the ions escape from the ion source; moreover, the ions that escape are a mixture of those deriving from the dopant atoms and from the other atoms of the compound. For this reason, it is necessary to introduce an analyzing system made of a third and fourth block, where the ions are filtered: the result of the filtering action is that all ions belonging to chemical species different from the dopant are eliminated, and the dopant ions whose kinetic energy is different from that prescribed by the process are eliminated as well. In summary, the filtering action must involve two dynamic properties: the mass, to distinguish the ions belonging to different chemical species, and the kinetic energy; or, equivalently, other two dynamic properties uniquely related to the above like, e.g., momentum and velocity. As shown below, the ions having the prescribed momentum are obtained from the third block, called *magnetic filter*; these ions then enter the fourth block, the *electromagnetic filter*, that eliminates the ions not possessing the prescribed velocity. The outcome is that only those ions having the correct momentum and velocity pass the two filters. After the filtering action is completed, the ions are oriented onto the target by suitable deflection electrodes. Prior to that, a further acceleration block may be present, with V_B the potential drop.

The action of the magnetic filter is described below, with the aid of Fig. 23.5. The magnetic induction is oriented upward with respect to the plane of the cross-section shown in the figure, and the initial velocity of the ion belongs to such a plane. As the orientation of the magnetic induction is the same as in Prob. 4.6, the ion's trajectory within the magnetic filter is a portion of circumference of radius $r = u/\omega = u\,m/(qB)$, whence

$$m\,u = r\,q\,B\,. \tag{23.31}$$

The ion escapes from the magnetic filter if $r = r_0$, with r_0 the filter's radius (Fig. 23.5); otherwise, it is intercepted by the filter's walls. Letting $r = r_0$, (23.31) shows that B can be calibrated to obtain a desired value of the ion's momentum $m\,u$.

The magnetic filter is followed by the electromagnetic filter. In the latter, the same magnetic induction $\mathbf{B}$ as before is present; in addition to it, a constant electric field $\mathbf{E}$ is applied as well through the voltage V shown in the figure. Let $\mathbf{u}$ be the velocity of the ion at the instant of entering the electromagnetic filter, that is, at the instant of escaping the magnetic filter; as before, the direction of $\mathbf{u}$ is normal to $\mathbf{B}$, while its modulus (with which the ion escapes the magnetic filter) is yet unknown, because (23.31) prescribes the product $m\,u$ only. Finally, the direction of $\mathbf{E}$ is normal to both $\mathbf{B}$ and $\mathbf{u}$. The structure of the apparatus is such that the ion escapes the electromagnetic filter only if it suffers no deviation; this happens if $\mathbf{F} = q\,(\mathbf{E} + \mathbf{u} \wedge \mathbf{B}) = 0$. Considering the orientation of the three vectors in the above, it follows $E = u\,B$; combining the latter with (23.31) provides $m = r_0\,q\,B^2/E$. Finally, collecting the above results one finds for the mass, velocity, and kinetic energy of the filtered ion,

$$m = r_0\, q\, \frac{B^2}{E}\,, \qquad u = \frac{E}{B}\,, \qquad T = \frac{1}{2}\, r_0\, q\, E\,. \tag{23.32}$$

This result shows that a suitable choice of E and B filters only the ions of a certain mass; other ions possibly present in the ion source are eliminated. Moreover, all filtered ions have the same kinetic energy.

From the description above it follows that the ion beam is formed by identical ions that move in the same direction with a common velocity of modulus u and a uniform concentration c_I. In principle, the direction of the ions' velocities should spread due to mutual repulsion; however, the beam is dilute, to the extent that this effect is negligible. Let the velocity of the ions be aligned with the x axis, and let A be the cross-section of the beam in the direction normal to x. From the concentration $c_I = \mathrm{d}n_I/(\mathrm{d}x\,\mathrm{d}A)$ of the ions in the beam, one finds the relation

$$u\, c_I = u\, \frac{\mathrm{d}n_I}{\mathrm{d}x\,\mathrm{d}A} = u\, \frac{\mathrm{d}n_I}{u\,\mathrm{d}t\,\mathrm{d}A} = \frac{\mathrm{d}n_I}{\mathrm{d}t\,\mathrm{d}A} = F\,, \tag{23.33}$$

with F the ion flux. The latter is constant due to the constancy of u and c_I. The current density of the ions is $J = q\,F$, and the corresponding current reads

$$I = \int_A J\,\mathrm{d}A = A\,J = A\,q\,F\,. \tag{23.34}$$

Letting t_P be the implant's duration, and using (23.49), the implanted dose is found from

$$Q = \int_0^{t_P} F\,\mathrm{d}t = \frac{1}{q\,A} \int_0^{t_P} I\,\mathrm{d}t = \frac{I}{q\,A}\, t_P\,. \tag{23.35}$$

In (23.35), the area A of the beam is given; as the current measurement is accomplished with a very high accuracy,[5] the dose is determined with the same accuracy by calibrating the duration t_P of the process. For this reason the accuracy of ion implantation is superior to that of thermal diffusion. From (23.32) and (23.35) one finds that, as mentioned before, the kinetic energy and the dose are controlled independently from each other.

23.7.2 Ion Trajectories

When it enters the semiconductor, the implanted ion undergoes a series of collisions with the nuclei of the semiconductor, and also with the electron cloud surrounding

[5] Remembering the description of the process given at the beginning of the section, the current that is measured is in fact that of the electrons extracted from the dopant atoms.

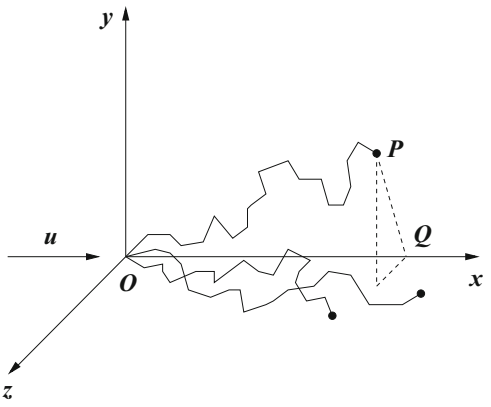


Fig. 23.6 Sketch of the trajectory of implanted ions

the nuclei. As indicated above, the ion gradually loses energy until it comes to rest; also, due to collisions, the ion's trajectory within the semiconductor is not a straight line: for this reason, it is convenient to use a curvilinear abscissa s along the trajectory, whose origin is in the position where the ion enters the semiconductor.

A sketch of such a trajectory is shown in Fig. 23.6, where the yz plane defines the semiconductor's surface; the semiconductor belongs to the half space $x > 0$, and the ion under consideration enters the semiconductor, with an initial velocity $\mathbf{u}$, in a position corresponding to the origin of the reference. The collisions with the nuclei are considered point-like, whence the trajectory is represented as a broken line; at each collision, the ion loses a fraction of its energy. Along the segment from one collision to the next, the ion still suffers an energy loss due the interaction with the electrons. Letting P indicate the final position of the ion, a number of parameters are of use: specifically, letting $r = PO$ (called *range*) be the modulus of the position vector at the end of the trajectory, the projection of r along the x axis, $r_P = QO$, is called *projected range*, while the distance between the final position and the x axis, $r_L = QP$, is called *lateral range*.

Due to the randomness of the collisions, ions entering the semiconductor at the same position and with the same velocity follow in general different trajectories; three of them are shown in Fig. 23.6 by way of example. It is then necessary to describe the outcome of the process by means of statistical parameters; one of them is the dose Q defined in (23.35), other parameters are the *average projected range* and the *average lateral range*, respectively defined as[6]

$$R_P = \langle r_P \rangle \,, \qquad R_L = \langle r_L \rangle \,. \tag{23.36}$$

[6]The masks used in an ion-implantation process are similar to those used in thermal diffusion. For the ions that enter the semiconductor far from the mask's edges, the lateral ranges r_L tend to compensate each other when calculating the average lateral range. For this reason, more attention is given here to the average range.

Table 23.1 Example of parameters of an ion implanter

Parameter	Symbol	Min	Max	Units
Current (low)	I	50	500	μA
Current (medium)	I	0.5	5	mA
Current (high)	I	5	50	mA
Dose	Q	10^{12}	10^{18}	cm^{-2}
Kinetic energy	T	1	1000	keV
Average projected range	R_P	10^{-2}	10	μm

The averages in (23.36) are taken over all ions. Some typical parameters characterizing ion implanters are given in Table 23.1; the range of the current qualifies the small, medium, or large equipment.

The loss $S = -\mathrm{d}T/\mathrm{d}s > 0$ of kinetic energy suffered by the ion along the element $\mathrm{d}s$ of trajectory is called *stopping power*;[7] due to the different mechanisms involved, it is convenient to separate the losses produced by the interaction with electrons from those due to collisions with nuclei:

$$S = S_e + S_n\,, \qquad S_e = -\left(\frac{\mathrm{d}T}{\mathrm{d}s}\right)_e\,, \qquad S_n = -\left(\frac{\mathrm{d}T}{\mathrm{d}s}\right)_n\,. \tag{23.37}$$

If the electronic and nuclear stopping powers are expressible in terms of the kinetic energy possessed by the ion when entering the element $\mathrm{d}s$ of trajectory, the relation defining the stopping power is readily separated to yield

$$\mathrm{d}s = -\frac{\mathrm{d}T}{S_e(T) + S_n(T)}\,, \qquad s_0 = \int_0^{r_0\,q\,E/2} \frac{\mathrm{d}T}{S_e(T) + S_n(T)}\,, \tag{23.38}$$

where the upper integration limit is taken from (23.32). In this way it is possible to express the length s_0 of the trajectory, although its form is not known. The expressions for S_e and S_n are given with different orders of approximation [130, Chap. 8], [88, Chap. 11]. When the kinetic energy of the ion is low, the electronic stopping power can be assimilated to a viscous drag force which, to first order, is proportional to the velocity of the ion; in this order of approximation one obtains $S_e(T)$ as

$$S_e \propto \sqrt{T}\,. \tag{23.39}$$

As for the nuclear stopping power, a qualitative analysis may be based on the calculation shown in Prob. 3.3, relative to the variation ΔT in the kinetic energy of a particle of mass m_1, that interacts with another particle of mass m_2 initially at rest; the interaction potential energy is of the Coulomb type, with $Z_1\,q$, $Z_2\,q$ the particles' charges and c the impact parameter. The calculation leads to

[7] The units of the stopping power are those of a force.

$$\Delta T = -\frac{4\,(m_1/m_2)\,T}{(1+m_1/m_2)^2 + (c/\alpha)^2\,T^2}\,, \qquad \alpha = \frac{Z_1\,Z_2\,q^2}{8\pi\,\varepsilon_0} > 0\,. \tag{23.40}$$

The above expression can be recast in a more manageable form as

$$\Delta K = -b\,\frac{K}{1+K^2}\,, \qquad K = \frac{m_2\,c/\alpha}{m_1+m_2}\,T\,, \qquad b = \frac{4\,m_1\,m_2}{(m_1+m_2)^2} \le 1\,. \tag{23.41}$$

In the problem in hand, the exchange of energy occurs over a path of the order of the lattice constant a; dividing by a both sides of the first relation in (23.41), and assuming that ΔK is small, one obtains $\mathrm{d}K/\mathrm{d}s \simeq -(b/a)\,K/(1+K^2)$. Combining the above with (23.39) after using K instead of T in the latter yields

$$-\,\mathrm{d}s = \frac{\mathrm{d}K}{S_e(K)+S_n(K)} = \frac{(1+K^2)\,\mathrm{d}K}{g\,\sqrt{K}\,(1+K^2)+(b/a)\,K}\,, \tag{23.42}$$

with g a suitable constant. The above treatment of the nuclear stopping power is only qualitative because it does not take screening into account; in fact, if the electronic stopping power was missing ($g = 0$), integration of (23.42) would yield a diverging length of the trajectory due to the long-range decay of the pure Coulomb potential (compare with Sect. 20.6.4). Also, an integration of (23.42) would imply that the impact parameter is the same for all collisions; this parameter appears in fact in one of the integration limits due to the second relation in (23.38). A more realistic analysis shows that the majority of the implanted ions[8] undergo a large number of collisions, so that many different impact parameters are involved in a single trajectory; it is then reasonable to assume that for such ions the same set of impact parameters, although in different order, are present. It follows that the same set of energy losses applies to all ions, whence the lengths of the trajectories are equal for all; this helps in building up a statistical theory of the position of the end points of the trajectories.

A sketch of the behavior of the nuclear and electronic stopping power is shown in Fig. 23.7, in arbitrary units and using $\sqrt{T}$ as abscissa. The energy T_c corresponding to the intersection is called *critical energy*. If $r_0\,q\,E/2 \le T_c$, as the ion enters the semiconductor a large part of the energy loss is due to nuclear collisions; some of the latter, in turn, displace the nuclei, so that the crystal damage starts from the superficial layers of the semiconductor. When, instead, it is $r_0\,q\,E/2 > T_c$, most of the initial energy loss is due to electronic collisions, which produce no damage; however, as the ion continues its motion into the semiconductor its kinetic energy decreases, and eventually reaches the critical value: as a consequence, the crystal damage is larger far from the semiconductor's surface.

[8] The ions not belonging to this majority are included in the analysis at a later stage (Sect. 23.7.3).

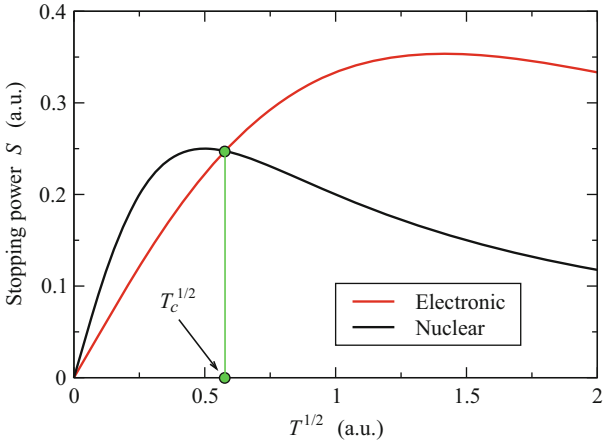


Fig. 23.7 The nuclear and electronic stopping power. The vertical line marks the critical energy

23.7.3 Implanted Profile

As mentioned in Sect. 23.7.2, for the majority of implanted ions the same set of energy losses apply, although in different order. This yields the same length s_0 of the trajectories; their end points, however, are different, and a statistical description is necessary. For a single ion, the important parameter is the projected range r_P, namely, the distance of the end point from the semiconductor surface in the direction x normal to the surface. The r_P values of the different ions are distributed around the average projected range R_P: it is therefore of interest to determine the form $N(r_P)$ of such a distribution (called *implanted profile*) around the average. Let $P(r_P)$ be the probability density of the implanted profile, such that $P(r_P)\,\mathrm{d}r_P$ is the probability that the end point of the trajectory of an ion lies between r_P and $r_P + \mathrm{d}r_P$. From the normalization condition $\int_0^\infty P(r_P)\,\mathrm{d}r_P = 1$ and from the definition (23.14) of the dose it follows $N(r_P) = Q\,P(r_P)$.

For the analysis of the implanted profile, the simplified form of the trajectory of an ion, shown in Fig. 23.8, is of help; the number of collisions shown is much smaller than in realistic cases, and a two-dimensional path is assumed. Each portion of the trajectory between two nuclear collisions is projected along the axis normal to the semiconductor surface, the projections being labeled $\eta_1, \eta_2, \ldots$; the projected and lateral ranges are indicated as well. In the analysis, the length of each projection is described by a probability density $p(\eta_i)$, such that $p(\eta_i)\,\mathrm{d}\eta_i$ is the probability that the ith projection lies between η_i and $\eta_i + \mathrm{d}\eta$. The projections are assumed to be independent from each other, with the same probability density applicable to all; the projected range is the sum $r_P = \eta_1 + \eta_2 + \cdots + \eta_n$, and the issue is to find the probability density $P(r_P)$. The conditions of the problem are those of the *random walk problem* (Sect. C.18); also, the number of collisions is sufficiently high to concentrate the majority of the projected ranges in the vicinity of the average.

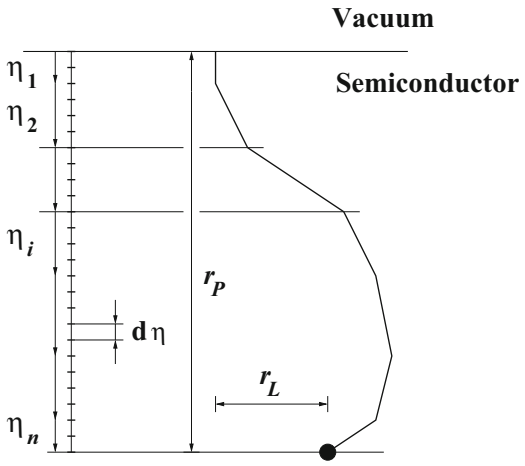


Fig. 23.8 Example of the random walk of an ion within the semiconductor

If the provisional assumption is made that the probability density $p(\eta_i)$ complies with the definition (C.169) of slowly varying function, the central part of the profile has a Gaussian form (Sect. C.19):

$$N(r_p) \simeq \frac{Q}{\sqrt{2\,\pi}\,\Delta R_p} \exp\left[-\frac{(r_p - R_p)^2}{2\,(\Delta R_p)^2}\right], \tag{23.43}$$

with ΔR_p the standard deviation. In principle, the normalization condition of (C.173) does not hold as it stands, because the implanted profile exists only for $x \geq 0$; however, the dopant concentration near the semiconductor surface is small enough to render $\int_0^\infty N(r_P)\,\mathrm{d}r_P = Q$.

23.7.4 Deviations from the Gaussian Profile

Experimental evidence shows that the actual implanted profile deviates from the Gaussian function derived in Sect. 23.7.3. Besides the approximations inherent in the derivation of (23.43), two additional sources of error must be considered. They are described with the aid of Fig. 23.9, representing the trajectories of three different ions; symbols c_1, c_2, ... indicate the impact parameters of each collision. For the trajectory on the left it is assumed that a large number of collisions occur; this trajectory partakes in the statistical treatment described above. One notes that the collision characterized by c_4, which is shorter than the other impact parameters, produces a dislocation of the nucleus: the latter moves from the original lattice position (the white dot) from a new position (the annular symbol). The trajectory in the center describes an ion that enters the semiconductor in a position that is practically symmetric with respect to the nuclei, so that the ion penetrates the

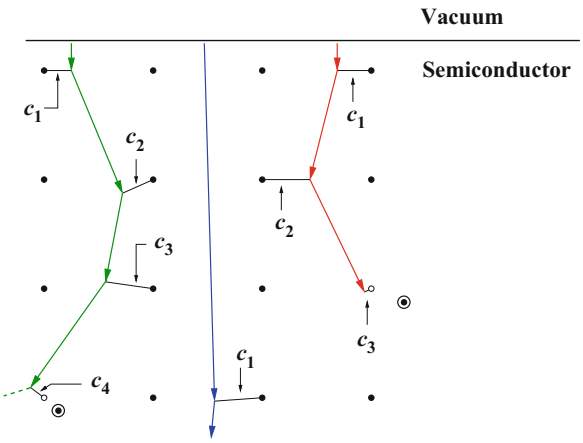


Fig. 23.9 For the trajectory on the left (green line) a large number of collisions occur. The central trajectory (blue line) describes a channeling ion. The trajectory on the right (red line) describes an ion that soon after entering the semiconductor suffers a nuclear collision in which most of its kinetic energy is lost

material for several lattice constants before suffering a sensible deviation due to nuclear collisions. This event, called *channeling*, pertains to a small fraction of the implanted ions; for such ions, the typical depth of the end points of the trajectories is much larger than $R_P + \Delta R_P$: thus, channeling builds up a tail of the implanted profile that cannot be tackled with the statistical treatment shown above.[9] Finally, the trajectory on the right describes an ion that soon after entering the semiconductor suffers a nuclear collision in which most of its kinetic energy is lost; such ions are a small fraction of the total ions implanted, whence the statistical treatment above is not applicable. The typical depth of their end points is much smaller than $R_P - \Delta R_P$: thus, the implanted profile near the semiconductor surface turns out to be richer than that predicted by the Gaussian distribution.

A sensible method to treat the deviations with respect to the theory of Sect. 23.7.3 is that of generalizing the Gaussian distribution by introducing more parameters in it. To this purpose one observes from (23.43) that a normalized Gaussian distribution, $P = N/Q$ in this case, is characterized by two parameters: the mean value R_P, which is the moment of order 1 of the distribution (Sect. C.18), and the standard deviation ΔR_P, which is obtained from the moment of order 2. It is expected that the modified distribution will depend also on moments of order higher than 2, or on parameters related to such moments. To this purpose one defines the *skewness* γ and *kurtosis* β of a distribution P by the relations[10]

[9] Some techniques have been devised to reduce channeling. The basic idea is disrupting the orderly arrangement "seen" by the ion when approaching the semiconductor. For instance, pre-amorphization of the surface layers of the wafer, or growth of an insulating layer (e.g., thermally grown silicon dioxide is amorphous). Another method consists in tilting the wafer with respect to the incident beam; tilting angles are typically less than 15°, with 7° the preferred choice [88, Chap. 11].

[10] It is assumed that $\Delta R_P \neq 0$. The units of γ and β are those of r_P.

$$\gamma = \int_{-\infty}^{+\infty} \left(\frac{r_P - R_P}{\Delta R_P} \right)^3 P(r_P)\, \mathrm{d}r_P\,, \qquad \beta = \int_{-\infty}^{+\infty} \left(\frac{r_P - R_P}{\Delta R_P} \right)^4 P(r_P)\, \mathrm{d}r_P\,. \tag{23.44}$$

The skewness is an indicator of the asymmetry of the distribution about the mean value; it may be negative. For a Gaussian distribution the integrand in the definition of γ is odd, whence $\gamma = 0$. The kurtosis is positive definite; for a Gaussian distribution one may let $u = (r_P - R_P)/(\sqrt{2}\,\Delta R_P)$ and limit the integration domain to the positive u axis, with the provision of multiplying the final result by 2. Then, letting $w = u^2$ yields $\beta = (4/\sqrt{\pi})\,\Gamma(5/2) = 3$, with Γ the Euler Gamma function (Sect. C.10).

To express a more general distribution in terms of the moments it is convenient to start from its differential form; to ease the notation one lets $s = r_P - R_P$, $\sigma = \Delta R_P$, whence the Gaussian distribution reads $P = (2\,\pi\,\sigma^2)^{-1/2}\,\exp[-s^2/(2\,\sigma^2)]$, with a differential form given by

$$\frac{\mathrm{d}P}{\mathrm{d}s} = -\frac{s}{\sigma^2}\,P\,. \tag{23.45}$$

The above expression is generalized by replacing $-s/\sigma^2$ with the ratio of two polynomials; of particular importance in this matter is the *Pearson family* of distributions, defined by the equation

$$\frac{\mathrm{d}P}{\mathrm{d}s} = \frac{s + a}{b_0 + b_1 s + b_2 s^2}\,P\,, \tag{23.46}$$

where the parameters a, b_0, b_1, and b_2 must be expressed in terms of the moments of order 2, 3, and 4, namely,[11] σ, γ, and β. Introducing the auxiliary parameter $w = 10\,\beta - 12\,\gamma^2 - 18$, one finds [130, Chap. 8]

$$b_0 = \sigma^2\,\frac{3\,\gamma^2 - 4\,\beta}{w}\,, \quad b_1 = a = \gamma\,\sigma\,\frac{\beta + 3}{w}\,, \quad b_2 = \frac{3\,\gamma^2 + 6 - 2\,\beta}{w}\,. \tag{23.47}$$

The Pearson distributions are thus defined in terms of their moments of order 1 through 4. For implanted profiles the most commonly used one is the so-called *Pearson IV* distribution, in which the coefficients satisfy the condition $0 < b_1^2/(4\,b_0\,b_2) < 1$. The values of b_0, b_1, b_2 are found by comparison with the experimental data.

[11] The moment of order 0 is the normalization condition, and that of order 1 specifies the position of the distribution along the horizontal axis. As a consequence, they do not enter this analysis. Note that the moments of order $m > 1$ are defined with reference to that of order 1, namely, from the average of $(r_P - R_P)^m$.

23.7.5 Annealing

Ion implantation produces lattice damage; in addition, for most of the implanted ions the end point of the trajectory does not coincide with a lattice site, namely, the ion is not in a substitutional position. A thermal process (*annealing*) becomes then necessary, to achieve *i)* removal of lattice damage by epitaxial regrowth of the damaged layers, and *ii)* electrical activation of the implanted ions (namely, by moving into the lattice sites the ions introduce electronic states in the expected positions of the energy gap; in this way, they act as dopants proper). Annealing also results in increasing carrier lifetimes and mobility. Different techniques are available: *thermal annealing*, in which the wafer is kept for about 1 h in a furnace at 900–1000°C; *laser annealing*, where local heating of the surface layers of the wafer is induced by laser irradiation; *rapid thermal annealing*, where the surface layers of the wafer are heated to about 1000°C for a relatively short time (1–20 s) using a tungsten halogen lamp. Thermal annealing modifies the existing dopant profiles due to thermal diffusion; this must be accounted for in the fabrication process of the integrated circuit. A problem with laser annealing is that constructive or destructive interferences occur when the laser wavelength is of the same order as the thicknesses of the layers subjected to irradiation; the local difference in temperature may force the layers to bend due to different thermal expansions. Rapid thermal annealing does not suffer the drawbacks of the other two methods.

23.8 Complements

23.8.1 Generation and Destruction of Particles

The discussion carried out in Sect. 23.2 about the continuity equation implies the possibility that particles may be generated or destroyed. To tackle this issue consider the problem "counting the time variation of students in a classroom." Assuming that the classroom has only one exit, to accomplish the task it suffices to count the students that cross the exit, say, every second. The students that enter (leave) the room are counted as a positive (negative) contribution.

Consider now a slightly modified problem: "counting the time variation of *non-sleeping* students in a classroom." To accomplish the task it does not suffice anymore to count the students that cross the exit. In fact, a student who is initially awake inside the classroom may fall asleep where she sits (one assumes that sleeping students do not walk); this provides a negative contribution to the time variation sought, without the need of crossing the exit. Similarly, an initially sleeping student may wake up, thus providing a positive contribution. Falling asleep (waking up) is equivalent to destruction (creation) of a non-sleeping student.

In the two examples above the objects to be counted are the same, however, in the second example they have an extra property that is not considered in the first one.

This shows that creation/destruction of a given type of objects may occur or not, depending on the properties that are considered. When particles instead of students are investigated, it is often of interest to set up a continuity equation for describing the time variation, in a given volume, of the particles *whose energy belongs to a specified range*. Due to their motion, the particles undergo collisions that change their energy. As a consequence a particle may enter, or leave, the specified energy range without leaving the spatial volume to which the calculation applies. In this example the origin of the net generation rate per unit volume W introduced in Sect. 23.2 is the extra property about the particles' energy.

23.8.2 Balance Relations

As indicated in Sect. 23.2, and with the provisions illustrated in Sect. 23.8.1, the continuity equation is a balance relation for the number of particles. Due to its intrinsic simplicity and generality, the concept of balance relation is readily extended to physical properties different from the number of particles; for instance, momentum, energy, energy flux, and so on. A detailed illustration of this issue is given in Chap. 19. It is also worth noting, in contrast, that the transport equation of the diffusion type (23.4), being based on a specific assumption about the transport mechanism, is less general than the continuity equation.

23.8.3 Lateral Diffusion

The treatment of predeposition and drive-in diffusion carried out in Sect. 23.5 is based on a one-dimensional model. This implies that the concentration of the dopant at the interface between the source and wafer regions is constant along the y and z directions. In the practical cases this is impossible to achieve, because the area over which the source is brought into contact with the wafer is finite. In fact, prior to the predeposition step the surface of the wafer is covered with a protective layer, called *mask*. As indicated in Sect. 24.1, in the current silicon technology the mask is typically made of thermally grown silicon dioxide. Next, a portion of the mask is removed to expose the silicon surface over a specific area, called *window*, through which the predeposition step takes place.

From the description above it follows that the initial condition N_0 of the predeposition step is constant only within the window, while it is equal to zero in the other parts of the y, z plane. This makes the hypothesis of a one-dimensional phenomenon inappropriate, and calls for the use of the three-dimensional solution (23.28). The subsequent drive-in diffusions must be treated in three dimensions as well, due to the form of their initial conditions. An important effect is the diffusion of the dopant underneath the edges of the mask. This phenomenon, called *lateral diffusion*, makes the area where the doping profile is present larger than the original mask, and must be accounted for in the design of the integrated circuit.

23.8.4 Alternative Expression of the Dose

The definition of the dose Q deriving from the one-dimensional model problem is (23.14). Letting $W = 0$ in (23.3), using its one-dimensional form $\partial N/\partial t = -\partial F/\partial x$ and observing that it is $F(+\infty, t) = 0$ due to the initial condition give the following expression for the time derivative of the dose:

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = -\int_0^{+\infty} \frac{\partial F(x,t)}{\partial x}\,\mathrm{d}x = F(0,t)\,. \tag{23.48}$$

Integrating (23.48) and remembering that the dose at $t = 0$ is equal to zero yield

$$Q(t') = \int_0^{t'} F(0,t)\,\mathrm{d}t\,. \tag{23.49}$$

The procedure leading from the original definition (23.14) of the dose to its alternative expression (23.49) is based solely on (23.3); hence, it does not depend on a specific transport model.

23.8.5 The Initial Condition of the Predeposition Step

The initial condition N_0 of the predeposition step is given by the first of (23.15). To carry out the solution of the diffusion equation it is necessary to recast N_0 in an integral representation of the Fourier type. However, (23.15) does not fulfill the condition (C.19) that is sufficient for the existence of such a representation.

Nevertheless the solution procedure leading to (23.13) is still applicable. In fact, remembering the definition (C.8) of the unit step function H, the initial condition can be recast as $N_0(\xi) = 2\,C\,[1 - H(\xi)]$. In turn, as shown in Appendix C.4, H can be represented in the required form.

Problems

23.1 A Gaussian doping profile $N = 2\,Q\,\exp(-x^2/c_1)/\sqrt{\pi c_1}$ undergoes a thermal-diffusion process at a temperature such that $D = 10^{-11}$ cm^2/s. Assuming $c_1 = 1.6 \times 10^{-7}$ cm^2, calculate the time that is necessary to reduce the peak value of the profile to $2/3$ of the initial value.

23.2 A Gaussian doping profile $N = 2\,Q\,\exp(-x^2/c_1)/\sqrt{\pi c_1}$, $c_1 = 9 \times 10^{-8}$ cm^2, undergoes a thermal-diffusion process with $c_2 = 16 \times 10^{-8}$ cm^2 yielding another Gaussian profile. Find the value $\bar{x}$ (in microns) where the two profiles cross each other.

23.3 A Gaussian doping profile $N = 2\,Q\,\exp(-x^2/c_1)/\sqrt{\pi\,c_1}$, $c_1 = 2.5 \times 10^{-6}$ cm^2, undergoes a 240 minute-long thermal-diffusion process at a temperature such that the diffusion coefficient is $D = 2.5 \times 10^{-10}$ cm^2 s^{-1}. Determine the ratio between the peak value of the final profile and that of the initial one.

23.4 A Gaussian doping profile $N = 2\,Q\,\exp(-x^2/c_1)/\sqrt{\pi\,c_1}$, $c_1 = 10^{-6}$ cm^2, undergoes a thermal-diffusion process in which $c_2 = 3{\times}10^{-8}$ cm^2. Find the position $\bar{x}$ (in microns) where the value of the initial doping profile equals the value that the final profile has in $x = 0$.

23.5 A Gaussian doping profile $N = 2\,Q\,\exp(-x^2/c_1)/\sqrt{\pi\,c_1}$ undergoes a thermal-diffusion process in which $c_2 = 10^{-8}$ cm^2. The value of the final profile in the origin is equal to that of the initial profile at $x_0 = 1.1 \times \sqrt{c_1}$. Find the value of c_1 in cm^2.

23.6 A Gaussian doping profile $N = 2\,Q\,\exp(-x^2/c_1)/\sqrt{\pi\,c_1}$, $c_1 = 1.8 \times 10^{-8}$ cm^2, undergoes a thermal-diffusion process whose duration is $t = 10$ min, with $D = 10^{-11}$ cm^2 s^{-1}. At the end of the process the concentration at some point x_0 is $N_1 = 3 \times 10^{16}$ cm^{-3}. If the process duration was 20 min, the concentration at the same point would be $N_2 = 3 \times 10^{17}$. Find the value of x_0 in microns.

23.7 The doping profile resulting from a predeposition process with $D = 10^{-11}$ cm^2 s^{-1} is $N(x) = N_S\,\mathrm{erfc}(x/\sqrt{c})$. The ratio between the dose and surface concentration is $Q/N_S = \lambda/\sqrt{\pi}$, $\lambda = 1{,}095$ nm. Find the duration t of the predeposition process, in minutes.

23.8 The initial condition of a drive-in diffusion is given by $N_0 = 2\,Q\,(h{-}x)/h^2$ for $0 \le x \le h$, and by $N_0 = 0$ elsewhere, where $Q > 0$ is the dose. Find the expression of the profile at $t > 0$.

Chapter 24
Thermal Oxidation—Layer Deposition

24.1 Introduction

High-quality oxide is essential in silicon technology. The most important applications of the oxide are the passivation of the wafer's surface, the isolation between metallizations, the formation of masks for, e.g., diffusion or implantation of dopants, the isolation between devices, and the formation of the gate insulator in MOS devices.

While the oxides used for passivation or isolation between metallization are typically obtained by chemical vapor deposition (Sect. 24.5), the oxide suitable for the other applications listed above is obtained by thermal oxidation. In fact, the extraordinary evolution of the VLSI technology in the last decades is due to a large extent to the excellent electrical properties of the thermally grown layers of silicon dioxide and to the reliability and control of the growth process.

In crystalline silicon dioxide (quartz), one silicon atom forms chemical bonds with four oxygen atoms, creating a tetrahedral structure. In turn, one oxygen atom forms chemical bonds with two silicon atoms. The tetrahedra are thus connected to form a structure with a stoichiometric ratio 1 : 2 and density $\varrho = 2.65$ g cm^{-3} (Fig. 24.1). In thermally grown SiO_2 not all tetrahedra are connected, because at the silicon–oxide interface chemical bonds must be created with the pre-existing silicon crystal, whose interatomic distance is different from that of SiO_2. As a consequence, the oxide has a shorter-range order giving rise to a more open (amorphous) structure with density $\varrho = 2.20$ g cm^{-3}. Because of this the diffusion of contaminants, in most cases Na and Li ions [4], or H_2O ions, is easier than in crystalline silicon dioxide.

Also, the need of adapting to the silicon crystal produces a mechanical stress in the oxide layer closer to the silicon surface, which in turn influences the concentration of substrate defects and the value of some electrical properties in MOS devices (typically, the threshold voltage, Sect. 22.8.2). The properties of the mechanically stressed layer are influenced by the process temperature T. At

M. Rudan, *Physics of Semiconductor Devices*, DOI 10.1007/978-3-319-63154-7_24

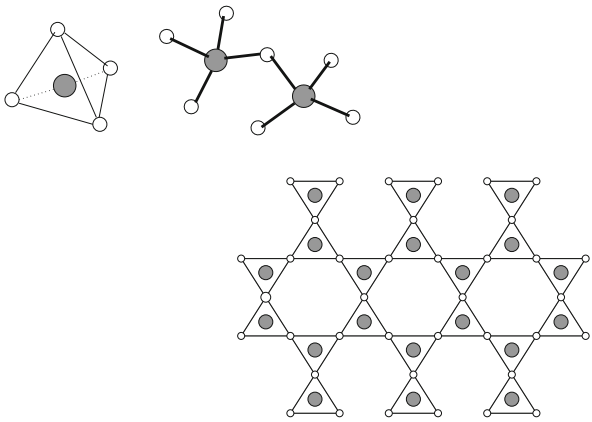


Fig. 24.1 Structure of quartz. Silicon atoms are represented in gray, oxygen atoms in white. Within the tetrahedron, the distance between two oxygen atoms is about 0.227 nm, that between the silicon atom and an oxygen atom is about 0.160 nm. The schematic representation in two dimensions is shown in the lower-right part of the figure

relatively low process temperatures, $T < 950°C$, the stressed layer is thinner and the mechanical stress in it is stronger; when $T > 950°C$, the stress distributes over a thicker layer and becomes locally weaker.

The growth of a thermal oxide's layer is obtained by inducing a chemical reaction between the silicon atoms belonging to the wafer and an oxidant species that is brought into contact with it. As mentioned above, another technique for obtaining an oxide layer is deposition. The latter process has actually a broader scope, in fact it is used for depositing several types of conducting or insulating materials that are necessary in the fabrication of the integrated circuits. Deposition differs from the thermal growth because the chemical reaction may be absent or, if present, it does not involve the species that are in the solid phase. One special type of deposition is epitaxy, that is used to grow a crystalline layer over another crystalline layer.

This chapter illustrates the oxidation of silicon, starting from the description of the chemical reactions involved in it, and deriving the relation between the thickness of the oxide layer and that of the silicon layer consumed in its growth. The kinetics of the oxide growth is analyzed, the *linear–parabolic model* is worked out, and its features are commented. Then, a brief description of the deposition processes is given, followed by the description of the chemical reaction involved in the epitaxial process and by the analysis of the epitaxial kinetics.

In the last part of this chapter a number of complementary issues are discussed. Specific data about the parameters governing the thermal oxidation, deposition, and epitaxial processes in semiconductors are in [61, Chap. 2], [129, Chap. 9], [130, Chap. 3], [88, Chap. 3]. Many carefully drawn illustrations are found in [93, Sec. 1.2].

24.2 Silicon Oxidation

Silicon exposed to air at room temperature oxidizes spontaneously and forms a shallow layer of SiO_2 of about 1 nm called *native oxide*. As soon as the native oxide is formed, the oxygen molecules of the air cannot reach the silicon surface anymore and the chemical reaction dies out. In fact, oxidation of silicon is caused by the inward motion of the oxidant. To activate the reaction and grow a layer of the desired thickness it is necessary to place the wafers at atmospheric pressure in a furnace (Fig. 24.2) kept at a temperature in the range $800°C \leq T \leq 1{,}200°C$. This increases the diffusion coefficient of the oxidant. The latter penetrates the already-formed oxide layer and reaches the silicon surface, where new SiO_2 molecules are formed.

The furnace is made of a quartz or polycrystalline-silicon tube heated by a resistance or by induction through a radiofrequency coil. To grow the oxide layer in a reproducible way it is necessary to control the temperature inside the furnace within $\pm 1°C$. The oxidant is introduced from one end of the furnace after being mixed with a carrier gas (typically, N_2 or Ar).

The chemical reactions involved in the growth of thermal oxide are different depending on the type of oxidant. The latter is either molecular oxygen (O_2) or steam (H_2O). The corresponding thermal growth is called, respectively, *dry oxidation* or *wet (steam) oxidation*. The reactions read

$$Si + O_2 \rightleftharpoons SiO_2\,, \qquad Si + 2H_2O \rightleftharpoons SiO_2 + 2H_2\,, \tag{24.1}$$

where the hydrogen molecules produced by the second reaction are eliminated by the carrier gas. The formation of SiO_2 molecules is accompanied by a change in volume. In fact, each newly formed SiO_2 molecule uses up a silicon atom initially belonging to the silicon crystal. On the other hand, the concentration of the silicon atoms in a silicon crystal is about $N_1 = 5.0 \times 10^{22}\ cm^{-3}$, while that of the silicon

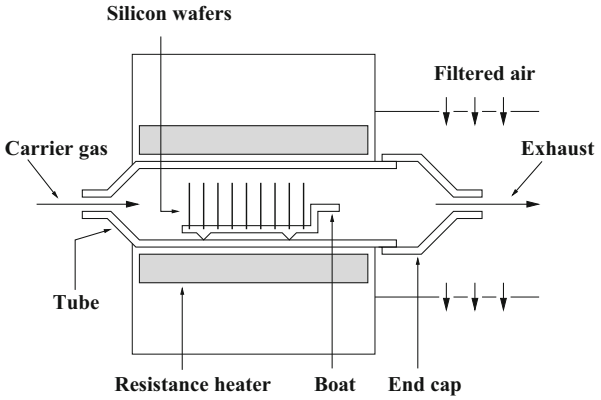


Fig. 24.2 Furnace for silicon oxidation. The intake of the carrier gas (O_2 or H_2O) is on the left end of the furnace, the exhaust on the right end. The tube, end cap, and boat are made of fused quartz to avoid contamination

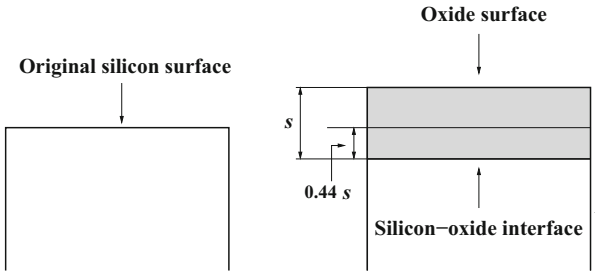


Fig. 24.3 The left part of the figure shows the position of the original silicon surface (prior to oxidation). The right part shows the position of the oxide's surface on of the silicon–oxide interface after an oxide layer of thickness s has been grown

atoms in a thermally grown SiO_2 layer is about $N_2 = 2.2\times10^{22}\,\mathrm{cm}^{-3}$. Thus the ratio between the volume V of SiO_2 and that of the silicon consumed for its formation is

$$\frac{V(\mathrm{SiO_2})}{V(\mathrm{Si})} = \frac{N_1}{N_2} \simeq 2.28\,. \tag{24.2}$$

As the oxide layer is free to expand only in the direction normal to the wafer, (24.2) is actually the ratio between the thickness s of the newly formed SiO_2 layer and that of the silicon layer consumed in the process. It follows

$$\frac{s(\mathrm{Si})}{s(\mathrm{SiO_2})} = \frac{V(\mathrm{Si})/A}{V(\mathrm{SiO_2})/A} = \frac{N_2}{N_1} \simeq 0.44\,, \tag{24.3}$$

with A the area of the oxidized region. In other terms, when an oxide layer of thickness s is grown, the silicon–oxide interface shifts by $0.44\,s$ with respect to the original position (Fig. 24.3). If the oxidation takes place uniformly over the whole area of the wafer, the shift of the interface is uniform as well. However, in many cases the oxidation involves only portions of the wafer's area. This makes the silicon–oxide interface nonplanar, because the shift is different from one portion to another.

24.3 Oxide-Growth Kinetics

Growth kinetics is modeled after Deal and Grove [38]. The model describes the sequence of steps by which the oxidant, initially in the gaseous phase, comes into contact with silicon and reacts with it. The steps are: the oxidant *i)* diffuses from the source region into the already-formed oxide, *ii)* crosses the oxide still by diffusion and reaches the silicon–oxide interface, *iii)* produces a chemical reaction that forms a new SiO_2 molecule.

The motion of the gas parallel to the wafer surface is not considered. As a consequence, the only nonvanishing component of the oxidant average velocity has the direction x normal to the wafer surface. The corresponding flux density is

$F = \mathbf{F}\cdot\mathbf{i}$, with $\mathbf{i}$ the unit vector parallel to x. The oxidant concentration N is assumed uniform over the wafer's surface, thus making N and F to depend on x and t only.

The concentration of the oxidant in the bulk of the gaseous phase, N_G, is a known boundary condition because it is regulated by the microprocessors controlling the furnace. In the gaseous region, at the gas–oxide interface, and in the oxide region, no generation or destruction of oxidant molecules occurs. The flux density is given by

$$F = -D_S \frac{\partial N}{\partial x}\,, \qquad F = -D_O \frac{\partial N}{\partial x}\,, \tag{24.4}$$

respectively in the source and oxide region. In (24.4), the symbol D_S (D_O) indicates the diffusion coefficient of the oxidant in the source (oxide) region. Each diffusion coefficient is taken independent of time and spatially constant in its own region. The matching conditions at the source–oxide interface are the same as in (23.20), namely,

$$N_O = k\,N_S\,, \qquad D_O \left(\frac{\partial N}{\partial x}\right)_O = D_S \left(\frac{\partial N}{\partial x}\right)_S\,, \tag{24.5}$$

where k is the gas–oxide segregation coefficient, while the index S (O) attached to the concentration or its derivative indicates that the function is calculated at the source–oxide interface on the side of the source (oxide). As one of the two phases is gaseous, it is $D_S \gg D_O$ whence $|(\partial N/\partial x)_S| \ll |(\partial N/\partial x)_O|$. The situation here is similar to that illustrated in Fig. 23.3. It follows that the interface concentration of the source region, N_S, is practically equal to the boundary condition N_G. The first of (24.5) then yields $N_O = k\,N_G$.

To proceed one observes that due to the thinness of the oxide layer, the oxidant concentration in it can be described by a linear approximation. Due to this, the flux density in the oxide layer (the second equation in (24.4)) becomes

$$F = -D_O \frac{N_I - N_O}{s} = D_O \frac{kN_G - N_I}{s}\,, \tag{24.6}$$

where N_I is the oxidant concentration on the oxide side of the silicon–oxide interface, and s the oxide thickness. Note that the flux density in (24.6) is constant with respect to x, whereas it is time dependent because s increases with time.

When the oxidant reaches the silicon–oxide interface it reacts with silicon, so that there is no flux density of the oxidant on the semiconductor's side of this interface. In fact, the oxidant's molecules are destroyed at the interface to form molecules of SiO_2. The flux density F_I entering the silicon–oxide interface gives the number of oxidant molecules destroyed per unit area and time which, to a first approximation, is taken proportional to the concentration N_I. It follows

$$F_I = v_r\,N_I\,, \tag{24.7}$$

where the constant v_r (m s^{-1}) is called *reaction velocity*. As F_I is just another symbol to denote the spatially constant flux density within the oxide, one combines (24.7) with (24.6) to obtain

$$N_I = \frac{D_O\, k\, N_G}{v_r s + D_O}\,. \tag{24.8}$$

At a given instant equation (24.8) expresses N_I in terms of the boundary condition N_G, the process parameters k, D_O, v_r, and the oxide thickness s.

24.4 Linear-Parabolic Model of the Oxide Growth

The relation between the oxidant's flux density F and the growth velocity $\mathrm{d}s/\mathrm{d}t$ of the oxide layer is found as follows. Letting A be the area of the oxidized region, the product $A\,F$ provides the number of oxidant molecules reaching the silicon–oxide interface per unit time. Each molecule, in turn, makes the volume V of the oxide layer to increase by a finite amount w. As a consequence, the volume increase per unit time of the oxide layer is $\mathrm{d}V/\mathrm{d}t = wA\,F$. As shown in Sect. 24.2, the oxide layer is free to expand only in the direction normal to the wafer, so that $\mathrm{d}V/\mathrm{d}t = A\,\mathrm{d}s/\mathrm{d}t$. Combining the above relations and using (24.7, 24.8) yield a differential equation in the unknown s:

$$\frac{\mathrm{d}s}{\mathrm{d}t} = w\,F_I = w\,v_r\,N_I = v_r\,D_O\,\frac{w\,k\,N_G}{v_r\,s + D_O}\,. \tag{24.9}$$

The above is readily separated as $(s/D_O + 1/v_r)\,\mathrm{d}s = w\,k\,N_G\,\mathrm{d}t$ and integrated from $t = 0$, to yield

$$\frac{1}{c_p}\left(s^2 - s_i^2\right) + \frac{1}{c_l}\,(s - s_i) = t\,, \qquad \begin{cases} c_p = 2w\,k\,N_G\,D_O \\ c_l = w\,k\,N_G\,v_r \end{cases}, \tag{24.10}$$

with $s_i = s(t = 0)$. The relation between s and t given by (24.10) is called *linear–parabolic model* of the oxide growth. The quantities c_p (m^2 s^{-1}) and c_l (m s^{-1}) are the *parabolic coefficient* and *linear coefficient*, respectively. The model is recast as

$$\frac{1}{c_p}\,s^2 + \frac{1}{c_l}\,s = t + \tau\,, \qquad \tau = \frac{1}{c_p}\,s_i^2 + \frac{1}{c_l}\,s_i\,. \tag{24.11}$$

Using (24.11) and the definitions (24.10) of c_p and c_l one finds two limiting cases of the $s(t)$ relation. Specifically, it is $s \simeq c_l\ (t + \tau)$ when the oxide thickness is such that $v_r\,s \ll 2\,D$, while it is $s \simeq [c_p\,(t + \tau)]^{1/2}$ when the oxide thickness is such that $v_r\,s \gg 2\,D$. Due to the form of c_p and c_l, in the first limiting case the oxide growth does not depend on the diffusion coefficient D_O, whereas in the second limiting case it does not depend on the reaction velocity v_r. This is easily understood if

one considers that the concentration $N_O = k N_G$ is prescribed. As a consequence, as long as the oxide thickness is small the derivative of the concentration (hence the flux density) is limited essentially by the flux density entering the silicon–oxide interface, (24.7); on the contrary, when the oxide thickness becomes large the flux density is limited essentially by the diffusion across the oxide because the value of the concentration N_I at the silicon–oxide interface becomes less important.

The linear-parabolic model (24.11) is recast in terms of dimensionless variables by multiplying both sides by c_l^2/c_p, thus yielding $(c_l\, s/c_p)^2 + c_l\, s/c_p = c_l^2\,(t+\tau)/c_p$. Inverting the latter yields the curve shown in Fig. 24.6; the two limiting cases are reported as well. The differential form of (24.11),

$$\frac{\mathrm{d}t}{\mathrm{d}s} = \frac{2}{c_p}\, s + \frac{1}{c_l}\,, \tag{24.12}$$

is a linear relation between $\mathrm{d}t/\mathrm{d}s$ and s. Such quantities can be measured independently from each other, thus providing a method for measuring c_p and c_l. Repeating the measurement at different temperatures shows that the temperature dependence of the parabolic and linear coefficients is given by

$$c_p = c_{p0}\, \exp[-E_{ap}/(k_B T)]\,, \qquad c_l = c_{l0}\, \exp[-E_{al}/(k_B T)]\,. \tag{24.13}$$

The form of (24.13) is due to the temperature dependence of $D \propto \exp[-E_{ap}/(k_B T)]$ and, respectively, $v_r \propto \exp[-E_{al}/(k_B T)]$. In fact, the parameters w, N_G that appear in the definitions (24.10) are independent of temperature, whereas the temperature dependence of the segregation coefficient k, that can be measured independently, is shown to be relatively weak.

The measurement of c_p and c_l allows one to determine also other properties of the oxidation process; for instance, the effect of carrying out a steam or dry oxidation, and the influence of the substrate orientation. As for the first issue one finds

$$k(\text{Steam}) > k(\text{Dry})\,, \qquad D_O(\text{Steam}) > D_O(\text{Dry})\,. \tag{24.14}$$

The Arrhenius plots of c_p and c_l are shown in Figs. 24.4 and 24.5, respectively. In each plot the upper (lower) continuous curve refers to the steam (dry) oxidation. As for the effect of the crystal orientation of the silicon wafer, one observes that the number of chemical reactions per unit time involved in the formation of SiO_2 molecules must depend on the surface density of silicon atoms at the silicon–oxide interface. Due to this, the reaction velocity is expected to depend on the orientation of the interface. The crystal planes that are typically used in the silicon technology are the (111) one, whose surface density is 11.8×10^{14} cm^{-2}, and those equivalent to the (100) one,[1] whose surface density is 6.8×10^{14} cm^{-2}. In fact the experiments show that

[1] The definitions of the crystal planes are given in Sect. 17.9.1.

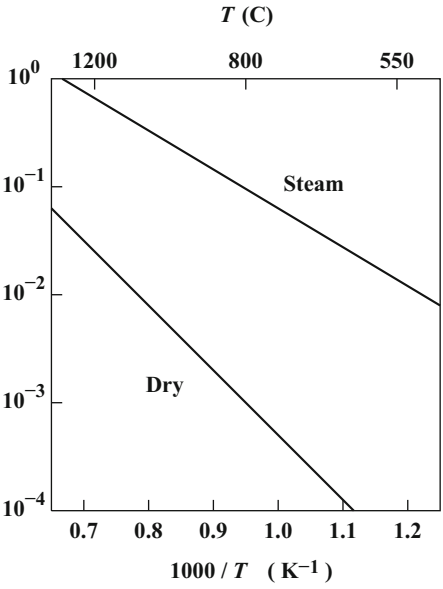


Fig. 24.4 Parabolic coefficient c_p as a function of $1{,}000/T$. The units are $\mu m^2\ h^{-1}$. The activation energy of the steam case is 0.71 eV, that of the dry case is 1.24 eV

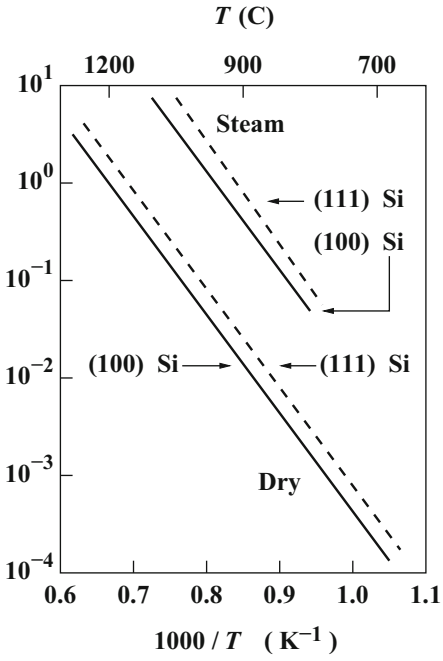


Fig. 24.5 Linear coefficient c_l as a function of $1{,}000/T$. The units are $\mu m\ h^{-1}$. The activation energy of the steam case is 2.05 eV, that of the dry case is 2.0 eV

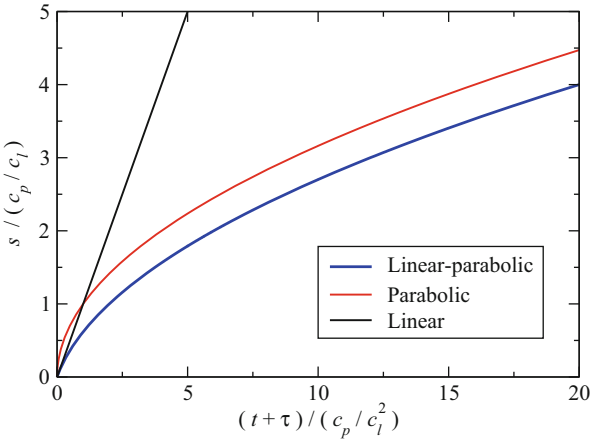


Fig. 24.6 The linear-parabolic model (24.11) is drawn using dimensionless variables (*blue line*). The linear (*black curve*) and parabolic (*red curve*) limiting cases are also shown

$$\frac{v_r[(111)]}{v_r[(100)]} = 1.68 \simeq \frac{11.8 \times 10^{14}\ \mathrm{cm}^{-2}}{6.8 \times 10^{14}\ \mathrm{cm}^{-2}}\,. \tag{24.15}$$

The effect on c_l of the crystal orientation is shown by the dotted curves in the Fig. 24.5.

24.5 Layer Deposition and Selective Oxide Growth

The deposition of films of different materials is necessary at several steps of the integrated-circuit fabrication. Conducting materials provide the electrical connections among the individual devices of the integrated circuit, while insulating materials provide the electrical insulation between the metal layers, and the protection from the environment. The majority of the deposition processes take place in the vapor phase under reduced-pressure or vacuum conditions. One exception is the deposition of resist, which is carried out in the liquid phase.

When the material to be deposited does not react chemically with other substances, the process is called *physical vapor deposition* (PVD). An example of PVD is the deposition of a metal by evaporation *in vacuo*. When the material to be deposited is the product of a chemical reaction that takes place over the wafer surface or in its vicinity, the process is called *chemical vapor deposition* (CVD). The materials that are most widely used in CVD processes are polycrystalline silicon (also termed *polysilicon*), silicon dioxide (SiO_2), and silicon nitride (Si_3N_4). Examples of CVD reactions are given in Table 24.1. More examples are found in [130, Sect. 6-2].

Table 24.1 Examples of CVD reactions

Product	Reaction[a]	Deposition temperature (°C)
Polysilicon	$SiH_4 \rightarrow Si + 2H_2$	575–650
Silicon dioxide	$SiH_4 + O_2 \rightarrow SiO_2 + 2H_2$	400–450
Silicon nitride	$3SiH_4 + 4NH_3 \rightarrow Si_3N_4 + 12H_2$	700–900

[a] SiH_4 and NH_3 are called *silane* and *ammonia*, respectively

The structure of the deposited layer depends on the substrate's properties and deposition conditions. In the manufacturing of integrated circuits the substrate is crystalline, that is, it has long-range order extending throughout the entire volume (Chap. 17). If the material to be deposited on a crystalline substrate is the same as that of the substrate, by means of a carefully controlled process it is possible to obtain a deposited layer that replicates the substrate's structure. Such a process is called *epitaxy* and, with reference to silicon, is described in Sect. 24.6.

The structure of silicon deposited on a different material is polycrystalline, that is, it has a long-range order only within small volumes. Such volumes, called *grains*, have an average diameter of about 1 μm and are oriented randomly with respect to each other. Polycrystalline silicon is used for fabricating the gate electrodes in MOS devices, for obtaining ohmic contacts to shallow junctions, and for producing resistors. To increase the gate's conductivity, a layer of metal or metal silicide (like tungsten or tantalum silicide) may be deposited over the polycrystalline silicon.

The structure of deposited SiO_2 or Si_3N_4 is amorphous, that is, it has a short-range order only. The applications of SiO_2 have been illustrated in Sect. 24.1. Silicon nitride Si_3N_4 provides a strong barrier to the diffusion of water, that corrodes the metallizations, and of other contaminants, like sodium, that make the devices unstable by changing their threshold voltage. In addition, Si_3N_4 is resistant to high temperatures and oxidizes slowly. For these reasons it is used for passivating the wafer and for producing the masks that are necessary for the selective oxidation of silicon. The latter process, also called *local oxidation* (LOCOS), consists in depositing and patterning a Si_3N_4 layer over the areas where the substrate's oxidation must be prevented. As oxidation is isotropic, a lateral penetration of the oxidized region occurs under the edge of Si_3N_4. This produces a characteristic profile of the oxide layer called *bird's beak* (Fig. 24.7). To compensate for the effect of the lateral penetration, the Si_3N_4 mask must be larger than the area whose oxidation is to be prevented.

A layer replicates the topography of the surface onto which it is deposited. For this reason it is important to avoid, or reduce, the formation of steps on the substrate. In fact, over a step the layer's thickness is smaller than on a flat surface (Fig. 24.8) which, in turn, may cause reliability problems in the final circuit. For instance, the nonuniform thickness of a metal line causes a nonuniform distribution of the current density. This may induce metal migration and the eventual breakdown of the metal connection.

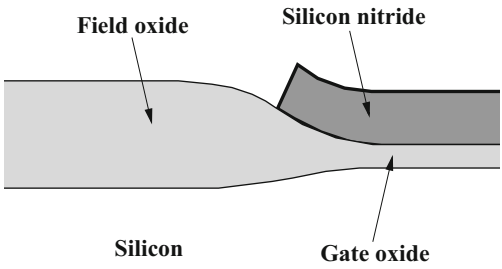


Fig. 24.7 Schematic cross-section of the transition from field oxide to gate oxide, showing the bird's beak profile

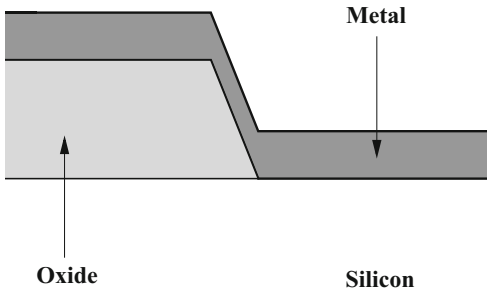


Fig. 24.8 Schematic description of the thinning of a metal layer deposited over a step

24.6 Epitaxy

Epitaxy (from the Greek verb *epitàsso*, "to deploy side by side") is used to grow a monocrystalline layer over another monocrystalline layer. Most epitaxial processes use the CVD method. When the epitaxial layer is made of the same material as the substrate, e.g., silicon over silicon, the term *homoepitaxy* is also used, while the term *heteroepitaxy* is reserved to the case where the materials are different. Heteroepitaxy is possible when the difference between the lattice constants[2] of the two materials is small. An example of heteroepitaxy is the silicon-on-sapphire (SOS) process that belongs to the silicon-on-insulator (SOI) technological family and consists in growing a thin layer of silicon (about 0.5 μm) on a wafer made of a sapphire crystal (Al_2O_3). Another application of heteroepitaxy is the fabrication of the heterojunctions that are necessary in optoelectronic devices.

In the silicon technology, epitaxy originated from the need of producing high-resistance layers in bipolar technology. This type of layers is necessary, e.g., for realizing the collector region of the bipolar junction transistor, whose dopant concentration must be substantially lower than that of the base region (Sect. 21.8). Due to the high temperature of the CVD process (about 1,200°C), during an epitaxy a diffusion occurs of the substrate dopant into the epitaxial layer and of the epitaxial-layer's dopant into the substrate. This effect must be accounted for, and compensated, at the design stage of the process.

[2]The definition of lattice constant is given in Sect. 17.6.4.

The fundamental reaction of epitaxy combines silicon tetrachloride $SiCl_4$ with molecular hydrogen in the vapor phase to obtain silicon in solid phase, while the hydrochloric acid HCl remains in the vapor phase and is eliminated:

$$SiCl_4 + 2H_2 \rightleftharpoons Si + 4HCl\,. \tag{24.16}$$

Reaction (24.16) is reversible: an excess of HCl removes silicon atoms from the wafer's surface and releases $SiCl_4$ and $2H_2$ in the vapor phases. This reaction is used in the first stages of the process to the purpose of cleaning the wafer's surface. Besides (24.16), a secondary reaction takes place as well, namely,

$$SiCl_4 + Si \rightarrow 2SiCl_2\,. \tag{24.17}$$

Reaction (24.17) removes silicon from the wafer's surface and releases silicon dichloride $SiCl_2$ in the vapor phase. For this reason, reactions (24.16) and (24.17) compete with each other. When the vapor concentration of $SiCl_4$ is sufficiently low the first reaction prevails and the thickness of the epitaxial layer increases with time. In contrast, at higher $SiCl_4$ concentrations the second reaction prevails and silicon is etched.

The epitaxial layer is doped by introducing hydrides of the dopants into the vapor phase. The hydride, e.g., *arsine* (AsH_3), *phosphine* (PH_3), or *diborane* (B_2H_6), is absorbed on the surface, decomposes, and is incorporated in the growing layer, e.g.,

$$2AsH_3 \rightarrow 2As + 3H_2\,. \tag{24.18}$$

24.7 Kinetics of Epitaxy

As in the case of the oxide-growth kinetics (Sect. 24.3), the motion of the vapor parallel to the wafer surface is not considered. As a consequence, the only nonvanishing component of the average velocity of the $SiCl_4$ molecules has the direction x normal to the wafer surface. The corresponding flux density is $F = \mathbf{F}\cdot\mathbf{i}$, with $\mathbf{i}$ the unit vector parallel to x. The $SiCl_4$ concentration N is assumed uniform over the wafer's surface, thus making N and F to depend on x and t only.

The $SiCl_4$ concentration in the bulk of the vapor phase, N_G, is a known boundary condition because it is regulated by the microprocessors controlling the furnace. The flux density is given by

$$F = -D\,\frac{\partial N}{\partial x} \simeq v_G\,(N_G - N_I)\,, \tag{24.19}$$

where D is the diffusion coefficient of $SiCl_4$ in the vapor phase and N_I the $SiCl_4$ concentration at the wafer's surface. The diffusion coefficient is taken independent of time and spatially constant. The form of the right-hand side of (24.19), where the parameter v_G (m s^{-1}) is called *gas-phase, mass-transfer coefficient*, is due to the observation that D is very large because the diffusion takes place in the vapor phase. As a consequence, the derivative $\partial N/\partial x$ is so small that a linear approximation for N is acceptable (the situation here is similar to that illustrated for the region on the left of the origin in Fig. 23.3). Note that the flux density in the vapor phase (24.19) is constant with respect to x. In principle it depends on time because the extension of the vapor phase decreases due to the growth of the epitaxial layer. However, this time dependence can be disregarded because the relative variation in the vapor-phase extension is negligible.

The flux density F_I entering the silicon surface gives the number of $SiCl_4$ molecules destroyed per unit area and time which, to a first approximation, is taken proportional to the concentration N_I. It follows

$$F_I = v_r N_I \,, \tag{24.20}$$

where the constant v_r (m s^{-1}), as in the case of the oxide-growth kinetics, is called *reaction velocity*. As F_I is just another symbol to denote the spatially constant flux density, one combines (24.20) with (24.19) to obtain

$$N_I = \frac{v_G}{v_r + v_G} N_G \,. \tag{24.21}$$

At a given instant equation (24.21) expresses N_I in terms of the boundary condition N_G and process parameters v_G, v_r.

The relation between the flux density F of $SiCl_4$ and the growth velocity $\mathrm{d}s/\mathrm{d}t$ of the epitaxial layer is found by the same reasoning as that used in Sect. 24.4 for the growth velocity of SiO_2. From (24.20, 24.21) it follows

$$\frac{\mathrm{d}s}{\mathrm{d}t} = wF_I = wv_rN_I = w\,\frac{v_rv_G}{v_r + v_G} N_G \tag{24.22}$$

whence, observing that $s(t = 0) = 0$,

$$s = c_l\, t \,, \qquad c_l = w\,\frac{v_rv_G}{v_r + v_G} N_G \,. \tag{24.23}$$

The $s(t)$ relation (24.23) is linear with respect to time. The growth velocity c_l of the epitaxial layer depends on the concentration N_G of $SiCl_4$ at the boundary and on the process parameters w, v_r, and v_G. The temperature dependence of the gas-phase mass-transfer coefficient v_G is weak. As w and N_G are independent of temperature, the temperature dependence of c_l is to be ascribed to v_r. It is found

$$v_r = v_{r0}\, \exp[-E_{al}/(k_BT)] \,. \tag{24.24}$$

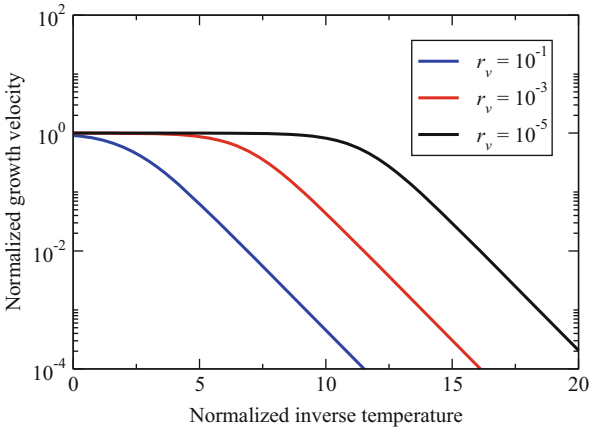


Fig. 24.9 Normalized growth velocity as a function of the normalized inverse temperature, as given by (24.23) and (24.24), at different values of the $r_v = v_G/v_{r0}$ ratio

When the temperature is such that $v_r \ll v_G$, which typically happens for $T < 1{,}150°\text{C}$, the second of (24.23) yields the limiting case $c_l \simeq w N_G v_r$, whence $c_l \propto \exp[-E_{al}/(k_B T)]$; when, instead, it is $v_r \gg v_G$, which typically happens for $T > 1{,}200°\text{C}$, the limiting case is $c_l \simeq w N_G v_G = \text{const}$. An Arrhenius plot of the normalized growth velocity $c_l/(w N_G v_G)$ as a function of the normalized inverse temperature $E_{al}/(k_B T)$ is shown in Fig. 24.9 for different values of the $r_v = v_G/v_{r0}$ ratio.

24.8 Complements

24.8.1 An Apparent Contradiction

In commenting (24.6) it was noted that the flux density F in the oxidation process is constant with respect to x, whereas it depends on time due to the time dependence of the oxide thickness s. This seems to bring about a contradiction. In fact, as the one-dimensional form of the continuity equation (23.3) with $W = 0$ yields $\partial N/\partial t + \partial F/\partial x = 0$, the constancy of F makes N independent of time. However, N does depend on time. This is demonstrated by Fig. 24.10 that shows the linear approximation of the oxidant concentration within the oxide at two different instants, t_1 and $t_2 > t_1$. The value $k N_G$ at the source–oxide interface is kept constant by the boundary condition as explained in Sect. 24.3, while the value N_I at the silicon–oxide interface changes with time due to (24.8), and the oxide's thickness changes as well.

The contradiction is eliminated by observing that the continuity equation (23.3) has been derived for the case where the boundary is fixed, whereas the growth of thermal oxide is a moving-boundary process. The motion of the boundary is not a rigid one (otherwise the problem could be fixed by moving the reference

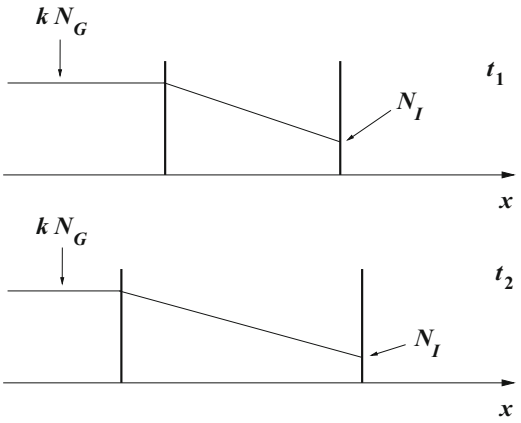


Fig. 24.10 Oxidant concentration within the oxide at two different instants, t_1 and $t_2 > t_1$

accordingly), because the oxide volume is actually expanding. In conclusion, equation (23.3) must not be used. In fact, the derivation of the linear–parabolic model of the oxide growth (24.10) is based solely on the definition of the flux density.

24.8.2 Elementary Contributions to the Layer's Volume

The relation $\mathrm{d}s/\mathrm{d}t = w F_I$ was used to connect, in Sect. 24.4, the growth velocity of the oxide layer to the flux density of the oxidant and, in Sect. 24.7, the growth velocity of the epitaxial layer to the flux density of $SiCl_4$. The coefficient w is the amount by which one oxidant or $SiCl_4$ molecule makes the volume of the layer to increase. To specify w for the oxidation process one must distinguish between the dry and steam cases. In the first one, each molecule of the oxidant produces one SiO_2 molecule. As a consequence, w is the volume of the SiO_2 molecule. In the steam case, two H_2O molecules are necessary for producing one SiO_2 molecule, hence w is half the volume of the latter. By the same token, in the epitaxial process w is the volume of a Si atom.

24.8.3 Features of the Oxide Growth and Epitaxial Growth

The quadratic term in the left-hand side of (24.10) becomes dominant at larger oxide thicknesses. This in turn slows down the growth rate, as shown by (24.9). A qualitative explanation of the phenomenon is easily obtained by considering that in order to reach the silicon–oxide interface, the oxidant must diffuse across the already-formed oxide. The slope of the oxidant concentration, hence its flux density, decreases with time because the thickness of the oxide region increases, while the value $k N_G$ at the source–oxide interface is kept constant by the boundary condition.

The decrease in the oxidant concentration N_I at the silicon–oxide interface, shown by (24.8), is not sufficient to contrast the decrease in the concentration's slope. The reasoning above does not apply to the epitaxial growth; in fact, in this case the chemical reaction occurs at the vapor–silicon interface and there is no intermediate layer to be crossed. As a consequence, the corresponding model (24.23) has no quadratic term.

In the analysis of the oxide-growth kinetics carried out in Sect. 24.3 it is assumed that the interface concentration in the source region, N_S, is practically equal to the boundary condition N_G. The simplification is used in the expression (24.6) of the flux density in the oxide layer. The calculation then proceeds by considering only the oxidant diffusion across the already-formed layer and the chemical reaction at the silicon–oxide interface. In this respect, the assumption $N_S = N_G$ has the mere effect of introducing a negligible change in (24.6). A similar approximation would not be possible in the analysis of the epitaxial growth. In fact, letting $N_I = N_G$ in (24.19) would set the flux density to zero. The difference between the two cases is that in the epitaxial growth the flux density exists only in the vapor phase, while in the oxide growth it exists both in the gaseous and solid phases. However, as $D_S \gg D_O$, only the diffusion in the solid phase plays a significant role in determining the kinetics of the oxidation process.

24.8.4 Reaction Velocity

The reaction velocity v_r is among the parameters used in the analysis of the oxide-growth kinetics and epitaxial kinetics. This parameter controls the flux density through (24.7) or (24.20), and is found to depend also on the concentration N_{dop} of dopant atoms in the silicon lattice. The dependence is negligible as long as $N_{\text{dop}} \leq n_i(T)$, where n_i (called *intrinsic concentration*, Sect. 18.3) is calculated at the process temperature. When $N_{\text{dop}} > n_i(T)$, the reaction velocity increases with N_{dop}. It should be noted that $n_i \simeq 10^{18}$ cm^{-3} at $T = 1{,}000$°C. As a consequence, the dependence of v_r on N_{dop} becomes important only at relatively high dopant concentrations.

24.8.5 Molecular Beam Epitaxy

Epitaxy can also be obtained by a process different from CVD, that is called *molecular beam epitaxy* (MBE) and is based on evaporation. The main advantages of MBE are the low-temperature processing and the fine control of the dopant distribution throughout the epitaxial layer. On the other side, MBE has a low throughput and a higher cost. As a consequence, CVD is used is in the majority of cases for growing epitaxial layers in the silicon technology [130, Sect. 6-3].

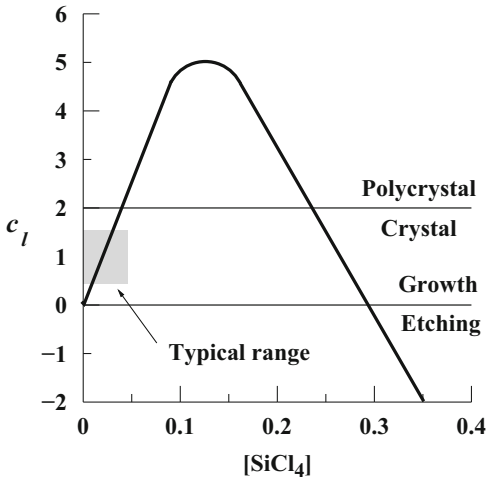


Fig. 24.11 Typical growth velocity c_l of an epitaxial process, expressed in microns per minute, as a function of the mole fraction of tetrachloride. The shaded area shows the typical operating range

24.8.6 Secondary Reaction in the Epitaxial Growth

The analysis of the epitaxial kinetics carried out in Sect. 24.7 is based on the hypothesis that only the fundamental reaction (24.16) is present. However, as mentioned in Sect. 24.7, the secondary reaction (24.17) also takes place, which removes silicon atoms from the wafer's surface and, therefore, competes with (24.16). At low concentrations of $SiCl_4$ the effect of the secondary reaction is negligible and the theory of Sect. 24.7 holds; in particular, as shown by (24.23), the growth velocity c_l is proportional to the concentration N_G of $SiCl_4$ in the bulk of the vapor phase. At larger tetrachloride concentrations the dependence of c_l on N_G becomes sublinear as shown in Fig. 24.11. Further increases in N_G make c_l to continuously decrease, and to eventually vanish when the secondary reaction balances the fundamental one. Further on, the secondary reaction prevails and c_l becomes negative, that is, silicon is etched.

The growth velocity also influences the structure of the epitaxial layer. When c_l is low the deposited silicon atoms match the preexisting crystalline structure, so that the newly formed layer is crystalline as well. Higher values of c_l make the matching more and more difficult. Beyond a critical value of c_l the epitaxial layer is polycrystalline, as sketched in Fig. 24.11. To avoid the growth of a polycrystal it is necessary to keep the $SiCl_4$ concentration in the range shown in the figure.

Problems

24.1 A silicon wafer covered with an $s_i = 0.21$ μm-thick thermal oxide undergoes a second thermal oxidation whose duration is 136 minutes. Using the values $c_p = 4.43 \times 10^{-2}$ μm^2 h^{-1} for the parabolic coefficient and $c_l = 8.86 \times 10^{-1}$ μm h^{-1} for the linear coefficient, calculate the silicon thickness consumed during the second thermal oxidation.

24.2 Consider a thermal-oxidation process of silicon where the parabolic and linear coefficients are, respectively, 4.43×10^{-2} μm^2 h^{-1} and 8.86×10^{-1} μm h^{-1}. At some instant the oxidant concentration N_I at the silicon–oxide interface is 1×10^{12} cm^{-3}. Find the gradient of the oxidant concentration within the oxide layer at the same instant, expressed in 10^{17} cm^{-4} units.

24.3 A silicon wafer covered with an $s_i = 105$-nm thick layer of thermal oxide undergoes a second thermal-oxidation process that consumes a 100-nm thick layer of silicon. Letting the parabolic and linear coefficients be $c_p = 4.43\times10^{-2}$ μm^2 h^{-1} and $c_l = 8.86 \times 10^{-1}$ μm h^{-1}, respectively, determine the duration in minutes of the second oxidation process.

24.4 Consider a thermal-oxidation process of silicon where the parabolic and linear coefficients are, respectively, 0.12 μm^2 h^{-1} and 3 μm h^{-1}. At some instant the oxide thickness is 20 nm, and the oxidant concentration in the oxide at the source–oxide interface is $N_O = 3\times10^{12}$ cm^{-3}. Find the oxidant concentration at the silicon–oxide interface, expressed in 10^{11} cm^{-3} units.

24.5 A silicon wafer covered with an $s_i = 5$-nm thick layer of thermal oxide undergoes a thermal-oxidation process that grows a $\Delta s_1 = 7$-nm thick oxide layer and a successive thermal-oxidation process that grows a $\Delta s_2 = 20$-nm thick oxide layer. In both processes the parabolic and linear coefficients are, respectively, 4.1×10^{-2} μm^2 h^{-1} and 8.5×10^{-1} μm h^{-1}. Determine the total duration in seconds of the two processes.

24.6 A silicon wafer covered with an $s_i = 80$-nm thick layer of thermal oxide undergoes a thermal-oxidation process whose linear coefficient is 1 μm h^{-1}. The ratio between the diffusion coefficient of the oxidant within the oxide and the reaction velocity at the silicon–oxide interface is $r = D_O/v_r = 50$ nm. Find how many minutes are necessary to reach a final oxide thickness equal to 150 nm.

24.7 A silicon wafer covered with an $s_i = 40$-nm thick layer of thermal oxide undergoes a thermal-oxidation process where the oxidant's diffusion coefficient in the oxide is $D_O = 4.5 \times 10^{-6}$ cm^2 s^{-1}, the reaction velocity is $v_r = 4$ cm s^{-1}, and the product of the parabolic coefficient and the process duration is $c_p t_P = 5 \times 10^{-11}$ cm^2. Calculate the final thickness of the oxide in nm.

24.8 A silicon wafer undergoes an epitaxial growth that produces a 12 μm-thick silicon layer. At the end of the process the wafer's weight has increased by 907 mg. Using $p_{\mathrm{Si}} = 2.33$ g cm^{-3} for the specific weight of silicon, determine the wafer's diameter in inches (1 in = 2.54 cm).

24.9 A 1.2 μm-thick epitaxial layer of silicon is grown by a 1 minute-long process in which the reaction velocity is $v_r = 10$ cm s^{-1}. Find the concentration of $SiCl_4$ at the silicon surface expressed in 10^{16} cm^{-3} units.

24.10 The flux density of $SiCl_4$ in an epitaxial process in silicon is 8.33×10^{16} $cm^{-2}s^{-1}$. Remembering that the concentration of the silicon atoms in the crystal lattice is 5×10^{22} cm^{-3}, determine how many minutes are necessary to grow a 2 μm-thick epitaxial layer.

24.11 Determine the reaction velocity v_r (in cm min^{-1}) of an epitaxial process in silicon that in 5 min grows an $s = 2$ μm thick layer. For the surface concentration of the silicon tetrachloride and the atomic volume of silicon use, respectively, the values 10^{16} cm^{-3} and $(5 \times 10^{22})^{-1}$ cm^3.

24.12 In an epitaxial process the ratio between the $SiCl_4$ concentration in the bulk of the vapor phase and at the wafer's surface is $a = N_G/N_I = 2$, while the ratio between the reaction velocity and growth velocity is $b = v_r/c_l = 4.87 \times 10^5$. Remembering that the concentration of the silicon atoms in the crystal lattice is 5×10^{22} cm^{-3}, determine the value of N_G in cm^{-3}.

Chapter 25
Measuring the Semiconductor Parameters

25.1 Introduction

A number of methods used for measuring the semiconductor parameters are illustrated here. Apart from the intrinsic usefulness, the methods are interesting because they show the connection with the theories worked out in other chapters. For example, the measurement of lifetimes exploits the features of the net thermal recombination and of optical generation, that are combined in a simplified form of the continuity equation for the minority carriers. Similarly, the measurement of mobility carried out with the Haynes-Shockley experiment is based on a clever use of the diffusion of optically generated carriers. The Hall effect, in turn, provides a powerful method to extract the information about the concentration and mobility of the majority carriers; the method exploits the effect of a magnetic field applied in the direction normal to that of the current density, and is widely used for determining, e.g., the dependence of concentration and mobility on the concentration of dopants and on temperature. The analysis of the Hall effect is enriched by a detailed treatment of the case where the standard theory is not applicable because the device is not sufficiently slender; a deeper analysis based on the concept of stream function shows that the equations describing the current-density field are in fact solvable for any aspect ratio of the device. This chapter is completed by the illustration of a method for measuring the doping profile in an asymmetric, reverse-biased, one-dimensional junction; the procedure is based on the observation that despite the fact that the relation between the applied voltage and the extension of the space-charge region is nonlinear, the differential capacitance of the junction has the same form as that of a parallel-plate, linear capacitor. Finally, the van der Pauw method for measuring the conductivity of a sample is illustrated, based on the use of a two-dimensional Green function introduced in an earlier chapter and on the conformal-mapping method shown in the Appendix.

M. Rudan, *Physics of Semiconductor Devices*, DOI 10.1007/978-3-319-63154-7_25

25.2 Lifetime Measurement

The lifetimes have been introduced in Sect. 20.2.3 with reference to the trap-assisted, thermal generation and recombination phenomena. A measurement method for the lifetimes is illustrated here with the aid of Fig. 25.1 that shows a uniformly doped, thin layer of semiconductor of length L and cross-section A. The method is able to measure the minority-carrier lifetime; in the example shown in the figure, which refers to an n-doped material, the outcome is τ_p. The x axis is taken normal to the external surface of the semiconductor, with the origin placed on such a surface. The latter, parallel to the y, z plane, is illuminated with a monochromatic radiation of frequency ν, whose intensity is uniform over the surface and constant in time. Remembering the expression (20.67) of the optical-generation term, and observing that due to uniformity the absorption coefficient does not depend on position, one finds

$$G_O = \eta \, \Phi_B \, k \, \exp(-k\,x) \,, \qquad k = k(\nu) \,. \tag{25.1}$$

At the same time, a constant voltage V is applied to the semiconductor, producing an electric field in the z direction. The device is thin in the x and y directions, and elongated in the z direction, to the extent that the flow lines of the current density are substantially parallel to the z axis; also, the small extension in the x direction makes it possible to neglect the x-dependence of G_O and let $G_O \simeq G_c = \eta \, \Phi_B \, k$. With these premises,[1] the material is spatially uniform also in the nonequilibrium case; it follows that the condition of local charge neutrality holds. For simplicity one assumes that the n-dopant concentration N_D is sufficiently low to ensure that nondegeneracy and complete ionization hold, $N_D^+ = N_D$; thus, local neutrality reads

$$n = p + N_D \,. \tag{25.2}$$

The nonequilibrium condition prescribed by the combination of illumination and bias is adjusted in such a way that a weak-injection condition holds (Sect. 20.2.3); in this case the hole-continuity equation (19.124) reads

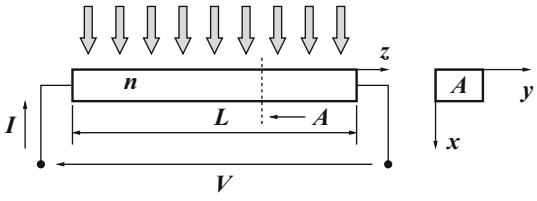


Fig. 25.1 Measurement scheme for the minority-carrier lifetime

[1] The required thinness of the device can be achieved by growing an n-type epitaxial layer over a p-type substrate, and keeping the layer-substrate junction reverse biased.

$$\frac{\partial p}{\partial t} + \frac{1}{q} \operatorname{div} \mathbf{J}_p = G - U \simeq G_c - \frac{p - p_{n0}}{\tau_p} . \tag{25.3}$$

In the steady-state, uniform condition considered here, (25.3) reduces to

$$\Delta p = p - p_{n0} = \tau_p \, G_c \, , \tag{25.4}$$

showing that optical generation is exactly balanced by thermal recombination. If, at time $t = 0$, the source of light is removed, G_c vanishes and recombination prevails; it follows that the semiconductor undergoes a transient to adapt to the new situation. On the other hand, spatial uniformity still holds: during the transient the spatial derivatives in (25.3) are still zero, and the equation reads

$$\frac{\mathrm{d}p}{\mathrm{d}t} = \frac{\mathrm{d}(p - p_{n0})}{\mathrm{d}t} = -\frac{p - p_{n0}}{\tau_p} \, , \qquad t > 0 \, . \tag{25.5}$$

From (25.4), the initial condition of (25.5) is $\Delta p(t = 0) = \tau_p \, G_c$; thus, the solution of (25.5) is found to be

$$\Delta p = \tau_p G_c \, \exp(-t/\tau_p) \, . \tag{25.6}$$

To determine the conductivity of the device it is necessary to find the electron concentration n or, equivalently, the difference $\Delta n = n - n_{n0}$. This is easily accomplished by observing that spatial uniformity holds also during the transient; as a consequence, the condition of local charge neutrality (25.2) applies. At the end of the transient, the latter becomes $n_{n0} = p_{n0} + N_D$ which, subtracted from (25.2), yields $\Delta n = \Delta p$ at all times. The semiconductor conductivity (19.134) can then be written

$$\sigma = \sigma_0 + \sigma_p \, \exp(-t/\tau_p) \, , \tag{25.7}$$

with $\sigma_0 = q \, (\mu_n \, n_{n0} + \mu_p \, p_{n0})$ and $\sigma_p = q \, (\mu_n + \mu_p) \, \tau_p \, G_c$. In the one-dimensional, uniform case considered here, the current is the product of the current density times the cross-section A of the device, and the electric field is the ratio of the applied voltage to the length L (Fig. 25.1). In conclusion, for $t > 0$ the current $I = V/R = V \, \sigma \, A/L$ is given by

$$I = I_0 + I_p \, \exp(-t/\tau_p) \, , \tag{25.8}$$

with $I_0 = \sigma_0 \, (A/L) \, V$, $I_p = \sigma_p \, (A/L) \, V$, namely, the current's decay time is the minority-carrier lifetime. The quantity to be measured is the current through the device which, for $t < 0$, is equal to the constant $I_0 + I_p$, while it decreases exponentially towards I_0 for $t > 0$ (Fig. 25.2). The measurement of τ_p is easily accomplished by observing that the tangent to the exponential branch drawn at the

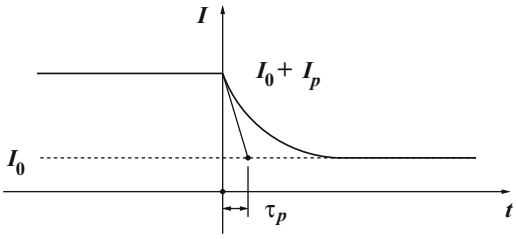


Fig. 25.2 Time dependence of the current flowing in the sample of Fig. 25.1 when the minority-carrier lifetime (τ_p in this case) is measured

origin intercepts[2] the asymptotic value I_0 at $t = \tau_p$. Note that the hypothesis of a monochromatic illumination is not essential; in fact, the analysis still holds if G_c in (25.3) is replaced with the integral of G_c over the frequencies.

25.2.1 Thermal Velocity and Capture Cross-Section

In an n-doped material the hole lifetime is related to the hole-transition coefficient α_p and to the trap concentration N_t by $\tau_p = \tau_{p0} = 1/(\alpha_p N_t)$ (Sect. 20.2.3). It follows that measuring τ_p after deliberately introducing a known concentration N_t of traps into the semiconductor provides the value of the transition coefficient α_p. The latter is often recast in a different form by defining the carrier *thermal velocity* u_{th} with the relation

$$\frac{1}{2}\, m^*\, u_{\text{th}}^2 = \frac{3}{2}\, k_B\, T\,, \tag{25.9}$$

where $m^* = m_e$ for electrons and $m^* = m_h$ for holes, with m_e, m_h the average effective masses. Using the data of Table 18.1 one finds, for silicon, $m_e \simeq 2.98\,10^{-31}$ kg and $m_h \simeq 3.20\,10^{-31}$ kg. In turn, the *capture cross-sections* of the traps, for electrons and holes respectively, are defined by

$$\sigma_e = \frac{\alpha_n}{u_{\text{th},e}}\,, \qquad \sigma_h = \frac{\alpha_p}{u_{\text{th},h}}\,. \tag{25.10}$$

From the above definitions it follows that in an n-doped material

$$\tau_p = \tau_{p0} = \frac{1}{\alpha_p\, N_t} = \frac{1}{\sigma_h\, u_{\text{th},h}\, N_t}\,. \tag{25.11}$$

In particular, in silicon at $T_L = 300$ K it is from (25.9) $u_{\text{th},e} \simeq u_{\text{th},h} \sim 2{\times}10^7$ cm s^{-1}. The measure of α_p thus provides for the cross-section the value $\sigma_h \simeq 5\,10^{-15}$ cm^2.

[2] See also Prob. 21.13.

A qualitative picture of the cross-section as a circle centered at the trap yields the definition of a radius r_h such that $\sigma_h = \pi\, r_h^2$. It turns out $r_h \simeq 4\,10^{-8}$ cm, namely, r_h is of the order of the atomic radius.[3] The measure of τ_n, α_n, and σ_e is carried out in a similar way, starting from a p-doped material.

25.3 Mobility Measurement—Haynes-Shockley Experiment

A measurement method for mobility is illustrated with the aid of Fig. 25.3, showing a uniformly doped layer of semiconductor to which a constant positive voltage V is applied; this produces an electric field in the x direction, $\mathbf{E} = E\,\mathbf{i}_1$. The method is able to measure the minority-carrier mobility; in the example shown in the figure, which refers to an n-doped material, the outcome is μ_p.

Holes are generated at some position x by, e.g., illuminating the material with a laser pulse. The electric field makes the holes to drift to the right, where they are eventually collected after crossing a distance Δx_1 in the direction parallel to $\mathbf{E}$. The lower part of Fig. 25.3 shows three profiles of the hole distribution at successive instants of time, from left to right.[4] The leftmost, thin profile corresponds to the

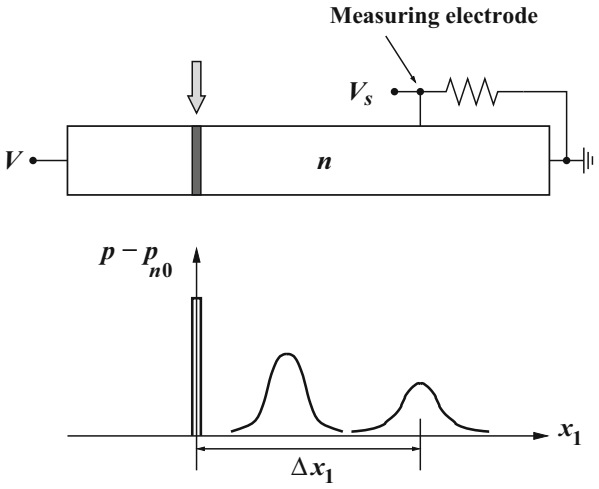


Fig. 25.3 Measurement scheme for mobility (Haynes-Shockley experiment)

[3]The fact that the size of the capture cross-section is sometimes similar, like in the present case, to the size of an atom may explain the popularity of expressing the capture phenomenon in terms of cross-sections. This, however, oversimplifies the problem: capture cross-sections, in fact, range from 10^{-25} to 10^{-12} cm^2 [12, Sect. 4.1.2].

[4]The generated electrons drift to the left and are absorbed by the left contact.

instant of the laser pulse; the other two profiles are shifted to the right because of the action of the field. The hole distribution becomes progressively wider and shorter because of diffusion; also, its area decreases with time due to recombination. When the profile crosses the section corresponding to the measuring electrode, the largest value of the measured voltage V_S corresponds to the profile's peak. This allows one to measure the time Δt necessary for the profile's peak to cover the distance from the section where the laser pulse is applied to that of the measuring electrode. Then, the average velocity of the peak is found from $\Delta x_1 / \Delta t$.

The analysis of the experiment is carried out assuming that the perturbation is sufficiently small, so that a weak-injection condition holds. The continuity equation for the minority carriers reads

$$\frac{\partial p}{\partial t} + \frac{p - p_{n0}}{\tau_p} + \operatorname{div}\left(\mu_p\, p\, \mathbf{E} - D_p\, \operatorname{grad} p\right) = 0\,, \tag{25.12}$$

with μ_p, D_p = const due to spatial uniformity. Remembering that $\operatorname{div}\mathbf{D} = \varrho$, in (25.12) it is

$$\operatorname{div}(p\,\mathbf{E}) = \mathbf{E}\cdot\operatorname{grad} p + p\,\frac{\varrho}{\varepsilon_{\mathrm{sc}}} \simeq \mathbf{E}\cdot\operatorname{grad} p\,, \tag{25.13}$$

on account of the fact that due to the weak-injection condition, the perturbation with respect to the local charge neutrality is small, whereas $\operatorname{grad} p$ is large. Using the auxiliary function $f = (p - p_{n0})\,\exp(t/\tau_p)$ transforms (25.12) into

$$\frac{\partial f}{\partial t} - D_p\,\nabla^2 f + \mu_p\,\mathbf{E}\cdot\operatorname{grad} f = 0\,. \tag{25.14}$$

The above equation is further simplified by applying a suitable change of the variables,

$$\mathbf{r}^\star = \mathbf{r}^\star(\mathbf{r}, t) = \mathbf{r} - \mathbf{v}\,t\,, \qquad t^\star = t^\star(\mathbf{r}, t) = t\,, \tag{25.15}$$

where $\mathbf{v}$ is a constant velocity, yet undefined. The relations between the spatial derivatives with respect to the old and new variables read

$$\frac{\partial f}{\partial x_i} = \sum_{j=1}^{3} \frac{\partial f}{\partial x_j^\star}\,\frac{\partial x_j^\star}{\partial x_i} = \frac{\partial f}{\partial x_i^\star}\,, \tag{25.16}$$

so that, using the star to indicate the operators acting on the new spatial variables,

$$\operatorname{grad}_\star f = \operatorname{grad} f\,, \qquad \nabla_\star^2 f = \nabla^2 f\,. \tag{25.17}$$

The time derivatives are treated in the same manner to find

$$\frac{\partial f}{\partial t} = \frac{\partial f}{\partial t^\star}\,\frac{\partial t^\star}{\partial t} + \sum_{j=1}^{3} \frac{\partial f}{\partial x_j^\star}\,\frac{\partial x_j^\star}{\partial t} = \frac{\partial f}{\partial t^\star} - \mathbf{v}\cdot \mathrm{grad}_\star f\,. \tag{25.18}$$

Replacing (25.17), (25.18) into (25.14) yields

$$\frac{\partial f}{\partial t^\star} = D_p\,\nabla_\star^2 f + \left(\mathbf{v} - \mu_p\,\mathbf{E}\right)\cdot \mathrm{grad}_\star f\,. \tag{25.19}$$

Exploiting the arbitrariness of $\mathbf{v}$ one lets $\mathbf{v} = \mu_p\,\mathbf{E}$, so that (25.19) simplifies to a diffusion equation, $\partial f/\partial t^\star = D_p\,\nabla_\star^2 f$. The solution of the latter is given by (23.28), namely,

$$f(\mathbf{r}^\star, t^\star) = \iiint_{-\infty}^{+\infty} f(\mathbf{s}, 0)\,\Delta(\mathbf{r}^\star - \mathbf{s}, t^\star)\;\mathrm{d}^3\xi\,, \tag{25.20}$$

with

$$\Delta(\mathbf{r}^\star - \mathbf{s}, t^\star) = \frac{1}{\left(4\,\pi\,D_p\,t^\star\right)^{3/2}}\,\exp\left(-\frac{|\mathbf{r}^\star - \mathbf{s}|^2}{4\,D_p\,t^\star}\right)\,, \tag{25.21}$$

and $f(\mathbf{s}, 0) = p(\mathbf{s}, 0) - p_{n0}$. Using again the old variables yields

$$p = p_{n0} + \exp(-t/\tau_p)\iiint_{-\infty}^{+\infty} f(\mathbf{s}, 0)\,\Delta(\mathbf{r} - \mathbf{v}\,t - \mathbf{s}, t)\;\mathrm{d}^3\xi\,. \tag{25.22}$$

The input pulse can be selected in such a way that $f(\mathbf{s}, 0) \approx c\,\delta(\mathbf{s})$, where c is a dimensionless constant. In conclusion,

$$p = p_{n0} + \frac{c\,\exp(-t/\tau_p)}{\left(4\,\pi\,D_p\,t\right)^{3/2}}\,\exp\left[-\frac{(x_1 - \mu_p\,E\,t)^2 + x_2^2 + x_3^2}{4\,D_p\,t}\right]\,, \tag{25.23}$$

which, apart from the additive constant p_{n0}, is a Gaussian whose amplitude decreases in time and whose peak moves along the x_1 direction with the constant velocity $v = \mu_p\,E$. Measuring the time Δt_1 needed for the peak to cross the distance Δx_1 finally yields the mobility

$$\mu_p = \frac{1}{E}\,\frac{\Delta x_1}{\Delta t_1}\,. \tag{25.24}$$

25.4 Hall-Voltage Measurement

The measurements based on the *Hall effect* are a powerful investigation tool that exploits the combined action of the electric and magnetic field. The Hall effect is the production of a voltage drop, transverse to the direction of the electric current, due to the application of a magnetic field. The qualitative features of the method are explained with the aid of Fig. 25.4. Consider a uniformly doped, prismatic block of semiconductor. The block is slender, and a constant voltage V is applied to it to produce an electric field $\mathbf{E}$ aligned with the longitudinal direction. As a consequence, one can assume that the flow lines of the current density be parallel to $\mathbf{E}$; due to spatial uniformity, such a current density is essentially due to the drift of majority carriers. At the same time, a constant magnetic-induction field $\mathbf{B}$ is applied, normal to one of the lateral faces. The upper part of the figure refers to an n-doped semiconductor; there, the majority carriers are electrons, whose average velocity is oriented opposite to the field; it follows that the Lorentz force $\mathbf{F} = \varrho^- \mathbf{v}^- \wedge \mathbf{B}$ (Sect. 4.11) is oriented as shown in the figure. The negative indices in the expression of the Lorentz force remind one that the charge density and average velocity are those of negative charges. The mobile electrons are pushed by the Lorentz force towards the lower face of the device, where they form a negative charge layer. The flow lines of the current density are still parallel to the longitudinal direction; however, their density is not uniform anymore. Due to the global charge neutrality, the negative charge layer is compensated by a positive charge layer that forms at the upper face. The two opposite layers, schematically indicated in the diagram in the upper-right part of Fig. 25.4, produce an electric field normal to the upper and lower faces; as a result, a measurable voltage drop (*Hall voltage*) between the two faces comes into existence: for the example in hand, the voltage of the upper face is larger than that of the lower face.

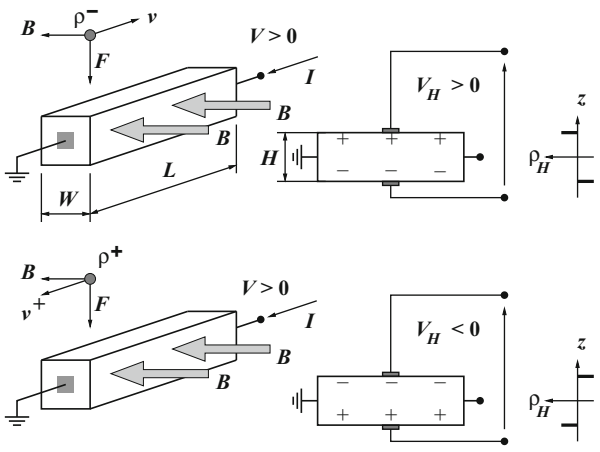


Fig. 25.4 Scheme of a Hall-voltage measurement

In a p-doped semiconductor (lower part of the figure), the majority carriers are holes, whose average velocity is oriented in the direction of the field; the Lorentz force $\mathbf{F} = \varrho^+ \mathbf{v}^+ \wedge \mathbf{B}$ is oriented as in the previous case, because both charge density and average velocity change sign with respect to the n-doped semiconductor. The consequence is that the mobile holes are pushed towards the lower face of the device, where they form a positive charge layer. In conclusion, the sign of the Hall voltage is opposite with respect to the case of the n-doped semiconductor.

The analysis of the experiment is based on the drift-diffusion equations incorporating the magnetic terms, (19.107), (19.123); the diffusion terms are neglected, whence

$$\mathbf{J}_n = q\,\mu_n\,n\,\mathbf{E} - q\,a_n\,\mu_n^2 n\,\mathbf{E}\wedge\mathbf{B}\,, \qquad \mathbf{J}_p = q\,\mu_p\,p\,\mathbf{E} + q\,a_p\,\mu_p^2\,p\,\mathbf{E}\wedge\mathbf{B}\,. \tag{25.25}$$

The total current density $\mathbf{J} = \mathbf{J}_n + \mathbf{J}_p$ then reads

$$\mathbf{J} = \sigma\,\mathbf{E} + r\,\sigma^2 \mathbf{E}\wedge\mathbf{B}\,, \tag{25.26}$$

where $\sigma = q\,\mu_p\,p + q\,\mu_n\,n$ is the electric conductivity and

$$r = \frac{q}{\sigma^2}\,\left(a_p\,\mu_p^2\,p - a_n\,\mu_n^2\,n\right) = \frac{a_p\,\mu_p^2\,p - a_n\,\mu_n^2\,n}{q\,\left(\mu_p\,p + \mu_n\,n\right)^2} \tag{25.27}$$

is the *Hall coefficient*. The two quantities σ and r can be measured independently as shown below; while σ is positive definite, r has a sign. In particular, the following limiting cases hold: for the p-type dopant it is $p \gg n$, whence $\sigma \simeq q\,\mu_p\,p$ and $r \simeq a_p/(q\,p) > 0$; thus,

$$p = \frac{a_p}{q\,r}\,, \qquad \mu_p = \frac{r}{a_p}\,\sigma \qquad (p \gg n)\,. \tag{25.28}$$

Similarly, for the n-type dopant it is $n \gg p$, whence $\sigma \simeq q\,\mu_n\,n$ and $r \simeq -a_n/(q\,n) < 0$; thus,

$$n = -\frac{a_n}{q\,r}\,, \qquad \mu_n = -\frac{r}{a_n}\,\sigma \qquad (n \gg p)\,. \tag{25.29}$$

From (25.28) and (25.29) it follows that the concentration and mobility of the majority carriers can be determined independently, provided σ and r are known. In turn, the measurement of σ and r is easily carried out by applying (25.26) to the prismatic sample of Fig. 25.4. Let $\mathbf{i}_L$ be the unit vector of the longitudinal direction, and $\mathbf{i}_W$ the unit vector parallel to $\mathbf{B}$, so that $\mathbf{B} = B\,\mathbf{i}_W$. Observing that $E_W = 0$, $J_W = 0$, and $\mathbf{E}\wedge\mathbf{B} = E_L\,B\,\mathbf{i}_H - E_H\,B\,\mathbf{i}_L$, it follows

$$J_L = \sigma\,E_L - r\,\sigma^2\,E_H\,B\,, \qquad J_H = \sigma\,E_H + r\,\sigma^2\,E_L\,B\,, \tag{25.30}$$

with $J_H = 0$. In turn, B is small enough to make the following approximations possible:

$$J = J_L \simeq \sigma E_L\,, \qquad E_H = -r\,\sigma\,E_L\,B \simeq -r\,J\,B\,. \tag{25.31}$$

On the other hand, it is $E_L \simeq V_L/L$, $E_H \simeq V_H/H$, and $J = I/(W\,H)$, where the block's length L, height H, and width W are indicated in Fig. 25.4; thus, $V_H = -r\,B\,I/W$ and $I/(W\,H) = \sigma\,V_L/L$. In conclusion,

$$\sigma = \frac{L\,I}{W\,H\,V_L}\,, \qquad r = -\frac{W\,V_H}{B\,I}\,, \tag{25.32}$$

namely, the two parameters are obtained by combining the Hall voltage with other known physical and geometrical parameters. Typical applications of the measurement scheme shown in this section are the measurements of concentrations and mobilities as functions of temperature and dopant concentration.

25.5 Hall Voltage—Arbitrary Aspect Ratio

The measurement technique based on the Hall effect, outlined in Sect. 25.4, is based on the hypothesis that the sample under investigation is slender, so that the flow lines of the current density are aligned with the longitudinal direction, whose unit vector is $\mathbf{i}_L$, and the contacts where the Hall voltage is measured are far away from the contacts onto which the external perturbation is applied. A typical arrangement is that shown in Fig. 25.4, where the contacts used for measuring the Hall voltage are in the mid section of the sample. In more realistic situations the device of interest exhibits geometrical features that do not allow for the approximations listed above. As a typical example, consider the device shown in Fig. 25.5, which is used for measuring the conductivity and the Hall voltage. The former quantity is measured by the four-contact, van der Pauw method outlined in Sect. 25.7: for this, the contacts labeled c_1, c_2, c_3, c_4, also termed *Hall probes*, are used, while the two lateral contacts k_1, k_2 are left open. In turn, the Hall voltage V_H is measured at a pair of Hall probes, e.g., c_1 and c_3, after applying the external perturbation to the lateral contacts and a magnetic induction $\mathbf{B}$ normal to the x, y plane. Here it is assumed for simplicity that the external perturbation is a constant current I forced into the device.

To decrease the relative error in the conductivity measurement, for a given length L of the device it is preferable to place the probes near the contacts, like in Fig. 25.5; on the other hand, considering that the contacts are equipotential, to measure the Hall voltage it would be preferable to place the probes far away from the contacts, to avoid a possible perturbation due to the proximity of the equipotential regions created by the contacts themselves. Clearly, the latter constraint is incompatible with the former. Even if a third pair of Hall probes was placed at the mid section

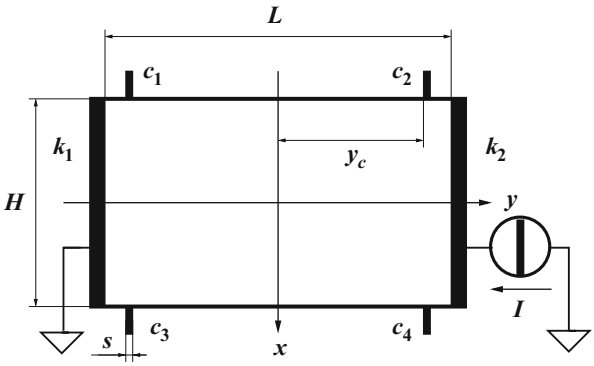


Fig. 25.5 Scheme for the combined conductivity and Hall-voltage measurements

of the sample, when the latter is not sufficiently elongated the ratio H/L influences the distribution of the current density, again due to the presence of the equipotential regions.

In the experimental setup for the measurement of conductivity and Hall voltage, the thickness h of the active region (i.e., in the z direction normal to the x, y plane of Fig. 25.5) is typically very small, and the dopant distribution is constant. At the same time, the carrier concentration in the active region is high enough to make the diffusive part of the current density negligible. By way of example, one may use a resistor fabricated within an epitaxial layer, or a MOSFET whose gate bias is sufficiently high to produce a strong-inversion regime in the channel. Considering that there are only two components of the current density and one component of the magnetic induction, $B = B_z$, one applies directly (19.95) so that, for the valley of index a, and assuming that the carriers are electrons, one finds

$$J_{1a} = J'_{1a} - \mu_{1a} B J_{2a} , \qquad J_{2a} = J'_{2a} + \mu_{2a} B J_{1a} , \tag{25.33}$$

with $J'_{1a} = q \mu_{1a} n_a E_1$ and $J'_{2a} = q \mu_{2a} n_a E_2$ due to the neglect of the diffusive component.[5] When adding up (25.33) over the valleys, the result depends on which, between J_{1a} and J_{2a}, is aligned with the x axis. To avoid this complicacy one takes a suitable average of μ_{1a} and μ_{2a}, indicated with μ_H; letting $\theta = \mu_H B$ one finds

$$\sigma E_x \simeq J_x + \theta J_y , \qquad \sigma E_y \simeq J_y - \theta J_x . \tag{25.34}$$

In (25.34) it is $J_x = \sum_a J_{1a}$ and $E_x = E_1$; similarly, it is $J_y = \sum_a J_{2a}$ and $E_y = E_2$; in turn, the conductivity is given by the standard expression $\sigma = q n \sum_a \mu_{1a}/M_C$.

[5] It is useful to remark that in a uniform resistor it is $J'_{3a} = q \mu_{3a} n_a E_3 = 0$, since $E_3 = 0$ due to uniformity; in a MOSFET's channel it is $J'_{3a} = k_B T \mu_{3a} \partial n_a / \partial x_3 + q \mu_{3a} n_a E_3 = 0$ because the drift and diffusion components balance each other. Also, in a MOSFET's channel the components of the current density parallel to the x, y plane depend also on z; however, one can neglect such a dependence in (25.33) by considering J_{1a} and J_{2a} as averages over h [115].

Note that thanks to the approximation leading to (25.34) it is not necessary to assume that B is small (compare with (19.98) and (19.99)).

The two-dimensional current field (25.33) is expressible as $\mathbf{J} = \mathrm{rot}(\psi\,\mathbf{k})$, where $\mathbf{k}$ is the unit vector of the z axis and ψ a function of x, y only, which is sometimes called *stream function*. The latter fulfills the Laplace equation (Prob. 25.1),

$$\nabla^2\psi = 0\,, \tag{25.35}$$

supplemented with suitable boundary conditions (Prob. 25.2). Once the current density is found, the components of the electric field are determined from (25.33); then, from $\mathbf{E} = -\,\mathrm{grad}\,\varphi$, the voltage drop between any two points of the domain, boundary included, is found from the negative path integral of $\mathbf{E}$ along any line connecting the two points and belonging to the domain.

25.5.1 Solution of the Stream-Function Equation

The solution of (25.35) is sought over the domain D of Fig. 25.5, namely, the rectangle $-H/2 \le x \le H/2$, $-L/2 \le y \le L/2$. The device has two contacts connected with a current generator, whose current is I; they are placed at the two ends $y = \pm L/2$ and have a common width H. The voltage measurements at the Hall probes are performed using a zero-admittance equipment, hence no current is absorbed there. However, the Hall probes may still affect the boundary condition of (25.35), because the portion of the boundary onto which a probe is applied becomes equipotential. Such an effect obviously disappears in the ideal case of point-like probes, and may still be considered negligible if the common width s of the probes is much smaller than L [8, 90]. As such a constraint is easily fulfilled by the present technology, the effect of the probes' size is neglected here. From Prob. 25.2, the boundary conditions of (25.35) are found to be[6]

$$\psi(-H/2, y) = \psi_0\,, \qquad \left[\frac{\partial\psi(x,y)}{\partial y}\right]_{y=-L/2} = -\frac{\theta\, I}{h\, H} \tag{25.36}$$

and

$$\psi(+H/2, y) = \psi_1\,, \qquad \left[\frac{\partial\psi(x,y)}{\partial y}\right]_{y=+L/2} = -\frac{\theta\, I}{h\, H}\,, \tag{25.37}$$

[6] With reference to Prob. 25.2, the second condition in (25.36) is derived by observing that along the side $y = -L/2$ the unit vector $\mathbf{t}$ is aligned with the side itself and points in the positive direction of x. As a consequence, $\mathbf{n}$ points in the positive direction of y, namely, $\mathbf{n} = \mathbf{j}$ and $\partial\psi/\partial y = \partial\psi/\partial n$. By the same token, the sign in the second condition of (25.37) is justified by observing that the currents at the contacts sum to zero, and by using the relation $\partial\psi/\partial y = -\partial\psi/\partial n$ that holds on the side $y = L/2$ because $\mathbf{n} = -\mathbf{j}$ there. It is assumed that the current density is uniformly distributed along the contact.

with $\psi_0 - \psi_1 = I/h$. As for the thickness h, it is sufficient to assume that h is small, in order to keep the validity of the two-dimensional analysis carried out here; as shown in Prob. 25.4, the actual value h does not influence the measurement's results. The solution of (25.35) is a harmonic function, fulfilling (25.36) and (25.37), which may be written as $\psi = P + Q$, with $P = (\psi_1 - \psi_0)\, x/H + (\psi_1 + \psi_0)/2$; clearly, P is harmonic and fulfills the same Dirichlet conditions as ψ at $x = \pm H/2$, whereas it fulfills the homogeneous Neumann conditions at $y = \pm L/2$. As a consequence, Q must be harmonic and fulfill the same Neumann conditions as ψ at $y = \pm L/2$, and the homogeneous Dirichlet conditions at $x = \pm H/2$:

$$Q(\pm H/2, y) = 0\,, \qquad \left[\frac{\partial Q(x,y)}{\partial y}\right]_{y=\pm L/2} = -\frac{\theta\, I}{h\, H}\,. \tag{25.38}$$

Condition (25.38) suggests to seek for a function that is even with respect to x and odd with respect to y, e.g., $q_n = \cos(k_n\, x)\, \sinh(k_n\, y)$ with $k_n = (2\, n + 1)\, \pi/H$; in fact, q_n is harmonic and fulfills by construction the homogeneous Dirichlet condition at $x = \pm H/2$. On the contrary, it does not fulfill the nonhomogeneous Neumann condition at $y = \pm L/2$; to achieve this result one forms a linear combination $Q = \sum_{n=0}^{\infty} \lambda_n\, q_n$ and suitably determine the coefficients λ_n. It is sufficient to investigate the case $y = L/2$ because $\partial Q/\partial y$ is even; the condition to be fulfilled then reads

$$-\frac{\theta\, I}{h\, H} = \left[\frac{\partial Q(x,y)}{\partial y}\right]_{y=L/2} = \sum_{n=0}^{\infty} \lambda_n\, k_n\, \cosh(k_n\, L/2)\, \cos(k_n\, x)\,. \tag{25.39}$$

Multiplying both sides of (25.39) by $\cos(k_m\, x)$, integrating from $-H/2$ to $+H/2$, and observing that the integral of $\cos(k_m\, x)\, \cos(k_n\, x)$ equals $(H/2)\, \delta_{mn}$, with δ_{mn} the Kronecker symbol (A.18), shows that the coefficients of the linear combination sought fulfill the relation $(-1)^m\, 4\, \theta\, I/(h\, H\, \lambda_m) = -k_m^2\, H\, \cosh(k_m\, L/2)$; as a consequence,

$$Q = -\frac{4\, \theta\, I}{h} \sum_{n=0}^{\infty} (-1)^n\, \frac{\cos(k_n\, x)}{H^2\, k_n^2}\, \frac{\sinh(k_n\, y)}{\cosh(k_n\, L/2)}\,. \tag{25.40}$$

Both differentiation and integration term-by-term of the series (25.40) are admissible (Prob. 25.3). As far as function P is concerned, one exploits the arbitrariness left in ψ (Prob. 25.2) by letting $\psi_1 = 0$, whence $\psi_0 = I/h$ and

$$P = \frac{I}{2\, h}\left(1 - \frac{2\, x}{H}\right)\,. \tag{25.41}$$

25.5.2 Local Hall Voltage

The analysis carried out in this section allows one to calculate the *local Hall voltage*, that is, the voltage drop $V_H(y) = \varphi(H/2, y) - \varphi(-H/2, y)$ between two points facing each other on the insulating boundaries $x = \pm H/2$ (Fig. 25.5). To proceed, one starts from the first of (25.34) and combines $\psi = P + Q$ with (25.40), (25.41), and with the results of Prob. 25.1, to obtain

$$\sigma E_x = J_x + \theta J_y = \frac{\partial Q}{\partial y} - \theta \left(\frac{\mathrm{d}P}{\mathrm{d}x} + \frac{\partial Q}{\partial x} \right) . \tag{25.42}$$

Using (25.42), the local Hall voltage reads

$$V_H(y) = \int_{+H/2}^{-H/2} E_x(x, y)\, \mathrm{d}x = \frac{\theta}{\sigma} (\psi_1 - \psi_0) + \frac{1}{\sigma} \frac{\mathrm{d}}{\mathrm{d}y} \int_{+H/2}^{-H/2} Q(x, y)\, \mathrm{d}x . \tag{25.43}$$

The term $\partial Q/\partial x$ does not contribute to the integral in (25.43) because Q fulfills the homogeneous Dirichlet boundary conditions at $x = \pm H/2$. Thanks to the uniform convergence (Prob. 25.3), the series (25.40) can be integrated term by term; in this calculation one uses $\int_{-H/2}^{+H/2} \cos(k_n x)\, \mathrm{d}x = 2\,(-1)^n/k_n$ and $\psi_0 - \psi_1 = I/h$, to eventually find

$$V_H(y) = V_{H0} \left[1 - \frac{8}{\pi^2} S(y) \right] , \tag{25.44}$$

where

$$V_{H0} = -\frac{\theta\, I}{\sigma\, h} = -\frac{\mu_H\, B\, I}{\sigma\, h} , \qquad S(y) = \sum_{n=0}^{\infty} \frac{\cosh(k_n y)}{(2\,n + 1)^2 \cosh(k_n\, L/2)} . \tag{25.45}$$

The above result is to be compared with that obtained in Sect. 25.4, where the standard Hall-voltage theory yielded $V_H = -r\, B\, I/W$; introducing into the latter expression the second relation of (25.29), that applies to the case where the electron current is dominant, shows that expression of the Hall voltage of the standard theory is made to coincide[7] with V_{H0} by letting $W = h$ and $a_n\, \mu_n = \mu_H$. For $y = 0$ the series in (25.45) attains the minimum, $S(0) = \sum_{n=0}^{\infty} [(2n + 1)^2 \cosh(k_n\, L/2)]^{-1}$, whose leading term is $1/\cosh[\pi\, L/(2\, H)]$; this shows that the contribution of the series at $y = 0$ is small when $H \ll L$. In this limiting case the result of the standard theory is recovered; in the other cases, instead, the measured Hall

[7] The difference in sign is related to the orientation of **B** with respect to the z axis. In the analysis of Sect. 25.5, such an orientation is not specified.

voltage differs significantly from V_{H0} [115]. Also, remembering from Prob. 25.3 that $\sum_{n=0}^{\infty}(2n+1)^{-2} = \pi^2/8$, it also follows $V_H(\pm L/2) = 0$, which is consistent with the condition $E_t = 0$ imposed by the contact.

25.6 Measurement of Doping Profiles

The calculation of the depletion capacitance for an arbitrary doping profile, in the one-dimensional case, has been carried out in Sect. 21.9.3; the analysis has yielded, among others, the two relations (21.124) and (21.125) that are reported below:

$$\mathrm{d}Q = \varrho(b)\ \mathrm{d}b\,, \qquad C = \frac{\mathrm{d}Q}{\mathrm{d}\psi} = \frac{\varepsilon_{\mathrm{sc}}}{b-a}\,, \tag{25.46}$$

where C is the differential capacitance per unit area of the space-charge region, whose boundaries are a and b, and $\psi = \varphi(b) - \varphi(a)$. For a strongly asymmetric junction it is, e.g., $b - a \simeq b$, and (25.46) become

$$C \simeq \frac{\varepsilon_{\mathrm{sc}}}{b}\,, \qquad \varepsilon_{\mathrm{sc}}\ \mathrm{d}\psi \simeq b\,\varrho(b)\ \mathrm{d}b = \frac{1}{2}\,\varrho(b)\ \mathrm{d}b^2 = \frac{1}{2}\,\varrho(b)\ \mathrm{d}\left(\frac{\varepsilon_{\mathrm{sc}}^2}{C^2}\right)\,. \tag{25.47}$$

As a consequence,

$$\varrho(b) = \frac{2}{\varepsilon_{\mathrm{sc}}}\left[\frac{\mathrm{d}\left(1/C^2\right)}{\mathrm{d}\psi}\right]^{-1}\,. \tag{25.48}$$

Basing upon (25.48), a measurement scheme can be devised, which proceeds as follows:

1. C is measured at a given bias V, and b is determined from $b = \varepsilon_{\mathrm{sc}}/C$.
2. C is measured again after slightly varying the bias from V to $V + \mathrm{d}V = V + \mathrm{d}\psi$.
3. A numerical calculation of the derivative yields $\varrho(b)$.

An example of application of the above scheme is given with reference to an asymmetric p-n junction with, e.g., $a = -l_p$, $b = l_n$, and $l_n \gg l_p$. In the reverse-bias condition, and using the full-depletion approximation (20.34), it is $\varrho \simeq q\,N$, whence

$$N(l_n) \simeq \frac{2}{q\,\varepsilon_{\mathrm{sc}}}\left[\frac{\mathrm{d}\left(1/C^2\right)}{\mathrm{d}\psi}\right]^{-1}\,. \tag{25.49}$$

This provides a method for measuring the dopant distribution.

25.7 Van der Pauw Method

The theory of the Green function in two dimensions, outlined in Sect. 4.12.4, lends itself to an interesting application to a technique for measuring the conductivity and Hall coefficient of a sample. The method, named after van der Pauw [139], is outlined here with reference to the conductivity measurement. The sample to be measured is approximately two dimensional (that is, its thickness is uniform and much smaller that the length and width) and simply connected. Apart from these constraints, its form is arbitrary. Four point-like contacts are placed at arbitrary points along the sample's periphery; two of them are connected with a current generator, whereas the other two contacts are used to measure the voltage drop produced by the current.

25.7.1 Solution over the Upper Half Plane

To begin with, one considers the case of a thin layer of material of uniform conductivity σ, extending over the upper half plane; due to uniformity, the current flow is described by Ohm's law $\mathbf{J} = \sigma\,\mathbf{E}$, with $\mathbf{E}$ a two-dimensional field. Charge neutrality is supposed to hold, so that the electric potential φ is found by solving the Laplace equation $\nabla^2\varphi = 0$. As for the boundary conditions, it is assumed that a current I_{ab} is forced by a generator through two contacts lying along the x axis and centered at $x = a$ and $x = b$, respectively. The current enters the layer at the contact centered at a and leaves it at the contact centered at b. The size s of both contacts is provisionally left undetermined. Let $J = k/s$ be the component of $\mathbf{J}$ normal to the x axis at the contact centered at a, with $k = \text{const} > 0$. Letting h be the thickness of the layer in the direction normal to the x, y plane, it is

$$I_{ab} = h \int_{a-s/2}^{a+s/2} J \, \mathrm{d}x = h\,k > 0\,, \tag{25.50}$$

independent of the contact's size. The normal component of the electric field at the contact centered at a turns out to be $E_a = J/\sigma = \eta/s$, with $\eta = k/\sigma = I_{ab}/(h\,\sigma) > 0$, while that at the contact centered at b reads $E_b = -J/\sigma = -\eta/s$. The remaining parts of the x axis are electrically isolated. Thus, the normal component of the current density and, consequently, that of the electric field are zero there. Due to the form of E it also follows

$$\int_{-\infty}^{+\infty} E(\xi)\,\mathrm{d}\xi = 0\,. \tag{25.51}$$

The assumptions above prescribe the normal derivative of φ everywhere along the x axis, and make it possible to determine φ in the upper half plane $y > 0$. It is

also possible to determine φ along the x axis away from the contacts. For this, with no loss of generality one may assume $a < b$. Letting $y = 0$ in (4.67), one splits the integral over the subintervals of the x axis; in so doing, one finds that only the integrals extended over the contacts are different from zero. The other parts of the x axis, in fact, do not contribute due to the vanishing normal component of E. Using $E_a = \eta/s$ and $E_b = -\eta/s$ yields

$$\varphi(x,0) = \varphi_0 - \frac{\eta/s}{2\,\pi}\left[\int_{a-s/2}^{a+s/2} G(x,0;\xi)\,\mathrm{d}\xi - \int_{b-s/2}^{b+s/2} G(x,0;\xi)\,\mathrm{d}\xi\right]. \tag{25.52}$$

Replacing into the first integral the expressions (4.66) of G, with $y = 0$, provides

$$\frac{\eta/s}{2\,\pi}\int_{a-s/2}^{a+s/2} 2\,\log|\xi - x|\,\mathrm{d}\xi \;\;\rightarrow\;\; \frac{\eta}{\pi}\,\log|a-x|\,, \tag{25.53}$$

where the limiting case holds for $s \to 0$. Combining with a similar result for the contact centered at b yields

$$\varphi(x,0) = \varphi_0 + \frac{\eta}{\pi}\,\log\frac{|b-x|}{|a-x|}\,. \tag{25.54}$$

The above result holds away from the contacts, namely, for $x \neq a$ and $x \neq b$. Selecting two more points $c > b$ and $d > c$ and using $\eta = I_{ab}/(h\,\sigma)$, the voltage difference $V_{dc} = \varphi(d,0) - \varphi(c,0)$ is determined from (25.54) and reads

$$V_{dc} = \frac{I_{ab}}{\pi\,h\,\sigma}\,\log\frac{(d-b)\,(c-a)}{(d-a)\,(c-b)}\,, \tag{25.55}$$

where all quantities in parentheses are positive. In (25.55) it is assumed that two point-like contacts are placed at $x = c$, $x = d$, and the voltage difference V_{dc} produced by I_{ab} is measured. Observing that the positions a, b, c, d of the contacts are prescribed, the only unknown left in (25.55) is the conductivity σ. Also, introducing the positive quantities $\lambda = b - a$, $\mu = c - b$, and $\nu = d - c$, transforms (25.55) into

$$\exp(-\varrho_{ab}^{dc}\,\sigma) = \frac{(\lambda+\mu+\nu)\,\mu}{(\lambda+\mu)\,(\mu+\nu)} < 1\,, \tag{25.56}$$

where $\varrho_{ab}^{dc} = \pi\,h\,V_{dc}/I_{ab}$ has the units of a resistivity. Due to the inequality in (25.56) it is found that $\varrho_{ab}^{dc} > 0$. A suitable change in the positions of the current generator and voltmeter provides

$$V_{ad} = \frac{I_{bc}}{\pi\,h\,\sigma}\,\log\frac{(c-a)\,(d-b)}{(b-a)\,(d-c)}\,, \tag{25.57}$$

corresponding to a cyclic permutation of indices, $a \to b \to c \to d \to a$. Letting $\varrho_{bc}^{ad} = \pi\, h\, V_{ad}/I_{bc}$ transforms (25.57) into

$$\exp(-\varrho_{bc}^{ad}\,\sigma) = \frac{\lambda\,\nu}{(\lambda+\mu)\,(\mu+\nu)} < 1\,, \tag{25.58}$$

which shows that ϱ_{bc}^{ad} is positive as well. Adding up (25.56) and (25.58) yields

$$\exp(-\varrho_{ab}^{dc}\,\sigma) + \exp(-\varrho_{bc}^{ad}\,\sigma) = 1\,. \tag{25.59}$$

The above shows that suitable measurements of the ratios ϱ_{ab}^{dc} and ϱ_{bc}^{ad} provide a relation for σ, in which the relative distances of the four contacts where the current generator or voltmeter are applied do not explicitly appear. In fact, the power of the method lies in its ability to accurately measure the conductivity of the sample by making the procedure independent of the positions of the contacts.

25.7.2 Solution over the Unit Circle

Another fundamental property of (25.59) is that it does not depend on the sample's geometry. Consider by way of example the conformal-mapping transformation $x, y \to u, v$, shown in Sect. B.7, that brings the upper half plane $y > 0$ onto the disk $u^2 + v^2 < 1$. As a consequence of the transformation, the solution of the Laplace equation with Neumann boundary conditions, given by (4.67), transforms into (B.65), (B.66), where the polar coordinates $u = \zeta \cos\vartheta$, $v = \zeta \sin\vartheta$ are used for convenience. The same boundary conditions used in the case of the upper half plane will be studied here. In particular, in that problem the component of the electric field normal to the x axis at the contact of size s centered at a was $E_a = \eta/s$. Similarly, the normal component at the contact of the same size centered at b was $E_b = -\eta/s$. The remaining part of the x axis was supposed to be electrically isolated. When the variable transformations $x, y \to u, v$ and $u, v \to \zeta, \vartheta$ are carried out, the boundary is specified by letting $y \to 0$, $u^2 + v^2 \to 1$, and $\zeta \to 1$, respectively; as a consequence, the first integral in (25.52) transforms as

$$\int_{a-s/2}^{a+s/2} \frac{\eta}{s}\, G(x, 0; \xi)\, \mathrm{d}\xi = \frac{\eta}{s} \int_{\gamma(a-s/2)}^{\gamma(a+s/2)} \frac{G'(1, \vartheta; \gamma)}{1 - \cos\gamma}\, \mathrm{d}\gamma\,, \tag{25.60}$$

with $\gamma(a \pm s/2) = -2 \arctan[1/(a \pm s/2)]$. Letting $\gamma_a = \gamma(a)$, the integral at the right-hand side of (25.60) tends to $\eta\, G'(1, \vartheta; \gamma_a)$ for $s \to 0$, as is easily found by observing that

$$\lim_{s\to 0} \frac{\gamma(a + s/2) - \gamma(a - s/2)}{s\,(1 - \cos\gamma_a)} = 1\,. \tag{25.61}$$

From the form of E it also follows $\int_0^{2\pi} E'(\gamma)\,\mathrm{d}\gamma = 0$. The angles ϑ and γ_a determine over the circumference C two points of coordinates $(\cos\vartheta, \sin\vartheta)$ and $(\cos\gamma_a, \sin\gamma_a)$, respectively; from the same procedure as that leading to (25.53) it follows that $\eta\, G'(1, \vartheta; \gamma_a) = \eta\, \log\left\{4\sin^2[(\vartheta - \gamma_a)/2]\right\}$. Combining the above with a similar result obtained at the other contact yields, due to (B.65),

$$\psi(1, \vartheta) = \varphi_0 + \frac{\eta}{\pi}\, \log \frac{|\sin[(\vartheta - \gamma_b)/2]|}{|\sin[(\vartheta - \gamma_a)/2]|}\,. \tag{25.62}$$

Using $\eta = I_{ab}/(h\,\sigma)$, the difference $V_{dc} = \psi(1, \vartheta = \gamma_d) - \psi(1, \vartheta = \gamma_c)$ then reads

$$V_{dc} = \frac{I_{ab}}{\pi\, h\, \sigma}\, \log \frac{\sin[(\gamma_d - \gamma_b)/2]\ \sin[(\gamma_c - \gamma_a)/2]}{\sin[(\gamma_d - \gamma_a)/2]\ \sin[(\gamma_c - \gamma_b)/2]}\,. \tag{25.63}$$

The inequalities $a < b < c < d$ in the x, y domain correspond to $0 < \gamma_a < \gamma_b < \gamma_c < \gamma_d < 2\,\pi$ in the ζ, ϑ domain; it follows $0 < (\gamma_d - \gamma_a)/2 < \gamma_d/2 < \pi$, and similarly for the other arguments of the trigonometric functions that appear in (25.63). As a consequence, such quantities are all strictly positive. Also, letting

$$\lambda = (\gamma_b - \gamma_a)/2\,, \qquad \mu = (\gamma_c - \gamma_b)/2\,, \qquad \nu = (\gamma_d - \gamma_c)/2\,, \tag{25.64}$$

provides for the resistivity $\varrho_{ab}^{dc} = \pi\, h\, V_{dc}/I_{ab}$ the expression

$$\exp(-\varrho_{ab}^{dc}\,\sigma) = \frac{\sin(\lambda + \mu + \nu)\ \sin\mu}{\sin(\lambda + \mu)\ \sin(\mu + \nu)}\,. \tag{25.65}$$

Using the identity $2\,\sin\delta\,\sin\kappa = \cos(\delta - \kappa) - \cos(\delta + \kappa)$, the trigonometric functions in (25.65) can be manipulated to show that $\exp(-\varrho_{ab}^{dc}\,\sigma) < 1$, whence ϱ_{ab}^{dc} is positive. A suitable change in the positions of the current generator and voltmeter provides

$$V_{ad} = \frac{I_{bc}}{\pi\, h\, \sigma}\, \log \frac{\sin[(\gamma_c - \gamma_a)/2]\ \sin[(\gamma_d - \gamma_b)/2]}{\sin[(\gamma_b - \gamma_a)/2]\ \sin[(\gamma_d - \gamma_c)/2]}\,, \tag{25.66}$$

corresponding to a cyclic permutation of indices, $a \to b \to c \to d \to a$. Using $\varrho_{bc}^{ad} = \pi\, h\, V_{ad}/I_{bc}$ and (25.64) transforms (25.66) into

$$\exp(-\varrho_{bc}^{ad}\,\sigma) = \frac{\sin\lambda\ \sin\nu}{\sin(\lambda + \mu)\ \sin(\mu + \nu)} < 1\,, \tag{25.67}$$

which shows that ϱ_{bc}^{ad} is positive as well. Combining (25.65) with (25.67) finally yields

$$\exp(-\varrho_{ab}^{dc}\,\sigma) + \exp(-\varrho_{bc}^{ad}\,\sigma) = 1\,, \tag{25.68}$$

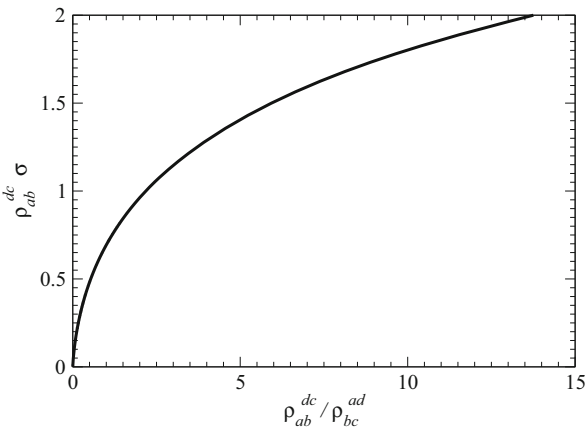


Fig. 25.6 Form of the $\chi = -\omega/\log[1-\exp(-\omega)]$ relation, with $\chi = \varrho_{ab}^{dc}/\varrho_{bc}^{ad}$ and $\omega = \varrho_{ab}^{dc}\,\sigma$, found in the analysis of the van der Pauw method

identical to (25.59). In other terms, despite the fact that the expressions appearing in the intermediate steps, leading either to (25.59) or (25.68), are different, the final expression is the same. New forms of the domain obtained by other conformal-mapping transformations would still lead to the same result. In conclusion, the relation connecting the sample's conductivity with the measurement results is independent not only of the contacts' positions, but also of the form of the sample itself. The latter can then be chosen in such a way as to minimize the experimental errors.

Using the auxiliary parameters $\chi = \varrho_{ab}^{dc}/\varrho_{bc}^{ad} > 0$ and $\omega = \varrho_{ab}^{dc}\,\sigma > 0$ gives (25.59) and (25.68) the form $\chi = -\omega/\log[1-\exp(-\omega)]$, whose inverse is shown in Fig. 25.6.

Problems

25.1 Show that in a simply connected, two-dimensional domain D of arbitrary shape, belonging to the x, y plane, the current field (25.33) is expressible as $\mathbf{J} = \mathrm{rot}(\psi\,\mathbf{k})$, where $\mathbf{k}$ is the unit vector of the z axis and ψ a function of x, y only. Show that ψ is harmonic, namely, $\nabla^2\psi = 0$.

25.2 Considering the domain D used in Prob. 25.1, assume that $n \geq 2$ contacts are placed along the boundary, with $n-1$ of them connected to current generators. Let I_i be the current through the ith contact, and s_i the length of the same contact. Find the boundary conditions to be associated with the solution of the $\nabla^2\psi = 0$ equation discussed in Prob. 25.1.

25.3 Prove that both differentiation and integration term-by-term of the series (25.40) are admissible.

25.4 With reference to the device of Fig. 25.5, calculate the product $\sigma\, h$ from the measurement of the voltage drop V_c between two Hall probes symmetrically placed along the same insulating boundary.

25.5 Starting from (25.45), find a form of the series $S(y)$ that lends itself to a fast calculation.

Appendix A
Vector and Matrix Analysis

A.1 Scalar Product

Consider two complex, n-dimensional column vectors

$$\mathbf{a} = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}, \tag{A.1}$$

whose entries are the components in an n-dimensional Cartesian reference and may depend on position, time, and other parameters. The *scalar product* of the two vectors is indicated with $\mathbf{a} \cdot \mathbf{b}$ and is defined as

$$\mathbf{a} \cdot \mathbf{b} = \sum_{i=1}^{n} a_i^* b_i \,. \tag{A.2}$$

with a_i^* the complex conjugate of a_i. Two nonvanishing vectors $\mathbf{a}$ and $\mathbf{b}$ are *orthogonal* if $\mathbf{a} \cdot \mathbf{b} = 0$. As $\mathbf{b} \cdot \mathbf{a} = (\mathbf{a} \cdot \mathbf{b})^*$, the order of the factors in the scalar product matters; in fact it becomes irrelevant only when the factors are real. The scalar product is distributive and bilinear; if, say, $\mathbf{a} = h_1\,\mathbf{p}_1 + h_2\,\mathbf{p}_2$, then

$$\mathbf{a}\cdot(k_1\,\mathbf{b}_1 + k_2\,\mathbf{b}_2) = h_1^*\,k_1\,\mathbf{p}_1\cdot\mathbf{b}_1 + h_2^*\,k_1\,\mathbf{p}_2\cdot\mathbf{b}_1 + h_1^*\,k_2\,\mathbf{p}_1\cdot\mathbf{b}_2 + h_2^*\,k_2\,\mathbf{p}_2\cdot\mathbf{b}_2 \,, \tag{A.3}$$

where h_1, h_2, k_1, k_2 are complex constants (in (A.3), the product $k_1\,\mathbf{b}_1$ is the vector of components $k_1\,b_{1i}$, and so on). The *modulus* of $\mathbf{a}$ is defined as

$$a = |\mathbf{a}| = \sqrt{\mathbf{a} \cdot \mathbf{a}} = \left(\sum_{i=1}^{n} |a_i|^2 \right)^{1/2} \geq 0 \,. \tag{A.4}$$

M. Rudan, *Physics of Semiconductor Devices*, DOI 10.1007/978-3-319-63154-7

A.2 Schwarz Inequality and Generalizations

Using (A.2, A.3, A.4) one proves the *Schwarz inequality*

$$|\mathbf{a}\cdot\mathbf{b}| \le a\,b\,. \tag{A.5}$$

The above is obvious if $\mathbf{a}=0$ or $\mathbf{b}=0$; let $\mathbf{b}\neq 0$ and define $\mathbf{c}=\mathbf{a}-(\mathbf{a}\cdot\mathbf{b})\,\mathbf{b}/b^2$, whence $\mathbf{c}\cdot\mathbf{b}=0$. It follows

$$a^2 = \left(\mathbf{c}+\frac{\mathbf{a}\cdot\mathbf{b}}{b^2}\,\mathbf{b}\right)\cdot\left(\mathbf{c}+\frac{\mathbf{a}\cdot\mathbf{b}}{b^2}\,\mathbf{b}\right) = c^2+\frac{|\mathbf{a}\cdot\mathbf{b}|^2}{b^2} \ge \frac{|\mathbf{a}\cdot\mathbf{b}|^2}{b^2}\,, \tag{A.6}$$

which is equivalent to (A.5). The strict equality in (A.5) holds if and only if $\mathbf{b}=k\,\mathbf{a}$, with k any complex constant. Observing that $|\mathbf{a}\cdot\mathbf{b}|^2=\Re^2(\mathbf{a}\cdot\mathbf{b})+\Im^2(\mathbf{a}\cdot\mathbf{b})$, from (A.5) one also derives the inequalities $-a\,b\le\Re(\mathbf{a}\cdot\mathbf{b})\le+a\,b$. Thanks to this, one defines the cosine of the angle ϑ between two nonvanishing vectors $\mathbf{a}$ and $\mathbf{b}$ as

$$\cos\vartheta = \frac{\Re(\mathbf{a}\cdot\mathbf{b})}{a\,b}\,. \tag{A.7}$$

Other types of products may be defined besides the scalar product, also involving higher-rank factors: for instance, $n\times n$ matrices of the second rank like

$$\mathbf{A}=\begin{bmatrix} A_{11} & A_{12} & \dots & A_{1n}\\ A_{21} & A_{22} & \dots & A_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ A_{n1} & A_{n2} & \dots & A_{nn}\end{bmatrix}\,,\qquad \mathbf{B}=\begin{bmatrix} B_{11} & B_{12} & \dots & B_{1n}\\ B_{21} & B_{22} & \dots & B_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ B_{n1} & B_{n2} & \dots & B_{nn}\end{bmatrix}\,, \tag{A.8}$$

and so on. Given a second-rank matrix $\mathbf{A}$ of entries A_{ij}, its *transpose* $\mathbf{Q}=\mathbf{A}^T$ is the matrix of entries $Q_{ij}=A_{ji}$. Transposition applies also to vectors: the transpose of the column vector $\mathbf{a}$ defined in (A.1) is the row vector $\mathbf{a}^T=[a_1,\dots,a_n]$. With these premises, given the column vectors $\mathbf{a}$, $\mathbf{b}$ and the matrices $\mathbf{A}$, $\mathbf{B}$, the products $\mathbf{A}\,\mathbf{B}$, $\mathbf{A}\,\mathbf{b}$, and $\mathbf{a}\,\mathbf{b}^T$ yield, respectively, an $n\times n$ matrix, an n-dimensional column vector, and an $n\times n$ matrix whose entries are

$$(\mathbf{A}\,\mathbf{B})_{ij}=\sum_{k=1}^{n}A_{ik}\,B_{kj}\,,\qquad (\mathbf{A}\,\mathbf{b})_i=\sum_{j=1}^{n}A_{ij}\,b_j\,,\qquad \left(\mathbf{a}\,\mathbf{b}^T\right)_{ij}=a_i\,b_j\,. \tag{A.9}$$

Applying definitions (A.9) one finds

$$(\mathbf{A}\,\mathbf{B})^T=\mathbf{B}^T\mathbf{A}^T\,,\qquad (\mathbf{A}\,\mathbf{b})^T=\mathbf{b}^T\mathbf{A}^T\,,\qquad \left(\mathbf{a}\,\mathbf{b}^T\right)^T=\mathbf{b}\,\mathbf{a}^T\,. \tag{A.10}$$

A.3 *Nabla* Operator

A further extension of the concepts introduced in this chapter consists in replacing one or more factors with an operator. An important example is that of the real, vector operator *nabla*,[1]

$$\nabla = \begin{bmatrix} \partial/\partial x_1 \\ \vdots \\ \partial/\partial x_n \end{bmatrix}, \tag{A.11}$$

where $x_1, \ldots, x_n$ are the coordinates of an n-dimensional Cartesian reference. The product of ∇ and a complex, scalar function $f(x_1, \ldots, x_n)$ is defined in the same manner as the product of a vector and a scalar quantity introduced above: ∇f is a vector of components $(\nabla)_i f$, namely,

$$\nabla f = \begin{bmatrix} \partial f/\partial x_1 \\ \vdots \\ \partial f/\partial x_n \end{bmatrix}. \tag{A.12}$$

In turn, the scalar product of ∇ and a complex vector $\mathbf{a}$ of the same dimension as ∇ yield

$$\nabla \cdot \mathbf{a} = \frac{\partial a_1}{\partial x_1} + \ldots + \frac{\partial a_n}{\partial x_n}. \tag{A.13}$$

The product defined by (A.12) is also called *gradient* of f, whereas the scalar product (A.13) is also called *divergence* of $\mathbf{a}$. The corresponding symbols are $\nabla f = \operatorname{grad} f$ and $\nabla \cdot \mathbf{a} = \operatorname{div} \mathbf{a}$, respectively. The scalar product of ∇ by itself is called *Laplacian operator*

$$\nabla^2 = \nabla \cdot \nabla = \frac{\partial^2}{\partial x_1^2} + \ldots + \frac{\partial^2}{\partial x_n^2}, \tag{A.14}$$

then,

$$\nabla^2 f = \frac{\partial^2 f}{\partial x_1^2} + \ldots + \frac{\partial^2 f}{\partial x_n^2}, \qquad \nabla^2 \mathbf{a} = \begin{bmatrix} \nabla^2 a_1 \\ \vdots \\ \nabla^2 a_n \end{bmatrix}. \tag{A.15}$$

Combining the above definitions yields the identities

[1] Symbol ∇ is *not* a Greek letter. However, the term *nabla* is a Greek word, meaning "harp."

$$\nabla^2 f = \nabla \cdot (\nabla f) = \operatorname{div} \operatorname{grad} f\,, \quad \nabla \cdot (f^* \mathbf{a}) = \operatorname{div}(f^* \mathbf{a}) = f^* \operatorname{div} \mathbf{a} + \operatorname{grad} f \cdot \mathbf{a}\,. \tag{A.16}$$

If, in turn, it is $\mathbf{a} = \operatorname{grad} g$, the second relation of (A.16) with the aid of the first one yields the identity

$$\operatorname{div}(f^* \operatorname{grad} g) = f^* \nabla^2 g + \operatorname{grad} f \cdot \operatorname{grad} g\,. \tag{A.17}$$

A.4 Dyadic Products

Sometimes it is convenient to adopt a notation that uses the basis set of real, mutually orthogonal unit vectors $\mathbf{i}_1, \ldots, \mathbf{i}_n$ associated with the axes of a Cartesian reference. By construction it is $\mathbf{i}_r \cdot \mathbf{i}_s = \delta_{rs}$, where the *Kronecker symbol* δ_{rs} is the entry of indices *rs* of a second-rank matrix defined as

$$\delta_{rs} = \begin{cases} 1 & s = r \\ 0 & s \neq r \end{cases} \tag{A.18}$$

The expression of vector $\mathbf{a}$ in terms of the basis vectors is $\mathbf{a} = a_1\,\mathbf{i}_1 + \ldots + a_n\,\mathbf{i}_n$. The notation applies also to the higher-rank objects; for instance, in this notation the matrix $\mathbf{A}$ of (A.8) reads

$$\mathbf{A} = A_{11}\,\mathbf{i}_1\,\mathbf{i}_1^T + A_{12}\,\mathbf{i}_1\,\mathbf{i}_2^T + \ldots + A_{n,n-1}\,\mathbf{i}_n\,\mathbf{i}_{n-1}^T + A_{nn}\,\mathbf{i}_n\,\mathbf{i}_n^T\,, \tag{A.19}$$

A group like $\mathbf{i}_r\,\mathbf{i}_s^T$ is also called *dyadic product*. Observing that $\mathbf{i}_r$ is an *n*-dimensional column vector whose *r*th entry is equal to 1 while all the other entries are equal to 0, the application of the third equation in (A.9) shows that $\mathbf{i}_r\,\mathbf{i}_s^T$ is an $n \times n$ matrix whose entry of indices *rs* is equal to 1, while all the other entries are equal to zero. As a consequence, the form (A.19) expresses $\mathbf{A}$ as a sum of matrices, each associated with an individual entry. Using this notation, a product like $\mathbf{A}\,\mathbf{b}$ reads $\sum_{rs} A_{rs}\,\mathbf{i}_r\,\mathbf{i}_s^T \sum_k b_k\,\mathbf{i}_k$. On the other hand, due to the second equation in (A.9), the same product is equal to $\sum_{rs} A_{rs}\,b_s\,\mathbf{i}_r$. This shows that $\mathbf{i}_r\,\mathbf{i}_s^T\,\mathbf{i}_k = \mathbf{i}_r\,\delta_{sk}$, that is, the juxtaposition of the right unit vector of the dyadic product with the next unit vector must be treated as a scalar product.

The relation defined by the second equation in (A.9) applies also when $\mathbf{b}$ is replaced with a vector operator, with the provision that the operator is meant to act towards the left. For instance, replacing $\mathbf{b}$ with ∇ yields $(\mathbf{A}\,\nabla)_i = \sum_{j=1}^n \partial A_{ij}/\partial x_j$. It follows that the divergence of a second-rank matrix is a column vector of the form

$$\operatorname{div} \mathbf{A} = \sum_{j=1}^{n} \frac{\partial A_{1j}}{\partial x_j}\,\mathbf{i}_1 + \ldots + \sum_{j=1}^{n} \frac{\partial A_{nj}}{\partial x_j}\,\mathbf{i}_n\,. \tag{A.20}$$

In turn, considering the product defined by the third equation in (A.9) and replacing $\mathbf{b}$ with ∇, still with the provision that the operator acts towards the left, yield $(\mathbf{a}\,\nabla^T)_{ij} = \partial a_i/\partial x_j$. It follows that the gradient of a column vector is a second-rank matrix of the form

$$\operatorname{grad}\mathbf{a} = \frac{\partial a_1}{\partial x_1}\,\mathbf{i}_1\,\mathbf{i}_1^T + \frac{\partial a_1}{\partial x_2}\,\mathbf{i}_1\,\mathbf{i}_2^T + \ldots + \frac{\partial a_n}{\partial x_{n-1}}\,\mathbf{i}_1\,\mathbf{i}_{n-1}^T + \frac{\partial a_n}{\partial x_n}\,\mathbf{i}_n\,\mathbf{i}_n^T \tag{A.21}$$

whence, from (A.20),

$$\operatorname{div}(f\,\mathbf{A}) = f\,\operatorname{div}\mathbf{A} + \mathbf{A}\,\operatorname{grad} f\,, \qquad \operatorname{div}(\mathbf{a}\,\mathbf{b}^T) = \mathbf{a}\,\operatorname{div}\mathbf{b} + (\operatorname{grad}\mathbf{a})\,\mathbf{b}\,. \tag{A.22}$$

A.5 Divergence Theorem

The *divergence theorem* (or *Gauss theorem*) states that

$$\int_V \operatorname{div}\mathbf{v}\,\mathrm{d}V = \int_S \mathbf{n}\cdot\mathbf{v}\,\mathrm{d}S\,, \tag{A.23}$$

where V is an n-dimensional volume, $\mathrm{d}V = \mathrm{d}x_1\ldots\mathrm{d}x_n$, S the $(n-1)$-dimensional surface enclosing V, and $\mathbf{n}$ the unit vector normal to the surface element $\mathrm{d}S$, oriented in the outward direction with respect to S. Letting $\mathbf{v} = f^*\,\operatorname{grad} g$ and using (A.17) yield the *first Green theorem*

$$\int_S f^*\,\frac{\partial g}{\partial n}\,\mathrm{d}S = \int_V \left(f^*\,\nabla^2 g + \operatorname{grad} f\cdot\operatorname{grad} g\right)\,\mathrm{d}V\,, \tag{A.24}$$

where $\partial g/\partial n = \mathbf{n}\cdot\operatorname{grad} g$ is the derivative of g in the direction of $\mathbf{n}$. It is easily found that (A.24) is the generalization to n dimensions of the integration by parts. Rewriting (A.24) after letting $\mathbf{v} = g\,\operatorname{grad} f^*$, and subtracting from (A.24), yield the *second Green theorem*

$$\int_S \left(f^*\,\frac{\partial g}{\partial n} - g\,\frac{\partial f^*}{\partial n}\right)\,\mathrm{d}S = \int_V \left(f^*\,\nabla^2 g - g\,\nabla^2 f^*\right)\,\mathrm{d}V\,. \tag{A.25}$$

A special case of the first Green theorem occurs when vector $\mathbf{b} = \operatorname{grad} g$ is constant; the relation (A.24) then reduces to

$$\int_S f^*\,\mathbf{n}\,\mathrm{d}S\cdot\mathbf{b} = \int_V \operatorname{grad} f\,\mathrm{d}V\cdot\mathbf{b}\,, \qquad \mathbf{b} = \text{const.} \tag{A.26}$$

As identity (A.26) holds for any choice of $\mathbf{b}$, the two integrals in it are equal to each other.

A.6 Vector Product

Another possible product between two vectors is the *vector product* $\mathbf{a} \wedge \mathbf{b}$, which yields a column vector. In contrast with the other products introduced in this section, the definition of the vector product will be limited to the three-dimensional case; it is given as the expansion of a determinant, namely,

$$\mathbf{a} \wedge \mathbf{b} = \begin{bmatrix} \mathbf{i}_1 & \mathbf{i}_2 & \mathbf{i}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix} \quad \Rightarrow \quad \mathbf{a} \wedge \mathbf{b} = \begin{bmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{bmatrix} . \tag{A.27}$$

From (A.27) it follows $\mathbf{b} \wedge \mathbf{a} = -\mathbf{a} \wedge \mathbf{b}$ and $\mathbf{a} \wedge \mathbf{a} = 0$. The latter also shows that if two nonvanishing vectors are parallel to each other, say, $\mathbf{b} = k\,\mathbf{a} \neq 0$, then $\mathbf{a} \wedge \mathbf{b} = 0$. When the vector product involves the unit vectors associated with the axes of a right-handed Cartesian reference, the following relations are found:

$$\mathbf{i}_1 \wedge \mathbf{i}_2 = \mathbf{i}_3 , \qquad \mathbf{i}_2 \wedge \mathbf{i}_3 = \mathbf{i}_1 , \qquad \mathbf{i}_3 \wedge \mathbf{i}_1 = \mathbf{i}_2 . \tag{A.28}$$

An intrinsic relation that provides the modulus of $\mathbf{a} \wedge \mathbf{b}$ is found by specifying (A.7) for the case of three-dimensional, real vectors, thus yielding

$$\cos^2 \vartheta = 1 - \sin^2 \vartheta = \frac{\left(\sum_{i=1}^{3} a_i \, b_i \right)^2}{a^2 \, b^2} . \tag{A.29}$$

As $\cos \vartheta = 1$ when the two vectors are parallel, $\mathbf{b} = k\,\mathbf{a}$, $k > 0$, while $\cos \vartheta = -1$ when they are antiparallel, $\mathbf{b} = k\,\mathbf{a}$, $k < 0$, the range of ϑ is $[0, \pi]$. Letting $r_{ij} = a_i \, b_j - a_j \, b_i$ and observing that $(\sum_{i=1}^{3} a_i^2)\,(\sum_{i=1}^{3} b_i^2) = (\sum_{i=1}^{3} a_i \, b_i)^2 + r_{23}^2 + r_{31}^2 + r_{12}^2$ provides

$$\sin^2 \vartheta = \frac{r_{23}^2 + r_{31}^2 + r_{12}^2}{a^2 \, b^2} = \frac{|\mathbf{a} \wedge \mathbf{b}|^2}{a^2 \, b^2} , \qquad |\mathbf{a} \wedge \mathbf{b}| = a \, b \, \sin \vartheta , \tag{A.30}$$

where $\sin \vartheta \geq 0$ due to the range of ϑ.

A.7 Mixed Product

The vector product $\mathbf{a} \wedge \mathbf{b}$ can in turn be scalarly multiplied by another vector $\mathbf{c}$, to yield a scalar quantity called *mixed product*. For the sake of simplicity, in the definition of the mixed product the three vectors will be considered real. From (A.2) one finds

$$\mathbf{a} \wedge \mathbf{b} \cdot \mathbf{c} = \sum_{i=1}^{3} (\mathbf{a} \wedge \mathbf{b})_i \, c_i = \begin{bmatrix} c_1 & c_2 & c_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} . \tag{A.31}$$

The two determinants in (A.31) are equal because they transform into each other by interchanging rows an even number of times. On the other hand, from their equality it follows $\mathbf{a} \wedge \mathbf{b} \cdot \mathbf{c} = \mathbf{a} \cdot \mathbf{b} \wedge \mathbf{c}$, namely, the mixed product is invariant upon interchange of the "wedge" and "dot" symbols.

Considering three nonvanishing vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, where $\mathbf{a}$ and $\mathbf{b}$ are not parallel to each other, and remembering the properties of determinants, one finds that the mixed product vanishes if $\mathbf{c}$ is parallel to $\mathbf{a}$ or parallel to $\mathbf{b}$. In fact,

$$\mathbf{a} \wedge \mathbf{b} \cdot \mathbf{a} = \mathbf{a} \wedge \mathbf{b} \cdot \mathbf{b} = 0 \, . \tag{A.32}$$

It follows that the vector product $\mathbf{a} \wedge \mathbf{b}$ is normal to both $\mathbf{a}$ and $\mathbf{b}$, namely, is normal to the plane defined by the two nonparallel vectors $\mathbf{a}$ and $\mathbf{b}$. If one associates the plane of $\mathbf{a}$ and $\mathbf{b}$ with that of the unit vectors $\mathbf{i}_1$ and $\mathbf{i}_2$, then, using (A.28), the vector product simplifies to $\mathbf{a} \wedge \mathbf{b} = (a_1 \, b_2 - a_2 \, a_1) \, \mathbf{i}_3$, that provides the information about the direction of $\mathbf{a} \wedge \mathbf{b}$. Finally, using (A.27) twice provides the expression for the *double vector product*

$$\mathbf{a} \wedge (\mathbf{b} \wedge \mathbf{c}) = \mathbf{a} \cdot \mathbf{c} \, \mathbf{b} - \mathbf{a} \cdot \mathbf{b} \, \mathbf{c} \, , \qquad (\mathbf{a} \wedge \mathbf{b}) \wedge \mathbf{c} = \mathbf{a} \cdot \mathbf{c} \, \mathbf{b} - \mathbf{b} \cdot \mathbf{c} \, \mathbf{a} \, . \tag{A.33}$$

A.8 Rotational of a Vector

The expressions involving the vector product can be extended to the case where one or two vectors are replaced with the nabla operator (A.11). The vector product

$$\nabla \wedge \mathbf{a} = \begin{bmatrix} \mathbf{i}_1 & \mathbf{i}_2 & \mathbf{i}_3 \\ \partial/\partial x_1 & \partial/\partial x_2 & \partial/\partial x_3 \\ a_1 & a_2 & a_3 \end{bmatrix} = \begin{bmatrix} \partial a_3/\partial x_2 - \partial a_2/\partial x_3 \\ \partial a_1/\partial x_3 - \partial a_3/\partial x_1 \\ \partial a_2/\partial x_1 - \partial a_1/\partial x_2 \end{bmatrix} \tag{A.34}$$

is also called *rotational* of $\mathbf{a}$, the corresponding symbol being $\nabla \wedge \mathbf{b} = \operatorname{rot} \mathbf{a}$. Combining (A.34) with the three-dimensional case of (A.12) and (A.13) shows that the following identities hold:

$$\operatorname{rot}(f \, \mathbf{a}) = f \, \operatorname{rot} \mathbf{a} + \operatorname{grad} f \wedge \mathbf{a} \, , \qquad \operatorname{rot} \operatorname{grad} f = 0 \, , \qquad \operatorname{div} \operatorname{rot} \mathbf{a} = 0 \, , \tag{A.35}$$

$$\operatorname{rot} \operatorname{rot} \mathbf{a} = \operatorname{grad} \operatorname{div} \mathbf{a} - \nabla^2 \mathbf{a} \, , \qquad \operatorname{div}(\mathbf{a} \wedge \mathbf{b}) = \mathbf{b} \cdot \operatorname{rot} \mathbf{a} - \mathbf{a} \cdot \operatorname{rot} \mathbf{b} \, . \tag{A.36}$$

Integrating the second equation in (A.36) over a three-dimensional volume V and using (A.23) yield the identity

$$\int_S \mathbf{n} \cdot \mathbf{a} \wedge \mathbf{b} \, \mathrm{d}S = \int_V (\mathbf{b} \cdot \operatorname{rot} \mathbf{a} - \mathbf{a} \cdot \operatorname{rot} \mathbf{b}) \, \mathrm{d}V \,. \tag{A.37}$$

A special case of (A.37) occurs when vector $\mathbf{a}$ is constant. In fact, noting that $\mathbf{n} \cdot \mathbf{a} \wedge \mathbf{b} = -\mathbf{n} \cdot \mathbf{b} \wedge \mathbf{a} = -\mathbf{n} \wedge \mathbf{b} \cdot \mathbf{a}$, (A.37) reduces to

$$\mathbf{a} \cdot \int_S \mathbf{n} \wedge \mathbf{b} \, \mathrm{d}S = \mathbf{a} \cdot \int_V \operatorname{rot} \mathbf{b} \, \mathrm{d}V \,, \qquad \mathbf{a} = \text{const.} \tag{A.38}$$

As identity (A.38) holds for any choice of $\mathbf{a}$, the two integrals in it are equal to each other.

A.9 Rotational Theorem

The *rotational theorem* (or *Stokes theorem*) states that

$$\int_S \mathbf{n} \cdot \operatorname{rot} \mathbf{v} \, \mathrm{d}S = \int_C \mathbf{t} \cdot \mathbf{v} \, \mathrm{d}C \,, \tag{A.39}$$

where C is the boundary curve of the open surface S, $\mathbf{t}$ the unit vector tangent to C, and $\mathbf{n}$ the unit vector normal to the surface element $\mathrm{d}S$. The direction of the unit vectors is such that the orientation of $\mathbf{b} = \mathbf{t} \wedge \mathbf{n}$ is external with respect to the curve (Fig. A.1).

A.10 Helmholtz Theorem

A vector $\mathbf{u}$ such that $\operatorname{rot} \mathbf{u} = 0$ is called *irrotational*. From the second identity in (A.35) one finds that if $\mathbf{u} = \operatorname{grad} f$, then $\mathbf{u}$ is irrotational. The inverse is not true in general; however, if the domain of $\mathbf{u}$ is simply connected, the condition $\operatorname{rot} \mathbf{u} = 0$ implies that $\mathbf{u}$ can be expressed as a gradient: $\mathbf{u} = \operatorname{grad} f$.

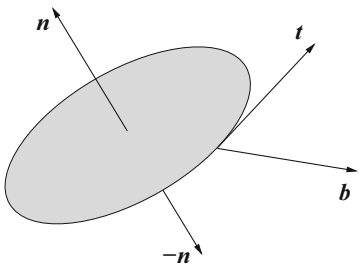


Fig. A.1 Rotational theorem (Sect. A.9): orientation of the unit vectors

A vector $\mathbf{v}$ such that $\operatorname{div}\mathbf{v} = 0$ is called *solenoidal*. From the third identity in (A.35) one finds that if $\mathbf{v} = \operatorname{rot}\mathbf{a}$, then $\mathbf{v}$ is solenoidal. The inverse is not true in general; however, if the domain of $\mathbf{v}$ is simply connected, the condition $\operatorname{div}\mathbf{v} = 0$ implies that $\mathbf{v}$ can be expressed as a rotational: $\mathbf{v} = \operatorname{rot}\mathbf{a}$.

The *Helmholtz theorem* states that a vector $\mathbf{w}$ defined in a simply connected domain can be expressed in a unique manner as the sum of an irrotational and a solenoidal vector:

$$\mathbf{w} = \operatorname{grad} f + \operatorname{rot}\mathbf{a}\,. \tag{A.40}$$

Scalar f is found by taking the divergence of both sides of (A.40) and using the identities $\operatorname{div}\operatorname{grad} f = \nabla^2 f$, $\operatorname{div}\operatorname{rot}\mathbf{a} = 0$. In turn, vector $\mathbf{a}$ is found by taking the rotational of both sides of (A.40) and using the first identity in (A.36) along with the auxiliary condition $\operatorname{div}\mathbf{a} = 0$. By this procedure it is found that f and $\mathbf{a}$ fulfill the relations

$$\nabla^2 f = \operatorname{div}\mathbf{w}\,, \qquad \nabla^2\mathbf{a} = -\operatorname{rot}\mathbf{w}\,. \tag{A.41}$$

The right-hand sides of (A.41) are known because $\mathbf{w}$ is prescribed. As a consequence, the problem of finding f and $\mathbf{a}$ is equivalent to solving a set of Poisson equations. The solution of (A.41) is unique provided that $\mathbf{w}$ vanishes at infinity faster than r^{-1} [82, Sect. XI.3]. Unless some additional prescriptions are imposed on f and $\mathbf{a}$, (A.40) still holds if one adds to f an arbitrary constant and, to $\mathbf{a}$, the gradient of an arbitrary scalar function.

A.11 Matrices

The following definitions and properties[2] apply to a real square matrix $\mathbf{A}$ of order N, whose entries are indicated with $a_{ij} = (\mathbf{A})_{ij}$. It is known that the determinant of a matrix changes its sign when two rows or columns are exchanged. It follows that if the matrix is non-singular, $\det\mathbf{A} \neq 0$, the matrix obtained by a sequence of such exchanges is still non-singular.

In the numerical solution of the differential equations that model a semiconductor device it is necessary to repeatedly solve an algebraic system of the form $\mathbf{A}\,\mathbf{x} = \mathbf{b}$, with $\mathbf{A}$ a non-singular, real square matrix of order N. This section illustrates a number of matrix properties that are useful for the problem under consideration; the application to the semiconductor-device equations is illustrated in Sect. A.13.

It is important to distinguish the solution of the algebraic system, which amounts to finding $\mathbf{x}$ given the data $\mathbf{b}$, from the inversion of $\mathbf{A}$, which amounts to finding $\mathbf{B} = \mathbf{A}^{-1}$. The second operation lends itself to solving the system as well: in fact,

[2] Several definitions and properties illustrated in this section have an analogue in the theory of operators. An illustration of the latter is given in Sects. 8.3, 8.4, and 8.6.

$\mathbf{x} = \mathbf{A}^{-1}\,\mathbf{b} = \mathbf{B}\,\mathbf{b}$; however, it is much more expensive than the first one.[3] To show this, let $\mathbf{b} = \mathbf{i}_k$, where $\mathbf{i}_k$ is a vector whose kth entry is equal to unity, while all other entries vanish; observing that the jth entry of $\mathbf{i}_k$ is equal to the Kronecker symbol δ_{jk} (A.18), the entries of $\mathbf{x} = \mathbf{B}\,\mathbf{i}_k$ are

$$x_i = \sum_{j=1}^{N} B_{ij}\,\delta_{jk} = B_{ik}\,. \tag{A.42}$$

Thus, the solution vector $\mathbf{x}$ is the kth column of $\mathbf{B} = \mathbf{A}^{-1}$; the whole inverse matrix $\mathbf{B}$ is found by repeating the solution process with $k = 1, 2, \ldots, N$. In conclusion, if $C(N)$ is the computational cost of solving the $N \times N$ algebraic system $\mathbf{A}\,\mathbf{x} = \mathbf{b}$, then the cost of inverting $\mathbf{A}$ is $N\,C(N)$.

The *main diagonal* of a square matrix of order N is the set of entries $a_{11}, \ldots, a_{NN}$. A real, square matrix of entries a_{ij} is *symmetric* when $a_{ji} = a_{ij}$; thus, a symmetric matrix is defined by prescribing the N diagonal elements and the $(N^2 - N)/2$ elements on one side of the diagonal. A real, square matrix is *anti-symmetric* or *skew-symmetric* when $a_{ji} = -a_{ij}$; thus, the diagonal elements of a skew-symmetric matrix vanish, and the matrix is defined by giving the $(N^2 - N)/2$ elements on one side of the diagonal.

A *diagonal matrix* is a square matrix in which all entries that do not belong to the main diagonal are equal to zero.

The *transpose* $\mathbf{A}^T$ of a square matrix $\mathbf{A}$ of order N is the matrix of entries $(\mathbf{A}^T)_{ij} = (\mathbf{A})_{ji}$; thus, a real, symmetric matrix is equal to its transpose, $\mathbf{A}^T = \mathbf{A}$. A non-singular, square matrix $\mathbf{A}$ is *orthogonal* if $\mathbf{A}^T = \mathbf{A}^{-1}$; for an orthogonal matrix it then follows $\mathbf{A}\,\mathbf{A}^T = \mathbf{A}^T\,\mathbf{A} = \mathbf{I}$, whence the scalar product of each column by itself is equal to unity, $a_{r1}^2 + \cdots + a_{rN}^2 = 1$, while the scalar product of a column by another column is equal to zero, $a_{r1}\,a_{s1} + \cdots + a_{rN}\,a_{sN} = 0$; the same property applies to rows. If $\mathbf{A}$ is orthogonal, the norm of vector $\mathbf{d} = \mathbf{A}\,\mathbf{c}$ is equal to that of vector $\mathbf{c}$; the demonstration is the same as for the unitary matrix, given by (A.43).

A real square matrix $\mathbf{A}$ is *real normal* if it commutes with its transpose, $\mathbf{A}^T\,\mathbf{A} = \mathbf{A}\,\mathbf{A}^T$. Symmetric, skew-symmetric, and orthogonal matrices matrices are real normal.

The above definitions are extended to the case of a complex square matrix $\mathbf{A}$ of order N, whose entries are still indicated with a_{ij}. The *conjugate transpose* $\mathbf{A}^\dagger$ of $\mathbf{A}$ is the matrix of entries $(\mathbf{A}^\dagger)_{ij} = (\mathbf{A})_{ji}^*$. Matrix $\mathbf{A}$ is *Hermitean* if $(\mathbf{A})_{ji} = (\mathbf{A})_{ij}^*$; thus, a Hermitean matrix is equal to its conjugate transpose, $\mathbf{A}^\dagger = \mathbf{A}$. From the definition it follows that in a Hermitean matrix the entries of the main diagonal are real. A non-singular, complex square matrix $\mathbf{A}$ is *unitary* if $\mathbf{A}^\dagger = \mathbf{A}^{-1}$; for a unitary matrix it then follows $\mathbf{A}\,\mathbf{A}^\dagger = \mathbf{A}^\dagger\,\mathbf{A} = \mathbf{I}$, whence the scalar product of each column by itself is equal to unity, $|a_{r1}|^2 + \cdots + |a_{rN}|^2 = 1$, while the scalar product of a column by another column is equal to zero, $a_{r1}^*\,a_{s1} + \cdots + a_{rN}^*\,a_{sN} = 0$. The same property

[3] The statement is true in the general case, that is, when $\mathbf{A}$ has no special structure.

applies to rows: $|a_{1r}|^2 + \cdots + |a_{Nr}|^2 = 1$ and $a_{1r}^* a_{1s} + \cdots + a_{Nr}^* a_{Ns} = 0$. If $\mathbf{A}$ is unitary, the norm of vector $\mathbf{d} = \mathbf{A}\,\mathbf{c}$ is equal to that of vector $\mathbf{c}$; in fact, observing that the product of two rows of a unitary matrix can be written in compact form as $\sum_{i=1}^{N} a_{ir}^* a_{is} = \delta_{rs}$, it is

$$\mathbf{d}^* \cdot \mathbf{d} = \sum_{i=1}^{N} \left(\sum_{r=1}^{N} a_{ir}\, c_r \right)^* \left(\sum_{s=1}^{N} a_{is}\, c_s \right) = \sum_{r\,s=1}^{N} c_r^*\, c_s \sum_{i=1}^{N} a_{ir}^*\, a_{is} = \mathbf{c}^* \cdot \mathbf{c}\,. \tag{A.43}$$

Matrix $\mathbf{A}$ is *skew-Hermitean* if $(\mathbf{A})_{ji} = -(\mathbf{A})_{ij}^*$; thus, a skew-Hermitean matrix is equal to the negative of its conjugate transpose, $\mathbf{A}^\dagger = -\mathbf{A}$. From the definition it follows that in a skew-Hermitean matrix the entries of the main diagonal are imaginary.

A complex square matrix $\mathbf{A}$ is *normal* if it commutes with its conjugate transpose, $\mathbf{A}^\dagger\,\mathbf{A} = \mathbf{A}\,\mathbf{A}^\dagger$. Hermitian, skew-Hermitian, and unitary matrices are normal. Instead, a unitary matrix is not necessarily Hermitean: for this to hold it must simultaneously be $\mathbf{A}^\dagger = \mathbf{A}^{-1}$ and $\mathbf{A}^\dagger = \mathbf{A}$, whence $\mathbf{A}$ must fulfill the constraint $\mathbf{A}^2 = \mathbf{I}$.

A.11.1 Eigenvalues

Given an $N \times N$ non-singular matrix $\mathbf{V}$, consider the algebraic system

$$\mathbf{V}\,\mathbf{e} = \lambda\,\mathbf{e}\,, \tag{A.44}$$

with λ an undetermined parameter. If λ exists, such as (A.44) has a nontrivial solution, then λ is an *eigenvalue* of $\mathbf{V}$ and the nontrivial solution thus found is the *eigenvector* corresponding to λ. For the existence of nontrivial solutions of (A.44) it is necessary that

$$\det(\mathbf{V} - \lambda\mathbf{I}) = 0\,. \tag{A.45}$$

The above is an algebraic equation of degree N, called *characteristic equation* or also *secular equation*, which has N solutions $\lambda_1, \ldots, \lambda_N$; such solutions are not necessarily distinct from each other, and are generally complex. Successively inserting $\lambda_1, \lambda_2, \ldots$ into (A.44) yields the eigenvectors $\mathbf{e}_1, \mathbf{e}_2, \ldots$, that are the only nontrivial solutions of (A.44).

The eigenvalues of a non-singular matrix are different from zero; in fact, if it were $\lambda_m = 0$, then it would be $\mathbf{V}\,\mathbf{e}_m = 0$; however, the latter equation can have a nontrivial solution only if $\mathbf{V}$ were singular, contrary to the hypothesis. The converse is also true: if $\mathbf{V}$ is singular, then there exists a nonvanishing vector $\mathbf{e}_m$ that makes the left-hand side of (A.44) equal to zero; this is compatible with (A.44) only if the corresponding eigenvalue is equal to zero.

From the homogeneity of (A.44) it follows that the eigenvectors are determined apart from a multiplicative constant. For each eigenvector $\mathbf{e}_r$, the constant can then be chosen such that the vector is normalized to unity: $|\mathbf{e}_r| = 1$. Another important property is that eigenvectors corresponding to different eigenvalues are different (that is, nonparallel) from each other. Let the eigenvalues be λ_r and $\lambda_s \neq \lambda_r$, and let $\mathbf{e}_r$ be an eigenvector corresponding to λ_r, and $\mathbf{e}_s$ an eigenvector corresponding to λ_s. Specifying (A.44) for index r and index s, and subtracting from each other the two relations thus found, yields $\mathbf{V}(\mathbf{e}_r - \mathbf{e}_s) = \lambda_r \mathbf{e}_r - \lambda_s \mathbf{e}_s$. If it were $\mathbf{e}_s = \mathbf{e}_r$ it would follow $(\lambda_r - \lambda_s)\,\mathbf{e}_r = 0$; however, the latter relation is impossible because $\mathbf{e}_r$, being an eigenvector, does not vanish, and the two eigenvalues are different by hypothesis. As a corollary, it follows that an eigenvector belongs to one and only one eigenvalue.

The *spectral radius* of a matrix is the maximum modulus of its eigenvalues, $\varrho = \max_i(|\lambda_i|)$. There are several theorems providing upper bounds for ϱ [94, Chap. III]; among these, the *Gershgorin theorem* states that

$$|\lambda| \leq \min(R, C)\,, \tag{A.46}$$

where the two real parameters R, C depend on the entries of $\mathbf{V}$ as shown below. To demonstrate (A.46) one starts from (A.44) and let λ be an eigenvalue of $\mathbf{V}$, with $\mathbf{e}$ an eigenvector corresponding to it. Take the entry of $\mathbf{e}$ whose modulus is maximum, say, e_i, and consider the ith row of (A.44), $\sum_{j=1}^{N} V_{ij}\, e_j = \lambda\, e_i$. Using the arbitrary multiplicative constant of the eigenvalues, normalize $\mathbf{e}$ so that $e_i = 1$; it follows $V_{ii} + \sum_{j \neq i} V_{ij}\, e_j = \lambda$ where, by construction, $|e_j| \leq 1$. Thus, letting $R_i = \sum_{j=1}^{N} |V_{ij}|$,

$$|\lambda| \leq |V_{ii}| + \left| \sum_{j \neq i} V_{ij}\, e_j \right| \leq |V_{ii}| + \sum_{j \neq i} |V_{ij}| \leq R_i\,. \tag{A.47}$$

This shows that the eigenvalue under consideration belongs to the disc of the complex plane centered in the origin and having R_i as radius. On the other hand, as the reasoning leading to (A.47) can be repeated for any eigenvalue, one concludes that the maximum of such discs contains all eigenvalues of $\mathbf{V}$, namely, $|\lambda| \leq R = \max_i(R_i)$. Repeating the procedure starting from the transpose of (A.44) yields $|\lambda| \leq C = \max_i(C_i)$ with $C_i = \sum_{j=1}^{N} |V_{ji}|$, namely, the moduli of the entries are added columnwise; since both constraints $|\lambda| \leq R$ and $|\lambda| \leq C$ hold simultaneously, (A.46) ensues.

Linear Independence of the Eigenvectors

From the properties demonstrated in Sect. A.11.1 it follows that two eigenvectors $\mathbf{e}_r$, $\mathbf{e}_s$ belonging to different eigenvalues are linearly independent. If the opposite were true, there would exist two nonvanishing constants σ_r, σ_s such that $\sigma_r\, \mathbf{e}_r + \sigma_s \mathbf{e}_s = 0$; this, however, is impossible because the two eigenvectors would be parallel to

each other. More generally, it is found that the eigenvectors of $\mathbf{V}$ belonging to $m > 2$ distinct eigenvalues are linearly independent. For this, assume that (A.45) has m distinct roots, with $2 < m \leq N$, and number such roots with indices ranging from 1 to m, so that the corresponding eigenvectors are $\mathbf{e}_1, \ldots, \mathbf{e}_m$. Now, if these eigenvectors were linearly dependent, there would exist m constants $\gamma_1, \ldots, \gamma_m$ such that

$$\gamma_1 \mathbf{e}_1 + \ldots + \gamma_m \mathbf{e}_m = 0\,, \tag{A.48}$$

with at least two of the constants different from zero. Multiplying (A.48) by $\mathbf{V}$ and using (A.44) provides

$$\gamma_1\, \lambda_1\, \mathbf{e}_1 + \ldots + \gamma_m\, \lambda_m\, \mathbf{e}_m = 0\,. \tag{A.49}$$

Now, for any index k such that $1 \leq k \leq m$, multiplying (A.48) by λ_k and subtracting the result from (A.49) yield

$$\sum_{j=1}^{m} \gamma_j\, (\lambda_k - \lambda_j)\, \mathbf{e}_j = 0 \tag{A.50}$$

where, apart from the summand with $j = k$, the differences in parentheses are different from zero because the eigenvalues are distinct; from (A.50) it then follows that the $m-1$ eigenvectors $\mathbf{e}_1, \ldots, \mathbf{e}_{k-1}, \mathbf{e}_{k+1}, \ldots, \mathbf{e}_m$ are linearly dependent. On the other hand, this reasoning can be repeated for all values of k between 1 and m; as a consequence, if (A.48) were true, then all sets made of $m-1$ eigenvalues extracted from $\mathbf{e}_1, \ldots, \mathbf{e}_m$ would fulfill relations similar to (A.48). Starting from this result one may repeat the reasoning to show that linear independence would be absent for all sets made of $m-2, m-3, \ldots$ eigenvectors, to finally reach all sets made of two eigenvectors. This, however, is absurd as shown above; in conclusion, (A.48) holds only if all constants $\gamma_1, \ldots, \gamma_m$ vanish, thus showing that eigenvectors belonging to distinct eigenvalues are linearly independent.[4]

If it is $m = N$, matrix $\mathbf{V}$ has N distinct eigenvalues, each of them associated with a different eigenvector. If, instead, it is $m < N$, thus implying that (A.45) has multiple eigenvalues, the above reasoning is not applicable to the indices larger than m; in fact, if one included in (A.50) also the summands relative to $\lambda_{m+1}, \lambda_{m+2}, \ldots$, the parentheses containing repeated eigenvalues would vanish, and the left-hand side of (A.50) would still be made of $m-1$ summands.

[4]Note that the conclusion is reached without assuming any special property of the matrix, apart from that of being non-singular.

Orthogonalization of a Set of Vectors

It is shown here that from a set of N linearly independent vectors $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_N$ it is possible to extract a set of N mutually orthogonal vectors $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_N$. The procedure, called *Gram-Schmidt orthogonalization*, forms linear combinations of the $\mathbf{p}_r$ vectors according to the scheme

$$\mathbf{w}_1 = \mathbf{p}_1\,, \qquad \mathbf{w}_2 = \mathbf{p}_2 + d_{21}\,\mathbf{w}_1\,, \qquad \mathbf{w}_3 = \mathbf{p}_3 + d_{31}\,\mathbf{w}_1 + d_{32}\,\mathbf{w}_2\,, \quad \ldots \tag{A.51}$$

or, more generally, $\mathbf{w}_r = \mathbf{p}_r + \sum_{k=1}^{r-1} d_{rk}\,\mathbf{w}_k$, where coefficients d_{rk} are provisionally left undetermined. Left multiplying by $\mathbf{w}_1^*$ the second relation in (A.51), and prescribing that $\mathbf{w}_2$ and $\mathbf{w}_1$ are orthogonal, yields $d_{21}\,|\mathbf{w}_1|^2 = -\mathbf{w}_1^* \cdot \mathbf{p}_2$. The left-hand side of the latter cannot vanish: if it did, it would also be $\mathbf{p}_1 = \mathbf{w}_1 = 0$ and, in this case, vectors $\mathbf{p}_r$ would not be linearly independent. Thus, extracting $d_{21} = -\mathbf{w}_1^* \cdot \mathbf{p}_2/|\mathbf{w}_1|^2$ and replacing it in the second relation in (A.51), provides $\mathbf{w}_2$ as a linear combination of $\mathbf{p}_1$ and $\mathbf{p}_2$; therefore, $\mathbf{w}_2$ does not vanish. Then, one proceeds by left multiplying the third relation in (A.51) by either $\mathbf{w}_1^*$ or $\mathbf{w}_2^*$, and prescribing that $\mathbf{w}_3$ be orthogonal to $\mathbf{w}_1$ in the first case, and to $\mathbf{w}_2$ in the second case. Remembering that $\mathbf{w}_1$ and $\mathbf{w}_2$ are orthogonal, one finds $d_{31}\,|\mathbf{w}_1|^2 = -\mathbf{w}_1^* \cdot \mathbf{p}_3$ and $d_{32}\,|\mathbf{w}_2|^2 = -\mathbf{w}_2^* \cdot \mathbf{p}_3$, where the vectors at the left-hand side are different from zero. Thus, one extracts the coefficients d_{31} and d_{32}, determines $\mathbf{w}_3$, and so on. In conclusion, the general expression for the coefficients reads $d_{rk}\,|\mathbf{w}_k|^2 = -\mathbf{w}_k^* \cdot \mathbf{p}_r$.

A.11.2 Properties of Hermitean Matrices

The eigenvalues and eigenvectors of Hermitean matrices have a number of properties that are worth illustrating. To begin one finds that if $\mathbf{V}$ is Hermitean, the scalar product $\mathbf{a}^* \cdot \mathbf{V}\,\mathbf{a}$ is real for any choice of vector $\mathbf{a}$; in fact, the product $\mathbf{a}^* \cdot \mathbf{V}\,\mathbf{a}$ reads

$$\sum_{ij=1}^{N} a_i^*\, V_{ij}\, a_j = \sum_{ji=1}^{N} a_j^*\, V_{ji}\, a_i = \sum_{ji=1}^{N} a_j^*\, V_{ij}^*\, a_i\,, \tag{A.52}$$

where the second form is obtained from the first one by exchanging the indices. From (A.52) it follows that the sum is equal to its conjugate, namely, it is real. A consequence of this finding is that the eigenvalues of a Hermitean matrices are real; to show this, let $\mathbf{e}_r$ be an eigenvector of $\mathbf{V}$ and λ_r the corresponding eigenvalue. A scalar multiplication on the left by $\mathbf{e}_r^*$ of the eigenvalue equation (A.44) yields $\mathbf{e}_r^\star \cdot \mathbf{V}\,\mathbf{e}_r = \lambda_r\,\mathbf{e}_r^\star \cdot \mathbf{e}_r$, where the left-hand side is real as demonstrated with (A.52). The right-hand side $\mathbf{e}_r^\star \cdot \mathbf{e}_r = \sum_{j=1}^{N} \left[\Re(e_{rj})^2 + \Im(e_{rj})^2\right] = |\mathbf{e}_r|^2$ is strictly positive because an eigenvector does not vanish; in conclusion,

$$\lambda_r = \frac{1}{|\mathbf{e}_r|^2} \sum_{j,k=1}^{n} e^{\star}_{rj} \, V_{jk} \, e_{rk} \tag{A.53}$$

is real. It may happen that the real product $\mathbf{a}^* \cdot \mathbf{V}\,\mathbf{a}$ is strictly positive (negative) for any choice of vector $\mathbf{a}$; in this case, $\mathbf{A}$ is called *positive definite* (*negative definite*). From (A.53) it then follows that the eigenvalues of a positive-definite Hermitean matrix are all positive; thus, as shown in Sect. A.11.1, such a matrix is non-singular.

Other relations are derived by considering eigenvectors with different indices, e.g., $\mathbf{e}_r$ and $\mathbf{e}_s$: left multiplying $\mathbf{V}\,\mathbf{e}_r = \lambda_r \, \mathbf{e}_r$ by $\mathbf{e}^{\star}_s$, right multiplying the conjugate of $\mathbf{V}\,\mathbf{e}_s = \lambda_s \, \mathbf{e}_s$ by $\mathbf{e}_r$, and subtracting the relations thus obtained, one finds that the left-hand sides delete each other due to the hermiticity of $\mathbf{V}$. In conclusion,

$$(\lambda_r - \lambda_s) \; \mathbf{e}^{\star}_s \cdot \mathbf{e}_r = 0 \, . \tag{A.54}$$

If all roots of the characteristic equation (A.45) are distinct, the term in parentheses in (A.54) is different from zero for any pair of indices r, s; it follows that all eigenvectors are *mutually orthogonal*. Remembering (Sect. A.11.1) that the eigenvectors are defined apart from a multiplicative constant, hence they are normalizable to unity, for a Hermitean matrix with distinct eigenvalues one can always assume

$$\mathbf{e}^{\star}_s \cdot \mathbf{e}_r = \delta_{sr} \, , \tag{A.55}$$

with δ_{sr} the Kronecker symbol (A.18). Eigenvalues that fulfill (A.55) are called *orthonormal*.[5]

If the characteristic equation (A.45) has multiple roots, there are pairs of indices such that the difference in parentheses of (A.54) vanishes; for this reason, the analysis cannot be based upon (A.54). However, as shown in Sect. A.11.1, there always exist N mutually independent eigenvectors; also, for Hermitean matrices it is always possible to find a set of N mutually orthonormal eigenvectors, regardless of the multiplicity of the eigenvalues (section "Diagonalization of Hermitean Matrices—Multiple Eigenvalues"). Finally, if $\mathbf{V}$ is real and symmetric the eigenvalue equation (A.44) is an algebraic system with real coefficients; it is then possible to select the normalization constants in a way that makes all eigenvectors real.

[5]Condition (A.55) provides a simpler demonstration, with respect to the general one shown in Sect. A.11.1, of the linear independence of the eigenvectors. If the eigenvectors were linearly dependent there would be N constants $\gamma_1, \ldots, \gamma_N$, not all vanishing, such that the relation $\sum_{r=1}^{N} \gamma_r \, \mathbf{e}_r = 0$ holds. Assuming for instance that $\gamma_s \neq 0$, left multiplying the above relation by $\mathbf{e}^{\star}_s$, and using (A.55), yields that all summands at the left-hand side vanish with the exception of the sth one, thus yielding the absurd conclusion $|\mathbf{e}_s|^2 = 0$.

Diagonalization of Hermitean Matrices

Let $\mathbf{V}$ be an $N \times N$ Hermitean matrix, whose eigenvalues and eigenvectors are λ_r and $\mathbf{e}_r$, respectively. Assume that all eigenvalues are distinct; it follows (section "Linear Independence of the Eigenvectors") that the eigenvectors are mutually independent; also, one can assume (section "Orthogonalization of a Set of Vectors") that the eigenvectors are mutually orthogonal and normalized to unity. With these premises, let $\mathbf{G}$ be an $N \times N$ matrix whose columns are the eigenvectors of $\mathbf{V}$, namely, the entry of indices $k\,r$ of $\mathbf{G}$ is the kth entry of $\mathbf{e}_r$. From the linear independence of the eigenvectors it follows that $\mathbf{G}$ is non-singular, whence its inverse $\mathbf{G}^{-1}$ exists. Left multiplying $\mathbf{G}$ by its conjugate transpose $\mathbf{G}^\dagger$, one finds that the entry of indices $s\,r$ of the product is $\sum_{k=1}^{N} (\mathbf{G}^\dagger)_{sk}\, G_{kr} = \sum_{k=1}^{N} G^\star_{ks}\, G_{kr} = \mathbf{e}^\star_s \cdot \mathbf{e}_r$. Remembering that the set of eigenvectors is orthonormal, one finds from (A.55) that $\mathbf{G}^\dagger \mathbf{G} = \mathbf{I}$, namely, from the definition of Sect. A.11, $\mathbf{G}$ is unitary.

Now, let $\mathbf{D}$ be a diagonal matrix of order N, whose diagonal entries are the eigenvalues λ_r, namely, $D_{kr} = \lambda_r\, \delta_{kr}$. From the above definitions, the N algebraic systems $\mathbf{V}\,\mathbf{e}_r = \lambda_r\,\mathbf{e}_r$ are recast in matrix form as $\mathbf{V}\,\mathbf{G} = \mathbf{G}\,\mathbf{D}$. Left multiplying the latter by $\mathbf{G}^{-1}$ yields

$$\mathbf{G}^{-1}\,\mathbf{V}\,\mathbf{G} = \mathbf{D}. \tag{A.56}$$

This result shows that the matrix made of the eigenvectors produces a similarity transformation (Sect. A.11.3) that brings the Hermitean matrix $\mathbf{V}$ into the matrix made of its eigenvalues. This type of similarity transformation is called *diagonalization*, because it yields a diagonal matrix, or also *unitary transformation*, because it is based upon a unitary matrix. An example of application to the diagonalization of a Hamiltonian function is shown in Sect. 3.10.

Diagonalization of Hermitean Matrices—Multiple Eigenvalues

This section examines the case of a Hermitean matrix whose characteristic equation (A.45) ha multiple roots. One starts by considering a unitary matrix $\mathbf{W}$, not necessarily made of the eigenvectors of $\mathbf{V}$. As shown in Sect. A.11, a unitary matrix is not Hermitean in general; instead, for any matrix $\mathbf{W}$ the property holds that if $\mathbf{V}$ is Hermitean, then $\mathbf{W}^\dagger\,\mathbf{V}\,\mathbf{W}$ is also Hermitean; in fact,

$$\left(\mathbf{W}^\dagger\,\mathbf{V}\,\mathbf{W}\right)_{jk} = \sum_{s\,r=1}^{N} W^\star_{sj}\, V_{sr}\, W_{rk}\,, \qquad \left(\mathbf{W}^\dagger\,\mathbf{V}\,\mathbf{W}\right)^\star_{kj} = \sum_{r\,s=1}^{N} W_{rk}\, V^\star_{rs}\, W^\star_{sj}, \tag{A.57}$$

where the right-hand sides are equal because $\mathbf{V}$ is Hermitean. Also, the product of two unitary matrices is unitary; observing in fact that for any pair of square matrices $\mathbf{W}_a$, $\mathbf{W}_b$ it is $(\mathbf{W}_a\,\mathbf{W}_b)^\dagger = \mathbf{W}^\dagger_b\,\mathbf{W}^\dagger_a$, for unitary matrices it is

$$(\mathbf{W}_a\,\mathbf{W}_b)^\dagger\,(\mathbf{W}_a\,\mathbf{W}_b) = (\mathbf{W}^\dagger_b\,\mathbf{W}^\dagger_a)\,(\mathbf{W}_a\,\mathbf{W}_b) = \mathbf{W}^\dagger_b\,(\mathbf{W}^\dagger_a\,\mathbf{W}_a)\,\mathbf{W}_b = \mathbf{I}\,. \tag{A.58}$$

Now, let $\lambda_1^{(N)}$ be an eigenvalue of (A.44), with $\mathbf{e}_1$ an eigenvector corresponding to it, normalized to unity. In the N-dimensional space it is always possible to find other $N-1$ vectors, linearly independent with respect to $\mathbf{e}_1$ and to each other; then, using the orthogonalization procedure of section "Orthogonalization of a Set of Vectors," one extracts $N-1$ vectors $\mathbf{w}_2, \ldots, \mathbf{w}_N$, also normalized to unity, that are orthogonal to $\mathbf{e}_1$ and with respect to each other. Let $\mathbf{W}_N$ be the unitary matrix having $\mathbf{e}_1$ as first column and $\mathbf{w}_2, \ldots, \mathbf{w}_N$ as the remaining columns. Remembering that $\mathbf{e}_1$ is an eigenvector of $\mathbf{V}$, the first column of the product $\mathbf{V}\,\mathbf{W}_N$ is $\lambda_1^{(N)}\mathbf{e}_1$, whence[6]

$$\mathbf{W}_N^\dagger\,\mathbf{V}\,\mathbf{W}_N = \begin{bmatrix} \lambda_1^{(N)} & 0 & \ldots & 0 \\ 0 & & & \\ \vdots & & \mathbf{V}_{N-1} & \\ 0 & & & \end{bmatrix} . \tag{A.59}$$

The $(N-1)\times(N-1)$ matrix $\mathbf{V}_{N-1}$ is Hermitean as well so that, iterating the procedure, one defines a unitary matrix $\mathbf{W}_{N-1}$ such that

$$\mathbf{W}_{N-1}^\dagger\,\mathbf{V}_{N-1}\,\mathbf{W}_{N-1} = \begin{bmatrix} \lambda_2^{(N-1)} & 0 & \ldots & 0 \\ 0 & & & \\ \vdots & & \mathbf{V}_{N-2} & \\ 0 & & & \end{bmatrix} , \tag{A.60}$$

with $\lambda_2^{(N-1)}$ an eigenvalue of $\mathbf{V}_{N-1}$. Letting

$$\mathbf{U}_N = \mathbf{W}_N\,, \qquad \mathbf{U}_{N-1} = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & & & \\ \vdots & & \mathbf{W}_{N-1} & \\ 0 & & & \end{bmatrix} , \tag{A.61}$$

one finds

$$\mathbf{U}_{N-1}^\dagger\,\mathbf{U}_N^\dagger\,\mathbf{V}\,\mathbf{U}_N\,\mathbf{U}_{N-1} = \begin{bmatrix} \lambda_1^{(N)} & 0 & 0 & \ldots & 0 \\ 0 & \lambda_2^{(N-1)} & 0 & \ldots & 0 \\ 0 & 0 & & & \\ \vdots & \vdots & & \mathbf{V}_{N-2} & \\ 0 & 0 & & & \end{bmatrix} . \tag{A.62}$$

[6] One notes in fact that the first entry in the first column of (A.59) reads $\lambda_1^{(N)}\,|\mathbf{e}_1|^2 = \lambda_1^{(N)}$, while the others are $\lambda_1^{(N)}\,\mathbf{e}_1 \cdot \mathbf{w}_r = 0$. Also, considering that $\mathbf{W}_N^\dagger\,\mathbf{V}\,\mathbf{W}_N$ is Hermitean, the entries of the first row apart from the first one must vanish as well.

The next iteration yields

$$\mathbf{W}_{N-2}^{\dagger}\,\mathbf{V}_{N-2}\,\mathbf{W}_{N-2} = \begin{bmatrix} \lambda_3^{(N-2)} & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & \mathbf{V}_{N-3} & \\ 0 & & & \end{bmatrix}, \quad \mathbf{U}_{N-2} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & \\ \vdots & \vdots & \mathbf{W}_{N-2} & \\ 0 & & & \end{bmatrix} \tag{A.63}$$

with $\lambda_3^{(N-2)}$ an eigenvalue of $\mathbf{V}_{N-2}$, and so on. At the end of the procedure one finds a relation similar to (A.62) whose right-hand side is a diagonal matrix $\mathbf{D}$; the elements of the latter are eigenvalues of Hermitean matrices, hence they are real. In turn, at the left-hand side one finds that matrix $\mathbf{V}$ is right multiplied by the unitary matrix $\mathbf{G} = \mathbf{U}_N\,\mathbf{U}_{N-1}\dots\mathbf{U}_2\,\mathbf{U}_1$ and left multiplied by the conjugate transpose of the above. One then finds that the procedure eventually leads to (A.56); in other terms, one can identify the entries $\lambda_1^{(N)}, \lambda_2^{(N-1)}, \dots, \lambda_N^{(1)}$ of $\mathbf{D}$ with the eigenvalues of $\mathbf{V}$, and the columns of $\mathbf{G} = \mathbf{U}_N\,\mathbf{U}_{N-1}\dots\mathbf{U}_2\,\mathbf{U}_1$ with its eigenvectors. In conclusion, without resorting to the hypothesis that the eigenvalues are simple, it has been demonstrated that a Hermitean matrix can be diagonalized with a unitary transformation; this, in turn, provides a set of N orthonormal eigenvectors. As shown in Sect. A.11.2, if matrix $\mathbf{V}$ is real and symmetric it is possible to choose the normalization constants such that the eigenvectors are real.

Eigenvalues of Other Special Matrices

Let $\mathbf{A}$ be a complex square matrix. If $\mathbf{A}$ is normal, and $\mathbf{e}_r$ is an eigenvector of $\mathbf{A}$ corresponding to eigenvalue λ_r, then $\mathbf{e}_r$ is also an eigenvector of $\mathbf{A}^{\dagger}$, corresponding to eigenvalue λ_r^*. To show this one defines $\mathbf{B} = \mathbf{A} - \lambda\,\mathbf{I}$ so that the eigenvalue equation (A.44) takes the form $\mathbf{B}\,\mathbf{e} = 0$. Matrix $\mathbf{B}$ is normal by construction, and $\mathbf{B}^{\dagger} = \mathbf{A}^{\dagger} - \lambda^*\,\mathbf{I}$. From $|\mathbf{B}\,\mathbf{e}|^2 = (\mathbf{B}\,\mathbf{e})^{\dagger}\,(\mathbf{B}\,\mathbf{e}) = 0$ one finds

$$(\mathbf{B}\,\mathbf{e})^{\dagger}\,(\mathbf{B}\,\mathbf{e}) = \mathbf{e}^{\dagger}\,\mathbf{B}^{\dagger}\,\mathbf{B}\,\mathbf{e} = \mathbf{e}^{\dagger}\,\mathbf{B}\,\mathbf{B}^{\dagger}\,\mathbf{e} = |\mathbf{B}^{\dagger}\,\mathbf{e}|^2 = 0\,. \tag{A.64}$$

The above is equivalent to $\mathbf{B}^{\dagger}\,\mathbf{e} = 0$, namely, to $\mathbf{A}^{\dagger}\,\mathbf{e} = \lambda^*\,\mathbf{e}$ as anticipated.

Other properties refer to the eigenvalues of skew-Hermitian matrices and of unitary matrices. The nonvanishing eigenvalues of a skew-Hermitian matrix $\mathbf{A} = -\mathbf{A}^{\dagger}$ are imaginary; to show this, one left multiplies by $\mathbf{e}^{\dagger}$ the eigenvalue equation $\mathbf{A}\,\mathbf{e} = \lambda\,\mathbf{e}$, to find

$$\lambda\,|\mathbf{e}|^2 = \mathbf{e}^{\dagger}\,\mathbf{A}\,\mathbf{e} = -\mathbf{e}^{\dagger}\,\mathbf{A}^{\dagger}\,\mathbf{e} = -\,(\mathbf{A}\,\mathbf{e})^{\dagger}\,\mathbf{e} = -\,(\lambda\,\mathbf{e})^{\dagger}\,\mathbf{e} = -\lambda^*\,|\mathbf{e}|^2\,, \tag{A.65}$$

with $|\mathbf{e}|^2 \neq 0$. It follows $\lambda = -\lambda^*$, namely, λ is either equal to zero or to an imaginary number.

If $\mathbf{A}$ is unitary, $\mathbf{A}^\dagger \mathbf{A} = \mathbf{I}$, the modulus of its eigenvalues is equal to unity. To show this, one left multiplies by $\mathbf{A}^\dagger$ the eigenvalue equation $\mathbf{A}\,\mathbf{e} = \lambda\,\mathbf{e}$ and exploits the property of normal matrices demonstrated with (A.64), to find

$$\mathbf{e} = \lambda\,\mathbf{A}^\dagger\,\mathbf{e} = \lambda\,\lambda^*\,\mathbf{e}\,, \tag{A.66}$$

with $\mathbf{e} \neq 0$. It follows $|\lambda|^2 = 1$.

Permutation Matrices

A given set of row and column exchanges to which $\mathbf{A}$ may be subjected is conveniently described by means of a *permutation matrix* $\mathbf{P}$ of order N; in each row and in each column of a permutation matrix there is one and only one entry equal to 1, while all the remaining entries are equal to 0; for instance, a possible permutation matrix of order 3 is

$$\mathbf{P} = \begin{bmatrix} 0\,1\,0 \\ 1\,0\,0 \\ 0\,0\,1 \end{bmatrix}\,. \tag{A.67}$$

If $\mathbf{A}$ is a 3×3 matrix, then $\mathbf{A}\,\mathbf{P}$, with $\mathbf{P}$ given by (A.67), is the 3×3 matrix obtained by exchanging the first two columns of $\mathbf{A}$, while $\mathbf{P}\,\mathbf{A}$ is the 3×3 matrix obtained by exchanging the first two rows of $\mathbf{A}$. It is easily found that $|\det \mathbf{P}| = 1$ for all permutation matrices; it follows that permutation matrices are non-singular.

A.11.3 Similarity Transformations

Let $\mathbf{V}$ be a square matrix of order N, and let $\mathbf{G}$ be a non-singular square matrix, still of order N. Using the above, one constructs a third matrix $\mathbf{G}^{-1}\,\mathbf{V}\,\mathbf{G}$. This operation, that applies to both real and complex matrices, is called *similarity transformation*; the two matrices $\mathbf{V}$ and $\mathbf{G}^{-1}\,\mathbf{V}\,\mathbf{G}$ are called *similar*. If two matrices are similar, they have the same determinant and the same eigenvalues. The first property is due to the fact that the determinant of a product is equal to the product of determinants,

$$\det(\mathbf{G}^{-1}\,\mathbf{V}\,\mathbf{G}) = \det(\mathbf{G}^{-1})\,\det(\mathbf{V})\,\det(\mathbf{G}) = \det(\mathbf{G}^{-1}\mathbf{G})\,\det(\mathbf{V}) = \det(\mathbf{V})\,, \tag{A.68}$$

The second property derives from the eigenvalue equation for $\mathbf{V}$. Letting λ_r be an eigenvalue and $\mathbf{I}$ the identity matrix, consider the matrix

$$\mathbf{G}^{-1}\,\mathbf{V}\,\mathbf{G} - \lambda_r\,\mathbf{I} = \mathbf{G}^{-1}\,\mathbf{V}\,\mathbf{G} - \lambda_r\,\mathbf{G}^{-1}\,\mathbf{I}\,\mathbf{G} = \mathbf{G}^{-1}\,(\mathbf{V} - \lambda_r\,\mathbf{I})\,\mathbf{G}\,, \tag{A.69}$$

where $\mathbf{I} = \mathbf{G}^{-1}\,\mathbf{G} = \mathbf{G}^{-1}\,\mathbf{I}\,\mathbf{G}$ has been used. From (A.68) it follows $\det(\mathbf{G}^{-1}\,\mathbf{V}\,\mathbf{G} - \lambda_r\,\mathbf{I}) = \det(\mathbf{V} - \lambda_r\,\mathbf{I})$. By definition, λ_r makes the right-hand side of the above to

vanish for all r; it follows that $\mathbf{V}$ and $\mathbf{G}^{-1}\,\mathbf{V}\,\mathbf{G}$ have the same eigenvalues. From the identity $\mathbf{V}\,\mathbf{e} = \mathbf{V}\,\mathbf{G}\,\mathbf{G}^{-1}\,\mathbf{e}$, which holds for any vector $\mathbf{e}$ of length N, it follows for the rth eigenvalue $\mathbf{G}^{-1}\,\mathbf{V}\,\mathbf{G}\,\mathbf{G}^{-1}\,\mathbf{e}_r = \lambda_r\,\mathbf{G}^{-1}\,\mathbf{e}_r$; as a consequence, $\mathbf{G}^{-1}\,\mathbf{e}_r$ is an eigenvector of $\mathbf{G}^{-1}\,\mathbf{V}\,\mathbf{G}$ for all r.

A.11.4 Doubly Stochastic Matrices

Consider a set of M square matrices of order M, $\mathbf{S}_1, \ldots, \mathbf{S}_M$, and a set of M real, nonnegative numbers θ_k such that $\theta_1 + \ldots + \theta_M = 1$. The matrix

$$\mathbf{S} = \sum_{k=1}^{M} \theta_k \mathbf{S}_k \tag{A.70}$$

is called *convex combination* of the $\mathbf{S}_k$ matrices.

The following theorem is easily proved: if the matrices $\mathbf{S}_k$ are doubly stochastic,[7] then $\mathbf{S}$ is doubly stochastic as well. In fact from the definition of $\mathbf{S}$ it is $(\mathbf{S})_{ij} = \sum_{k=1}^{M} \theta_k (\mathbf{S}_k)_{ij}$ whence, adding the terms row-wise,

$$\sum_{j=1}^{M} (\mathbf{S})_{ij} = \sum_{k=1}^{M} \theta_k \sum_{j=1}^{M} (\mathbf{S}_k)_{ij} = \sum_{k=1}^{M} \theta_k = 1\,. \tag{A.71}$$

The same result is obtained when summing column-wise. As permutation matrices are doubly stochastic, from the above theorem the special case follows: a convex combination of permutation matrices is a doubly stochastic matrix. The inverse property also holds: a doubly stochastic matrix is a convex combination of permutation matrices [11].

A.11.5 Diagonally Dominant Matrices and Irreducible Matrices

The following definitions are of use. Let $\mathbf{A}$ be a real, square matrix of order N, whose entries are a_{ij}; its main diagonal is said to be *strongly dominant* if

$$|a_{ii}| > \sum_{j \neq i} |a_{ij}|\,, \qquad i = 1\,, \ldots\,, N\,. \tag{A.72}$$

[7] The definition of doubly stochastic matrix is given in Sect. 7.6.1.

Conversely, if (A.72) holds for at least one index i while, for all indices for which (A.72) does not hold it is

$$|a_{ii}| \geq \sum_{j \neq i} |a_{ij}| \,, \tag{A.73}$$

the diagonal is said to be *weakly dominant*.

A square matrix $\mathbf{A}$ of order N is called *reducible* if a permutation matrix $\mathbf{P}$ exists such that

$$\mathbf{P}\,\mathbf{A}\,\mathbf{P}^{-1} = \begin{bmatrix} \mathbf{F} & \mathbf{O} \\ \mathbf{H} & \mathbf{G} \end{bmatrix} , \tag{A.74}$$

where $\mathbf{F}$, $\mathbf{G}$ are square matrices, $\mathbf{O}$ is a matrix whose entries are equal to zero, and $\mathbf{P}^{-1}$ is the inverse of $\mathbf{P}$. If no permutation matrix exists that makes (A.74) to hold, then $\mathbf{A}$ is *irreducible*. If $\mathbf{A}$ is reducible, let N_a be the order of $\mathbf{F}$ and $N_b = N - N_a$ that of $\mathbf{G}$; then, assume that an algebraic system $\mathbf{A}\,\mathbf{u} = \mathbf{b}$ is to be solved, with $\mathbf{u}$, $\mathbf{b}$ the unknown and data vectors, respectively. Left multiplying the system by $\mathbf{P}$, and observing that $\mathbf{P}^{-1}\,\mathbf{P} = \mathbf{I}$, with $\mathbf{I}$ the identity matrix, yields $\mathbf{P}\,\mathbf{A}\,\mathbf{P}^{-1}\,\mathbf{P}\,\mathbf{u} = \mathbf{P}\,\mathbf{b}$. Defining the vectors $\mathbf{w} = \mathbf{P}\,\mathbf{u}$ and $\mathbf{g} = \mathbf{P}\,\mathbf{b}$, let $\mathbf{w}_a$ and $\mathbf{g}_a$ be the vectors made of the first N_a entries of $\mathbf{w}$ and $\mathbf{g}$, respectively, and let $\mathbf{w}_b$ and $\mathbf{g}_b$ be the vectors made of the remaining N_b entries. The algebraic system to be solved then becomes

$$\mathbf{F}\,\mathbf{w}_a = \mathbf{g}_a \,, \qquad \mathbf{H}\,\mathbf{w}_a + \mathbf{G}\,\mathbf{w}_b = \mathbf{g}_b \,. \tag{A.75}$$

If $\mathbf{A}$ is non-singular, $\mathbf{P}\,\mathbf{A}\mathbf{P}^{-1}$ is non-singular as well,[8] whence $\mathbf{w}$ is unique. It follows that the solution of the first equation in (A.75), whose unknown vector is made of the first N_a entries of $\mathbf{w}$, is also unique; as a consequence, $\mathbf{F}$ is non-singular. Solving for $\mathbf{w}_a$ the first equation in (A.75), and replacing the result in the second one, yields an algebraic system in the unknown vector $\mathbf{w}_b$. By the same token one finds that $\mathbf{G}$ is also non-singular, whence $\mathbf{w}_b$ is found. Thus, the solution of the original algebraic system $\mathbf{A}\,\mathbf{u} = \mathbf{b}$ of order N has been reduced to that of two systems of lower order.

The irreducibility property is amenable to an interesting graphic representation: given an $N \times N$ matrix, with $N > 1$, choose N different points of the plane and arbitrarily number them from 1 to N; then, for each nonvanishing entry a_{ij} of the matrix, $j \neq i$, draw an arc oriented from the ith to the jth point.[9] The construction yields an *oriented graph*; the latter is called *connected* if, for any choice of two different points i, j, one or both of the following occur: A) there exists an arc that connects points i and j; B) there exists an *oriented path*, made of more than one arc,

[8] If $\mathbf{A}$, $\mathbf{B}$ are two square matrices of order N, it is $\det(\mathbf{A}\,\mathbf{B}) = (\det \mathbf{A})\,(\det \mathbf{B})$. From $\mathbf{P}\,\mathbf{P}^{-1} = \mathbf{I}$ it follows $\pm 1 \times \det(\mathbf{P}^{-1}) = 1$, namely, $\det \mathbf{P} = \det(\mathbf{P}^{-1})$; thus, $\det(\mathbf{P}\,\mathbf{A}\mathbf{P}^{-1}) = \det \mathbf{A}$.

[9] The diagonal entries a_{ii} are of no interest in this construction; for the sake of completeness one might draw a closed arc at each point i for which $a_{ii} \neq 0$.

that connects point i to point α, point α to β, ..., point μ to j. It can be shown that if the oriented graph of $\mathbf{A}$ is connected, then $\mathbf{A}$ is irreducible, and vice versa [147, Par. 2.5].

A.11.6 Properties of Diagonally Dominant Matrices

Still considering a real square matrix $\mathbf{A}$ of order N, and basing upon the definitions given in Sect. A.11.5, the following theorems are now proven:

1. *If the main diagonal of* $\mathbf{A}$ *is strongly dominant, then* $\mathbf{A}$ *is non-singular.*
2. *If* $\mathbf{A}$ *is irreducible and its main diagonal is weakly dominant, then* $\mathbf{A}$ *is non-singular and all entries* a_{ii} *of the main diagonal are different from zero.*

The demonstration of the first theorem starts from (A.72), which shows that $a_{ii} \neq 0$, $i = 1, \ldots, N$; next, consider the auxiliary, homogeneous system $\mathbf{A}\,\mathbf{v} = 0$, whose ith row reads $a_{ii}\, v_i + \sum_{j \neq i} a_{ij}\, v_j = 0$. Letting $b_{ij} = 0$ for $j = i$ and $b_{ij} = -a_{ij}/a_{ii}$ for $j \neq i$ transforms the latter into

$$v_i = \sum_{j=1}^{N} b_{ij}\, v_j\,, \qquad i = 1\,, \ldots\,, N\,. \tag{A.76}$$

Using the new symbols transforms (A.72) into $\sum_{j=1}^{N} |b_{ij}| < 1$ for $i = 1, \ldots, N$. Now, define $M = \max_{1 \leq i \leq N} |v_i|$, and let k be an index such that $|v_k| = M$; thus, fixing $i = k$ in (A.76),

$$M = |v_k| = \left| \sum_{j=1}^{N} b_{kj}\, v_j \right| \leq \sum_{j=1}^{N} |b_{kj}|\, |v_j|\,. \tag{A.77}$$

If $\mathbf{A}$ were singular, there would be a nontrivial solution to $\mathbf{A}\,\mathbf{v} = 0$, whence it would be $M > 0$. In this case, multiplying by M both sides of $\sum_{j=1}^{N} |b_{ij}| < 1$, letting $i = k$, and subtracting from (A.77) would provide

$$\sum_{j=1}^{N} |b_{kj}|\, (|v_j| - M) > 0\,. \tag{A.78}$$

This inequality is false because, due to the definition of M, the quantity in parentheses in (A.78) is non-positive for all indices. It follows $M = 0$, namely, the only possible solution of the homogeneous system $\mathbf{A}\,\mathbf{v} = 0$ is the trivial one, whence $\mathbf{A}$ is non-singular.

The demonstration of the second theorem starts from the trivial case $N = 1$; it must be $|a_{11}| > 0$ due to the definition of weakly dominant diagonal, whence

the theorem for the trivial case is proven. Consider now the nontrivial case $N > 1$ and assume that $a_{ii} = 0$ for some index i; from (A.73) it follows $a_{ij} = 0$ for all j, namely, all entries of the i row vanish. By suitable permutations it is then possible to transform this row into the first one; comparing with (A.74) shows that $\mathbf{A}$ is reducible, contrary to the hypothesis. The rest of the proof follows the same reasoning as for the first theorem: the condition that the main diagonal is weakly dominant prescribes that the strict inequality $\sum_{j=1}^{N} |b_{ij}| < 1$ holds for at least one index i, whereas for all indices i for which the strict inequality does not hold it is

$$\sum_{j=1}^{N} |b_{ij}| \leq 1 \,. \tag{A.79}$$

If $\mathbf{A}$ were singular it would be $M > 0$, so that (A.78) would be replaced by

$$\sum_{j=1}^{N} |b_{kj}| \, (|v_j| - M) \geq 0 \,, \tag{A.80}$$

where M and k have the same meaning as in the first theorem. Due to the definition of M, the quantity in parentheses in (A.80) is non-positive for all indices; it follows that $|v_j| = M$ for all values of j such that $b_{kj} \neq 0$. On the other hand, $\mathbf{A}$ is irreducible, whence (remembering the construction of the oriented graph illustrated above) for any pair $j \neq k$ it is $b_{kj} \neq 0$, or there exist indices $\alpha, \beta, \ldots, \mu$ such that $b_{i\alpha} \, b_{\alpha\beta} \ldots b_{\mu j} \neq 0$. Due to (A.80), this is equivalent to $|v_j| = M$ or $|v_\alpha| = |v_\beta| = \ldots = |v_j| = M$, both of which imply $|v_j| = M$ for all $j \neq k$. Now, let m be one of the indices such that

$$\sum_{j=1}^{N} |b_{mj}| < 1 \,. \tag{A.81}$$

On the other hand, it is also

$$M = |v_m| \leq \sum_{j=1}^{N} |b_{mj}| \, |v_j| = \sum_{j=1}^{N} |b_{mj}| \, M \,. \tag{A.82}$$

The two relations (A.81), (A.82) are incompatible for $M > 0$; it follows $M = 0$, namely, also in this case the only possible solution of the homogeneous system $\mathbf{A}\,\mathbf{v} = 0$ is the trivial one, whence $\mathbf{A}$ is non-singular.

A.11.7 Solution of a Tridiagonal Algebraic System

In problems deriving from the numerical discretization of the semiconductor equations in one dimension (Sect. A.13), one must solve an algebraic system $\mathbf{A}\,\mathbf{x} = \mathbf{b}$ made of N equations in N unknowns, whose matrix $\mathbf{A}$ has a *tridiagonal form*, namely,

$$\mathbf{A} = \begin{bmatrix} r_1 & z_1 & 0 & \dots & 0 \\ w_2 & r_2 & z_2 & \dots & 0 \\ 0 & w_3 & r_3 & z_3 & \dots \\ & & \vdots & & \\ 0 & \dots & w_{N-1} & r_{N-1} & z_{N-1} \\ 0 & \dots & 0 & w_N & r_N \end{bmatrix} . \tag{A.83}$$

It is assumed that all terms w_i, r_i, z_i are different from zero, so that the matrix is irreducible (Sect. A.11.5), and also that the main diagonal is dominant, namely, $|r_1| > |z_1|$, $|r_N| > |z_N|$, and $|r_i| > |z_i| + |w_i|$ for $i = 2, \dots, N-1$. It follows that matrix $\mathbf{A}$ is non-singular (Sect. A.11.6).

L U Decomposition

One of the possible methods for solving an algebraic system (not necessarily tridiagonal) is the so-called $\mathbf{A} = \mathbf{L}\,\mathbf{U}$ *decomposition*, where $\mathbf{L}$ is a *lower-triangular matrix* and $\mathbf{U}$ an *upper-triangular matrix*.[10] For a matrix not having a special structure, the decomposition is in general expensive from the computational viewpoint; however, for a tridiagonal matrix it is simple, and the form of $\mathbf{L}$ and $\mathbf{U}$ turns out to be:

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ \lambda_2 & 1 & 0 & \dots & 0 \\ 0 & \lambda_3 & 1 & 0 & \dots \\ & & \vdots & & \\ 0 & \dots & \lambda_{N-1} & 1 & 0 \\ 0 & \dots & 0 & \lambda_N & 1 \end{bmatrix} , \qquad \mathbf{U} = \begin{bmatrix} g_1 & z_1 & 0 & \dots & 0 \\ 0 & g_2 & z_2 & \dots & 0 \\ 0 & 0 & g_3 & z_3 & \dots \\ & & \vdots & & \\ 0 & \dots & 0 & g_{N-1} & z_{N-1} \\ 0 & \dots & 0 & 0 & g_N \end{bmatrix} . \tag{A.84}$$

[10] A lower(upper)-triangular matrix is a square matrix in which all entries above (below) the main diagonal are equal to zero.

It follows that the decomposition requires only to determine the entries λ_i and g_i. The determinant of $\mathbf{L}$ is equal to unity, that of $\mathbf{U}$ is equal to $g_1 \dots g_N$. It can be shown that $\mathbf{U}$ also is non-singular; as a consequence the diagonal entries g_i are nonzero.[11]

The relations between the entries of $\mathbf{A}$ and those of $\mathbf{L}$ and $\mathbf{U}$ are found from (A.83), (A.84) by performing the matrix multiplication. Considering the first row of $\mathbf{A}$ one finds $r_1 = g_1$; the relations for the other rows are $r_i = \lambda_i z_{i-1} + g_i$ and $w_i = \lambda_i g_{i-1}$, with $i = 2, \dots, N$. Next, replacing $\mathbf{A}$ with $\mathbf{L}\,\mathbf{U}$ in the original system $\mathbf{A}\,\mathbf{x} = \mathbf{b}$, one splits the latter into two subsystems:

$$\mathbf{L}\,\mathbf{e} = \mathbf{b}\,, \qquad \mathbf{U}\,\mathbf{x} = \mathbf{e}\,. \tag{A.85}$$

The first of (A.85) is equivalent to $e_1 = b_1$ and $\lambda_i e_{i-1} + e_i = b_i$, with $i = 2, \dots, N$, that is, to a sequence of steps that provides the entries e_i by *forward substitution*, while the second of (A.85) is equivalent to $g_N x_N = e_N$ and $g_i x_i + z_i x_{i+1} = e_i$, with $i = N-1, \dots, 1$, that is, to a sequence of steps that provides the entries x_i by *backward substitution*. In the algorithm used to solve (A.85) it is not necessary to store the vector of entries λ_i: unless such a vector is necessary for other purposes, one can use a single scalar λ. Proceeding in this way one finds $g_1 = r_1$ and $e_1 = b_1$; then, for $i = 2, \dots, N$,

$$\lambda = \frac{g_{i-1}}{w_i}\,, \qquad g_i = r_i - \lambda\, z_{i-1}\,, \qquad e_i = b_i - \lambda\, e_{i-1}\,, \tag{A.86}$$

that combines the calculation of the entries g_i with the forward substitution that yields the entries e_i. Then, the backward substitution proceeds as

$$x_N = \frac{e_N}{g_N}\,, \qquad x_i = \frac{e_i - z_i\, x_{i+1}}{g_i}\,, \qquad i = N-1, \dots, 1\,. \tag{A.87}$$

The divisions by w_i or g_i in (A.86), (A.87) are possible because it is $w_i \neq 0$ and $g_i \neq 0$ as shown above. The method requires the storage of two vectors, $\mathbf{g}$ and $\mathbf{e}$, besides those necessary for the storage of $\mathbf{A}$ and $\mathbf{b}$. The computational cost of the solution of $\mathbf{A}\,\mathbf{x} = \mathbf{b}$ is found from (A.86), (A.87), and turns out to be $6\,(N-1)$ multiplications and $3\,(N-1)$ additions.[12]

A Cheaper Solution Scheme

The tridiagonal structure lends itself to another solution methods, that has some advantages over the $\mathbf{L}\,\mathbf{U}$ decomposition. The method is outlined here, with reference

[11] From $\det \mathbf{L} = 1$ and $\mathbf{L}^{-1}\,\mathbf{L} = \mathbf{I}$ it follows in fact $\det \mathbf{U} = \det\,(\mathbf{L}^{-1}\,\mathbf{A}) = \det \mathbf{L}^{-1}\, \det \mathbf{A} = \det \mathbf{A}$.

[12] It is mentioned in Sect. A.13.1 that the cost of solving an order-N system is $C \simeq \text{const} \times N^a$, with a ranging between 2 and 3. This, however, refers to matrices having no special structure.

to the algebraic system deriving from the discretization of Poisson's equation; for consistency with the symbols of (A.110), here the unknown is indicated with **u** instead of **x**. To begin, one considers the simplest form of (A.110), where $a = 1$ and all elements are equal, $h_1 = \ldots = h_N = h$. Remembering that $\Omega_i = (h_i + h_{i+1})/2$, this yields

$$-u_{i-1} + 2\,u_i - u_{i+1} = c_i\,, \qquad c_i = h^2\,C_i\,, \qquad i = 1, 2, \ldots, N\,. \tag{A.88}$$

As a consequence, the entries of matrix **A** in (A.83) become $r_1 = \ldots = r_N = 2$ and $w_2 = \ldots w_N = z_1 = \ldots = z_{N-1} = -1$. The equations corresponding to nodes 1 and 2 read

$$u_1 - u_0 = c_1 + u_2 - u_1\,, \qquad u_2 - u_1 = c_2 + u_3 - u_2\,; \tag{A.89}$$

iterating (A.89) up to node $i - 1$ yields

$$u_{i-1} - u_{i-2} = c_{i-1} + u_i - u_{i-1}\,. \tag{A.90}$$

If relations like (A.89) and (A.90) are added up, the data c_i cumulate while the majority of the unknowns cancel each other; letting $p_i = \sum_{j=1}^{i} c_j$ for $i = 1, \ldots, N$, one finds in fact

$$u_1 - u_0 = p_{i-1} + u_i - u_{i-1}\,, \qquad i = 2, \ldots, N + 1\,. \tag{A.91}$$

Rewriting (A.91) for $i = 2, 3$ provides

$$u_1 - u_0 = p_1 + u_2 - u_1\,, \qquad u_1 - u_0 = p_2 + u_3 - u_2\,, \tag{A.92}$$

the first of which is obviously a replica of the first relation in (A.89). Again, if relations like (A.91) and (A.92) are added up, the terms p_i cumulate while the majority of the unknowns cancel each other; one finds

$$(i - 1)\,(u_1 - u_0) = y_{i-1} + u_i - u_1\,, \qquad i = 2, \ldots, N + 1\,, \tag{A.93}$$

where $y_1 = p_1$, $y_2 = p_1 + p_2$, and $y_i = \sum_{j=1}^{i} p_j$. As u_0 is a boundary condition, (A.93) contains the two unknowns u_1 and u_i; one now exploits the second boundary condition u_{N+1} by letting $i = N + 1$ in (A.93), thus providing a relation for u_1, namely, $N\,(u_1 - u_0) = y_N + u_{N+1} - u_1$. The latter is recast in a more compact form as

$$u_1 = u_0 + R\,, \qquad R = \frac{u_{N+1} - u_0 + y_N}{N + 1}\,. \tag{A.94}$$

Replacing (A.94) into (A.93) finally yields

$$u_i = u_0 + i\,R - y_{i-1}\,, \qquad i = 2, \ldots, N\,. \tag{A.95}$$

An example of FORTRAN programming of the algorithm is given in Table A.1, with $u_L = u_{N+1}$. The method requires the storage of one vector, $\mathbf{y} = (y_1, y_2, \ldots)$; the storage of $\mathbf{p} = (p_1, p_2, \ldots)$ is not necessary, as shown in the table. The computational cost of the solution is found from (A.95) and from the calculation of p_i and y_i; it turns out to be $N - 1$ multiplications and $4\,(N - 1)$ additions. Recasting (A.93) as $u_i = u_0 + i\,(u_1 - u_0) - y_{i-1}$ one finds the discrete counterpart of

$$u(x) = u_0 + u_0'\,x - \int_0^x (x - s)\,C(s)\,\mathrm{d}s\,, \tag{A.96}$$

which is in turn the solution of $-u'' = C$ in the continuous case.[13] In fact, one finds

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n-1} \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 2 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ n-1 & n-2 & \cdots & 0 \\ n & n-1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_{n-1} \\ c_n \end{bmatrix}, \tag{A.97}$$

where the matrix corresponds to the $x - s$ kernel of (A.96).

Table A.1 FORTRAN program for solving a tridiagonal system with (A.95)

```
c
      q = y(1) = c(1)
c
      do 100, i = 2,N
         q = q + c(i)
         y(i) = y(i-1) + q
100   continue
c
      R = ( uL - u0 + y(N) ) / (N+1)
      u(1) = u0 + R
c
      do 200, i = 2,N
         u(i) = u0 + i * R - y(i-1)
200   continue
c
```

[13] Solution (A.96) holds when the given boundary conditions are u_0 and u_0'; if, instead, u_0 and $u_L = u(x = L)$ are given, one determines u_0' in terms of u_L by letting $x = L$ in (A.96).

The procedure leading to (A.95) is readily extended to the case of a nonuniform grid and of position-dependent material properties. The starting point is (A.109), where $\Omega_i\, C_i$ may be replaced by another expression if a different interpolation scheme is used prior to discretization (e.g., the Numerov process of Sect. A.13.3). Rewriting $S_i - S_{i+1} = \Omega_i\, C_i$ with $i = 1, 2, \ldots$ and adding up the results yield $S_1 - S_{i+1} = P_i$, with $P_i = \sum_{j=1}^{i} \Omega_j\, C_j$ and $i \leq N$. Still from (A.109) one obtains $u_{i+1} - u_i = (h_{i+1}/a_{i+1})\, S_{i+1}$; adding up such relations yields

$$u_{i+1} - u_0 = \frac{h_1}{a_1}\, S_1 + \frac{h_2}{a_2}\, (S_1 - P_1) + \cdots + \frac{h_{i+1}}{a_{i+1}}\, (S_1 - P_i)\,. \tag{A.98}$$

Letting $Q_i = \sum_{j=1}^{i} (h_j/a_j)$, $Y_1 = 0$, and $Y_{i+1} = \sum_{j=2}^{i+1} (h_j/a_j)\, P_{i-1}$ transforms (A.98) into

$$u_{i+1} = u_0 + Q_{i+1}\, S_1 - Y_{i+1}\,, \qquad i = 0, \ldots, N\,, \tag{A.99}$$

with u_0 a boundary condition. Using the other boundary condition u_{N+1}, one extracts S_1 from (A.99) by letting $i = N$; it is found $S_1 = (u_{N+1} - u_0 + Y_{N+1})/Q_{N+1}$, with $Q_{N+1} \geq \text{const} > 0$ by construction. In this way, all terms at the right-hand side of (A.99) are defined; as a consequence, (A.99) provides the solution sought. The method requires the storage of two vectors, $\mathbf{Q} = (Q_1, Q_2, \ldots)$ and $\mathbf{Y} = (Y_1, Y_2, \ldots)$, while no vector is necessary to store $\mathbf{P} = (P_1, P_2, \ldots)$. The computational cost is $3\,(N-1)$ multiplications and $6\,(N-1)$ additions, to be compared with the $6\,(N-1)$ multiplications and $3\,(N-1)$ additions of the $\mathbf{L}\,\mathbf{U}$ decomposition. Another advantage of the method illustrated in this section is that each nodal value (A.95) or (A.99) can be calculated independently from the others; in contrast, to calculate the ith unknown with the backward substitution (A.87) it is necessary to calculate first all the other unknowns whose index is larger than i.

A.12 Wronskian Determinant

The Wronskian determinant provides the condition of linear independence of functions [66, Sect. 5.2]. Although its properties hold for any number of functions, they will be discussed here for the case of two functions only, say, u and v defined on some interval of the independent variable x. It is convenient to seek for the condition of linear dependence first. If u, v are linearly dependent, then two nonvanishing constants c_1, c_2 exist such that

$$c_1\, u + c_2\, v = 0 \tag{A.100}$$

for all x in the interval. If (A.100) holds, it is easily found that both c_1 and c_2 must differ from zero. Also, as the function at the left-hand side of (A.100) vanishes

identically, its derivative vanishes as well. Such a derivative exists because u and v are supposed to be solutions of a second-order differential equation. Then,

$$c_1\, u' + c_2\, v' = 0 \tag{A.101}$$

for all x in the interval. As (A.100, A.101) hold together, for all x the two constants c_1, c_2 are the nontrivial solution of a homogeneous algebraic system. Now, if the nontrivial solution of the algebraic system exists for all x, the determinant $W = u\, v' - u'\, v$ must vanish identically. That is, the condition $W = 0$ (identically) is necessary for the linear dependence of u, v. As a consequence, the condition $W \neq 0$ (identically) is sufficient for the linear independence of u, v.

A.13 Numerical Solution of the Semiconductor Equations

As mentioned in Sect. 22.11.7, several steps in the design of integrated circuits are carried out by computer simulation. In many cases, this entails the solution of differential equations; as the form of the latter is seldom amenable to an analytic solution, one must resort to numerical methods. In this section, a brief account is given of methods useful for solving the drift-diffusion model in a semiconductor region, in the form worked out in Sect. 19.5.5; the equations read

$$\operatorname{div} \mathbf{D} = q\,(p - n + N)\,, \qquad \mathbf{D} = -\varepsilon_{\mathrm{sc}}\, \operatorname{grad} \varphi = \varepsilon_{\mathrm{sc}}\, \mathbf{E}\,, \tag{A.102}$$

$$\frac{\partial n}{\partial t} - \frac{1}{q}\, \operatorname{div} \mathbf{J}_n = W_n\,, \qquad \mathbf{J}_n = q\, \mu_n\, n\, \mathbf{E} + q\, D_n\, \operatorname{grad} n\,, \tag{A.103}$$

$$\frac{\partial p}{\partial t} + \frac{1}{q}\, \operatorname{div} \mathbf{J}_p = W_p\,, \qquad \mathbf{J}_p = q\, \mu_p\, p\, \mathbf{E} - q\, D_p\, \operatorname{grad} p\,. \tag{A.104}$$

Apart from the constants q, $\varepsilon_{\mathrm{sc}}$, the model's coefficients are the mobilities μ_n, μ_p and the diffusion coefficients D_n, D_p; the latter are proportional to the corresponding mobilities through the Einstein relations (19.102). The data are the dopant distribution N and the generation-recombination rates W_n, W_p. The model (A.102), (A.103), (A.104) can be viewed as a set of six equations of the first order with respect to the spatial variables; alternatively, inserting the expression of $\mathbf{D}$, $\mathbf{J}_n$, or $\mathbf{J}_p$, appearing on the right, into the divergence operator belonging to the same line, yields a set of three equations of the second order. The latter form is the most common one for the application of numerical methods, because only three scalar unknowns appear in it (φ, n, and p). The equations must be supplemented with suitable boundary conditions. In dynamic conditions, the equations are also of the first order with respect to time, and the initial condition must be supplemented as well.

A suitable definition of the symbols recasts the set of equations (A.102), (A.103), (A.104) in the more compact form

$$-\operatorname{div}\mathbf{S} = C\,, \qquad \mathbf{S} = a\,\operatorname{grad} u + \mathbf{b}\,u\,. \tag{A.105}$$

In fact, (A.105) reduces to (A.102) by letting $u = q\,\varphi/(k_B\,T)$, $a = \varepsilon_{\rm sc}$, $\mathbf{b} = 0$, $\mathbf{S} = -q\,\mathbf{D}/(k_B\,T)$, $C = q\,\varrho/(k_B\,T)$; in turn, it reduces to (A.103) by letting $u = n$, $a = D_n$, $\mathbf{b} = \mu_n\,\mathbf{E}$, $\mathbf{S} = \mathbf{J}_n/q$, $C = W_n - \partial n/\partial t$; finally, it reduces to (A.104) by letting $u = p$, $a = -D_p$, $\mathbf{b} = \mu_p\,\mathbf{E}$, $\mathbf{S} = \mathbf{J}_p/q$, $C = -W_p + \partial p/\partial t$. As mentioned above, elimination of $\mathbf{S}$ transforms the pair of first-order equations (A.105) into a single, second-order equation, whose unknown is u.

A.13.1 Decoupled/Coupled Solution

Apart from the simpler case where the drift term is negligible, the problem to be solved is nonlinear due to the presence of the $n\,\mathbf{E}$, $p\,\mathbf{E}$ products in the transport equations. Also, the mobilities and generation-recombination terms depend in general on the problem's unknowns (besides depending on parameters like, e.g., lattice temperature); such dependences may be nonlinear as well. To handle the presence of nonlinear terms one may resort to different solution strategies; the simplest one, called *decoupled solution*, works as follows: a tentative solution $n^{(0)}(\mathbf{r})$, $p^{(0)}(\mathbf{r})$ is prescribed for the carrier concentrations, and (A.102) is solved to yield the first-iteration electric potential $\varphi^{(1)}(\mathbf{r})$. Next, (A.103) is solved for n using $\varphi^{(1)}$ and $p^{(0)}$ as data, to yield the first-iteration electron concentration $n^{(1)}$; finally, (A.104) is solved for p using $\varphi^{(1)}$ and $n^{(1)}$ as data, to yield the first-iteration hole concentration $p^{(1)}$. In general, the spatial dependence of $n^{(1)}$, $p^{(1)}$ differs from that of $n^{(0)}$, $p^{(0)}$; then, one solves again (A.102) using $n^{(1)}$, $p^{(1)}$ at the right-hand side, to find the second-iteration electric potential $\varphi^{(2)}(\mathbf{r})$, and so on. If the iterative scheme converges, the process is brought to an end when one or more suitable error indicators become lower than a prescribed limit; examples of such indicators are the norms of the increments between two successive iterations,

$$||\varphi^{(k+1)} - \varphi^{(k)}||\,, \qquad ||n^{(k+1)} - n^{(k)}||\,, \qquad ||p^{(k+1)} - p^{(k)}||\,, \tag{A.106}$$

with k the iteration index. The advantage of the decoupled solution lies in the fact that the problem is decomposed into simpler problems, whose computational cost is relatively small; on the other hand, the precision of the method is limited: from a qualitative standpoint the power of the method is of the first order, namely, the improvement in the solution is about one digit per iteration.

In contrast, the *coupled solution* prescribes the tentative solution $\varphi^{(0)}(\mathbf{r})$, $n^{(0)}(\mathbf{r})$, $p^{(0)}(\mathbf{r})$ to all unknowns, and linearizes the whole set of equations around the tentative solution. In this way, the actual unknowns of the linearized system are the

increments $\delta\varphi(\mathbf{r})$, $\delta n(\mathbf{r})$, $\delta p(\mathbf{r})$; once found, they are added to the tentative solution to yield the first-iteration solution:

$$\varphi^{(1)} = \varphi^{(0)} + \delta\varphi\,, \qquad n^{(1)} = n^{(0)} + \delta n\,, \qquad p^{(1)} = p^{(0)} + \delta p\,. \tag{A.107}$$

The norm of the increments is monitored at each iteration to check the convergence behavior. By way of example, consider the equilibrium condition, in which the set of equations (A.102), (A.103), (A.104) reduces to the Poisson equation alone, $-\varepsilon_{\rm sc}\,\nabla^2\varphi = \varrho$ where (compare with (18.56) and (18.57)) it is $\varrho = q\,[p(\varphi) - n(\varphi) + N]$. Linearizing the equation yields, for the kth iteration,

$$-\,\varepsilon_{\rm sc}\,\nabla^2\delta\varphi + q\left(\frac{\mathrm{d}n}{\mathrm{d}\varphi} - \frac{\mathrm{d}p}{\mathrm{d}\varphi}\right)^{(k)}\delta\varphi = \varepsilon_{\rm sc}\,\nabla^2\varphi^{(k)} + \varrho^{(k)}\,. \tag{A.108}$$

At convergence, both sides of (A.108) vanish; in fact, the increment $\delta\varphi$ tends to zero while $\varphi^{(k)}$ at the right-hand side tends to the equation's solution. The advantage of the coupled-solution method is that its power is of the second order; on the other hand, its computational cost is significantly higher than that of the decoupled solution.[14] As a rule of thumb, the more accurate coupled scheme is used when the solution is nearly reached, so that a small number of iterations is sufficient to complete the calculation; the cheaper decoupled scheme is used instead during the initial stages. It must be remarked that the convergence of the iteration methods (either coupled or decoupled) is not guaranteed: it depends on the coefficients of the equations and on the choice of the tentative solution. Although the discussion about the iteration methods is beyond the scope of this book, a hint about their general structure is given in Sect. B.8.

It is also important to note that the above description of the decoupled- and coupled-solution schemes is referred to the dependence of the unknowns on the spatial coordinates, at a given instant of time t. If the problem depends on time, when the spatial solution is completed, the process must be repeated after updating the time variable. Considering that the differential problem with respect to time is of the first order, one has to deal with an open integration. As before, the convergence of the integration process depends on the structure of the equations and is not guaranteed in general.

[14] As shown below, the differential equation to be solved is transformed into an algebraic system whose matrix has the size $N \times N$, with N the number of spatial points where the unknown function is sought. If $C(N)$ is the computational cost of solving an $N \times N$ algebraic system, the cost of the iteration in a decoupled solution is $3\,C(N)$. When, instead, the coupled solution is used, each iteration entails the solution of a $3N \times 3N$ algebraic system, thus yielding a cost per iteration equal to $C(3N)$. If the matrix has no special structure, the dependence of C on N is of the type $C \simeq \text{const} \times N^a$, with a ranging between 2 and 3 depending on the solution method and the type of matrix. It follows $C(3N)/[3\,C(N)] = 3^{a-1}$.

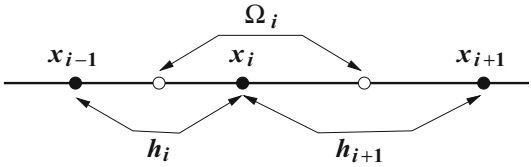


Fig. A.2 Illustration of the concepts of node, element, and cell used in the discretization of the semiconductor equations

A.13.2 Discretization Scheme in One Dimension

An example of the solution method of (A.105) in one dimension is given here. The solution is sought within a finite interval $[x_0, x_{N+1}]$, with prescribed boundary conditions $u_0 = u(x_0)$, $u_{N+1} = u(x_{N+1})$. The interval is subdivided into $N + 1$ regions (*elements*), whose end points are called *nodes*. The number of internal nodes is N, their positions are $x_1 < x_2 < \ldots < x_N$. The element bounded by x_i and x_{i+1} is indicated with h_{i+1}; the same symbol indicates the length of the element, $h_{i+1} = x_{i+1} - x_i$. Finally, a subdomain Ω_i, called *cell*, is associated with each node. For the internal nodes $x_1, \ldots, x_N$ the cell is bounded by $x_i - h_i/2$ and $x_i + h_{i+1}/2$. The same symbol is used to indicate also the cell length, $\Omega_i = (h_i + h_{i+1})/2$. The left boundary x_0 is associated with the cell Ω_0 of length $h_1/2$ placed on the right of x_0, while the right boundary x_{N+1} is associated with the cell Ω_{N+1} of length $h_{N+1}/2$ placed on the left of x_{N+1}. This procedure yields the same tessellation (or *grid*) as that used in Sect. 13.6.6 with reference to the Schrödinger equation (Fig. A.2). The lengths of the elements are not necessarily equal to each other; indeed, the choice of the number of nodes, and of their placement along the domain, is by no means an easy task. The general rule for constructing a grid is that the nodes should be denser where the curvature of the unknown function is larger; unfortunately, the information about curvature is not available *a priori*. An excessively large number of nodes makes the solution procedure more expensive; nevertheless, a uniform grid might in some cases be preferable to a coarser, nonuniform grid, because uniform grids lend themselves to the application of more efficient discretization methods (A.13.3).

Discretization of the Poisson Equation

To proceed, consider (A.105) in one dimension with $\mathbf{b} = 0$, namely, the Poisson equation made of the pair $-\mathrm{d}S/\mathrm{d}x = C$ and $S = a\,\mathrm{d}u/\mathrm{d}x$. Aim of the calculation is to determine u at the nodes, given C and the boundary conditions. In the Poisson equation, coefficient a represents the material's permittivity; if the domain $[x_0, x_{N+1}]$ extends over different materials, the permittivity changes. To consider this possibility, one must place a node at the interface between the two materials, and ascribe the correct value to a on the two sides of the interface. More generally, one

can also assume that a takes a different value over each element:[15] to indicate this, one associates the element's index with a. Now, the integration of $-\mathrm{d}S/\mathrm{d}x = C$ over the ith cell yields $S_i - S_{i+1} = \int_{\Omega_i} C\,\mathrm{d}x$, where the integral at the right-hand side can be calculated to various degrees of approximation, depending on the information available about C; the zero-order approximation consists in replacing the integral with the product $\Omega_i\, C_i$, with C_i the nodal value of C. The next step is the elimination of S by means of $S = a\,\mathrm{d}u/\mathrm{d}x$; to this purpose, one must express S_i and S_{i+1} in terms of the nodal values of u. The simplest way is using the difference quotients over each element; in summary, one obtains

$$S_i - S_{i+1} \simeq \Omega_i\, C_i\,, \qquad S_i \simeq a_i\,\frac{u_i - u_{i-1}}{h_i}\,, \qquad S_{i+1} \simeq a_{i+1}\,\frac{u_{i+1} - u_i}{h_{i+1}}. \tag{A.109}$$

Replacing S_i, S_{i+1} and collecting the nodal values of u yield

$$-\frac{a_i}{h_i}\,u_{i-1} + \left(\frac{a_i}{h_i} + \frac{a_{i+1}}{h_{i+1}}\right) u_i - \frac{a_{i+1}}{h_{i+1}}\,u_{i+1} = \Omega_i\, C_i\,, \qquad i = 1, 2, \ldots, N, \tag{A.110}$$

namely, an algebraic system made of N equations in the N unknowns $u_1, \ldots, u_N$. The $N \times N$ system matrix is given a more compact form by introducing the symbols $\alpha_i = \beta_i + \gamma_i$, where $\beta_i = a_i/h_i$ and $\gamma_i = a_{i+1}/h_{i+1}$, with $i = 1, \ldots, N$. The first ($i = 1$) and last ($i = N$) equations embed the boundary conditions; the latter can be moved to the right-hand side, to yield

$$\alpha_1\, u_1 - \gamma_1\, u_2 = \Omega_1\, C_1 + \beta_1\, u_0\,, \quad -\beta_N\, u_{N-1} + \alpha_N\, u_N = \Omega_N\, C_N + \gamma_N\, u_{N+1}. \tag{A.111}$$

From the form of (A.110) it follows that the entries $\alpha_1, \ldots, \alpha_N$ are placed along the main diagonal of the matrix, while $\beta_2, \ldots, \beta_N$ are placed right below the main diagonal and $\gamma_1, \ldots, \gamma_{N-1}$ are placed right above it; as all the remaining entries vanish, the matrix is *tridiagonal*. Also, from the definition it follows $\gamma_i = \beta_{i+1}$, namely, the matrix is symmetric. In conclusion, letting

$$\mathbf{A} = \begin{bmatrix} \alpha_1 & -\gamma_1 & 0 & \cdots & 0 & 0 & 0 \\ -\beta_2 & \alpha_2 & -\gamma_2 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -\beta_{N-1} & \alpha_{N-1} & -\gamma_{N-1} \\ 0 & 0 & 0 & \cdots & 0 & -\beta_N & \alpha_N \end{bmatrix}, \qquad \mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_{N-1} \\ u_N \end{bmatrix}, \tag{A.112}$$

[15] Coefficient a eventually becomes a factor of the entries of the matrix into which the differential operator $-\mathrm{d}^2/\mathrm{d}x^2$ is transformed. As is apparent from the discussion in Sect. A.11, it is necessary that $|a| \geq \text{const} > 0$, namely, a must be either strictly positive or strictly negative. When the Poisson equation is concerned, this condition is fulfilled because a stands for permittivity. As for the drift-diffusion equation (discussed below), a is proportional to the carrier mobility: in this case, care must be taken not to use models in which the degradation due to parameters (like, e.g., dopant concentration) makes mobility to become arbitrarily small.

the algebraic system to be solved reads

$$\mathbf{A}\,\mathbf{u} = \mathbf{c} + \mathbf{e}\,, \qquad \mathbf{c} = \begin{bmatrix} \Omega_1\, C_1 \\ \Omega_2\, C_2 \\ \vdots \\ \Omega_{N-1}\, C_{N-1} \\ \Omega_N\, C_N \end{bmatrix}, \qquad \mathbf{e} = \begin{bmatrix} \beta_1\, u_0 \\ 0 \\ \vdots \\ 0 \\ \gamma_N\, u_{N+1} \end{bmatrix}. \tag{A.113}$$

Remembering the definitions given in Sect. A.11, the diagonal of $\mathbf{A}$ is weakly dominant (in fact, strict inequality holds only for the first and last rows). Also, $\mathbf{A}$ is irreducible: in fact, using for a moment the two-index notation, it is $-\gamma_i = (\mathbf{A})_{i,i+1} \neq 0$, $-\beta_{i+1} = (\mathbf{A})_{i+1,i} \neq 0$, with $i = 1, \ldots, N-1$; remembering the construction based on the oriented graph illustrated in Sect. A.11.5, one finds that all points of the graph are pairwise connected in both directions, whence the graph is connected. As a consequence of the second theorem in Sect. A.11.6, $\mathbf{A}$ is non-singular and the solution of (A.113) is unique.[16] It must be remarked that the tridiagonal structure of $\mathbf{A}$ is due to the fact that the nodes are numbered in a specific order. On the other hand, any permutation in the node numbering is equivalent to multiplying $\mathbf{A}$ by one or more permutation matrices and, as shown in Sect. A.11, a permutation multiplies $\det \mathbf{A}$ by ± 1. Thus, if $\det \mathbf{A} \neq 0$ for a specific node numbering, it will be so for any other numbering.

Discretization of the Linearized Poisson Equation

The solution method for the Poisson equation is readily extended to the linearized form (A.108); in fact one lets, as before, $a = \varepsilon_{\mathrm{sc}}$, $\mathbf{b} = 0$, and

$$u = \frac{q}{k_B\, T}\, \delta\varphi\,, \qquad \mathbf{S} = -\frac{q}{k_B\, T}\, \delta\mathbf{D}\,, \tag{A.114}$$

$$C = \frac{q}{k_B\, T} \left[\varepsilon_{\mathrm{sc}}\, \nabla^2 \varphi^{(k)} + \varrho^{(k)} \right], \qquad \chi = q \left(\frac{\mathrm{d}n}{\mathrm{d}\varphi} - \frac{\mathrm{d}p}{\mathrm{d}\varphi} \right)^{(k)}, \tag{A.115}$$

to find, in one dimension,

$$-\frac{\mathrm{d}S}{\mathrm{d}x} = C - \chi\, u\,, \qquad S = a\, \frac{\mathrm{d}u}{\mathrm{d}s}\,. \tag{A.116}$$

[16] The solution of (A.113), $\mathbf{u} = \mathbf{A}^{-1}\,\mathbf{c} + \mathbf{A}^{-1}\,\mathbf{e}$, is the sum of two contributions: the first of them derives only from the data, and the second one derives only from the boundary condition. Compare with the comments in Sect. 4.12.3.

In principle, the term χu could be incorporated into the definition of C; this, however, is not convenient because of its explicit dependence on the unknown. The discretization scheme of (A.116) leads again to an algebraic system of the form (A.112), (A.113), the only difference being in the definition of the diagonal entries of $\mathbf{A}$; one finds in fact $\alpha_i = \beta_i + \gamma_i + \Omega_i \chi_i$ where, as before, $\beta_i = a_i/h_i$ and $\gamma_i = a_{i+1}/h_{i+1}$, $i = 1, \ldots, N$.

An important observation is that $\chi_i > 0$; this is easily found in the approximation of parabolic bands and nondegenerate conditions, where the equilibrium concentrations are given by (18.62); one finds in fact $\mathrm{d}(n-p)/\mathrm{d}\varphi = q\,(n+p)/(k_B\,T) > 0$. The result is the same also in the general case (Sect. 18.5), where the equilibrium concentrations are given by

$$n = \int_{E_C}^{E_{CU}} \gamma(E)\, P(E,\varphi)\, \mathrm{d}E\,, \qquad p = \int_{E_{VL}}^{E_V} \gamma(E)\, [1 - P(E,\varphi)]\, \mathrm{d}E\,, \tag{A.117}$$

with $\gamma(E) > 0$ the density of states per unit volume and $0 < P(E,\varphi) < 1$ the position-dependent Fermi-Dirac statistics (18.54). In fact, the dependence on φ in (18.54) is such that $\mathrm{d}P/\mathrm{d}\varphi = q\,P\,(1-P)/(k_B\,T)$, whence

$$\frac{\mathrm{d}n}{\mathrm{d}\varphi} - \frac{\mathrm{d}p}{\mathrm{d}\varphi} = \int_{E_C}^{E_{CU}} \gamma(E)\, \frac{P\,(1-P)}{k_B\,T/q}\, \mathrm{d}E + \int_{E_{VL}}^{E_V} \gamma(E)\, \frac{P\,(1-P)}{k_B\,T/q}\, \mathrm{d}E > 0\,. \tag{A.118}$$

Observing that $\beta_i, \gamma_i > 0$ the addition of the positive term $\Omega_i\,\chi_i$ to the diagonal makes the latter strongly dominant (Sect. A.11.5). This does not change the nature of the system's matrix, which would be non-singular even if all χ_i vanished; however, when an algebraic system is tackled with an iterative-solution method, the more dominant the diagonal, the faster the convergence.

Discretization of the Drift-Diffusion Equation

Consider now the case of (A.105) with $\mathbf{b} \neq 0$, still in one dimension; this yields the pair of normalized continuity and drift-diffusion transport equations $-\mathrm{d}S/\mathrm{d}x = C$ and $S = a\,\mathrm{d}u/\mathrm{d}x + b\,u$. Aim of the calculation is again to determine u at the nodes, given C and the boundary conditions. As symbol u is used here to indicate the carrier concentration, a change in notation is necessary with respect to the case of Poisson's equation; specifically, one lets $\sigma = q\,\varphi/(k_B\,T)$ in the electron-transport equation, and $\sigma = -q\,\varphi/(k_B\,T)$ in the hole-transport equation: with this provision, both equations take the form $S = a\,(\mathrm{d}u/\mathrm{d}x - u\,\mathrm{d}\sigma/\mathrm{d}x)$. Integration of $-\mathrm{d}S/\mathrm{d}x = C$ over the ith cell yields, like in the case of Poisson's equation, $S_i - S_{i+1} \simeq \Omega_i\,C_i$. The transport equation is more suitably recast in the self-adjoint form $S\,\exp(-\sigma) = a\,\mathrm{d}\,[u\,\exp(-\sigma)]\,/\mathrm{d}x$; integration of the latter over the ith element h_i, with $S = S_i$ and $a = a_i$, provides

$$S_i\,Y_i \simeq a_i\,[u_i\,\exp(-\sigma_i) - u_{i-1}\,\exp(-\sigma_{i-1})]\,, \qquad Y_i = \int_0^{h_i} \exp(-\sigma)\,\mathrm{d}x > 0\,. \tag{A.119}$$

Letting $w_i = u_i\,\exp(-\sigma_i)$ gives the first relation in (A.119) the more compact form $S_i\,Y_i \simeq a_i\,(w_i - w_{i-1})$. Replacing S_i, S_{i+1} and collecting the nodal values of u yield

$$-\frac{a_i}{Y_i}\,w_{i-1} + \left(\frac{a_i}{Y_i} + \frac{a_{i+1}}{Y_{i+1}}\right) w_i - \frac{a_{i+1}}{Y_{i+1}}\,w_{i+1} = \Omega_i\,C_i\,, \qquad i = 1, 2, \ldots, N\,, \tag{A.120}$$

structurally identical to (A.110). The auxiliary unknown $w = u\,\exp(-\sigma)$ has a physical meaning: remembering in fact that the drift-diffusion equations in (A.103) and (A.104) hold within the approximations of parabolic bands and nondegenerate conditions (Sect. 19.5.2), the electron and hole concentrations can be expressed by means of the quasi-Fermi potentials φ_n, φ_p like in (19.140); it follows that $w = n_i\,\exp[-q\,\varphi_n/(k_B\,T)]$ for the electrons, and $w = n_i\,\exp[q\,\varphi_p/(k_B\,T)]$ for the holes. Given the boundary conditions u_0, u_{N+1} of the concentration (Sect. 19.5.6), the algebraic system to be solved has the form (A.112), (A.113), with $\alpha_i = \beta_i + \gamma_i$, $\beta_i = a_i/Y_i$ and $\gamma_i = a_{i+1}/Y_{i+1}$, $i = 1, \ldots, N$; in particular, the system matrix is tridiagonal and symmetric. The matrix is also irreducible with a weakly dominant diagonal; as a consequence of the second theorem in Sect. A.11.6, it is non-singular: in conclusion, the solution $w_1, \ldots, w_N$ is unique.

To calculate the denominator Y_i, one must provide the explicit form of the electric potential over the corresponding element h_i. To this purpose one observes that in the solution of the Poisson equation carried out earlier, the derivative of the electric potential over each element is approximated with the difference quotient; this is equivalent to assuming that the electric potential is piecewise constant over the elements. The same approximation is used here to yield $\sigma = \sigma_{i-1} + (\sigma_i - \sigma_{i-1})\,x/h_i$ over h_i. From the integral in (A.119) it then follows

$$\frac{Y_i}{h_i} = \frac{\exp(-\sigma_{i-1}) - \exp(-\sigma_i)}{\sigma_i - \sigma_{i-1}}\,, \qquad \frac{\exp(-\sigma_i)}{Y_i} = \frac{B(\sigma_i - \sigma_{i-1})}{h_i}\,, \tag{A.121}$$

with B the Bernoulli function (Sect. C.17). Similarly, one finds $h_i\,\exp(-\sigma_{i-1}) = Y_i\,B(\sigma_{i-1} - \sigma_i)$. The calculation can now be completed in two alternative ways: if the quantities w_i are used as unknowns, one extracts Y_i from the first relation in (A.121) and replaces it into (A.120). If, on the contrary, the original unknowns u_i are to be kept, one replaces $\exp(-\sigma_i)/Y_i$ from the second relation in (A.121); in this case, letting $B_j^k = B(\sigma_j - \sigma_k)$, the algebraic system for determining the carrier concentration takes the form[17]

[17] The discretization method leading to (A.122) was proposed in [117] and is referred to as *exponential fitting* or *Scharfetter-Gummel discretization scheme*.

$$-\frac{a_i\,B_{i-1}^{i}}{h_i}\,u_{i-1}+\left(\frac{a_i\,B_i^{i-1}}{h_i}+\frac{a_{i+1}\,B_i^{i+1}}{h_{i+1}}\right)u_i-\frac{a_{i+1}\,B_{i+1}^{i}}{h_{i+1}}\,u_{i+1}=\Omega_i\,C_i\,, \qquad \text{(A.122)}$$

with $i = 1, 2, \ldots, N$. The structure of the algebraic system generated by (A.122) is still tridiagonal, and the matrix is irreducible; however, it is not symmetric: in fact, the matrix would be symmetric if it were $B_i^{i+1} = B_{i+1}^{i}$, namely, $B(\sigma_i - \sigma_{i+1}) = B(\sigma_{i+1} - \sigma_i)$. Remembering from Sect. C.17 that $B(-\lambda) - B(\lambda) = \lambda$, the symmetry condition does not hold in general.[18] Also, the main diagonal of the matrix generated by (A.122) is not necessarily dominant. To show this, observing that the ratios a_i/h_i, $i = 1, 2, \ldots$ have all the same sign, assume that for some index i it is $\sigma_i = 0$, $\sigma_{i-1} = \sigma_{i+1} = \sigma$; (A.122) then transforms into

$$B(\sigma)\left[-\frac{a_i}{h_i}\,u_{i-1}+\exp(\sigma)\left(\frac{a_i}{h_i}+\frac{a_{i+1}}{h_{i+1}}\right)u_i-\frac{a_{i+1}}{h_{i+1}}\,u_{i+1}\right]=\Omega_i\,C_i\,, \qquad \text{(A.123)}$$

whose diagonal term is strongly dominant if $\sigma > 0$, weakly dominant if $\sigma = 0$, nondominant if $\sigma < 0$ (these cases correspond, respectively, to an electric potential that is concave, constant, or convex at the ith node). More generally, letting $\lambda^- = \sigma_{i-1} - \sigma_i$, $\lambda^+ = \sigma_{i+1} - \sigma_i$, and using $B(-\lambda) - B(\lambda) = \lambda$, the difference between the diagonal and non-diagonal terms of (A.122) reads $(a_i/h_i)\,\lambda^- + (a_{i+1}/h_{i+1})\,\lambda^+$; as λ^- and λ^+ may have either sign, the latter sum may turn out to be positive, null, or negative.

A.13.3 The Numerov Process

Considering the discretized form (A.110) of Poisson's equation, a few comments are useful. First, the right-hand side $\Omega_i\,C_i$ is one of the possible approximations of $\int_{\Omega_i} C\,\mathrm{d}x$; other forms of the right-hand side are possible if a different interpolation scheme was used. Also, the left-hand side of (A.110) may be viewed as an interpolation of the unknown function u over the three nodes x_{i-1}, x_i, x_{i+1}, that determines one of the nodal values when the other two are given; it is therefore of interest to investigate the order of the approximation of the left-hand side. For this, one takes the case[19] $a = 1$ and seeks for a parabolic interpolation of the unknown u over a triad of consecutive nodes, $u \leftarrow f(x) = f_0\,x^2 + f_1\,x + f_2$. After fixing $x_i = 0$, let the abscissæ of the right and left node be h_R and $-h_L$, respectively; the corresponding nodal values of the unknown are u_0, u_R, and u_L, while the cell's

[18] Symmetry holds only when $\lambda \to 0$ over each element. In fact, $\lim_{\lambda\to 0} B(\lambda) = 1$; this case corresponds to $\sigma = \text{const}$ (diffusive-only transport), where the equation to be solved becomes identical to Poisson's.

[19] The restriction $a = 1$ is taken here for the sake of simplicity. As shown below, the analysis would be the same in the general case $a = a(x)$.

length is $(h_L + h_R)/2$. Here the equation to be solved reads $-\mathrm{d}^2u/\mathrm{d}x^2 = C$: in the parabolic approximation its left-hand side becomes $-2f_0$, namely, a constant; one then chooses such a constant as $C_0 = C(0)$. Then, the coefficients of the parabolic interpolation are readily found by imposing that f equals the nodal values of u; in particular, one finds $f_2 = f(0) = u_0$, so that the other two nodal values fulfill the relations

$$f_0\, h_R^2 + f_1\, h_R = u_R - u_0\,, \qquad f_0\, h_L^2 - f_1\, h_L = u_L - u_0 \tag{A.124}$$

whence

$$h_R\,(u_0 - u_L) + h_L\,(u_0 - u_R) = -f_0\,(h_R^2\, h_L + h_L^2\, h_R)\,. \tag{A.125}$$

Letting $-f_0 = C_0/2$ and dividing both sides by $h_R\, h_L$ renders $(u_0 - u_L)/h_L + (u_0 - u_R)/h_R = C_0\,(h_L + h_R)/2$, identical to the special case of (A.110) where a is set equal to unity. This shows that the left-hand side of (A.110) is in fact a parabolic interpolation.

As for the calculation of the right-hand side, the accuracy of the interpolation can be increased, with respect to that of (A.110), without changing the number of nodes involved in it, and with a negligible increase in the computational cost, by better exploiting the form of the equation. This is achieved by the so-called *Numerov process* (the original reference is [100], cited, e.g., in [22]), which is illustrated below. Prior to that, one disposes of the simplifying assumption $a = 1$ as follows: consider the general form of the equation, $-\mathrm{d}S/\mathrm{d}x = C$, $S = a\,\mathrm{d}u/\mathrm{d}x$, and assume that $a = a(x)$ is a given function of position; changing the independent variable from x to ξ, such that

$$\xi = \int_{x_0}^{x} \frac{\mathrm{d}x'}{a(x')}\,, \qquad \mathrm{d}x = a(x)\,\mathrm{d}\xi\,, \tag{A.126}$$

yields

$$-\frac{\mathrm{d}S}{\mathrm{d}\xi} = B\,, \qquad B = a\,C\,, \qquad S = \frac{\mathrm{d}u}{\mathrm{d}\xi}\,, \tag{A.127}$$

where u, S, a, and C are now functions of ξ. With no loss of the generality one can then base the description of the Numerov process upon the form $-\mathrm{d}^2u/\mathrm{d}\xi^2 = B$ where, due to the change in the independent variable, the new elements of the discretization grid will be indicated with s_i. To proceed, it is necessary to assume that the elements are all equal, $s_1 = \ldots = s_{N+1} = s$; a series expansion of u at ξ_i, using s as increment,[20] combined with $u'' = -B$, provides

[20] It must be assumed that s is smaller than the convergence radius of the series.

$$u_{i+1} = u_i + u_i' \, s - B_i \, \frac{s^2}{2} - B_i' \, \frac{s^3}{6} - B_i'' \, \frac{s^4}{24} \quad \cdots \tag{A.128}$$

By the same token one expresses u_{i-1} using $-s$ as increment; this yields a new series whose odd-degree terms have the opposite sign with respect to those of (A.128). Adding up the two series, and leaving out the terms with the derivatives of u of the sixth order or higher, yields

$$2\,u_i - u_{i+1} - u_{i-1} \simeq B_i \, s^2 + B_i'' \, \frac{s^4}{12} \, . \tag{A.129}$$

The second derivative of (A.129), after leaving out the term with the sixth derivative of u and using again $u'' = -B$, reads

$$-\,2\,B_i + B_{i+1} + B_{i-1} \simeq B_i'' \, s^2 \, . \tag{A.130}$$

Multiplying both sides of (A.130) by $s^2/12$, and replacing into (A.129), eliminates B_i''; in conclusion, one finds[21]

$$-\,\frac{u_{i-1}}{s} + \frac{2\,u_i}{s} - \frac{u_{i+1}}{s} = s \, \frac{B_{i-1} + 10\,B_i + B_{i+1}}{12} \, . \tag{A.131}$$

As shown above, the essence of the method is the elimination of the derivatives of odd order, and the exploitation of the form of the original equation to eliminate the fourth derivative of u. In this respect, the method is applicable to other classes of equations that provide the form of the second derivative of the unknown function, without involving the first derivative; an example of this is the time-independent Schrödinger equation, which is in fact amenable to the application of the Numerov process (Sect. 13.6.6 and Prob. 13.3). One also notes that if the last term at the right-hand side of (A.129) was neglected, the discretized form of the Poisson equation would become identical to (A.109), the latter with $a = 1$ and $h_1 = \ldots = h_{N+1} = h$. Due to this, and considering also that the derivative of the fifth order is eliminated from the beginning, one finds that the interpolation of the right-hand side provided by the Numerov process is more accurate by three orders with respect to the standard one [19, 20].[22]

By a suitable generalization of the variable transformation (A.126), the Numerov process can be extended to the discretization of the continuity and drift-diffusion equations. As shown above, these equations can be recast as $-\mathrm{d}S/\mathrm{d}x = C$ and, respectively, $S \exp(-\sigma) = a \, \mathrm{d}w/\mathrm{d}x$, with $w = u \exp(-\sigma)$. By analogy with (A.126) one changes the independent variable from x to η, such that

[21] Using the new unknown $v = 12\,u/s^2$ in (A.131) makes the grid's details invisible.

[22] The elimination of the odd-order derivatives requires the uniformity of the elements' size; this drawback, however, is largely compensated by the superior accuracy of the method, which allows for coarser grids.

$$\eta = \int_{x_0}^{x} \exp[-\sigma(x')] \, \frac{\mathrm{d}x'}{a(x')} \,, \qquad \mathrm{d}x = a(x) \, \exp[\sigma(x)] \, \mathrm{d}\eta \,; \tag{A.132}$$

this yields

$$-\frac{\mathrm{d}S}{\mathrm{d}\eta} = D \,, \qquad D = a \, \exp(\sigma) \, C \,, \qquad S = \frac{\mathrm{d}w}{\mathrm{d}\eta} \,, \tag{A.133}$$

where w, σ, S, a, and D are now functions of η. With no loss of generality one can then base the description of the Numerov process upon the form $-\mathrm{d}^2 w/\mathrm{d}\eta^2 = D$. Note that, since the nodal values of σ are involved in the calculation, one should take a uniform grid with the same elements $s_1 = \ldots = s_{N+1} = s$ as for the Poisson equation. This, in general, is not possible because the variable transformations (A.126) and (A.132) are different; for this reason, an interpolation of σ from one grid to the other is necessary.[23] The solution of (A.133) is then completed as for the Poisson equation, leading to an equation identical to (A.131) where w and D replace u and B.

[23] If the Poisson equation is solved over a domain where the permittivity a is constant, then the grid transformation (A.126) becomes trivial. However, the interpolation to the other grid is still necessary.

Appendix B
Coordinates

B.1 Spherical Coordinates

When the problem in hand has a spherical symmetry it is convenient to describe the position of a particle by means of the *spherical coordinates*. With reference to Fig. B.1, the transformation relations between the Cartesian (x, y, z) and spherical (r, ϑ, φ) coordinates are

$$\begin{cases} x = r \sin\vartheta \cos\varphi \\ y = r \sin\vartheta \sin\varphi \\ z = r \cos\vartheta \end{cases} \qquad \begin{cases} r^2 = x^2 + y^2 + z^2 \\ \cos\vartheta = z/r \\ \tan\varphi = y/x \end{cases} \tag{B.1}$$

that are a special case of (1.26). The limits of the spherical coordinates are $0 \leq r < \infty$, $0 \leq \vartheta \leq \pi$, $0 \leq \varphi < 2\pi$. The 3×3 matrix of the partial derivatives of the Cartesian coordinates with respect to the spherical ones, expressed in terms of the latter (*Jacobian matrix*), is

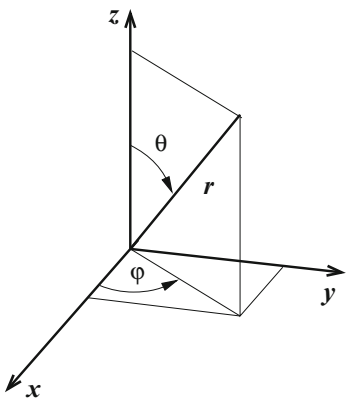


Fig. B.1 Cartesian (x, y, z) and spherical (r, ϑ, φ) coordinates

M. Rudan, *Physics of Semiconductor Devices*, DOI 10.1007/978-3-319-63154-7

$$\frac{\partial(x,y,z)}{\partial(r,\vartheta,\varphi)} = \begin{bmatrix} \sin\vartheta\,\cos\varphi & r\,\cos\vartheta\,\cos\varphi & -r\,\sin\vartheta\,\sin\varphi \\ \sin\vartheta\,\sin\varphi & r\,\cos\vartheta\,\sin\varphi & r\,\sin\vartheta\,\cos\varphi \\ \cos\vartheta & -r\,\sin\vartheta & 0 \end{bmatrix}, \tag{B.2}$$

where the left-hand side is a short-hand notation for a matrix whose elements are $J_{11} = \partial x/\partial r$, $J_{12} = \partial x/\partial\vartheta$, and so on. The *Jacobian determinant* is

$$J = \det\frac{\partial(x,y,z)}{\partial(r,\vartheta,\varphi)} = r^2\,\sin\vartheta\,. \tag{B.3}$$

The matrix of the partial derivatives of the spherical coordinates with respect to the Cartesian ones, expressed in terms of the former, is

$$\frac{\partial(r,\vartheta,\varphi)}{\partial(x,y,z)} = \begin{bmatrix} \sin\vartheta\,\cos\varphi & \sin\vartheta\,\sin\varphi & \cos\vartheta \\ (1/r)\,\cos\vartheta\,\cos\varphi & (1/r)\,\cos\vartheta\,\sin\varphi & -(1/r)\,\sin\vartheta \\ -(1/r)\,\sin\varphi/\sin\vartheta & (1/r)\,\cos\varphi/\sin\vartheta & 0 \end{bmatrix}, \tag{B.4}$$

whence

$$\det\frac{\partial(r,\vartheta,\varphi)}{\partial(x,y,z)} = \frac{1}{r^2\,\sin\vartheta} = \frac{1}{J}\,. \tag{B.5}$$

To calculate (B.4) consider, e.g., the last term of the second row, $(\partial\vartheta/\partial z)_{xy} = -(1/r)\,\sin\vartheta$. The second line of the second group of (B.1) yields $(\partial\cos\vartheta/\partial z)_{xy} = 1/r - z^2/r^3$, where $(\partial r/\partial z)_{xy} = z/r$ has been used, that in turn derives from the first line of the second group of (B.1). The relation $z = r\cos\vartheta$ then yields $(\partial\cos\vartheta/\partial z)_{xy} = (1/r)\,\sin^2\vartheta$. On the other hand, the same quantity can also be written as $(\partial\cos\vartheta/\partial z)_{xy} = -\sin\vartheta\,(\partial\vartheta/\partial z)_{xy}$. Comparing the two expressions above yields the result sought.

Differentiating with respect to time the first of (B.1) yields the relations

$$\begin{cases} \dot{x} = \dot{r}\,\sin\vartheta\,\cos\varphi + r\,\dot{\vartheta}\,\cos\vartheta\,\cos\varphi - r\,\dot{\varphi}\,\sin\vartheta\,\sin\varphi \\ \dot{y} = \dot{r}\,\sin\vartheta\,\sin\varphi + r\,\dot{\vartheta}\,\cos\vartheta\,\sin\varphi + r\,\dot{\varphi}\,\sin\vartheta\,\cos\varphi \\ \dot{z} = \dot{r}\,\cos\vartheta - r\,\dot{\vartheta}\,\sin\vartheta \end{cases} \tag{B.6}$$

that express the components of the velocity in the Cartesian reference as functions of the generalized coordinates r, ϑ, φ and generalized velocities $\dot{r}, \dot{\vartheta}, \dot{\varphi}$ of the spherical reference. From (B.6) the expression of the kinetic energy in spherical coordinates follows:

$$T = \frac{1}{2}m\,(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) = \frac{1}{2}m\,\left(\dot{r}^2 + r^2\dot{\vartheta}^2 + r^2\dot{\varphi}^2\,\sin^2\vartheta\right)\,. \tag{B.7}$$

B.2 Polar Coordinates

To describe the motion of particles confined over a plane one may adopt, instead of the Cartesian coordinates x, y, the *polar coordinates* r, φ. The relations between the two groups of coordinates are

$$\begin{cases} x = r\cos\varphi \\ y = r\sin\varphi \end{cases} \qquad \begin{cases} r^2 = x^2 + y^2 \\ \tan\varphi = y/x \end{cases} \tag{B.8}$$

The limits of the polar coordinates are $0 \le r < \infty$, $0 \le \varphi < 2\pi$. The Jacobian matrix and the Jacobian determinant are, respectively,

$$\frac{\partial(x,y)}{\partial(r,\varphi)} = \begin{bmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{bmatrix}, \qquad J = \det\frac{\partial(x,y)}{\partial(r,\varphi)} = r\,. \tag{B.9}$$

Differentiating with respect to time the first of (B.8) yields the relations

$$\begin{cases} \dot{x} = \dot{r}\cos\varphi - r\dot{\varphi}\sin\varphi \\ \dot{y} = \dot{r}\sin\varphi + r\dot{\varphi}\cos\varphi \end{cases} \tag{B.10}$$

that express the components of the velocity in the Cartesian reference as functions of the generalized coordinates r, φ and generalized velocities $\dot{r}, \dot{\varphi}$ of the polar reference. From (B.10) the expression of the kinetic energy in polar coordinates follows:

$$T = \frac{1}{2}m\,(\dot{x}^2 + \dot{y}^2) = \frac{1}{2}m\,\left(\dot{r}^2 + r^2\dot{\varphi}^2\right)\,. \tag{B.11}$$

B.3 Coordinate Rotation

Consider a coordinate transformation that consists in a rotation around the origin, bringing a right-handed system of coordinates $\mathbf{x} = (x_1, x_2, x_3)$ into another right-handed system $\mathbf{s} = (s_1, s_2, s_3)$. The transformation is described by the linear relations

$$\begin{cases} s_1 = a_{11}\,x_1 + a_{12}\,x_2 + a_{13}\,x_3 \\ s_2 = a_{21}\,x_1 + a_{22}\,x_2 + a_{23}\,x_3 \\ s_3 = a_{31}\,x_1 + a_{32}\,x_2 + a_{33}\,x_3 \end{cases} \tag{B.12}$$

which can be recast in the matrix form $\mathbf{s} = \mathbf{A}\mathbf{x}$. It is known that a matrix describing this type of transformation is *orthogonal* ([56, Sect. 4.2] and Sect. A.11), namely,

$$\sum_{i=1}^{3} a_{ij}\, a_{ik} = \delta_{jk}\,, \qquad \det \mathbf{A} = 1\,, \qquad \mathbf{A}^{-1} = \mathbf{A}^{T}\,, \qquad j,k = 1,2,3\,, \tag{B.13}$$

where apex T indicates the transpose. From (B.13) it follows $\left(\mathbf{A}^{-1}\right)^{T} = \mathbf{A}$. As a consequence, the effect of the rotation onto the modulus of a particle's velocity is found from

$$\left(\dot{\mathbf{x}}^{T},\, \dot{\mathbf{x}}\right) = \left[\left(\mathbf{A}^{-1}\,\dot{\mathbf{s}}\right)^{T},\, \mathbf{A}^{-1}\,\dot{\mathbf{s}}\right] = \left(\dot{\mathbf{s}}^{T}\mathbf{A},\, \mathbf{A}^{-1}\,\dot{\mathbf{s}}\right) = \left(\dot{\mathbf{s}}^{T},\, \mathbf{A}\,\mathbf{A}^{-1}\,\dot{\mathbf{s}}\right) = \left(\dot{\mathbf{s}}^{T},\, \dot{\mathbf{s}}\right)\,. \tag{B.14}$$

In (B.14) the symbol $(\mathbf{a}^{T}, \mathbf{b})$ denotes the scalar product between the vectors $\mathbf{a}$ and $\mathbf{b}$, namely, it is equivalent to $\mathbf{a}\cdot\mathbf{b}$. The above calculation shows that $u^2 = (\dot{\mathbf{x}}^{T}, \dot{\mathbf{x}})$ is invariant under rotation of the coordinate system. The same reasoning applies to the modulus of position $r^2 = (\mathbf{x}^{T},\, \mathbf{x}) = (\mathbf{s}^{T},\, \mathbf{s})$.

B.4 Differential Operators Under Coordinate Transformations

Consider the coordinate transformation between the two sets x_i, ξ_i, $i = 1, 2, \ldots, n$:

$$\xi_i = \xi_i(x_1, \ldots, x_n)\,, \qquad x_i = x_i(\xi_1, \ldots, \xi_n)\,. \tag{B.15}$$

If a function f is transformed using the above, the following hold:

$$\frac{\partial f}{\partial x_i} = \sum_{j=1}^{n} \frac{\partial f}{\partial \xi_j}\,\frac{\partial \xi_j}{\partial x_i}\,, \qquad \frac{\partial^2 f}{\partial x_i^2} = \sum_{j=1}^{n}\left(\frac{\partial f}{\partial \xi_j}\,\frac{\partial^2 \xi_j}{\partial x_i^2} + \sum_{k=1}^{n} \frac{\partial^2 f}{\partial \xi_j \partial \xi_k}\,\frac{\partial \xi_j}{\partial x_i}\,\frac{\partial \xi_k}{\partial x_i}\right)\,. \tag{B.16}$$

Adding up over i in the second of (B.16) yields

$$\nabla^2 f = \sum_{j=1}^{n}\left(\frac{\partial f}{\partial \xi_j}\,\nabla^2 \xi_j + \sum_{k=1}^{n} \frac{\partial^2 f}{\partial \xi_j \partial \xi_k}\,\nabla \xi_j \cdot \nabla \xi_k\right)\,, \tag{B.17}$$

where symbols ∇ and ∇^2 indicate, respectively, the gradient and the Laplacian operator with respect to the coordinates x_i. By way of example consider the transformation (B.1) from Cartesian to spherical coordinates. Remembering (B.4) one finds

$$\nabla r \cdot \nabla r = 1\,, \qquad \nabla \vartheta \cdot \nabla \vartheta = \frac{1}{r^2}\,, \qquad \nabla \varphi \cdot \nabla \varphi = \frac{1}{r^2\,\sin^2 \vartheta}\,, \tag{B.18}$$

$$\nabla r \cdot \nabla \vartheta = 0\,, \qquad \nabla \vartheta \cdot \nabla \varphi = 0\,, \qquad \nabla \varphi \cdot \nabla r = 0\,, \tag{B.19}$$

whence

$$\nabla^2 f = \frac{\partial f}{\partial r}\,\nabla^2 r + \frac{\partial^2 f}{\partial r^2} + \frac{\partial f}{\partial \vartheta}\,\nabla^2 \vartheta + \frac{\partial^2 f}{\partial \vartheta^2}\,\frac{1}{r^2} + \frac{\partial f}{\partial \varphi}\,\nabla^2 \varphi + \frac{\partial^2 f}{\partial \varphi^2}\,\frac{1}{r^2\,\sin^2\vartheta}\,. \tag{B.20}$$

In turn, letting $\gamma = \sin\vartheta\cos\vartheta/r$, $\zeta = \sin\varphi\cos\varphi/r$, $\sigma = \sin^2\vartheta$, the terms $\nabla^2 r$, $\nabla^2\vartheta$, $\nabla^2\varphi$ are found from

$$\begin{bmatrix} \partial^2 r/\partial x^2 & \partial^2 r/\partial y^2 & \partial^2 r/\partial z^2 \\ \partial^2\vartheta/\partial x^2 & \partial^2\vartheta/\partial y^2 & \partial^2\vartheta/\partial z^2 \\ \partial^2\varphi/\partial x^2 & \partial^2\varphi/\partial y^2 & \partial^2\varphi/\partial z^2 \end{bmatrix} = \frac{1}{r}\times \tag{B.21}$$

$$\times \begin{bmatrix} 1-\sigma\cos^2\varphi & 1-\sigma\sin^2\varphi & \sigma \\ \gamma\,(\sin^2\varphi/\sigma - 2\cos^2\varphi) & \gamma\,(\cos^2\varphi/\sigma - 2\sin^2\varphi) & 2\gamma \\ 2\zeta/\sigma & -2\zeta/\sigma & 0 \end{bmatrix}, \tag{B.22}$$

whence

$$\nabla^2 r = \frac{2}{r}\,, \qquad \nabla^2\vartheta = \frac{1}{r^2}\,\frac{\cos\theta}{\sin\theta}\,, \qquad \nabla^2\varphi = 0\,, \tag{B.23}$$

$$\nabla^2 f = \frac{\partial f}{\partial r}\,\frac{2}{r} + \frac{\partial^2 f}{\partial r^2} + \frac{\partial f}{\partial \vartheta}\,\frac{1}{r^2}\,\frac{\cos\theta}{\sin\theta} + \frac{\partial^2 f}{\partial \vartheta^2}\,\frac{1}{r^2} + \frac{\partial^2 f}{\partial \varphi^2}\,\frac{1}{r^2\,\sin^2\vartheta} = \tag{B.24}$$

$$= \frac{1}{r}\,\frac{\partial^2}{\partial r^2}(rf) + \frac{1}{r^2\,\sin\vartheta}\,\frac{\partial}{\partial\vartheta}\left(\sin\vartheta\,\frac{\partial f}{\partial\vartheta}\right) + \frac{1}{r^2\,\sin^2\vartheta}\,\frac{\partial^2 f}{\partial\varphi^2}\,. \tag{B.25}$$

B.5 Density of States

Consider a function s that depends on the coordinates u, v, w, and on one or more additional parameters that will collectively be indicated with σ. Let

$$S(\sigma) = \iiint s(u,v,w,\sigma)\,\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}w\,, \tag{B.26}$$

where the integration is carried out over the whole domain of u, v, w. Next, consider the transformation from the original variables to the new variables α, β, η,

$$\alpha = \alpha(u,v,w)\,, \quad \beta = \beta(u,v,w)\,, \quad \eta = \eta(u,v,w)\,, \quad J = \frac{\partial(u,v,w)}{\partial(\alpha,\beta,\eta)}\,, \tag{B.27}$$

which is assumed to be invertible, so that the Jacobian determinant J does not vanish. After the transformation is carried out, (B.26) takes the form

$$S(\sigma) = \iiint s(\alpha, \beta, \eta, \sigma)\, |J|\, \mathrm{d}\alpha\, \mathrm{d}\beta\, \mathrm{d}\eta\,. \tag{B.28}$$

It may happen that one is interested in the dependence of s on one of the new variables, say, η, rather than in the details about its dependence on the whole set of new variables. In this case one first carries out the integrals with respect to α and β in (B.28), to find

$$h(\eta, \sigma) = \iint s(\alpha, \beta, \eta, \sigma)\, |J|\, \mathrm{d}\alpha\, \mathrm{d}\beta\,. \tag{B.29}$$

Then one defines

$$b(\eta) = \iint |J|\, \mathrm{d}\alpha\, \mathrm{d}\beta\,. \tag{B.30}$$

A function like $b(\eta)$ plays an important role in many physical problems (e.g., Sects. 14.6, 15.8.1, 15.8.2). For this reason, although its derivation in this section is of a purely formal nature, $b(\eta)$ will be called *density of states in* η. Note that the density of states depends only on the structure of the variable transformation (B.27) and (at most) on η. The form of (B.29) and (B.30) shows that the ratio $\bar{s} = h/b$ is a weighed average of $s(\alpha, \beta, \eta, \sigma)$ over the two variables α and β, that uses $|J|$ as weight. Introducing the definition of $\bar{s}$ into (B.28) gives the latter form

$$S(\sigma) = \int b(\eta)\, \bar{s}(\eta, \sigma)\, \mathrm{d}\eta\,. \tag{B.31}$$

If s happens to be independent of α and β, definition (B.29) yields $h = s\, b$, whence $\bar{s}(\eta, \sigma) = s(\eta, \sigma)$. The derivation of b is not limited to a three-dimensional case; in fact it applies to any number of dimensions. In the following, a few examples in one, two, and three dimensions are given, in which one of the transformation relations (B.27), namely, $\eta = \eta(u, v, w)$, has a quadratic form; these examples are in fact particularly significant for the physical applications, where η stands for the energy and u, v, w stand for the generalized coordinates.

Considering a one-dimensional case with $\eta = u^2$, one finds[1]

$$u = \pm \eta^{1/2}\,, \qquad |J| = \left| \frac{\partial u}{\partial \eta} \right| = \frac{1}{2}\, \eta^{-1/2}\,, \qquad b(\eta) = 2\, \frac{1}{2}\, \eta^{-1/2} = \eta^{-1/2}\,. \tag{B.32}$$

[1] Factor 2 in the last expression of (B.32) accounts for the fact that both positive and negative parts of the segment $[-\eta^{1/2}, +\eta^{1/2}]$ must be considered.

This case is straightforward because there are no other variables involved in the transformation. Instead, in the two-dimensional case with $\eta = u^2 + v^2$, a convenient transformation involving the second variable is of the polar type (B.8), specifically, $u = \eta^{1/2} \cos\varphi$, $v = \eta^{1/2} \sin\varphi$. One finds

$$|J| = \frac{1}{2}\,, \qquad b(\eta) = \int_0^{2\pi} \frac{1}{2}\,\mathrm{d}\varphi = \pi\,. \tag{B.33}$$

In the three-dimensional case with $\eta = u^2 + v^2 + w^2$, a convenient transformation involving the other two variables is of the spherical type (B.1), specifically, $u = \eta^{1/2} \sin\vartheta \cos\varphi$, $v = \eta^{1/2} \sin\vartheta \sin\varphi$, $w = \eta^{1/2} \cos\vartheta$. One finds

$$|J| = \frac{1}{2}\,\eta^{1/2} \sin\vartheta\,, \qquad b(\eta) = \int_0^{2\pi} \int_0^{\pi} \frac{1}{2}\,\eta^{1/2} \sin\vartheta\,\mathrm{d}\vartheta\,\mathrm{d}\varphi = 2\,\pi\,\eta^{1/2}\,. \tag{B.34}$$

The above examples show that despite the fact that the $\eta = \eta(u, v, w)$ relation is quadratic in all cases, the form of $b(\eta)$ changes depending on the number of spatial dimensions.

Still considering the case where one of the transformation relations (B.27) has a quadratic form, the analysis can be extended to arbitrary values of the number of spatial dimensions. As a starting point, and considering provisionally the three-dimensional case, one notes from (B.30) that the following equality holds:[2]

$$B = \int b(\eta)\,\mathrm{d}\eta = \iiint |J|\,\mathrm{d}\alpha\,\mathrm{d}\beta\,\mathrm{d}\eta = \iiint \mathrm{d}u\,\mathrm{d}v\,\mathrm{d}w\,. \tag{B.35}$$

Remembering the definition of b, it follows that B is the number of states in the domain of u, v, w. Due to the last integral in (B.34), B is also equal to the volume of such a domain; in turn, due to the first integral, B can be thought of as the sum of the volumes of elementary shells of thickness $\mathrm{d}\eta$, with $b(\eta)$ the area of each shell (that is, the area of the two-dimensional surface $\eta = \text{const}$). These observations provide the key to extending the analysis to the case where η is a quadratic form in an arbitrary number of dimensions,

$$u_1^2 + u_2^2 + \ldots + u_n^2 = \eta\,, \qquad \eta = g^2\,. \tag{B.36}$$

Letting $\eta = \text{const}$, (B.36) is the equation of an $(n-1)$-dimensional sphere of radius $g \geq 0$ immersed into the n-dimensional space. The problem is thus reduced to expressing the area of the sphere in terms of η; although it can be solved by using a

[2] In the practical applications of the concepts illustrated here, the integrands in (B.35) embed a constant factor Q_0, called *density of states in the u, v, w space* which, besides describing some properties of the physical problem under investigation, makes B dimensionless. Here, all variables involved are dimensionless, and Q_0 is set equal to unity.

generalization of the spherical coordinates to n dimensions, a more elegant approach consists in finding a recursive expression involving also the sphere's volume.

To this purpose, let V_n indicate the volume of a sphere of an n-dimensional space, and let S_{n-1} indicate the surface area of the same sphere. When $n = 1$, the sphere is a segment whose volume is the length $V_1 = 2\,g$; for $n = 2$, the sphere is a circle whose volume is the area $V_2 = \pi\,g^2$; for $n = 3$ it is $V_3 = (4/3)\,\pi\,g^3$; for $n = 4$ it is $V_4 = \pi^2\,g^4/2$, and so on; in turn, for $n = 2$ the surface is a circumference whose area is the length $S_1 = 2\,\pi\,g$; for $n = 3$ it is $S_2 = 4\,\pi\,g^2$; for $n = 4$ it is $S_3 = 2\,\pi^2\,g^3$, and so on. Consistently with the expression of B as the integral of b given by (B.35), one finds from the above values the general relation

$$V_n = \frac{g}{n}\,S_{n-1}\,. \tag{B.37}$$

Combining (B.37) with $V_1 = 2\,g$ also yields $S_0 = 2$, that is, the "surface" of the segment considered above; such a surface is made of the segment's endpoints $\{-1, +1\}$. From (B.37) it also follows that $V_n \propto g^n$ and $S_{n-1} \propto g^{n-1}$, whence $S_n \propto g\,V_{n-1}$ and $V_0 = \text{const}$. From the values found above one finds $S_2/(g\,V_1) = S_3/(g\,V_2) = 2\,\pi$; it follows that $S_n = 2\,\pi\,g\,V_{n-1}$ and $V_0 = 1$. The latter is the "volume" of a sphere in a zero-dimensional space. The recursive relation involving the volumes then reads

$$V_n = \frac{g}{n}\,S_{n-1} = \frac{g}{n}\,2\,\pi\,g\,V_{n-2} = \frac{2\,\pi\,g^2}{n}\,V_{n-2}\,, \qquad V_0 = 1\,, \qquad V_1 = 2\,g\,. \tag{B.38}$$

The above can further be improved by observing that the sequence $V_0, V_1, \ldots$ embeds Euler's Gamma function of half-integer order; in fact, combining (B.37) and (B.38) with the definitions of Sect. C.10, yields

$$V_n = \frac{\pi^{n/2}}{\Gamma(n/2+1)}\,g^n\,, \qquad S_{n-1} = \frac{n\,\pi^{n/2}}{\Gamma(n/2+1)}\,g^{n-1}\,. \tag{B.39}$$

The last step consists in expressing the result in terms of η. This is accomplished by noting that $b(\eta)\,\mathrm{d}\eta = S_{n-1}(g)\,\mathrm{d}g$, where $g = \sqrt{\eta}$ and $\mathrm{d}g = \mathrm{d}\sqrt{\eta} = \mathrm{d}\eta/(2\,\sqrt{\eta})$; then, one finds

$$b(\eta)\,\mathrm{d}\eta = \frac{n\,\pi^{n/2}\,\eta^{(n-1)/2}}{\Gamma(n/2+1)}\,\frac{\mathrm{d}\eta}{2\,\eta^{1/2}}\,, \qquad b(\eta) = \frac{n\,\pi^{n/2}}{2\,\Gamma(n/2+1)}\,\eta^{n/2-1}\,. \tag{B.40}$$

Letting $n = 1, 2, 3$ in the second expression of (B.40) renders (B.32), (B.33), (B.34), respectively.

B.6 Constrained Extrema—Lagrange Method

In several occasions it is necessary to determine an extremum point of a function under one or more additional constraints (*constrained extremum*). To begin, one considers a function f that depends on two variables, x and y. It is implied that f is differentiable in its domain of definition, D; for the sake of conciseness, the partial derivatives with respect to the variables are indicated here with f_x and f_y. The problem to be solved is finding an extremum of f in D subjected to a constraint of the form $F(x, y) = 0$, where it is assumed that F is differentiable in D and that if the extremum exists, either F_x, or F_y, or both, are different from zero there.

To proceed, let x_0, y_0 be the extremum's coordinates; since $F = 0$ everywhere in D, it must be $\mathrm{d}F = 0$ in (x_0, y_0), namely,

$$F_x(x_0, y_0)\,\mathrm{d}x + F_y(x_0, y_0)\,\mathrm{d}y = 0 \tag{B.41}$$

where, by hypothesis, at least one of the two derivatives is different from zero. Assume that $F_y(x_0, y_0) \neq 0$; then, using suffix "0" to indicate that the derivatives are calculated in the extremum, from (B.41) the relation $\mathrm{d}y = -(F_x/F_y)_0\,\mathrm{d}x$ follows. Such a relation provides the increment $\mathrm{d}y$ corresponding to an arbitrary increment $\mathrm{d}x$, in other terms it provides the derivative at (x_0, y_0) of the function $y = y(x)$ implicitly defined by $F = 0$. The same reasoning applies when $F_x(x_0, y_0) \neq 0$, yielding in this case the inverse function $x = x(y)$.

In the extremum it is also $\mathrm{d}f = 0$, namely,

$$f_x(x_0, y_0)\,\mathrm{d}x + f_y(x_0, y_0)\,\mathrm{d}y = 0\,, \tag{B.42}$$

where the relation between $\mathrm{d}x$ and $\mathrm{d}y$ is prescribed by (B.41). For (B.41) and (B.42) to be compatible it must be

$$\det\begin{bmatrix} f_x(x_0, y_0) & f_y(x_0, y_0) \\ F_x(x_0, y_0) & F_y(x_0, y_0) \end{bmatrix} = 0\,. \tag{B.43}$$

The determinant vanishes if, e.g., the rows are a linear combination of each other, say, $f_x = \lambda\, F_x, f_y = \lambda\, F_y$, with $\lambda \neq 0$ an undetermined parameter. Combining the above with the constraint $F = 0$ yields

$$f_x(x_0, y_0) = \lambda\, F_x(x_0, y_0)\,, \quad f_y(x_0, y_0) = \lambda\, F_y(x_0, y_0)\,, \quad F(x_0, y_0) = 0\,, \tag{B.44}$$

namely, a system of three equations in the unknowns x_0, y_0, λ, whose solution provides the coordinates of the extremum and the parameter. The latter is called *Lagrangian multiplier*. It is worth noting that the first two equations in (B.44) are equivalent to

$$\frac{\partial}{\partial x}(f - \lambda F) = 0\,, \qquad \frac{\partial}{\partial y}(f - \lambda F) = 0\,. \tag{B.45}$$

Thus, the calculation of a constrained extremum entails the construction of the auxiliary function $f - \lambda F$. If no constraint exists, the procedure is brought back to the calculation of the unconstrained extremum of f, which requires the solution of the system

$$f_x(x_0, y_0) = 0\,, \qquad f_y(x_0, y_0) = 0 \tag{B.46}$$

in the two unknowns x_0, y_0; such equations may be thought of as derived from (B.45) by letting $\lambda = 0$ there.

In the example using a two-variable function shown above it would be impossible to add another constraint, say, $G = 0$: in fact, the increments $\mathrm{d}y = -(F_x/F_y)_0\,\mathrm{d}x$ and $\mathrm{d}y = -(G_x/G_y)_0\,\mathrm{d}x$ are different in general. It follows that the number of constraints must be smaller than the number of variables. Some other examples of the procedure are shown below.

Consider the Lagrange method applied to a function of three variables x, y, z, to which one constraint of the form $F(x, y, z) = 0$ is applied. Following the same reasoning as in the two-variable case one obtains

$$f_x\,\mathrm{d}x + f_y\,\mathrm{d}y + f_z\,\mathrm{d}z = 0\,, \qquad F_x\,\mathrm{d}x + F_y\,\mathrm{d}y + F_z\,\mathrm{d}z = 0\,, \tag{B.47}$$

where the derivatives are calculated at the extremum point (x_0, y_0, z_0). Assuming that $F_z(x_0, y_0, z_0) \neq 0$ and eliminating $\mathrm{d}z$ from (B.47) yields

$$(f_x F_z - f_z F_x)\,\mathrm{d}x + \left(f_y F_z - f_z F_y\right)\,\mathrm{d}y = 0\,. \tag{B.48}$$

On account of the arbitrariness of $\mathrm{d}x$ and $\mathrm{d}y$ in (B.48), the corresponding coefficients must vanish, leading to the system

$$f_x = \lambda F_x\,, \qquad f_y = \lambda F_y\,, \qquad f_z = \lambda F_z\,, \qquad F = 0 \tag{B.49}$$

in the four unknowns x_0, y_0, z_0, λ.

As a final example consider the case of the constrained extremum of a function that depends on three variables, subjected to two constraints, namely, $F = 0, G = 0$. Letting $\mathrm{d}f = 0$, $\mathrm{d}F = 0$, and $\mathrm{d}G = 0$ in the extremum provides, as in the case of (B.43),

$$\det \begin{bmatrix} f_x(x_0, y_0, z_0) & f_y(x_0, y_0, z_0) & f_z(x_0, y_0, z_0) \\ F_x(x_0, y_0, z_0) & F_y(x_0, y_0, z_0) & F_z(x_0, y_0, z_0) \\ G_x(x_0, y_0, z_0) & G_y(x_0, y_0, z_0) & G_z(x_0, y_0, z_0) \end{bmatrix} = 0\,. \tag{B.50}$$

The determinant vanishes if, e.g., the rows are a linear combination of each other, say, $f_x = \lambda F_x + \mu G_x, f_y = \lambda F_y + \mu G_y, f_z = \lambda F_z + \mu G_z$, with $\lambda \neq 0$, $\mu \neq 0$ the two Lagrangian multipliers. In conclusion, the system to be solved is made of the two constraints $F = 0$, $G = 0$ and of

$$\frac{\partial}{\partial x}(f - \lambda F - \mu G) = 0\,, \qquad \frac{\partial}{\partial y}(f - \lambda F - \mu G) = 0\,, \qquad \frac{\partial}{\partial z}(f - \lambda F - \mu G) = 0\,. \tag{B.51}$$

The solution of the above yields the coordinates x_0, y_0, z_0 of the extremum and the two multipliers λ, μ.

B.7 Conformal Mapping

A *conformal mapping* is a variable transformation that preserves the angles, both in amplitude and orientation [7]. It is introduced here to show how the solution of a differential equation over a domain can be transformed into that over a different domain. To this purpose, consider a variable transformation in two dimensions defined by

$$u = u(x, y)\,, \qquad v = v(x, y)\,, \tag{B.52}$$

where x, y belong to a domain D and u, v to a domain D' and such that for all points of D,

$$J = \begin{bmatrix} \partial u/\partial x & \partial u/\partial y \\ \partial v/\partial x & \partial v/\partial y \end{bmatrix} \neq 0\,. \tag{B.53}$$

As a consequence, transformation (B.52) is invertible, namely, there exist the relations $x = x(u, v)$ and $y = y(u, v)$. It follows that starting from a function $\varphi(x, y)$ defined in D and letting $\psi(u, v) = \varphi[x(u, v), y(u, v)]$, it is

$$\varphi(x, y) = \psi[u(x, y), v(x, y)]\,. \tag{B.54}$$

Let φ be harmonic in D. Using the ∇ operator (Sect. A.3), it follows

$$\nabla^2\varphi = \frac{\partial\psi}{\partial u}\nabla^2 u + \frac{\partial\psi}{\partial v}\nabla^2 v + \frac{\partial^2\psi}{\partial u^2}|\nabla u|^2 + \frac{\partial^2\psi}{\partial v^2}|\nabla v|^2 + 2\,\frac{\partial^2\psi}{\partial u\partial v}\nabla u\cdot\nabla v = 0\,. \tag{B.55}$$

The conditions on the variable transformations are sought for which ψ is harmonic with respect to u, v. They are easily found as $\nabla^2 u = \nabla^2 v = 0$, $\nabla u \cdot \nabla v = 0$, and $|\nabla u|^2 = |\nabla v|^2 \neq 0$. The first condition shows that u and v must be harmonic. The second one yields

$$\frac{\partial u/\partial x}{\partial v/\partial y} = -\frac{\partial u/\partial y}{\partial v/\partial x} = \omega(x, y)\,, \tag{B.56}$$

with ω yet undetermined. As a consequence, the third condition becomes $|\nabla u|^2 = \omega^2 |\nabla v|^2$, thus yielding $\omega = \pm 1$ for all x, y. Taking $\omega = 1$ provides the Cauchy-Riemann relations of the first kind

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}\,, \qquad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}\,, \tag{B.57}$$

by which u and v are harmonic conjugate and the complex function defined by $u + jv$ is analytical in D. Alternatively, taking $\omega = -1$ provides the Cauchy-Riemann relations of the second kind

$$\frac{\partial u}{\partial x} = -\frac{\partial v}{\partial y}\,, \qquad \frac{\partial u}{\partial y} = \frac{\partial v}{\partial x}\,. \tag{B.58}$$

by which u and $-v$ are harmonic conjugate and the complex function defined by $u - jv$ is analytical in D. An invertible transformation fulfilling the Cauchy-Riemann relations determines a conformal mapping between the domains D and D' [7, 142].

By way of example, let $w = x^2 + (y + 1)^2$ and

$$u = (x^2 + y^2 - 1)/w\,, \qquad v = -2\,x/w\,. \tag{B.59}$$

It is easily found that u and v are harmonic and fulfill the Cauchy-Riemann relations of the first kind (B.57). In particular,

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} = 4\,\frac{x\,(y + 1)}{w^2}\,, \qquad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x} = 2\,\frac{(y + 1)^2 - x^2}{w^2}\,. \tag{B.60}$$

From the above the Jacobian determinant of the variable transformation turns out to be $|\nabla u|^2 = 4\,w^2 > 0$. To find the inverse relations one observes that from the above definitions,

$$u - 1 = -2\,(y + 1)/w\,, \qquad v = -2\,x/w\,. \tag{B.61}$$

Squaring (B.61), adding term by term, and using the definition of w yield $(u-1)^2 + v^2 = 4/w = -2\,v/x$ whereas, dividing term by term, yields $(u-1)/v = (y+1)/x$. Combining the above results and letting $m = (u - 1)^2 + v^2$, one finds

$$x = -2\,v/m\,, \qquad y = 2\,(1 - u)/m - 1\,. \tag{B.62}$$

In particular, letting $y = 0$ provides $u^2 + v^2 = 1$, thus showing that the variable transformation (B.59) carries the x axis of the x, y space onto the circumference C of unit radius centered in the origin of the u, v space. In turn, the upper half plane

$y > 0$ is carried onto the disk $u^2 + v^2 < 1$ whose boundary is C. In fact, inside the disk it is $m = u^2 - 2u + 1 + v^2 < 2(1-u)$, whence $y = 2(1-u)/m - 1 > 0$. At the same time, letting $u = 1 - \epsilon$, $v = \pm\epsilon$, yields $m = 2\epsilon^2$, whence $x = \mp 2/\epsilon$, showing that the values of $|x|$ may become arbitrarily large as u, v span the disk. A similar reasoning shows that the exterior of the disk, $u^2 + v^2 > 1$, is carried onto the lower half plane $y < 0$. When the point $Q = (\xi, 0)$ runs over the x axis from left to right, like, e.g., it does in the expression (4.67) of the electric potential, the corresponding point

$$Q' = \left(\frac{\xi^2 - 1}{\xi^2 + 1} , \frac{-2\xi}{\xi^2 + 1} \right) \tag{B.63}$$

makes a counterclockwise rotation around C, starting from (1,0). Letting for simplicity $\xi = -\cot(\gamma/2)$, in the u, v coordinates (4.67) becomes

$$\psi(u, v) = \varphi_0 - \frac{1}{2\pi} \int_0^{2\pi} E'(\gamma) \frac{G'(u, v; \gamma)}{1 - \cos\gamma} \, \mathrm{d}\gamma \,, \tag{B.64}$$

where $E'(\gamma) = E[\xi(\gamma)]$ and $G'(u, v; \gamma) = G[x(u, v), y(u, v); \xi(\gamma)]$. In conclusion, the function (B.64) obtained from $\varphi(x, y)$ by a transformation using harmonic-conjugate functions, is harmonic inside the disk of boundary C. Due to the form of the boundary it is more convenient to turn to the polar coordinates $u = \zeta \cos\vartheta$, $v = \zeta \sin\vartheta$, with $0 \le \zeta \le 1$, and $0 \le \vartheta \le 2\pi$. For the sake of simplicity the functions expressed in polar coordinates are indicated by the same symbols as before. From (B.64) it then follows

$$\psi(\zeta, \vartheta) = \varphi_0 - \frac{1}{2\pi} \int_0^{2\pi} E'(\gamma) \frac{G'(\zeta, \vartheta; \gamma)}{1 - \cos\gamma} \, \mathrm{d}\gamma \,, \tag{B.65}$$

$$G' = \log\left[1 + \zeta^2 - 2\zeta \cos(\vartheta - \gamma)\right] . \tag{B.66}$$

Using the same procedure as in Sect. 4.12.4 shows that the Neumann boundary conditions used there are left unaltered when the variable transformation is applied.

B.8 Contraction Mapping

Before illustrating the *contraction mapping* it is necessary to provide some preliminary concepts. Given a set X of elements $x, y, z, \ldots$, one defines a *distance* by means of a nonnegative, single-valued real function ϱ having the following properties: for any x, y, and z of X it is

1. $\rho(x, y) = 0$ if and only if $x = y$,
2. $\varrho(x, y) = \varrho(y, x)$ (*symmetry*),
3. $\varrho(x, z) \le \varrho(x, y) + \varrho(y, z)$ (*triangle inequality*).

A set equipped with a distance is called *metric space* and is indicated with R; its elements are called *points*. In general, there are several functions ϱ that fulfill the requisites of distance; each of them defines a different metric space.

A sequence $x_1\,, x_2\,, \ldots$, of points of a metric space R (also indicated with $\{x_n\}$) fulfills the *Cauchy criterion* if, given $\varepsilon > 0$, there exists an index n_ε such that $\rho(x_n\,, x_m) < \varepsilon$ for all $n\,, m > n_\varepsilon$; a sequence that fulfills the Cauchy criterion is called *Cauchy sequence*. A metric space R is *complete* if every Cauchy sequence $\{x_n\}$ of points of R converges to an element of R; otherwise, R is *incomplete*.[3] Examples of complete metric spaces are the space of real numbers with distance $\varrho(x\,, y) = |x-y|$, and the space of N-tuples $x = (x_1\,, x_2\,, \ldots\,, x_N)$ with Euclidean distance

$$\varrho(x\,, y) = \left(\sum_{k=1}^{N} |x_k - y_k|^n \right)^{1/n} , \qquad n = 2\,. \tag{B.67}$$

The triangle inequality for (B.67) is readily checked using the Schwarz inequality (A.5). The Euclidean distance (B.67) is generalized by letting $n = 3, 4, \ldots$. As n increases, the difference $x_k - y_k$ having the largest modulus becomes dominant and, in the $n \to \infty$ limit, (B.67) yields

$$\varrho(x\,, y) = \max_k |x_k - y_k|\,. \tag{B.68}$$

Now, consider a correspondence A between a point y of a metric space R and another point z of R. Such a correspondence is indicated with $z = A\,y$. For the time being, these symbols are used without a specific reference to the objects of interest; for instance, R could be the real line, an N-dimensional space, a space of functions, and so on; correspondingly, A could be a function proper, a matrix, an operator, and so on. Multiple applications of correspondence A are indicated with powers, e.g., $A^2 y$ means $A(A\,y)$. Given these premises, a point x of R is called *fixed point* of A if A brings x on itself, namely,

$$x = A\,x\,. \tag{B.69}$$

In turn, correspondence A is called *contraction mapping* or *contraction* if, for any pair of points x, y of R there exists a positive number $\alpha < 1$ such that

$$\varrho(A\,x\,, A\,y) \le \alpha\, \varrho(x\,, y)\,. \tag{B.70}$$

From (B.70) it follows that a contraction is continuous. The theorem holds [81, Sect. 2-8.1]: *Any contraction defined over a complete metric space has one and only one*

[3] For instance, in the metric space of rational numbers Z with distance $\rho(x\,, y) = |x - y|$, consider the sequence whose nth element has the integer part equal to unity and the decimal part made of the first n decimal figures of $\sqrt{2}$: $x_1 = 1.4$, $x_2 = 1.41$, $x_3 = 1.414$, and so on. This sequence does not converge in Z.

fixed point. The theorem, called *fixed-point theorem* or *successive-approximations theorem*, can be exploited to prove the existence and uniqueness of the solution of several classes of equations. As shown in the examples below, it also provides practical methods for calculating the solution.

B.8.1 Determining the Zero of a Function

Let φ be a function defined over the interval $I = [a, b]$, such that $\varphi(x)$ belongs to I as well; also, for any pair p, q of I, let φ fulfill the *Lipschitz condition*

$$|\varphi(p) - \varphi(q)| \leq K\ |p - q|\ , \qquad 0 < K < 1\ . \tag{B.71}$$

As $|p-q|$ is a distance proper, φ is a contraction and, given the initial point x_0 within I, the sequence $x_0\ ,\ \ x_1 = \varphi(x_0)\ ,\ \ x_2 = \varphi(x_1)\ , \ldots$, converges to the unique root of the equation $\varphi(x) = x$. If, in particular, φ is differentiable in I, for the mean-value theorem there exists $\xi \in [p, q]$ such that $[\varphi(p) - \varphi(q)]/(p - q) = \varphi'(\xi)$. Observing that (B.71) applies to any pair p, q, one may recast it as

$$\left|\varphi'\right| \leq K < 1 \qquad \text{in} \quad I\ . \tag{B.72}$$

The example shows that to determine the zero of an equation of the form $f(x) = 0$ within an interval I, one may resort to an auxiliary equation $\varphi(x) = x$ such that φ is a contraction in I and its fixed point coincides with the zero of f. For a given function f there are in general several possibilities to define such a φ; one of these, particularly interesting for the applications, provides the so-called *Newton's method*, extensively used in the numerical solution of the semiconductor equations (Sect. A.13.1).

B.8.2 Solving an Algebraic System

Let R be a real, N-dimensional space, over which the following distance between two points x, $\tilde{x}$ is selected:

$$\varrho(x\,, \tilde{x}) = \max_{1 \leq i \leq N} |x_i - \tilde{x}_i|\ . \tag{B.73}$$

Over the metric space thus defined, consider the correspondence $y = A\,x$ given by the set of linear relations

$$y_i = \sum_{j=1}^{N} c_{ij}\, x_j + b_i\,, \qquad i = 1\,, \ldots\,, N\,. \tag{B.74}$$

To determine the conditions that make A a contraction, consider any pair of points $x, \tilde{x}$ and the corresponding points $y = A\,x$, $\tilde{y} = A\,\tilde{x}$. From $y_i - \tilde{y}_i = \sum_{j=1}^{N} c_{ij}\,(x_j - \tilde{x}_j)$ one obtains, for every index i,

$$|y_i - \tilde{y}_i| \le \sum_{j=1}^{N} |c_{ij}|\ |x_j - \tilde{x}_j| \le \varrho(x\,,\tilde{x})\ \sum_{j=1}^{N} |c_{ij}|\ . \tag{B.75}$$

It follows

$$\rho(y\,,\tilde{y}) = \max_{1\le i\le N} |y_i - \tilde{y}_i| \le \rho(x\,,\tilde{x})\ \max_{1\le i\le N} \sum_{j=1}^{N} |c_{ij}|\ . \tag{B.76}$$

Comparing with (B.70) one finds

$$\alpha = \max_{1\le i\le N} \sum_{j=1}^{N} |c_{ij}| < 1\,, \qquad \sum_{j=1}^{N} |c_{ij}| \le \alpha < 1\,, \qquad i = 1\,,\dots\,,N\,. \tag{B.77}$$

If (B.77) holds, then the sequence $x_1 = A\,x_0\,,\ \ x_2 = A\,x_1 = A^2\,x_0\,,\ \ \dots$, with x_0 an arbitrary initial point, converges to the fixed point of A. Indicating the latter with z, it is

$$z_i = \sum_{j=1}^{N} c_{ij}\,z_j + b_i\,, \qquad i = 1\,,\dots\,,N \tag{B.78}$$

so that, letting $c_{ii} = 1 - a_{ii}$ and $c_{ij} = -a_{ij}$ for $j \ne i$, one finds that z is the solution of the algebraic system[4]

$$\sum_{j=1}^{N} a_{ij}\,x_j = b_i\,, \qquad i = 1\,,\dots\,,N\,. \tag{B.79}$$

The uniqueness of the solution implies that the matrix of (B.79) is non-singular. It is interesting to note that by a suitable manipulation one can always make the diagonal entries a_{ii} negative, whence $|c_{ii}| = |1 - a_{ii}| = 1 + |a_{ii}|$; from (B.77) it follows

$$\sum_{j=1}^{N} |a_{ij}| < \sum_{j=1}^{N} |c_{ij}| \le \alpha < 1\,, \qquad i = 1\,,\dots\,,N\,. \tag{B.80}$$

[4]It is worth mentioning that other definitions, different from (B.73), are possible for the distance between x and $\tilde{x}$. It follows that condition (B.77) is sufficient, but not necessary, for the existence and uniqueness of the solution of the algebraic system (B.79).

Comparing with the definitions and results of Sect. A.11.1 one finds that if the spectral radius of a non-singular matrix is smaller than unity, the contraction mapping described above is *unconditionally convergent*, namely, it converges for any choice of the initial point x_0.

Appendix C
Special Integrals

C.1 Sine Integral

Define the two functions

$$\mathrm{si}(t) = -\frac{\pi}{2} + \int_0^t \frac{\sin x}{x}\,\mathrm{d}x\,, \qquad N(a) = \int_0^\infty \frac{\sin(ax)}{x}\,\mathrm{d}x\,. \tag{C.1}$$

The first of them is called *sine integral* and fulfills the limit $\lim_{t\to\infty} \mathrm{si} = 0$, whence $N(1) = \pi/2$. To demonstrate the above one starts from the functions

$$F(y) = \int_0^\infty \exp(-x)\,\frac{\sin(xy)}{x}\,\mathrm{d}x\,, \quad G(y) = \int_0^\infty \exp(-xy)\,\frac{\sin x}{x}\,\mathrm{d}x\,, \quad y \geq 0\,. \tag{C.2}$$

The following hold true: $F(0) = 0$, $G(0) = N(1)$, $F(1) = G(1)$, and

$$\frac{\mathrm{d}F}{\mathrm{d}y} = \int_0^\infty \exp(-x)\,\cos(xy)\,\mathrm{d}x\,, \qquad \frac{\mathrm{d}G}{\mathrm{d}y} = \int_0^\infty \exp(-xy)\,\sin x\,\mathrm{d}x\,. \tag{C.3}$$

Integrating (C.3) by parts twice yields $\mathrm{d}F/\mathrm{d}y = 1/(1+y^2)$, $\mathrm{d}G/\mathrm{d}y = -1/(1+y^2)$ whence

$$F(y) = \arctan y + F(0)\,, \quad G(y) = -\arctan y + G(0)\,, \quad 0 \leq y < \frac{\pi}{2}\,. \tag{C.4}$$

It follows $F(1) = F(0) + \pi/4 = \pi/4$ and $F(1) = G(1) = G(0) - \pi/4 = N(1) - \pi/4$. Combining the above yields the result sought. This implicitly proves the convergence of the integrals in (C.1). The calculation of the second of (C.1) is now straightforward and yields

M. Rudan, *Physics of Semiconductor Devices*, DOI 10.1007/978-3-319-63154-7

$$N(a) = \begin{cases} -\pi/2\,, & a < 0\,, \\ 0\,, & a = 0\,, \\ +\pi/2\,, & a > 0\,. \end{cases} \tag{C.5}$$

The integrand in the second of (C.1) is even with respect to x. It follows that an integration carried out from $-\infty$ to $+\infty$ yields $2\,N(a)$. Basing on this one also finds

$$\int_{-\infty}^{+\infty} \frac{\exp(\mathrm{i}\,ax)}{\mathrm{i}\,x}\,\mathrm{d}x = 2\,N(a) + \int_{-\infty}^{+\infty} \frac{\cos(ax)}{\mathrm{i}\,x}\,\mathrm{d}x = 2\,N(a)\,. \tag{C.6}$$

When calculating the second integral in (C.6) one must let $z = \pm ax$, $\epsilon, Z > 0$ and use the principal part. In fact, observing that the integrand is odd one obtains

$$\int_{-\infty}^{+\infty} \frac{\cos(ax)}{\mathrm{i}\,x}\,\mathrm{d}x = \pm\mathrm{i} \lim_{\substack{\epsilon\to 0 \\ Z\to\infty}} \left(\int_{-Z}^{-\epsilon} \frac{\cos z}{z}\,\mathrm{d}z + \int_{+\epsilon}^{+Z} \frac{\cos z}{z}\,\mathrm{d}z \right) = 0\,. \tag{C.7}$$

Combining (C.7) with (C.6) provides an integral representation of the Fourier type for the *step function*

$$H(a) = \left\{ \begin{matrix} 0 \\ 1/2 \\ 1 \end{matrix} \quad \begin{matrix} a < 0 \\ a = 0 \\ a > 0 \end{matrix} \right\} = \frac{1}{2} + \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{\exp(\mathrm{i}ax)}{\mathrm{i}x}\,\mathrm{d}x\,. \tag{C.8}$$

Still from (C.6), using the identity $2\,\mathrm{i}\,\sin x = \exp(\mathrm{i}\,x) - \exp(-\mathrm{i}\,x)$, one finds

$$\int_{-\infty}^{+\infty} \frac{\sin x}{x}\,\exp(-\mathrm{i}\,a\,x)\,\mathrm{d}x = N(-a+1) - N(-a-1) = \begin{cases} 0 & |a| > 1 \\ \pi/2 & a = \pm 1 \\ \pi & |a| < 1 \end{cases} \tag{C.9}$$

From (C.9) one derives integrals of a similar form, where $\sin x/x$ is replaced with $\sin^n x/x^n$, $n = 2, 3, \ldots$. The example with $n = 2$ is given below: one starts from

$$\frac{\mathrm{d}}{\mathrm{d}a} \int_{-\infty}^{+\infty} \frac{\sin^2 x}{x^2}\,\exp(-\mathrm{i}\,a\,x)\,\mathrm{d}x = \int_{-\infty}^{+\infty} \frac{\sin^2 x}{\mathrm{i}\,x}\,\exp(-\mathrm{i}\,a\,x)\,\mathrm{d}x\,, \tag{C.10}$$

and uses the identity $2\,\sin^2 x = 1 - \cos(2\,x)$ to find

$$\int_{-\infty}^{+\infty} \frac{1-\cos(2\,x)}{2\,\mathrm{i}\,x}\,\exp(-\mathrm{i}\,a\,x)\,\mathrm{d}x = N(-a) + \int_{-\infty}^{+\infty} \frac{\cos(2\,x)}{2\,x}\,\sin(a\,x)\,\mathrm{d}x\,, \tag{C.11}$$

where $N(-a)$ derives from (C.6) and the integral on the right-hand side is obtained by eliminating the odd part of the integrand. From the identity $\sin[(a+2)\,x] + \sin[(a-2)\,x] = 2\,\sin(a\,x)\,\cos(2\,x)$ such an integral transforms into

$$\int_{-\infty}^{+\infty} \frac{\sin[(a+2)\,x] + \sin[(a-2)\,x]}{4\,x}\,\mathrm{d}x = \frac{1}{2}\,N(a+2) + \frac{1}{2}\,N(a-2)\,, \qquad \text{(C.12)}$$

where the second definition in (C.1) has been used. Combining (C.10), (C.11), and (C.12) yields

$$\frac{\mathrm{d}}{\mathrm{d}a}\int_{-\infty}^{+\infty} \frac{\sin^2 x}{x^2}\,\exp(-\mathrm{i}\,a\,x)\,\mathrm{d}x = \begin{cases} \pi/2 & -2 < a < 0 \\ -\pi/2 & 0 < a < 2 \\ 0 & |a| > 2 \end{cases} \qquad \text{(C.13)}$$

This result shows that the derivative with respect to a of the integral sought is piecewise constant in the interval $-2 < a < +2$, and vanishes elsewhere. The integral is also continuous with respect to a and should not diverge, because $|\sin^2 x/x^2| \le |\sin x/x|$ and (C.9) converges. This reasoning allows one to fix the integration constants, to finally obtain

$$\int_{-\infty}^{+\infty} \frac{\sin^2 x}{x^2}\,\exp(-\mathrm{i}\,a\,x)\,\mathrm{d}x = \begin{cases} (\pi/2)\,(a+2) & -2 < a < 0 \\ -(\pi/2)\,(a-2) & 0 < a < 2 \\ 0 & |a| > 2 \end{cases} \qquad \text{(C.14)}$$

By a procedure similar to that used to prove (C.14) one finds

$$\frac{\mathrm{d}}{\mathrm{d}a}\int_{-\infty}^{+\infty} \frac{\sin^2(a\,x)}{x^2}\,\mathrm{d}x = 2\,N(a)\,, \qquad \int_{-\infty}^{+\infty} \frac{\sin^2(a\,x)}{x^2}\,\mathrm{d}x = \begin{cases} \pi\,a\,, & a > 0 \\ -\pi\,a\,, & a < 0 \end{cases} \qquad \text{(C.15)}$$

C.2 Fourier Transform

Let $f(x)$ be a function defined over the entire x axis. Its Fourier transform is defined as the integral

$$G(k) = \mathscr{F}_x f = \frac{1}{\sqrt{2\,\pi}}\int_{-\infty}^{+\infty} f(x)\,\exp(-\mathrm{i}\,k\,x)\,\mathrm{d}x\,. \qquad \text{(C.16)}$$

In turn, the Fourier antitransform is defined as

$$f(x) = \mathscr{F}_x^{-1} G = \frac{1}{\sqrt{2\,\pi}}\int_{-\infty}^{+\infty} G(k)\,\exp(\mathrm{i}\,k\,x)\,\mathrm{d}k\,. \qquad \text{(C.17)}$$

Combining (C.16) and (C.17) provides a representation of f in the form

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp(\mathrm{i}\, k\, x) \left[\int_{-\infty}^{+\infty} f(\xi)\, \exp(-\mathrm{i}\, k\, \xi)\, \mathrm{d}\xi \right] \mathrm{d}k\,. \tag{C.18}$$

A sufficient condition for the representation (C.18) is

$$\int_{-\infty}^{+\infty} |f(x)|\, \mathrm{d}x < \infty\,. \tag{C.19}$$

If f is discontinuous of the first kind at some point x_0, the left-hand side of (C.18) must be replaced with $[f(x_0^+) + f(x_0^-)]/2$. As the condition (C.19) is sufficient, but not necessary, there are functions that admit an integral representation like (C.18) without fulfilling (C.19). An important example is the unit step function shown in Sect. C.1.

If f depends also on one or more parameters, $f = f(x, u, v, \ldots)$, then it is $G = G(k, u, v, \ldots)$. In an n-dimensional space, defining the vectors $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{k} = (k_1, \ldots, k_n)$, the Fourier transform reads

$$G(\mathbf{k}) = \mathscr{F}_{\mathbf{x}} f = \frac{1}{(2\pi)^{n/2}} \int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty} f(\mathbf{x})\, \exp(-\mathrm{i}\, \mathbf{k} \cdot \mathbf{x})\, \mathrm{d}x_1 \ldots \mathrm{d}x_n\,. \tag{C.20}$$

A useful relation is found by differentiating both sides of (C.17). To this purpose, one must assume that the conditions for exchanging the derivative with the integral are fulfilled. It is found

$$\frac{\mathrm{d}f}{\mathrm{d}x} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathrm{i}\, k\, G(k)\, \exp(\mathrm{i}kx)\, \mathrm{d}k\,. \tag{C.21}$$

Iterating the procedure yields

$$\frac{\mathrm{d}^n f}{\mathrm{d}x^n} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} (\mathrm{i}\, k)^n\, G(k)\, \exp(\mathrm{i}kx)\, \mathrm{d}k\,, \tag{C.22}$$

showing that if $G(k)$ is the Fourier transform of $f(x)$, then the Fourier transform of $\mathrm{d}^n f/\mathrm{d}x^x$ is $(\mathrm{i}\, k)^n\, G(k)$. Relations like (C.21) and (C.22) are useful, for instance, in the solution of linear differential equations with constant coefficients, because they turn differential relations into polynomial relations (compare with the solution of the diffusion equation carried out in Sect. 23.4).

C.3 Gauss Integral

The relation

$$I_G = \int_0^{+\infty} \exp(-x^2)\, \mathrm{d}x = \int_{-\infty}^{0} \exp(-x^2)\, \mathrm{d}x\,. \tag{C.23}$$

is called *Gauss integral* or *Poisson integral*. To calculate its value one may start from the double integral

$$F(R) = \iint_{\Sigma(R)} \exp[-(x^2 + y^2)]\, \mathrm{d}x\mathrm{d}y\,, \tag{C.24}$$

where $\Sigma(R)$ is a circle of radius R centered on the origin. Using the polar coordinates (B.8) yields

$$F(R) = \int_0^{2\pi} \mathrm{d}\vartheta \int_0^R \exp(-\rho^2)\, \rho\, \mathrm{d}\rho = \pi\, [1 - \exp(-R^2)]\,, \tag{C.25}$$

whence $\lim_{R\to\infty} F(R) = \pi$. On the other hand, due to (C.24) it is also

$$\lim_{R\to\infty} F(R) = \iint_{-\infty}^{+\infty} \exp[-(x^2 + y^2)]\, \mathrm{d}x\mathrm{d}y = \lim_{a\to\infty} \left(\int_{-a}^{+a} \exp(-x^2)\, \mathrm{d}x \right)^2 . \tag{C.26}$$

Combining (C.25, C.26) with (C.23) provides

$$\int_{-\infty}^{+\infty} \exp(-x^2)\, \mathrm{d}x = \sqrt{\pi}\,, \qquad I_G = \frac{\sqrt{\pi}}{2}\,. \tag{C.27}$$

From (C.27) it follows that for any $\lambda > 0$ it is

$$I_0(\lambda) = \int_0^{\infty} \exp(-\lambda x^2)\, \mathrm{d}x = \frac{1}{2}\sqrt{\frac{\pi}{\lambda}}\,. \tag{C.28}$$

Another integral generated by $\exp(-\lambda x^2)$ is

$$I_1(\lambda) = \int_0^{\infty} x\, \exp(-\lambda x^2)\, \mathrm{d}x = \frac{1}{2\lambda}\,. \tag{C.29}$$

Thanks to (C.28) and (C.29) it is possible to calculate all integrals of the form

$$I_n(\lambda) = \int_0^{\infty} x^n\, \exp(-\lambda x^2)\, \mathrm{d}x\,, \qquad n \geq 0\,. \tag{C.30}$$

In fact, using the recursive relation

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} I_n = \int_0^{\infty} \frac{\partial}{\partial\lambda} x^n\, \exp(-\lambda x^2)\, \mathrm{d}x = -\int_0^{\infty} x^{n+2}\, \exp(-\lambda x^2)\, \mathrm{d}x = -I_{n+2}\,, \tag{C.31}$$

in combination with (C.29) yields all the integrals whose index is odd,

$$I_{2m+1} = \frac{m!}{2}\, \lambda^{-(m+1)}\,, \qquad m = 0\,, 1\,, 2\,, \ldots\,. \tag{C.32}$$

Similarly, combining (C.31) with (C.28) yields all the integrals whose index is even,

$$I_{2m}(\lambda) = \frac{(2m-1)!!}{2^{m+1}}\,\lambda^{-(m+1/2)}\,\sqrt{\pi}\,, \qquad m = 0\,,1\,,2\,,\ldots\,, \tag{C.33}$$

where

$$(2m-1)!! = (2m-1)(2m-3)\ldots 3\cdot 1\,, \qquad (-1)!! = 1\,, \tag{C.34}$$

Finally, observing that the integrand of (C.30) is even (odd) if n is even (odd), one finds

$$\int_{-\infty}^{+\infty} x^{2m}\,\exp(-\lambda x^2)\,\mathrm{d}x = 2\,I_{2m}(\lambda)\,, \qquad \int_{-\infty}^{+\infty} x^{2m+1}\,\exp(-\lambda x^2)\,\mathrm{d}x = 0\,. \tag{C.35}$$

The results of this section still hold for a complex λ with $\Re\lambda > 0$.

C.4 Dirac's δ

Consider a function $\Delta_B(x,a)$ defined as follows:

$$\Delta_B = \begin{cases} 1/a & -a/2 \le x \le +a/2 \\ 0 & x < -a/2\,, \quad x > a/2 \end{cases} \tag{C.36}$$

with $a > 0$. The above definition yields

$$\lim_{a\to 0}\Delta_B = \begin{cases} 0 & x \neq 0 \\ +\infty & x = 0 \end{cases}\,, \qquad \int_{-\infty}^{+\infty}\Delta_B(x,a)\,\mathrm{d}x = \frac{1}{a}\int_{-a/2}^{+a/2}\mathrm{d}x = 1\,. \tag{C.37}$$

As the value of the integral in (C.37) is independent of a, the integral is equal to unity also in the limit $a \to 0$. Figure C.1 shows how the form of Δ_B changes with a: the width of the peak decreases as a decreases, while its height increases so that the area subtending the function remains constant. Note that the procedure depicted above gives a different result if one carries out the integration after calculating the limit. In other terms, the integration and the limit are to be carried out in a specific order (integration first). For a continuous function $f(x)$ the mean-value theorem provides

$$\int_{-\infty}^{+\infty}\Delta_B(x,a)\,f(x)\,\mathrm{d}x = \frac{1}{a}\int_{-a/2}^{+a/2} f(x)\,\mathrm{d}x = f(\bar{x})\,, \tag{C.38}$$

with $-a/2 < \bar{x} < +a/2$. As a consequence,

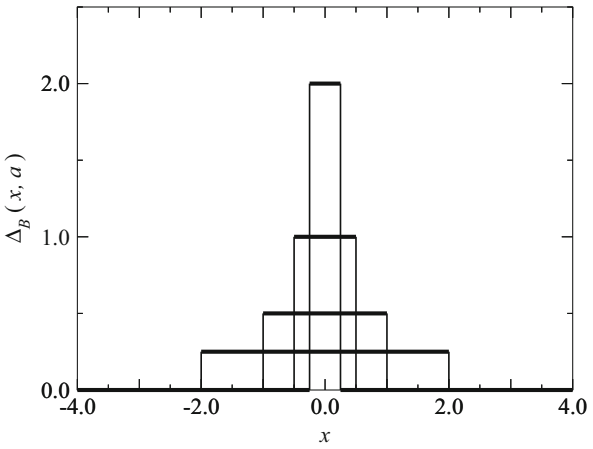


Fig. C.1 Generation of a Dirac δ using a barrier-like function. The peak's width is equal to a

$$\lim_{a\to 0} \int_{-\infty}^{+\infty} \Delta_B(x,a) f(x)\,\mathrm{d}x = f(0)\,. \tag{C.39}$$

This result is expressed in a more compact form by defining a linear functional $\delta(x)$ (called *Dirac's symbol*) such that

$$\int_{-\infty}^{+\infty} \delta(x) f(x)\,\mathrm{d}x = f(0)\,. \tag{C.40}$$

The functional associates the number $f(0)$ with the function $f(x)$. If the reasoning leading to (C.40) is repeated after shifting Δ_B from the origin to another point x_0, one finds the generalization of (C.40)

$$\int_{-\infty}^{+\infty} \delta(x-x_0) f(x)\,\mathrm{d}x = f(x_0)\,. \tag{C.41}$$

From (C.41) and (C.16) one obtains

$$\int_{-\infty}^{+\infty} \delta(x-x_0)\,\mathrm{d}x = 1\,, \qquad \mathscr{F}_x \delta(x-x_0) = \frac{1}{\sqrt{2\,\pi}} \exp(-\mathrm{i}\,k\,x_0)\,. \tag{C.42}$$

The antitransform (C.17) then reads

$$\delta(x-x_0) = \frac{1}{2\,\pi} \int_{-\infty}^{+\infty} \exp[\mathrm{i}\,k\,(x-x_0)]\,\mathrm{d}k\,, \tag{C.43}$$

that provides an integral representation of the Dirac δ. However, it is important to note that (C.43) has no meaning unless it is used within an integral like, e.g., (C.41). With this provision, one can consider the Dirac δ as the "derivative" of the step

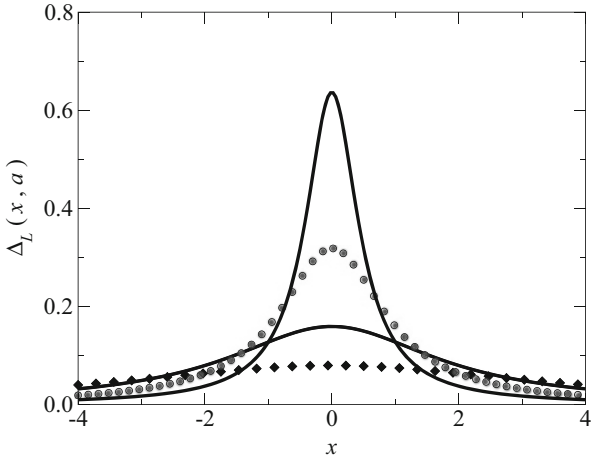


Fig. C.2 Generation of a Dirac δ using a Lorentzian function. The peak's width is proportional to a

function; in fact, after a suitable change in the symbols, one finds that the integral at the right-hand side of (C.43) is the derivative with respect to x of the integral at the right-hand side of (C.8). More details about the integral representation of the Dirac δ are given in Sect. C.5.

The function $\Delta_B(x, a)$ defined above is an example of *generating function* of the Dirac δ. Several other examples may be given, as shown below. In all cases, if the generating function $\Delta(x, x_0, a)$ is centered at some point x_0, it is even with respect to x_0 and has the properties $\lim_{a\to 0} \Delta = 0$ if $x \neq x_0$ and $\lim_{a\to 0} \Delta = +\infty$ if $x = x_0$. Consider for instance the *Lorentzian function* (centered at $x_0 = 0$)

$$\Delta_L = \frac{a/\pi}{a^2 + x^2}\,, \qquad \int_{-\infty}^{+\infty} \Delta_L \,\mathrm{d}x = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{\mathrm{d}}{\mathrm{d}x} \arctan\left(\frac{x}{a}\right) \mathrm{d}x = 1\,, \tag{C.44}$$

with $a > 0$. Apart from the limiting case $a \to 0$ the function has only one maximum that occurs at $x = 0$ and equals $1/(a\pi)$. For $x = \pm a$ the function's value halves with respect to the maximum, so $2a$ is conventionally taken as the width of Δ_L. The product $2/\pi$ of the maximum value by the conventional width is independent of a and is of order unity (Figs. C.2 and C.3). Finally, for a continuous function $f(x)$ it is

$$\lim_{a\to 0} \int_{-\infty}^{+\infty} \Delta_L(x, a) f(x)\,\mathrm{d}x = f(0)\,. \tag{C.45}$$

Another example of a δ-generating function is the *parameterized Gaussian function* (centered at $x_0 = 0$)

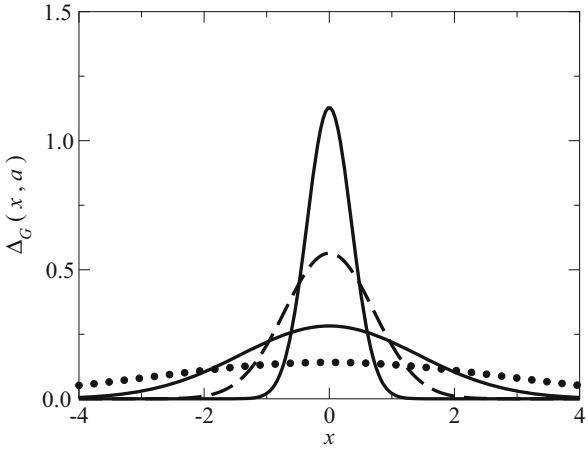


Fig. C.3 Generation of a Dirac δ using a parameterized Gaussian function. The peak's width is proportional to a

$$\Delta_G = \frac{\exp(-x^2/a^2)}{a\sqrt{\pi}}\,, \quad a > 0\,, \qquad \int_{-\infty}^{+\infty} \Delta_G(x,a)\,\mathrm{d}x = 1 \tag{C.46}$$

(more details about this function and integrals related to it are given in Sects. C.3 and C.8). The function has only one maximum that occurs at $x = 0$ and equals $1/(a\sqrt{\pi})$. For $x = \pm a\sqrt{\log 2} \simeq \pm 0.833\,a$ the function's value halves with respect to the maximum, thus yielding a conventional width of $2a\sqrt{\log 2}$. The product $2\sqrt{\log 2}/\sqrt{\pi}$ of the maximum value by the conventional width is independent of a and of order unity. For a continuous function $f(x)$ it is

$$\lim_{a\to 0} \int_{-\infty}^{+\infty} \Delta_G(x,a) f(x)\,\mathrm{d}x = f(0)\,. \tag{C.47}$$

A final example of a δ-generating function is the negative derivative of the *Fermi function* (centered at $x_0 = 0$)

$$\Delta_F = -\frac{\mathrm{d}}{\mathrm{d}x}\,\frac{1}{\exp(x/a)+1} = \frac{\exp(x/a)}{a\,[\exp(x/a)+1]^2}\,, \qquad a > 0\,, \tag{C.48}$$

$$\int_{-\infty}^{+\infty} \Delta_F(x,a)\,\mathrm{d}x = \int_{+\infty}^{-\infty} \frac{\mathrm{d}}{\mathrm{d}x}\,\frac{1}{\exp(x/a)+1}\,\mathrm{d}x = 1\,. \tag{C.49}$$

(more details about this function and integrals related to it are given in Sect. C.13). The function has only one maximum that occurs at $x = 0$ and equals $1/(4a)$. For $x = \pm a \log(3+\sqrt{8}) \simeq \pm 1.76\,a$ the function's value halves with respect to the maximum, thus yielding a conventional width of $2a\log(3+\sqrt{8})$. The product $(1/2)\,\log(3+\sqrt{8})$ of the maximum value by the conventional width is independent of a and of order unity (Fig. C.4). For a continuous function $f(x)$ it is

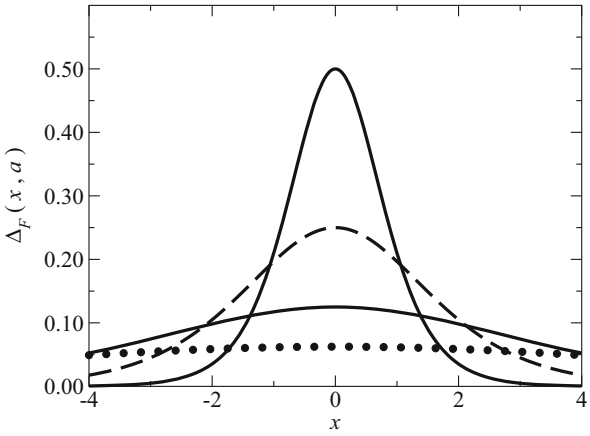


Fig. C.4 Generation of a Dirac δ using a Fermi function. The peak's width is proportional to a

$$\lim_{a\to 0}\int_{-\infty}^{+\infty} \Delta_F(x,a) f(x)\,\mathrm{d}x = f(0)\,. \tag{C.50}$$

The δ-generating functions Δ vanish for $x \to \pm\infty$, otherwise they would not be integrable from $-\infty$ to $+\infty$. Assuming that Δ is differentiable with respect to x yields, after integrating by parts,

$$\int_{-\infty}^{+\infty} f(x)\,\frac{\mathrm{d}\Delta(x,a)}{\mathrm{d}x}\,\mathrm{d}x = [\Delta(x,a) f(x)]_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} \Delta(x,a)\,\frac{\mathrm{d}f}{\mathrm{d}x}\,\mathrm{d}x\,, \tag{C.51}$$

with f a differentiable function. In (C.51) the integrated part is zero because Δ vanishes at infinity. Taking the limit $a \to 0$ at both sides of (C.51) and using (C.40) yields

$$\int_{-\infty}^{+\infty} f(x)\,\frac{\mathrm{d}\delta(x)}{\mathrm{d}x}\,\mathrm{d}x = -\int_{-\infty}^{+\infty} \delta(x)\,\frac{\mathrm{d}f}{\mathrm{d}x}\,\mathrm{d}x = -f'(0)\,, \tag{C.52}$$

which is used as the definition of the derivative of δ. Such a definition generalizes to

$$\int_{-\infty}^{+\infty} f(x)\,\frac{\mathrm{d}^n\delta(x)}{\mathrm{d}x^n}\,\mathrm{d}x = (-1)^n f^{(n)}(0)\,. \tag{C.53}$$

One notes in passing that the Fermi-Dirac statistics $P(x) = 1/[\exp(x/a)+1]$ fulfills the relation $P(-x) = 1 - P(x)$, whence $P(-x) - 1/2 = -[P(x) - 1/2]$, namely, $H(x) = P(x) - 1/2$ is an odd function of x. As a consequence, the derivative of P is even; in particular, one finds $-a\,\mathrm{d}P/\mathrm{d}x = P\,(1-P) = P(x)\,P(-x)$. These properties are exploited, e.g., in Prob. 15.2.

C.5 Some Properties of Dirac's δ

An integral representation of δ is derived from (C.18) after rearranging it as

$$f(x) = \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{+\infty} \frac{\exp[\mathrm{i}\, k\, (x-\xi)]}{2\,\pi}\, \mathrm{d}k \right] f(\xi)\, \mathrm{d}\xi \tag{C.54}$$

and comparing with (C.41):

$$\delta(\xi - x) = \int_{-\infty}^{+\infty} \frac{\exp[\mathrm{i}\, k\, (x-\xi)]}{2\,\pi}\, \mathrm{d}k\,. \tag{C.55}$$

Replacing k with $-k$ in (C.55) shows that δ is even with respect to its argument, $\delta(x-\xi) = \delta(\xi - x)$. Also, comparing (C.55) with (C.16) shows that $\delta(\xi - x)$ is the Fourier transform of $\exp(\mathrm{i}\, k\, x)/\sqrt{2\,\pi}$. The generalization of (C.55) to more than one dimension is immediate; e.g., the three-dimensional case reads

$$\delta(\mathbf{g} - \mathbf{x}) = \iiint_{-\infty}^{+\infty} \frac{\exp[\mathrm{i}\, \mathbf{k} \cdot (\mathbf{x} - \mathbf{g})]}{(2\,\pi)^3}\, \mathrm{d}^3 k\,. \tag{C.56}$$

The discrete-case analogue of (C.56) is given by (C.126, C.130), where the generalization of the Kronecker symbol is given. Note that the latter is dimensionless, whereas the units of Dirac's δ depend on its argument: by way of example, the integral $\int_{-\infty}^{+\infty} \delta(\xi - x)\, \mathrm{d}\xi = 1$ shows that the units of $\delta(\xi - x)$ are the inverse of those of $\mathrm{d}\xi$; similarly, the integral $\iiint_{-\infty}^{+\infty} \delta(\mathbf{g} - \mathbf{x})\, \mathrm{d}^3 g = 1$ shows that the units of $\delta(\mathbf{g} - \mathbf{x})$ are the inverse of those of $\mathrm{d}^3 g$, and so on.

A generalization of Dirac's δ is found by replacing $\delta(x)$ with $\delta[q(x)]$, with $q(x)$ a function having one or more zeros. Let x_1 be a simple zero of q, namely, $q'(x_1) \neq 0$, and consider the contribution of it to the integral $\int_{-\infty}^{+\infty} \delta[q(x)]\, \mathrm{d}x$. Observing that in a finite neighborhood I_1 of x_1 there are no other zeros, one can determine such a contribution by replacing $q(x)$ with $q'(x_1)\,(x - x_1)$; in this way, to bring the calculation back to the standard form one may provisionally scale the differential $\mathrm{d}x$ by $1/q'(x_1)$. However, if the scaling factor was negative, the evenness of δ would be violated; thus, the correct scaling factor is $|q'(x_1)|$, and

$$\int_{I_1} \delta[q(x)] f(x)\ \mathrm{d}x = \frac{1}{|q'(x_1)|} f(x_1)\,. \tag{C.57}$$

If q has n simple zeros, from (C.57) it follows

$$\int_{-\infty}^{+\infty} \delta[q(x)] f(x)\ \mathrm{d}x = \frac{1}{|q'(x_1)|} f(x_1) + \ldots + \frac{1}{|q'(x_n)|} f(x_n)\,. \tag{C.58}$$

C.6 Moments Expansion

For a given function $f(k)$ consider the integral

$$M_n = \int_{-\infty}^{+\infty} k^n f(k)\,\mathrm{d}k\,, \qquad n = 0, 1, \ldots \tag{C.59}$$

It is assumed that the integral converges for any n. This implies that f vanishes at infinity with a strength larger than any power. As the present considerations apply to a distribution function, the vanishing of f is typically of the exponential type. The quantity M_n is called *moment of order n* of function f. Thanks to its properties, f can be Fourier transformed; let

$$g(y) = \mathscr{F} f = \frac{1}{\sqrt{2\,\pi}} \int_{-\infty}^{+\infty} f(k)\, \exp(-\mathrm{i}\, y\, k)\,\mathrm{d}k\,. \tag{C.60}$$

Using the Taylor expansion $\exp(-\mathrm{i}\, y\, k) = \sum_{n=0}^{\infty} (-\mathrm{i}\, y\, k)^n / n!$ yields

$$g(y) = \sum_{n=0}^{\infty} \frac{1}{n!}\, \frac{(-\mathrm{i})^n\, M_n}{\sqrt{2\,\pi}}\, y^n\,. \tag{C.61}$$

The above is the Taylor expansion of g around the origin; it follows

$$\frac{(-\mathrm{i})^n\, M_n}{\sqrt{2\,\pi}} = \left(\frac{\mathrm{d}^n g}{\mathrm{d}y^n} \right)_0\,. \tag{C.62}$$

The above analysis shows that if the moments M_n of $f(k)$ are known, from them one constructs the Fourier transform $g(y) = \mathscr{F} f$ by means of a Taylor series. Then, one recovers the original function from the inverse transform $f(k) = \mathscr{F}^{-1} g$. In conclusion, the knowledge of the set of moments of f is equivalent to the knowledge of f. The result holds true also in the multi-dimensional case $f = f(\mathbf{k})$, where

$$M_{l+m+n} = \iiint_{-\infty}^{+\infty} k_1^l\, k_2^m\, k_3^n f(\mathbf{k})\,\mathrm{d}^3 k\,, \qquad l, m, n = 0, 1, \ldots \tag{C.63}$$

is the moment of order $l + m + n$ of f.

If only the lower-order moments are used, then the Taylor series for the Fourier transform is truncated and provides an approximation $\tilde{g}$ for g. As a consequence of this approximation, the inverse transform $\tilde{f} = \mathscr{F}^{-1} \tilde{g}$ provides an approximate form of the original function f.

An extension of the above concepts is obtained by replacing the monomial expression $k_1^l\, k_2^m\, k_3^n$ with a function $\alpha(\mathbf{k})$, that can be expressed by a polynomial interpolation. In this case, in fact, the integral of $\alpha(\mathbf{k}) f(\mathbf{k})$ is a combination of

moments of f. A further generalization consists in considering f, α, or both, as functions of other variables besides $\mathbf{k}$:

$$M_\alpha(\mathbf{r},t) = \iiint_{-\infty}^{+\infty} \alpha(\mathbf{r},\mathbf{k},t) f(\mathbf{r},\mathbf{k},t)\, \mathrm{d}^3k\,. \tag{C.64}$$

If $f(\mathbf{r},\mathbf{k},t)$ is the solution of a differential equation generated by an operator $\mathscr{A}$, say, $\mathscr{A}f = 0$, one can derive a set of moments from such an equation by selecting different forms of α:

$$\iiint_{-\infty}^{+\infty} \alpha\, \mathscr{A}f\, \mathrm{d}^3k = 0\,. \tag{C.65}$$

Each moment depends on the other variables $\mathbf{r}, t$. If operator $\mathscr{A}$ contains the derivatives with respect to $\mathbf{r}$, t, or both, then the moment of $\mathscr{A}f = 0$ is a differential equation in $\mathbf{r}$, t, or both.

C.7 Error Function

The *error function* and the *complementary error function* are defined, respectively, as

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-\xi^2)\, \mathrm{d}\xi\,, \qquad \mathrm{erfc}(x) = 1 - \mathrm{erf}(x)\,. \tag{C.66}$$

From the definitions (C.66) and from the Gauss integral (C.23) the following properties are derived:

$$\frac{\mathrm{d}}{\mathrm{d}x}\,\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}}\,\exp(-x^2)\,, \qquad \mathrm{erf}(-x) = -\,\mathrm{erf}(x)\,, \tag{C.67}$$

$$\mathrm{erf}(-\infty) = -1\,, \qquad \mathrm{erf}(0) = 0\,, \qquad \mathrm{erf}(+\infty) = 1\,, \tag{C.68}$$

$$\mathrm{erfc}(-\infty) = 2\,, \qquad \mathrm{erfc}(0) = 1\,, \qquad \mathrm{erfc}(+\infty) = 0\,. \tag{C.69}$$

Integrating by parts yields

$$\int_0^x \mathrm{erfc}(\xi)\, \mathrm{d}\xi = x\,\mathrm{erfc}(x) + \frac{1}{\sqrt{\pi}}\left[1 - \exp(-x^2)\right]. \tag{C.70}$$

Applying the de l'Hôpital rule shows that the first term at the right-hand side of (C.70) vanishes for $x \to +\infty$. It follows

$$\int_0^{+\infty} \operatorname{erfc}(x)\,\mathrm{d}x = \frac{1}{\sqrt{\pi}}\,. \tag{C.71}$$

Still applying the de l'Hôpital rule shows that

$$\lim_{x\to 0}\frac{\operatorname{erf}(x)}{x} = \frac{2}{\sqrt{\pi}}\,, \qquad \lim_{x\to+\infty}\frac{\operatorname{erfc}(x)}{\exp(-x^2)} = \lim_{x\to+\infty}\frac{1/\sqrt{\pi}}{x}\,, \tag{C.72}$$

whence

$$\operatorname{erf}(x) \simeq \frac{2}{\sqrt{\pi}}\,x \quad \text{for} \quad |x| \ll 1\,, \qquad \operatorname{erfc}(x) \simeq \frac{1}{\sqrt{\pi}}\,\frac{\exp(-x^2)}{x} \quad \text{for} \quad x \gg 1\,. \tag{C.73}$$

Other applications of the integration by parts yield

$$Y = \int_0^x \xi\,\operatorname{erfc}(\xi)\,\mathrm{d}\xi = x^2\,\operatorname{erfc}(x) - Y - \frac{1}{\sqrt{\pi}}\int_0^x \xi\left[\frac{\mathrm{d}}{\mathrm{d}\xi}\exp(-\xi^2)\right]\mathrm{d}\xi = \tag{C.74}$$

$$= \frac{1}{2}\,x^2\,\operatorname{erfc}(x) + \frac{1}{4}\,\operatorname{erf}(x) - \frac{1}{2\sqrt{\pi}}\,x\exp(-x^2)\,. \tag{C.75}$$

C.8 Parametrized Gaussian Function

The relations introduced in Sects. C.3 and C.7 are useful for investigating the properties of function

$$\Delta(x-\xi,a) = \frac{\exp[-(x-\xi)^2/(4\,a)]}{\sqrt{4\pi\,a}}\,, \qquad a > 0\,. \tag{C.76}$$

The behavior of Δ in the limit $a \to 0$ depends on the argument $x-\xi$, namely

$$\lim_{a\to 0}\Delta(x-\xi,a) = \begin{cases} 0 & \xi \neq x \\ +\infty & \xi = x \end{cases} \tag{C.77}$$

In contrast, its integral over ξ is independent of x and a. In fact, using (C.23) after letting $\mu = (x-\xi)/\sqrt{4a}$ yields

$$\int_{-\infty}^{+\infty}\Delta(x-\xi,a)\,\mathrm{d}\xi = \frac{1}{\sqrt{\pi}}\int_{-\infty}^{+\infty}\exp(-\mu^2)\,\mathrm{d}\mu = 1\,. \tag{C.78}$$

Adopting the same variable change leading to (C.78) and using (C.23, C.66) yields

$$\int_{-\infty}^{0} \Delta(x-\xi,a)\,\mathrm{d}\xi = \frac{1}{2}\,\mathrm{erfc}\left(\frac{x}{\sqrt{4a}}\right). \tag{C.79}$$

The relations (C.78, C.79) hold also in the limit for $a \to 0$, provided the limit is calculated *after* the integration. This property is typical of the functions that generate the Dirac δ (Sect. C.4). In fact it can be shown that for a continuous function $g(x)$ the following holds:

$$\lim_{a\to 0} \int_{-\infty}^{+\infty} g(\xi)\,\Delta(x-\xi,a)\,\mathrm{d}\xi = g(x)\,. \tag{C.80}$$

Other examples of δ-generating functions are given in Sect. C.4. This section is concluded by showing that $\Delta(x-\xi,a)$ admits an integral representation of the form

$$\Delta(x-\xi,a) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp[\mathrm{i}k(x-\xi) - ak^2]\,\mathrm{d}k\,. \tag{C.81}$$

To prove (C.81) one recasts the argument of the exponential by means of the identity

$$\mathrm{i}k(x-\xi) - ak^2 = -\frac{(x-\xi)^2}{4a} - a\left[k - \frac{\mathrm{i}(x-\xi)}{2a}\right]^2, \tag{C.82}$$

and uses (C.23) with $\sqrt{a}\,[k - \mathrm{i}(x-\xi)/(2a)]$ as the integration variable. It is interesting to note in passing that letting $\xi = 0$, $a = \sigma^2/2$ in (C.81) yields

$$\exp[-x^2/(2\,\sigma^2)] = \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp(-\sigma^2\,k^2/2)\,\exp(\mathrm{i}\,k\,x)\,\mathrm{d}k\,, \tag{C.83}$$

namely, the Gaussian function is the Fourier transform of itself.

C.9 Euler's Beta Function

The function defined by the integral

$$B(\lambda\,,\,\mu) = \int_0^1 x^{\lambda-1}\,(1-x)^{\mu-1}\,\mathrm{d}x\,, \tag{C.84}$$

with λ, μ complex numbers such that $\Re(\lambda) > 0$, $\Re(\mu) > 0$, is called *Euler's Beta function* or *Euler's integral of the first kind* [82]. Letting $x = y/(y+1)$ and replacing y with x gives (C.84) the equivalent form

$$B(\lambda\,,\,\mu) = \int_0^{+\infty} x^{\lambda-1}\,(1+x)^{-(\lambda+\mu)}\,\mathrm{d}x\,. \tag{C.85}$$

Limiting the variables' range to $0 < \Re(\lambda)\,, \Re(\mu) < 1$ and letting

$$\mu = 1-\lambda\,, \qquad T_0(\lambda) = B(\lambda\,, 1-\lambda) \tag{C.86}$$

yields

$$T_0(\lambda) = \int_0^{+\infty} \frac{x^{\lambda-1}}{1+x}\,\mathrm{d}x = \frac{\pi}{\sin(\lambda\pi)}\,. \tag{C.87}$$

The last equality is demonstrated by applying *Cauchy's residue theorem* [142, Sect. 64] to the function $f(z) = z^{\lambda-1}/(1+z)$, with z complex, that over the real axis reduces to the integrand of (C.87). To proceed, one must first observe that $f(z)$ is a multivalued function; this is readily seen by letting $z = r\,\exp(\mathrm{i}\,\alpha)$ and recasting f in the form $f(z) = (1+z)^{-1}\,\exp[(\lambda-1)\,(\log r + \mathrm{i}\,\alpha)]$. One then considers only one branch of f, letting for instance $0 \le \alpha < 2\,\pi$. Also, it is found that f is continuous in the complex plane with the exception of the positive part of the real axis; in fact, in the vicinity of a point x of it, it is $\alpha \to 0$ for $\Im(z) > 0$ and $\alpha \to 2\,\pi$ for $\Im(z) < 0$.

The corresponding limits of f are

$$\lim_{z\to x} f(z) = \frac{x^{\lambda-1}}{1+x}\,, \qquad \lim_{z\to x} f(z) = \frac{x^{\lambda-1}}{1+x}\,\exp(\mathrm{i}\,2\,\lambda\,\pi)\,. \tag{C.88}$$

The residue theorem states that $\int_C f(z)\,\mathrm{d}z = \mathrm{i}\,2\,\pi\,R$, with C the integration path and R the residue. To apply it to the present case one selects an integration path made of two circumferences A and B, centered in the origin and having radii $a < 1$ and $b > 1$, respectively, and of segment $Q = [a\,, b]$ (Fig. C.5): circumference A is traced clockwise starting from $z = a$; next, segment Q is traced from a to b along the upper edge, where $f(z)$ is given by the first expression in (C.88). Then, circumference B is traced counterclockwise, followed by segment Q from b to a along the lower edge, where $f(z)$ is given by the second expression in (C.88). In the interior of the integration path f has the simple pole $z_0 = -1 = \exp(\mathrm{i}\,\pi)$, whose residue is

$$R = \lim_{z\to z_0} (z-z_0) f(z) = \lim_{\substack{r\to 1\\ \alpha\to\pi}} \exp[(\lambda-1)\,(\log r + \mathrm{i}\,\alpha)] = -\exp(\mathrm{i}\,\lambda\,\pi)\,, \tag{C.89}$$

whence

$$\int_A f(z)\,\mathrm{d}z + \int_B f(z)\,\mathrm{d}z + [1-\exp(\mathrm{i}\,2\,\lambda\,\pi)] \int_a^b \frac{x^{\lambda-1}}{1+x}\,\mathrm{d}x = -\mathrm{i}\,2\,\pi\,\exp(\mathrm{i}\,\lambda\,\pi)\,. \tag{C.90}$$

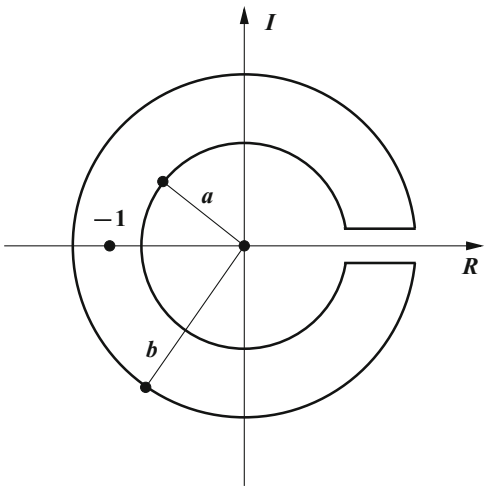


Fig. C.5 Integration path of (C.87)

To evaluate (C.87) one takes the limits $a \to 0$ and $b \to \infty$ in (C.90); as shown below, the limit makes the first two integrals at the left-hand side of (C.90) to vanish, this leading directly to (C.87).

Relation (C.87) can be exploited for calculating other integrals. For instance, for real values of λ one lets $\lambda = 1/(2\mu)$, $x = y^{2\mu}$ to find

$$\int_0^{+\infty} \frac{1}{1+y^{2\mu}}\,\mathrm{d}y = \frac{\pi/(2\mu)}{\sin[\pi/(2\mu)]}\,, \qquad \mu > \frac{1}{2}\,. \tag{C.91}$$

To show that the first two integrals at the left-hand side of (C.90) vanish in the limit, one lets $\lambda = p + \mathrm{i}\,q$ and observes that over circumference A it is $|1+z| \geq 1-a$ and $|z^{\lambda-1}| = a^{p-1}\,\exp(-\alpha\,q)$; the first of the above is obvious, the second one is found by letting $\sigma + \mathrm{i}\,\omega = (\lambda-1)\,(\log r + \mathrm{i}\,\alpha)$ and observing that $|z^{\lambda-1}| = \exp(\sigma)$. Indicating with M the maximum of $\exp(-\alpha\,q)$ with α varying between 0 and $2\,\pi$, the following inequalities hold over A:

$$|f(z)| \leq M\,\frac{a^{p-1}}{1-a}\,, \qquad \left|\int_A f(z)\,\mathrm{d}z\right| \leq 2\,\pi\,M\,\frac{a^p}{1-a}\,. \tag{C.92}$$

By the same token one finds over B:

$$|f(z)| \leq M\,\frac{b^{p-1}}{b-1}\,, \qquad \left|\int_B f(z)\,\mathrm{d}z\right| \leq 2\,\pi\,M\,\frac{b^p}{b-1}\,. \tag{C.93}$$

Remembering that $0 < p = \Re(\lambda) < 1$, one finally finds

$$\lim_{a\to 0}\frac{a^p}{1-a} = \lim_{b\to\infty}\frac{b^p}{b-1} = 0\,, \tag{C.94}$$

C.10 Euler's Gamma Function

The function defined by the integral

$$\Gamma(\lambda) = \int_0^{+\infty} x^{\lambda-1} \exp(-x)\, \mathrm{d}x\,, \tag{C.95}$$

with λ a complex number such that $\Re(\lambda) > 0$, is called *Euler's Gamma function* or *Euler's integral of the second kind* [44, Sect. 1.3].[1] The negative of its derivative $\Gamma' = \mathrm{d}\Gamma/\mathrm{d}\lambda$ calculated for $\lambda = 1$ is called *Euler's constant*, $\gamma = -\Gamma'(1) = \int_0^{+\infty} \exp(-x)\, \log(x)\, \mathrm{d}x \simeq 0.5772$. From (C.95) one finds $\Gamma(1) = 1$ and, after integrating by parts,

$$\Gamma(\lambda + 1) = \lambda\, \Gamma(\lambda)\,. \tag{C.96}$$

If $\lambda = n = 1, 2, \ldots$ (C.96) yields

$$\Gamma(n + 1) = n\, \Gamma(n) = n(n-1)\, \Gamma(n-1) = \ldots = n!\,. \tag{C.97}$$

The definition of Γ is extended by analytic continuation to the complex plane with the exception of the points $\lambda = 0, -1, -2, \ldots, -n, \ldots$. At each negative integer $-n$, the function Γ has a simple pole with a residue equal to $(-1)^n/n!$ [82], namely,

$$\lim_{\lambda \to -n} (\lambda + n)\, \Gamma(\lambda) = \frac{(-1)^n}{n!}\,, \qquad n = 0, 1, 2 \ldots \tag{C.98}$$

A straightforward calculation shows that the Beta and Gamma functions are connected by the relation [82]

$$\Gamma(\lambda)\, \Gamma(\mu) = \Gamma(\lambda + \mu)\, B(\lambda\,,\, \mu)\,. \tag{C.99}$$

Thanks to (C.99) one extends the definition of B to the complex plane with the exception of the points $\lambda\,,\, \mu\,,\, \lambda + \mu = 0\,, -1\,, -2\,, \ldots\,, -n\,, \ldots$. Moreover, limiting the variables' range to $0 < \Re(\lambda)\,, \Re(\mu) < 1$ and letting $\mu = 1 - \lambda$ so that $\Gamma(\lambda + \mu) = \Gamma(1) = 1$, from (C.87) one finds

$$\Gamma(\lambda)\, \Gamma(1 - \lambda) = \int_0^{+\infty} \frac{x^{\lambda-1}}{1 + x}\, \mathrm{d}x = T_0(\lambda)\,, \qquad 0 < \Re(\lambda) < 1\,. \tag{C.100}$$

For $\lambda = 1/2$ (C.100) yields

[1] As remarked in [44], Legendre's notation $\Gamma(\lambda)$ is unfortunate because the argument that appears at the right-hand side of the definition is $\lambda - 1$. Gauss used the notation $\Pi(\lambda - 1)$ for the left-hand side of (C.95).

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi} \tag{C.101}$$

whence, thanks to (C.96),

$$\Gamma\left(\frac{3}{2}\right) = \frac{1}{2}\,\Gamma\left(\frac{1}{2}\right) = \frac{1}{2}\,\sqrt{\pi}\,, \quad \Gamma\left(\frac{5}{2}\right) = \frac{3}{2}\,\Gamma\left(\frac{3}{2}\right) = \frac{3}{4}\,\sqrt{\pi}\,, \quad \ldots \tag{C.102}$$

Iterating (C.102) and comparing with (C.33) show that $\Gamma(m + 1/2) = 2\,I_{2m}(1)$, $m = 0\,,1\,,2\,,\ldots$

C.11 Gamma Function's Asymptotic Behavior

Euler's Gamma function introduced in Sect. C.10, considered for real values of λ, lends itself to a significant application of the asymptotic analysis. Specifically, one seeks another function $f(\lambda)$, expressible through elementary functions, such that $\lim_{\lambda\to\infty}[\Gamma(\lambda + 1)/f(\lambda)] = 1$. The asymptotic analysis applied to the Γ function shows that [36]

$$\lim_{\lambda\to\infty} \frac{\Gamma(\lambda + 1)}{\lambda^{\lambda+1/2}\exp(-\lambda)} = \sqrt{2\pi}\,, \tag{C.103}$$

namely, the function sought is $f(\lambda) = \sqrt{2\pi}\,\lambda^{\lambda+1/2}\exp(-\lambda)$. Equation (C.103) is called *Stirling's formula*. Remembering (C.97) one has $\Gamma(\lambda + 1) = \Gamma(n + 1) = n!$ when λ is a natural number. From (C.103) it follows

$$n! \simeq \sqrt{2\pi}\,n^{n+1/2}\,\exp(-n) = \sqrt{2\pi\,n}\,(n/\mathrm{e})^n\,, \tag{C.104}$$

that provides an approximation to the factorial for $n \gg 1$. Letting by way of example $n = 10$, the rounded value of the right-hand side of (C.104) turns out to be 3 598 696, that differs from $10! = 3\,628\,800$ by less than 1%.

The asymptotic value of the derivative of $\log\Gamma$ is also of interest, for instance when determining the equilibrium distribution of particles in statistical mechanics (Sects. 6.4, 15.8.1, 15.8.2). Using (C.103) one finds

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\log\Gamma(\lambda + 1) \simeq \frac{1}{2\,\lambda} + \log\lambda \simeq \log\lambda\,, \qquad \lambda \gg 1\,. \tag{C.105}$$

C.12 Integrals Related to the Harmonic Oscillator

Consider the integral

$$I(s) = \int_0^1 \frac{\mathrm{d}\xi}{\sqrt{1-\xi^s}}\,, \tag{C.106}$$

where s is a real parameter, $s > 0$. Letting $u = \xi^s$ one finds $1/\sqrt{1-\xi^s} = (1-u)^{1/2-1}$, $\mathrm{d}\xi = u^{1/s-1}\,\mathrm{d}u/s$ whence, using (C.84, C.99, C.101),

$$I(s) = \frac{1}{s}\,B(1/s, 1/2) = \frac{\sqrt{\pi}}{s}\,\frac{\Gamma(1/s)}{\Gamma(1/s+1/2)}\,. \tag{C.107}$$

By way of example $I(2) = \pi/2$, which can also be derived directly from (C.106). When $s \to \infty$ one can use (C.98) with $n = 0$. It follows

$$\lim_{s\to\infty} I(s) = 1\,. \tag{C.108}$$

Now consider the integral

$$J(s) = \int_0^1 \frac{\mathrm{d}\xi}{\sqrt{1/\xi^s - 1}}\,, \tag{C.109}$$

still with $s > 0$. The same procedure used for calculating $I(s)$ yields

$$J(s) = \frac{1}{s}\,B(1/s+1/2, 1/2) = \frac{\sqrt{\pi}}{s}\,\frac{\Gamma(1/s+1/2)}{\Gamma(1/s+1)} = \frac{\pi}{s}\,\frac{1}{I(s)}\,, \tag{C.110}$$

and $\lim_{s\to\infty} J(s) = 0$. By way of example $J(1) = \pi/2$, which can also be derived directly from (C.109). The integrals (C.107), (C.110) appear in the theory of the harmonic oscillator (Sect. 3.3 and problems 3.1, 3.2).

C.13 Fermi Integrals

The *Fermi integral* of order α is defined as

$$\Phi_\alpha(\xi) = \frac{1}{\Gamma(\alpha+1)} \int_0^\infty \frac{x^\alpha}{1+\exp(x-\xi)}\,\mathrm{d}x\,, \qquad \alpha > -1\,, \tag{C.111}$$

where Γ is defined by (C.95) and α is a real parameter. The constraint $\alpha > -1$ guarantees the convergence of the integral. If $-\xi \gg 1$ one has $\exp(x-\xi) \geq \exp(-\xi) \gg 1$ and, from (C.95),

$$\Phi_\alpha(\xi) \simeq \frac{\exp(\xi)}{\Gamma(\alpha+1)} \int_0^\infty x^\alpha \exp(-x)\,\mathrm{d}x = \exp(\xi)\,, \qquad \xi \gg -1\,. \tag{C.112}$$

A relation between Fermi integral of different order is found by considering, for some $\alpha > 0$, the integral of order $\alpha - 1$:

$$\frac{1}{\Gamma(\alpha)} \int_0^\infty \frac{x^{\alpha-1}}{1+\exp(x-\xi)}\,\mathrm{d}x = \frac{1}{\alpha\,\Gamma(\alpha)} \int_0^\infty \frac{x^\alpha \exp(x-\xi)}{[1+\exp(x-\xi)]^2}\,\mathrm{d}x\,, \tag{C.113}$$

where the right-hand side is derived through an integration by parts. Observing that $\alpha\,\Gamma(\alpha) = \Gamma(\alpha+1)$ and using again (C.111) shows that the right-hand side of (C.113) is equal to $\mathrm{d}\Phi_\alpha/\mathrm{d}\xi$. Then,

$$\frac{\mathrm{d}\Phi_\alpha}{\mathrm{d}\xi} = \Phi_{\alpha-1}\,, \qquad \frac{\mathrm{d}\log\Phi_\alpha}{\mathrm{d}\xi} = \frac{\Phi_{\alpha-1}}{\Phi_\alpha}\,. \tag{C.114}$$

The Fermi integrals are positive by construction; from the first relation in (C.114) it then follows that the Fermi integrals are monotonically increasing functions of the argument ξ. The Fermi integral of order 0 is expressed in terms of elementary functions,

$$\Phi_0 = \log\left[\exp(\xi)+1\right]\,. \tag{C.115}$$

In the applications to the semiconductor theory the Fermi integrals of small half-integer order (1/2, 3/2) are the most important ones (Sects. 18.2, 19.6.4). Remembering (C.101, C.102), they read

$$\Phi_{1/2}(\xi) = \int_0^\infty \frac{2\,x^{1/2}/\sqrt{\pi}}{1+\exp(x-\xi)}\,\mathrm{d}x\,, \quad \Phi_{3/2}(\xi) = \int_0^\infty \frac{(4/3)\,x^{3/2}/\sqrt{\pi}}{1+\exp(x-\xi)}\,\mathrm{d}x. \tag{C.116}$$

Approximations for the Fermi integrals are found, e.g., in [12, App. C]; there it is shown, for instance, that for the integral of order 1/2 the curve

$$\Phi_{1/2}(\xi) \simeq \frac{\exp(\xi)}{1+0.27\,\exp(\xi)}\,, \qquad \xi < 1.3 \tag{C.117}$$

gives an error within $\pm 3\%$ in the indicated range. Still considering the integral of order 1/2, [12, App. C] shows that for larger values of the argument other expressions are applicable: e.g., the curve

$$\Phi_{1/2}(\xi) \simeq \frac{4}{3\sqrt{\pi}} \left(\xi^2 + \frac{\pi^2}{6}\right)^{3/4}\,, \qquad \xi \geq 0.7 \tag{C.118}$$

gives an error within $\pm 3\%$ in the indicated range. The behavior of (C.112), (C.117), and (C.118) is shown in Fig. C.6.

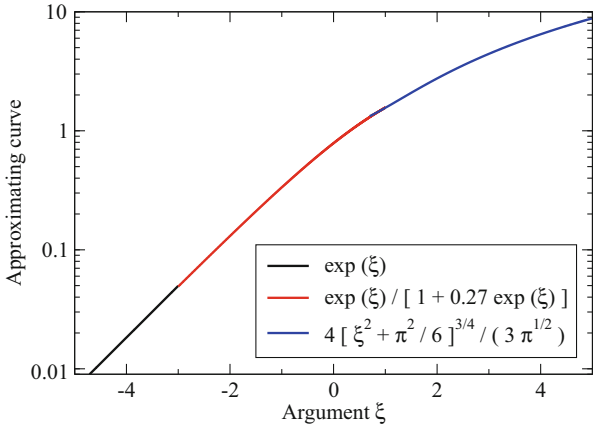


Fig. C.6 Approximations to the Fermi integrals: the *black line* represents (C.112), applicable to the Fermi integrals of any order for $\xi \ll -1$; the *red line* represents (C.117) that renders $\Phi_{1/2}$ within $\pm 3\%$ when $\xi < 1.3$; the *blue line* represents (C.118) that renders $\Phi_{1/2}$ within $\pm 3\%$ when $\xi \geq 0.7$

The Fermi integrals embed the Fermi-Dirac statistics (18.54), whose form is $P(\lambda) = 1/[\exp(\lambda) + 1]$. It is easily found that the latter fulfills the relation $P(\lambda) + P(-\lambda) = 1$, whence $-P(-\lambda) + 1/2 = P(\lambda) - 1/2$, namely, $K(\lambda) = P(\lambda) - 1/2$ is an odd function of λ. It follows that the set of coefficients in the Taylor expansion of $P(\lambda) = K(\lambda) + 1/2$ is made of those of $K(\lambda)$, whose index is odd, and of the zero-order coefficient $1/2$.

C.14 Hölder's Inequality

Hölder's inequality states that for any pair of real constants $b, c > 1$ such that $1/b + 1/c = 1$ it is

$$\int_\eta |F\,G|\,\mathrm{d}x \leq \left(\int_\eta |F|^b\,\mathrm{d}x\right)^{1/b} \left(\int_\eta |G|^c\,\mathrm{d}x\right)^{1/c}, \tag{C.119}$$

where F, G are any complex functions defined over the real interval η and such that the integrals in (C.119) converge. The inequality is proven starting from the function $\varphi(r) = r^b - b\,r + b - 1$, $r > 0$, $b > 1$, whose first derivative is $\varphi'(r) = b\,r^{b-1} - b$ and the second one $\varphi'' = b\,(b-1)\,r^{b-2}$. As a consequence, for $r > 0$ the function has only one minimum, located at $r = 1$. The inequality $r^b + b \geq b\,r + 1$ then holds, whence

$$\frac{r^{b-1}}{b} + \frac{1}{c\,r} \geq 1\,, \quad c = \frac{b}{b-1} > 1\,. \tag{C.120}$$

Let $F_1(x)$ and $G_1(x)$ be any two complex functions defined over η and fulfilling the normalization condition

$$\int_\eta |F_1|^b \,\mathrm{d}x = \int_\eta |G_1|^c \,\mathrm{d}x = 1\,. \tag{C.121}$$

Letting $r^{b-1} = |F_1|^{b-1}/|G_1|$ and replacing in (C.120) yields

$$\frac{|F_1|^b}{b} + \frac{|G_1|^c}{c} - |F_1\,G_1| \geq 0\,, \qquad \frac{1}{b} + \frac{1}{c} = 1\,. \tag{C.122}$$

Since the function at the left-hand side of (C.122) is nonnegative, its integral is nonnegative as well. Integrating (C.122) over η and using the normalization condition (C.121) yields

$$\int_\eta |F_1\,G_1| \,\mathrm{d}x \leq \frac{1}{b} + \frac{1}{c} = 1\,. \tag{C.123}$$

On the other hand, the normalization condition also yields

$$\left(\int_\eta |F_1|^b \,\mathrm{d}x\right)^{1/b} = \left(\int_\eta |G_1|^c \,\mathrm{d}x\right)^{1/c} = 1\,, \tag{C.124}$$

whence

$$\int_\eta |F_1 G_1| \,\mathrm{d}x \leq \left(\int_\eta |F_1|^b \,\mathrm{d}x\right)^{1/b} \left(\int_\eta |G_1|^c \,\mathrm{d}x\right)^{1/c}\,. \tag{C.125}$$

As (C.125) is homogeneous, it still holds after replacing F_1, G_1 with $F = \lambda\,F_1$ and $G = \mu\,G_1$, where λ, μ are arbitrary positive real numbers. This proves Hölder's inequality (C.119).

C.15 Integrals Related to the Electromagnetic Modes

In several applications (e.g., calculations related to the modes of the electromagnetic field, Sect. 5.5) one must evaluate integrals of the form

$$Y = \int_V \exp[\mathrm{i}\,(\mathbf{k} \pm \mathbf{k}') \cdot \mathbf{r}] \,\mathrm{d}^3 r\,, \tag{C.126}$$

where $\mathbf{k} = \mathbf{k}(n_1, n_2, n_3)$ is given by

$$\mathbf{k} = n_1 \frac{2\pi}{d_1} \mathbf{i}_1 + n_2 \frac{2\pi}{d_2} \mathbf{i}_2 + n_3 \frac{2\pi}{d_3} \mathbf{i}_3 , \qquad n_i = 0, \pm 1, \pm 2, \ldots , \tag{C.127}$$

$\mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3$ being the unit vectors parallel to the coordinate axes. The integration domain in (C.126) is a box whose sides d_1, d_2, d_3 are aligned with the axes and start from the origin (Fig. 5.1). The volume of the box is $V = d_1\, d_2\, d_3$. As $(\mathbf{k} \pm \mathbf{k}') \cdot \mathbf{r} = (k_1 \pm k_1')\, x_1 + (k_2 \pm k_2')\, x_2 + (k_3 \pm k_3')\, x_3$, where the upper (lower) signs hold together, the integral becomes $Y = Y_1\, Y_2\, Y_3$, with

$$Y_i = \int_0^{d_i} \exp[\mathrm{i}\,(k_i \pm k_i')\, x_i]\, \mathrm{d}x_i = \frac{\exp[\mathrm{i}\,(k_i \pm k_i')\, d_i] - 1}{\mathrm{i}\,(k_i \pm k_i')} . \tag{C.128}$$

Letting $\theta_i = (k_i \pm k_i')\, d_i/2 = \pi\, (n_i \pm n_i')$, (C.128) becomes

$$Y_i = d_i\, \exp(\mathrm{i}\,\theta_i)\, \frac{\exp(\mathrm{i}\,\theta_i) - \exp(-\mathrm{i}\,\theta_i)}{2\,\mathrm{i}\,\theta_i} = d_i\, \exp(\mathrm{i}\,\theta_i)\, \frac{\sin\theta_i}{\theta_i} . \tag{C.129}$$

It follows that $Y_i = 0$ if $n_i \pm n_i' \neq 0$, while $Y_i = d_i$ if $n_i \pm n_i' = 0$. Combining the three integrals shows that it is $Y = 0$ if $\mathbf{k} \pm \mathbf{k}' \neq 0$, while it is $Y = V$ if $\mathbf{k} \pm \mathbf{k}' = 0$. The result is recast in a compact form by means of the three-dimensional extension of the Kronecker symbol (A.18):

$$Y = V\, \delta[\mathbf{k} \pm \mathbf{k}', 0] = V\, \delta[\mathbf{k} \pm \mathbf{k}'] , \tag{C.130}$$

where the last form is obtained by dropping the zero for the sake of conciseness. Compare (C.126, C.130) with (C.56) and the comments therein.

C.16 Riemann's Zeta Function

The function defined by

$$\zeta(\lambda\,, a) = \sum_{k=1}^{\infty} \frac{1}{(k+a)^{\lambda}} , \tag{C.131}$$

where λ is a complex number with $\Re(\lambda) > 1$ and $a \geq 0$ is real, is called *Riemann's Zeta function*. It can be represented in integral form by combining it with the Gamma function (C.95): letting $x = (k+a)\, y$ in the latter, then replacing y back with x, yields

$$\Gamma(\lambda) = (k+a)^{\lambda} \int_0^{+\infty} x^{\lambda-1}\, \exp[-(k+a)\, x]\, \mathrm{d}x . \tag{C.132}$$

Dividing (C.132) by $(k+a)^{\lambda}$, letting $k = 1, 2, \ldots$, and adding over k provides

$$\Gamma(\lambda) \sum_{k=1}^{\infty} \frac{1}{(k+a)^{\lambda}} = \int_0^{+\infty} x^{\lambda-1} \exp(-ax) \left[\sum_{k=1}^{\infty} \exp(-kx) \right] \mathrm{d}x \,, \tag{C.133}$$

where $\sum_{k=1}^{\infty} \exp(-k\,x) = \exp(-x)\,[1+\exp(-x)+\exp(-2\,x)+\ldots] = 1/[\exp(x)-1]$ so that, from (C.131),

$$\zeta(\lambda\,, a) = \frac{1}{\Gamma(\lambda)} \int_0^{+\infty} \frac{x^{\lambda-1}}{\exp(x)-1} \exp(-a\,x)\,\mathrm{d}x \,, \qquad \Re(\lambda) > 1 \,. \tag{C.134}$$

Remembering (C.96) one finds that (C.134) fulfills the recursive relation

$$\frac{\partial}{\partial a} \zeta(\lambda\,, a) = -\lambda\,\zeta(\lambda+1\,, a) \,. \tag{C.135}$$

Also, letting $a = 0$ and $\lambda = 2\,m$, with $m = 1, 2, \ldots$ transforms (C.134) into

$$\int_0^{+\infty} \frac{x^{2\,m-1}}{\exp(x)-1}\,\mathrm{d}x = \Gamma(2\,m)\,\zeta(2\,m\,, 0) = \frac{(2\,\pi)^{2\,m}}{4\,m}\,|B_{2m}| \,, \tag{C.136}$$

with $B_{2m} = (-1)^{m+1}\,|B_{2m}|$, $m \geq 1$ the *Bernoulli number* of order $2\,m$ ([59] and Sect. C.17). Thanks to (C.136) one calculates integrals used in different applications. For instance, letting $m = 2$ and using $B_4 = -1/30$, one finds

$$\int_0^{+\infty} \frac{x^3}{\exp(x)-1}\,\mathrm{d}x = \frac{1}{15}\pi^4 \,, \tag{C.137}$$

that is used in (15.78) to calculate the Lagrangian multiplier in the equilibrium statistics for photons. From (C.134) one derives another important class of integrals; in fact, replacing x with $2\,x$ in the denominator of (C.134) yields

$$\int_0^{+\infty} \frac{x^{\lambda-1}}{\exp(2\,x)-1} \exp(-a\,x)\,\mathrm{d}x = 2^{-\lambda}\,\Gamma(\lambda)\,\zeta(\lambda\,, a/2) \,, \qquad \Re(\lambda) > 1 \tag{C.138}$$

whence, using the identity $2/[\exp(2\,x)-1] = 1/[\exp(x)-1] - 1/[\exp(x)+1]$ within (C.134), (C.138) provides

$$\int_0^{+\infty} \frac{x^{\lambda-1}}{\exp(x)+1} \exp(-a\,x)\,\mathrm{d}x = \Gamma(\lambda)\,\left[\zeta(\lambda\,, a) - 2^{1-\lambda}\,\zeta(\lambda\,, a/2)\right] \,. \tag{C.139}$$

Letting $a = 0$ and $\lambda = 2m$, $m = 1, 2, \ldots$ in the latter, and using (C.136), transforms (C.139) into

$$\int_0^{+\infty} \frac{x^{2m-1}}{\exp(x)+1}\,\mathrm{d}x = \frac{\pi^{2m}}{2m}\left(2^{2m-1}-1\right)|B_{2m}|\,. \tag{C.140}$$

For instance, for $m = 1$ and $m = 2$, (C.140) provides

$$\int_0^{+\infty} \frac{x}{\exp(x)+1}\,\mathrm{d}x = \frac{1}{12}\,\pi^2\,, \qquad \int_0^{+\infty} \frac{x^3}{\exp(x)+1}\,\mathrm{d}x = \frac{7}{120}\,\pi^4\,. \tag{C.141}$$

C.17 Bernoulli Function

The definition of the *Bernoulli function* and its Taylor expansion around the origin are

$$B(x) = \frac{x}{\exp(x)-1}\,, \qquad B(x) = \sum_{k=0}^{\infty} B_k\,\frac{x^k}{k!}\,, \qquad B_k = \left.\frac{\mathrm{d}^k B}{\mathrm{d}x^k}\right|_{x=0}\,. \tag{C.142}$$

The form of the function is shown in Fig. C.7; it is not singular in the origin, in fact it is $B(0) = B_0 = 1$, however, when considered in the complex plane z, it has poles for $z = 2\,\pi\,n\,\mathrm{i}$, $n \neq 0$. It follows that the expansion (C.142) converges for $|x| < 2\,\pi$. The coefficients B_k are called *Bernoulli numbers*. It is easily found that the only nonvanishing Bernoulli number of odd index is $B_1 = -1/2$; in fact, the Bernoulli function fulfills the relation $B(-x) - B(x) = x$, whence $B(-x) - x/2 = B(x) + x/2$,

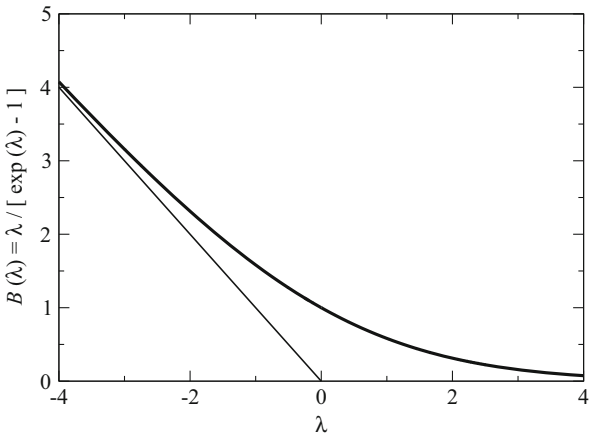


Fig. C.7 Bernoulli function (*thick line*) and its asymptote for $x \to -\infty$ (*thin line*)

namely, $H(x) = B(x) + x/2$ is an even function of x. It follows that the set of coefficients in the expansion of $B(x) = H(x) - x/2$ is made of those of $H(x)$, which are of even index, and B_1. To calculate the coefficients of even index it is convenient to use a recursive relation, starting from the definition of $B(x)$ written in the form $B(x)\,[\exp(x) - 1]/x$, and replacing each factor with the corresponding expansion:

$$\left(B_0 + B_1\,\frac{x}{1!} + B_2\,\frac{x^2}{2!} + B_3\,\frac{x^3}{3!} + \ldots\right)\left(1 + \frac{x}{2!} + \frac{x^2}{3!} + \frac{x^3}{4!} + \ldots\right) = 1\,. \tag{C.143}$$

As the two series in (C.143) are absolutely convergent, the product series converges absolutely to the product of the two for any ordering of the factors. Grouping the latter according to the powers of equal degree yields $B_0 = 1$ and

$$\frac{1}{2!\,0!}\,B_0 + \frac{1}{1!\,1!}\,B_1 = 0\,, \quad \frac{1}{3!\,0!}\,B_0 + \frac{1}{2!\,1!}\,B_1 + \frac{1}{1!\,2!}\,B_2 = 0\,, \quad \ldots \tag{C.144}$$

For any index $n \geq 1$ (C.144) becomes

$$\frac{B_0}{n!\,0!} + \frac{B_1}{(n-1)!\,1!} + \cdots + \frac{B_k}{(n-k)!\,k!} + \cdots + \frac{B_{n-1}}{1!\,(n-1)!} = 0\,. \tag{C.145}$$

Multiplying (C.145) by $n!$ yields for any $n \geq 1$

$$\sum_{k=0}^{n-1} \binom{n}{k}\,B_k = 0\,, \qquad \binom{n}{k} = \frac{n!}{(n-k)!\,k!}\,. \tag{C.146}$$

By way of example, (C.146) yields for $n = 3\,, 5\,, 7$

$$B_0 + 3\,B_1 + 3\,B_2 = 0$$

$$B_0 + 5\,B_1 + 10\,B_2 + 10\,B_3 + 5\,B_4 = 0$$

$$B_0 + 7\,B_1 + 21\,B_2 + 35\,B_3 + 35\,B_4 + 21\,B_5 + 7\,B_6 = 0$$

Using $B_0 = 1$, $B_1 = -1/2$, the first relation above yields $B_2 = 1/6$. The second and third relation, remembering that $B_3 = B_5 = 0$, provide $B_4 = -1/30$ and $B_6 = 1/42$, and so on. From this construction it is also found that the Bernoulli numbers are all rational, and that the signs of the coefficients of even index starting from B_2 alternate: $B_{2m} = (-1)^{m+1}\,|B_{2m}|$, $m \geq 1$.

The expressions (C.142) are readily generalized by considering the Taylor expansion of $B(x)\,\exp(x\,y)$, which reads

$$B(x)\,\exp(x\,y) = \sum_{k=0}^{\infty} \tilde{B}_k(y)\,\frac{x^k}{k!}\,, \qquad |x| < 2\pi\,, \tag{C.147}$$

where $\tilde{B}_k(y)$ is the kth derivative with respect to x of the left-hand side calculated for $x = 0$. Since each derivative of $\exp(x\,y)$ entails a multiplication by y, the coefficient $\tilde{B}_k(y)$ turns out to be a polynomial of degree k, called *Bernoulli polynomial*.[2] The relation between the Bernoulli polynomials and the Bernoulli numbers if found by comparing (C.147) with (C.142), thus yielding $B_k = \tilde{B}_k(0)$. If one integrates both sides of (C.147) with respect to y from any real number a to $a + 1$, the left-hand side yields $\exp(a\,x)$. Expanding the latter in a Taylor series and equating with the integral of the right-hand side yield[3]

$$\exp(a\,x) = \sum_{k=0}^{\infty} \frac{a^k\,x^k}{k!} = \sum_{k=0}^{\infty} \frac{x^k}{k!} \int_a^{a+1} \tilde{B}_k(y)\,\mathrm{d}y\,, \qquad \int_a^{a+1} \tilde{B}_k(y)\,\mathrm{d}y = a^k\,. \tag{C.148}$$

In turn, differentiating the second relation in (C.148) provides $\tilde{B}_k(a+1) - \tilde{B}_k(a) = k\,a^{k-1}$. The left-hand side of the latter is equal to the integral $\int_a^{a+1} \tilde{B}'_k(y)\,\mathrm{d}y$; on the other hand, from the second relation in (C.148) one also derives $a^{k-1} = \int_a^{a+1} \tilde{B}_{k-1}(y)\,\mathrm{d}y$. Combining the above one finally finds the recursive relation $\tilde{B}'_k(y) = k\,\tilde{B}_{k-1}(y)$. It is also of interest to consider the Fourier expansion of $\tilde{B}'_k(y)$, that reads [1, Sect. 23.1.16], [40, Sect. 65]

$$B_k(y) = -2\,\frac{k!}{(2\,\pi)^k} \sum_{r=1}^{\infty} r^{-k}\,\cos(2\,\pi\,r\,y - \pi\,k/2)\,. \tag{C.149}$$

Expansion (C.149) holds for $0 < y < 1$ if $k = 1$, and for $0 \le y \le 1$ if $k > 1$. Letting $y = 0$ and choosing for k any odd value larger than unity yields, as expected, the identity $0 = 0$; for any even value $k = 2\,m$, instead, (C.149) provides

$$\sum_{r=1}^{\infty} \frac{1}{r^{2\,m}} = \frac{(-1)^{m+1}\,(2\,\pi)^{2\,m}}{2\,(2\,m)!}\,B_{2\,m} = \frac{(2\,\pi)^{2\,m}}{2\,(2\,m)!}\,|B_{2\,m}|\,, \qquad m \ge 1\,, \tag{C.150}$$

that is, the sum of the even powers of the inverse of the natural numbers (compare with (C.136)); letting, e.g., $m = 1\,, 2\,, 3$ yields

$$\sum_{r=1}^{\infty} \frac{1}{r^2} = \frac{\pi^2}{6}\,, \qquad \sum_{r=1}^{\infty} \frac{1}{r^4} = \frac{\pi^4}{90}\,, \qquad \sum_{r=1}^{\infty} \frac{1}{r^6} = \frac{\pi^6}{945}\,. \tag{C.151}$$

[2] Another expression of these polynomials is $B_k(y) = \sum_{i=0}^{k} \binom{k}{i} B_i\,y^{k-i}$, $n \ge 0$ [40, Sect. 65].

[3] Term-by-term integration of (C.147) is possible thanks to the uniform convergence of the series.

C.18 Random Walk Problem

The problem considered in this section goes under the general name of *random walk problem*. In its simplest formulation, the problem considers an object performing a one-dimensional motion made of steps of equal length, in which each step has the same probability $0 < p < 1$ of being made in the positive direction; the probability for the negative direction is therefore $q = 1 - p$. Given the initial position of the object, the issue is determining the probability that after n steps the position has a specified value. A generalization of the problem is considering a motion in two or more dimensions; in this case, the issue is to find the probability that after n steps the object is located at a specified distance from the origin.

Another generalization, which is considered below, is that of a one-dimensional motion where the length η of each step is a continuous variable. In this case one must assign a probability density $p(\eta)$ such that $p(\eta)\,\mathrm{d}\eta$ is the probability that the displacement at the ith step lies between η and $\eta + \mathrm{d}\eta$; function p is the same for all steps, and the outcome of each step is independent from that of the other steps. The total displacement after n steps is $x = \eta_1 + \cdots + \eta_n$; letting $P(x)$ be the probability density such that $P(x)\,\mathrm{d}x$ is the probability that the total displacement lies between x and $x + \mathrm{d}x$, the issue is determining P, given p.

The problem may be tackled by provisionally considering the simpler example of a die tossed four times, thus giving the sequence of four numbers $\{m_1, m_2, m_3, m_4\}$, $1 \le m_i \le 6$. The probability[4] $\tilde{P}(x)$ is sought that the sum $x = m_1 + m_2 + m_3 + m_4$ equals a given number, e.g., $x = 7$. As the probability that the ith event yields m_i is $\tilde{p}_i(m_i) = 1/6$, it follows that the probability of each sequence $\{m_1, m_2, m_3, m_4\}$ of four independent events is $\tilde{p}_1(m_1) \cdots \tilde{p}_4(m_4) = 1/6^4 = 1/1{,}296$. Clearly,

$$\sum_{m_1=1}^{6} \sum_{m_2=1}^{6} \sum_{m_3=1}^{6} \sum_{m_4=1}^{6} \tilde{p}_1(m_1)\, \tilde{p}_2(m_2)\, \tilde{p}_3(m_3)\, \tilde{p}_4(m_4) = 1\,. \tag{C.152}$$

The probability of interest is then obtained as

$$\tilde{P} = \sum_{m_1=1}^{6} \sum_{m_2=1}^{6} \sum_{m_3=1}^{6} \sum_{m_4=1}^{6} \nu\; \tilde{p}_1(m_1)\, \tilde{p}_2(m_2)\, \tilde{p}_3(m_3)\, \tilde{p}_4(m_4)\,, \tag{C.153}$$

where $\nu = 1$ when $m_1 + m_2 + m_3 + m_4 = 7$ and $\nu = 0$ otherwise, namely, using the Kronecker symbol (A.18), $\nu = \delta[m_1 + m_2 + m_3 + m_4, 7] = \delta[m_1 + m_2 + m_3 + m_4 - 7, 0] = \delta[m_1 + m_2 + m_3 + m_4 - 7]$. The sequences thus selected are four of the type $\{1114\}$, four of the type $\{1222\}$, and twelve of the type $\{1123\}$, whence $\tilde{P}(7) = (4 + 4 + 12)/1{,}296 = 5/324$.

[4]The symbols used in the example are different because they indicate probabilities rather than probability densities.

Although this example may seem trivial, it provides a hint for solving the original problem. As the displacements are independent from each other, the probability of a particular sequence $\eta_1, \eta_2, \ldots, \eta_n$ is given by $p(\eta_1)\, p(\eta_2) \ldots p(\eta_n)$, where p fulfills the normalization condition

$$\int_{-\infty}^{+\infty} p(\eta_i)\, \mathrm{d}\eta_i = 1\,. \tag{C.154}$$

The *moment* of order m of p is defined[5] like in (C.59); in particular, the *mean value* of η is $\langle \eta \rangle = M_1$, whence $\langle \eta - \langle \eta \rangle \rangle = 0$. In turn, the *dispersion* of η is

$$(\Delta\eta)^2 = \langle (\eta - \langle \eta \rangle)^2 \rangle = \langle \eta^2 \rangle - \langle \eta \rangle^2 \geq 0\,. \tag{C.155}$$

Finally, the *standard deviation* of η is $\Delta\eta = \sqrt{(\Delta\eta)^2}$. Now, to find the probability density $P(x)$, with $x = \sum_{i=1}^{n} \eta_i$, one starts from the counterpart of (C.153) in the continuous case and lets

$$P(x) = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \delta(\xi)\, p(\eta_1)\, \mathrm{d}\eta_1 \cdots p(\eta_n)\, \mathrm{d}\eta_n\,, \tag{C.156}$$

with $\xi = -x + \sum_{i=1}^{n} \eta_i$. Using the integral expression (C.43) of $\delta(\xi)$ transforms (C.156) into

$$P(x) = \frac{1}{2\,\pi} \int_{-\infty}^{+\infty} \exp(-\mathrm{i}\, k\, x) \times \tag{C.157}$$

$$\times \left[\int_{-\infty}^{+\infty} p(\eta_1)\, \exp(\mathrm{i}\, k\, \eta_1)\, \mathrm{d}\eta_1 \cdots \int_{-\infty}^{+\infty} p(\eta_n)\, \exp(\mathrm{i}\, k\, \eta_n)\, \mathrm{d}\eta_n \right] \mathrm{d}k\,.$$

The quantity in brackets of (C.157) is the product of n replicas of the same integral; in conclusion, the probability distribution P is found to be

$$P(x) = \frac{1}{2\,\pi} \int_{-\infty}^{+\infty} q^n(k)\, \exp(-\mathrm{i}\, k\, x)\, \mathrm{d}k\,, \quad q(k) = \int_{-\infty}^{+\infty} p(\eta)\, \exp(\mathrm{i}\, k\, \eta)\, \mathrm{d}\eta\,. \tag{C.158}$$

Letting $n = 1$ in the first relation in (C.158) shows that $q(k)$ vanishes asymptotically due to the properties of the Fourier transform (Sect. C.2); also, comparing the second relation with (C.154) provides $q(0) = 1$. Other relations are

$$q' = \frac{\mathrm{d}q}{\mathrm{d}k} = \int_{-\infty}^{+\infty} \mathrm{i}\, \eta\, p(\eta)\, \exp(\mathrm{i}\, k\, \eta)\, \mathrm{d}\eta\,, \qquad q'(0) = \mathrm{i}\, \langle \eta \rangle\,, \tag{C.159}$$

[5] It is assumed that the decay at infinity of the probability density p is such that the integrals defining the moments M_m converge for all $m \geq 0$.

$$q'' = \frac{\mathrm{d}^2 q}{\mathrm{d}k^2} = \int_{-\infty}^{+\infty} -\eta^2\, p(\eta)\, \exp(\mathrm{i}\, k\, \eta)\, \mathrm{d}\eta\,, \qquad q''(0) = -\langle \eta^2 \rangle\,. \tag{C.160}$$

The above results help in finding useful relations between moments of p and P; in fact,

$$\int_{-\infty}^{+\infty} P(x)\, \mathrm{d}x = \int_{-\infty}^{+\infty} q^n(k) \left[\frac{1}{2\,\pi} \int_{-\infty}^{+\infty} \exp(-\mathrm{i}\, k\, x)\, \mathrm{d}x \right] \mathrm{d}k = q^n(0) = 1\,, \tag{C.161}$$

where the integral in brackets equals $\delta(k)$ due to (C.43). Then,

$$\langle x \rangle = \int_{-\infty}^{+\infty} x\, P(x)\, \mathrm{d}x = \frac{1}{2\,\pi} \int_{-\infty}^{+\infty} \left[x \int_{-\infty}^{+\infty} q^n(k)\, \exp(-\mathrm{i}\, k\, x)\, \mathrm{d}k \right] \mathrm{d}x\,. \tag{C.162}$$

Observing that $x\, \exp(-\mathrm{i}\, k\, x)$ is the derivative of $\mathrm{i}\, \exp(-\mathrm{i}\, k\, x)$ with respect to k, the integral over k in (C.162) can be calculated by parts using the asymptotic behavior of q. This yields

$$\langle x \rangle = -\frac{\mathrm{i}}{2\,\pi} \iint_{-\infty}^{+\infty} n\, q^{n-1}(k)\, q'(k)\, \exp(-\mathrm{i}\, k\, x)\, \mathrm{d}k\, \mathrm{d}x\,. \tag{C.163}$$

Integrating over x first, using again (C.43), and inserting the second relation of (C.159), transforms (C.163) into

$$\langle x \rangle = -\mathrm{i}\, n \int_{-\infty}^{+\infty} q^{n-1}(k)\, q'(k)\, \delta(k)\, \mathrm{d}k = -\mathrm{i}\, n\, q^{n-1}(0)\, q'(0) = n\, \langle \eta \rangle\,. \tag{C.164}$$

Similarly,

$$\langle x^2 \rangle = -n\, q''(0) - (n^2 - n)\, \big(q'(0)\big)^2 = n\, \langle \eta^2 \rangle + (n^2 - n)\, \langle \eta \rangle^2 \tag{C.165}$$

whence, observing that $(\Delta x)^2 = \langle x^2 \rangle - \langle x \rangle^2$ (compare with (C.155)),

$$(\Delta x)^2 = n\, \langle \eta^2 \rangle + (n^2 - n)\, \langle \eta \rangle^2 - n^2 \langle \eta \rangle^2 = n\, \langle \eta^2 \rangle - n\, \langle \eta \rangle^2 = n\, (\Delta \eta)^2\,. \tag{C.166}$$

The above results provide interesting information about the sum of statistically independent variables; in particular, if $\langle \eta \rangle \neq 0$ it is

$$\frac{\Delta x}{\langle x \rangle} = \frac{\sqrt{n}\, \Delta \eta}{n\, \langle \eta \rangle} = \frac{1}{\sqrt{n}}\, \frac{\Delta \eta}{\langle \eta \rangle}\,. \tag{C.167}$$

This shows that when n increases, the fractional deviation $\Delta x / \langle x \rangle$ of the values of x around the mean becomes increasingly negligible.

C.19 Central Limit Theorem

The results of Sect. C.18 have been obtained without the need of specifying the form of the probability density $p(\eta)$; in fact, the only hypothesis is that the events are mutually independent. It has also been proven that when the number n of events becomes large, the distribution of the sum x tends to concentrate around the mean value $\langle x \rangle$. This suggests that when the number of events is large, it is sensible to seek for an approximate description of the central part of $P(x)$ (namely, that corresponding to the values of x sufficiently near to $\langle x \rangle$), because the majority of the $P(x)$ values belong there.

To begin, one remembers that $q(0) = 1$ and $\lim_{|k|\to\infty} q(k) = 0$. This means that for $|k|$ sufficiently large it is $|q(k)| < q(0) = 1$. From the second relation in (C.158) one finds that the rapidness with which q vanishes for $k \neq 0$ depends on the form of $p(\eta)$; if $p(\eta)$ varies slowly in a range $\eta_a < \eta < \eta_b$ in which many oscillations of $\exp(\mathrm{i}\, k\, \eta)$ occur, the oscillations almost cancel each other and the corresponding integral tends to vanish:

$$\int_{\eta_a}^{\eta_b} p(\eta)\, \exp(\mathrm{i}\, k\, \eta)\, \mathrm{d}\eta \simeq p(\eta_a) \int_{\eta_a}^{\eta_b} \exp(\mathrm{i}\, k\, \eta)\, \mathrm{d}\eta \simeq 0\,. \tag{C.168}$$

The condition that makes p a slowly varying function in the given range is $\Delta p = |\mathrm{d}p/\mathrm{d}\eta|\, (\eta_b - \eta_a) \ll p$; in turn, the condition that the range contains many oscillations reads $|k|\, (\eta_b - \eta_a) \gg 1$; combining the two conditions yields

$$\frac{1}{|k|} \left| \frac{\mathrm{d}p}{\mathrm{d}\eta} \right| \ll p\,, \tag{C.169}$$

where the range of $\eta_b - \eta_a$ has been eliminated. For all values of k that fulfill (C.169), $|q|$ rapidly decays to zero as k departs from zero. The decay is much more rapid for q^n when n is large; this suggests to calculate $P(x)$ in (C.158) by expanding q^n, or some function of it, into a Taylor series around $k = 0$ and truncating the series to some order. To proceed, one recasts the second relation in (C.158) by expanding the exponential, to find (compare with (C.61))

$$q(k) = \int_{-\infty}^{+\infty} p(\eta) \sum_{m=0}^{\infty} \frac{1}{m!}\, (\mathrm{i}\, k\, \eta)^m\, \mathrm{d}\eta = \sum_{m=0}^{\infty} \frac{\langle \eta^m \rangle}{m!}\, (\mathrm{i}\, k)^m\,. \tag{C.170}$$

As indicated above, q^n decays rapidly to zero for a large n; it is then preferable to use the logarithm by letting $q(k) = 1 + \vartheta$, where $\vartheta = \langle \eta \rangle\, \mathrm{i}\, k - (1/2)\, \langle \eta^2 \rangle\, k^2 + \cdots$ is the part of the series starting from $m = 1$. It follows $q^n = \exp[n\, \log(1 + \vartheta)]$, where $|\vartheta| < 1$ because the values of k near $k = 0$ are considered: this makes it possible to use the series expansion $\log(1 + \vartheta) = \vartheta - (1/2)\, \vartheta^2 + (1/3)\, \vartheta^3 + \cdots$. Finally, one replaces the expression of ϑ in the last series and retains only the terms that are first or second order in k: one finds that ϑ contributes a term of the first order and

another term of the second order; next, $(1/2)\,\vartheta^2$ contributes a term of the second order, which is kept, and higher-order terms, which are neglected; the terms from $(1/3)\,\vartheta^3$ on are not considered. Thus,

$$n\,\log(1+\vartheta) \simeq n\,\langle\eta\rangle\,\mathrm{i}\,k - \frac{n}{2}\,\left(\langle\eta^2\rangle - \langle\eta\rangle^2\right)\,k^2 = \mathrm{i}\,\langle x\rangle\,k - \frac{1}{2}\,(\Delta x)^2\,k^2\,, \qquad \text{(C.171)}$$

where (C.164) and (C.166) have been used. Taking the exponential of the above and replacing it into the first relation of (C.158) yield

$$P(x) \simeq \frac{1}{2\,\pi}\,\int_{-\infty}^{+\infty} \exp[\mathrm{i}\,(\langle x\rangle - x)\,k - (\Delta x)^2\,k^2/2]\,\mathrm{d}k\,, \qquad \text{(C.172)}$$

which is identical to (C.81). It follows

$$P(x) \simeq \frac{1}{\sqrt{2\,\pi}\,\Delta x}\,\exp\left[-\frac{(x-\langle x\rangle)^2}{2\,(\Delta x)^2}\right]\,, \qquad \int_{-\infty}^{+\infty} P(x)\,\mathrm{d}x = 1\,, \qquad \text{(C.173)}$$

where the second relation derives from (C.78). In conclusion, for a sufficiently large set of mutually independent events, having a slowly varying probability density $p(\eta)$, the central part of the probability density of the sum $x = \eta_1 + \cdots + \eta_n$ is well approximated by a Gaussian. This result, called *central-limit theorem*, is very important because its conditions, being relatively unrestrictive, apply to a large variety of practical cases.

Appendix D
Tables

Table D.1 Fundamental constants

Quantity	Symbol	Value[a]	Units
Vacuum permittivity	ε_0	8.85419×10^{-12}	F m^{-1}
Speed of light	c	2.99792×10^{8}	m s^{-1}
Electron charge	q	1.60219×10^{-19}	C
Electron rest mass	m_0	9.10953×10^{-31}	kg
Proton rest mass	M_0	1.67265×10^{-27}	kg
Boltzmann constant	k_B	1.38066×10^{-23}	J K^{-1}
Stefan-Boltzmann const.	σ	5.67037×10^{-12}	W cm^{-2} K^{-4}
Planck constant	h	6.62616×10^{-34}	J s
Reduced Planck const.	$\hbar$	1.05459×10^{-34}	J s
Atomic radius	r_a	$\sim 10^{-10}$	m
Electron radius	r_e	2.81794×10^{-15}	m

[a] The ratio between the proton and electron rest masses is $M_0/m_0 \simeq 1836$
The vacuum permeability is found from $\mu_0 = 1/(c^2\varepsilon_0)$

M. Rudan, *Physics of Semiconductor Devices*, DOI 10.1007/978-3-319-63154-7

Table D.2 Greek alphabet

Small	Capital[a]	Name	Small	Capital	Name
α	A	alpha	ν	N	nu, ni
β	B	beta	ξ	Ξ	xi
γ	Γ	gamma	o	O	omicron
δ	Δ	delta	π	Π	pi
ε	E	epsilon	ϱ	P	rho
ζ	Z	zeta	σ	Σ	sigma
η	H	eta	τ	T	tau
θ, ϑ	Θ	theta	υ	Υ	upsilon
ι	I	iota	ϕ, φ	Φ	phi
κ	K	kappa	χ	X	chi
λ	Λ	lambda	ψ	Ψ	psi
μ	M	mu, mi	ω	Ω	omega

[a]Symbol ∇ is not a Greek letter. However, its name *nabla* is a Greek word, meaning "harp"

Solutions

Problems of Chap. 1

1.1 The distance between A and B is a functional of y:

$$G[y] = \int_{AB} \sqrt{\mathrm{d}x^2 + \mathrm{d}y^2} = \int_a^b \sqrt{1 + \dot{y}^2}\, \mathrm{d}x\,.$$

As $g(y, \dot{y}, x) = \sqrt{1 + \dot{y}^2}$ it is $\partial g/\partial y = 0$, whence the Euler-Lagrange equation reads

$$0 = \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial g}{\partial \dot{y}} = \frac{\mathrm{d}}{\mathrm{d}x}\frac{2\dot{y}}{2g} = \frac{\ddot{y}g - \dot{y}(2\dot{y}\ddot{y}/2g)}{g^2} = \frac{\ddot{y}}{g^3}\,(g^2 - \dot{y}^2) = \frac{\ddot{y}}{g^3}\,,$$

that is, $\ddot{y} = 0$, $y = c_1 x + c_2$. The two constants are found from $c_1 a + c_2 = y_a$, $c_1 b + c_2 = y_b$.

1.2 Letting $H = E$ one finds

$$\frac{x^2}{a^2} + \frac{p^2}{b^2} = 1\,, \qquad a = \sqrt{2E/c}\,, \quad b = \sqrt{2mE}\,.$$

The curves are ellipses whose axes are proportional to $\sqrt{E}$. The area of each ellipse is $\pi\, a\, b = 2\pi E/\omega$, with $\omega = \sqrt{c/m}$. As shown in Sect. 3.3, ω is the angular frequency of the oscillator, so that the area becomes $E\,T$, with $T = 2\pi/\omega$ the period. Examples of constant-energy curves, in arbitrary units, are shown in Fig. D.1, where the parameters have been set to $m = 0.5$, $c = 1$, and $E = 2$, $E = 4.5$, $E = 8$. The x coordinate ranges between $-x_M$ and x_M, with $x_M = \sqrt{2E/c}$, the p coordinate ranges between $-p_M$ and p_M, with $p_M = \sqrt{2mE}$. As time evolves, the phase point follows the curve in the clockwise direction; in fact, as the phase point reaches the maximum elongation $x_M > 0$ from the left, the momentum at x_M changes from positive to negative.

M. Rudan, *Physics of Semiconductor Devices*, DOI 10.1007/978-3-319-63154-7

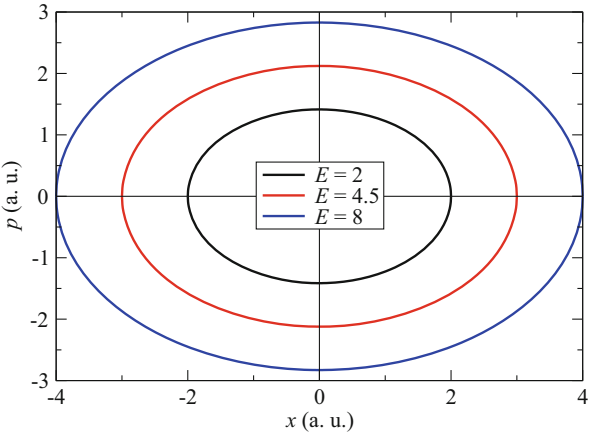


Fig. D.1 Constant-energy curves of the linear harmonic oscillator discussed in Prob. 1.2

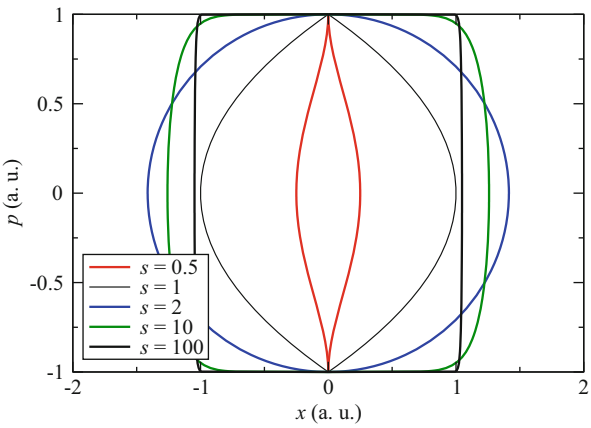


Fig. D.2 Constant-energy curves of the nonlinear harmonic oscillator discussed in Prob. 1.3

1.3 Letting $H = E$ one finds for the maximum elongation $x_M = (s\,E/c)^{1/s}$. Note that the units of c depend on the value of s. Examples of constant-energy curves, in arbitrary units, are shown in Fig. D.2, where the parameters have been set to $m = 1$, $c = 1$, $E = 1$, and $s = 0.5$, $s = 1$, $s = 2$, $s = 10$, $s = 100$. The x coordinate ranges between $-x_M$ and x_M, and the p coordinate ranges between $-p_M$ and p_M, with $p_M = \sqrt{2\,m\,E}$. The form of the constant-energy curves becomes more and more rectangular as s increases. Compare with Prob. 3.1, where it is shown that the $s \to \infty$ limit yields the case of a square well. As in Prob. 1.2, the phase point follows the curve in the clockwise direction.

1.4 The state trajectory of the linear harmonic oscillator is sketched in Fig. D.3. Assume that the particle is initially at A; its position x coincides with the negative extremum $-x_M$ of the trajectory, its momentum is zero, and the force acting on the

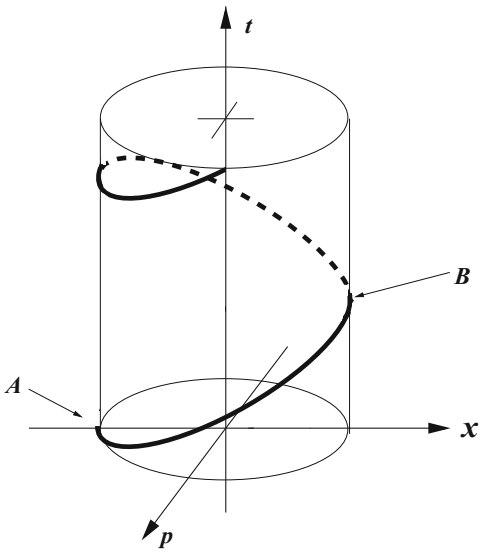


Fig. D.3 State trajectory of the linear harmonic oscillator

particle is positive. When the particle reaches the positive extremum x_M (point B in Fig. D.3), half a period has elapsed, corresponding to the difference between the positions of B and A along the vertical axis. At B the motion reverses (dashed line starting from B). Note that the orientations of the x and p axes are different from those used in Probs. 1.2 and 1.3; for this reason, here the motion of the phase point viewed from the positive side of the t axis is counterclockwise.

1.5 The state trajectory of the harmonic oscillator of the general form used in Prob. 1.3, in the $s \to \infty$ limit, is sketched in Fig. D.4. The problem reduces to describing the motion of a particle within a square well. Assume that the particle is initially at A; its position x coincides with the negative edge $-x_M$ of the well, and its momentum is positive. The particle's motion at A is oriented in the positive direction of the x axis; when the particle reaches the positive edge x_M (point B in Fig. D.4), a time has elapsed corresponding to the difference between the positions of B and A along the vertical axis. As shown by the first portion of the state trajectory (thick line between A and B), the $x(t)$ relation is linear; in fact, the particle's velocity p/m is constant. At point B, the particle is reflected by the right edge of the well, so that its momentum reverses in zero time: this is described by the segment between B and C, the two extrema of which have the same position along the x axis and also along the t axis. After the reflection has occurred, the particle continues its motion with a constant (negative) momentum towards the left edge of the well (dashed line starting from C). Note that the orientations of the x and p axes are different from those used in Probs. 1.2 and 1.3; for this reason, here the motion of the phase point viewed from the positive side of the t axis is counterclockwise.

1.6 The generating function of the functional is the three-dimensional form of the left-hand side of (1.77), $g = \mu^2 (w_x^2 + w_y^2 + w_z^2)/(2m) + V w^2$, with $w_x = \partial w/\partial x$

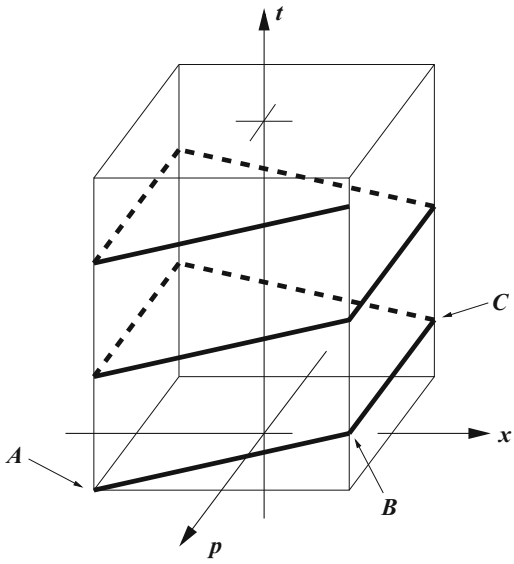


Fig. D.4 State trajectory of the harmonic oscillator of the general form

and the like for w_y, w_z. The generating function of the constraint is $g_0 = w^2$; applying (4.5) to $g_E = g - E\,g_0$ yields

$$\frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial g_E}{\partial w_x} + \frac{\mathrm{d}}{\mathrm{d}y}\frac{\partial g_E}{\partial w_y} + \frac{\mathrm{d}}{\mathrm{d}z}\frac{\partial g_E}{\partial w_z} = \frac{\partial g_E}{\partial w}\,,$$

with

$$\frac{\partial g_E}{\partial w} = 2\,(V - E)\,w\,, \qquad \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial g_E}{\partial w_x} = \frac{\mathrm{d}}{\mathrm{d}x}\frac{\mu^2}{m}\,w_x = \frac{\mu^2}{m}\,w_{xx}\,,$$

and the like for w_y, w_z. Combining the above and observing that $w_{xx} + w_{yy} + w_{zz} = \nabla^2 w$ yield

$$-\frac{\mu^2}{2\,m}\,\nabla^2 w + V\,w = E\,w\,.$$

Problems of Chap. 2

2.1 From (2.49) one finds

$$J(E) = \sqrt{m\,c}\oint \sqrt{2\,E/c - x^2}\;\mathrm{d}x\,,$$

where the integration path is the ellipse described in problem 1.2. Letting $x = \sqrt{2\,E/c}\,\sin\varphi$ transforms the above into

$$J(E) = 2\,\sqrt{\frac{m}{c}}\,E\int_0^{2\,\pi}\cos^2\varphi\,\mathrm{d}\varphi = \frac{2\,\pi}{\omega}\,E\,,\qquad \omega = \sqrt{\frac{c}{m}}\,.$$

The first of (2.51) then yields $\nu = \dot{w} = \partial H/\partial J = \partial E/\partial J = \omega/(2\,\pi)$.

Problems of Chap. 3

3.1 Like in problem 1.3, letting $H = E > 0$ one finds for the maximum elongation $x_M = (s\,E/c)^{1/s}$, where the units of c depend on the value of s. The motion is limited to the interval $[-x_M, +x_M]$ and the potential energy is symmetric with respect to the origin. Using (2.47) and exploiting the symmetry yield

$$T = 4\,\sqrt{\frac{m}{2}}\int_0^{x_M}\frac{\mathrm{d}x}{\sqrt{E - V(x)}} = \sqrt{\frac{8\,m}{E}}\int_0^{x_M}[1 - (x/x_M)^s]^{-1/2}\;\mathrm{d}x\,.$$

Letting $\xi = x/x_M$ and using (C.106, C.107) yield

$$T = \sqrt{\frac{8\,m}{E}}\,x_M\int_0^1\frac{\mathrm{d}\xi}{\sqrt{1-\xi^s}} = \sqrt{8\,\pi\,m}\;\frac{(1/s)\,\Gamma(1/s)}{\Gamma(1/s+1/2)}\left(\frac{s}{c}\right)^{1/s}E^{1/s-1/2}\,.$$

The result shows that the case $s = 2$, namely, that of the linear harmonic oscillator, is special. In fact, the period does not depend on the total energy, whereas for $s \neq 2$ it does. Still in the case $s = 2$ one finds $T = 2\,\pi/\omega$, $\omega = \sqrt{c/m}$, as should be (compare with the results of Sect. 3.3). In turn, the case $s \to \infty$ yields $s^{1/s} \to 1$, $c^{1/s} \to 1$ whence, using (C.108), $\lim_{s\to\infty} T = \sqrt{8\,m/E}$. The above is the period in a square well of length 2 (compare with the description of Sect. 3.2). In fact, as $s \to \infty$, the potential energy $c|x|^s/s$ transforms into a square well with $x_M = 1$. The potential energy is shown in Fig. D.5 for some values of s. Thanks to the result of this problem one may tune the form of the potential energy to make the period proportional to a chosen power $h = 1/s - 1/2 \geq -1/2$ of the energy. For instance, letting $s = 2/3$ makes T proportional to E, namely, $T = \sqrt{m/(3c^3)}\,2\pi E$.

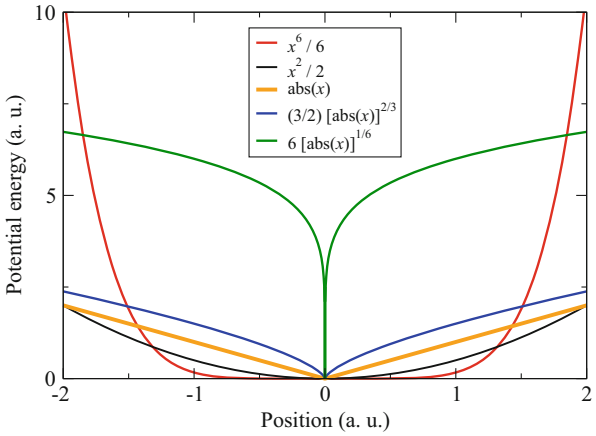


Fig. D.5 Form of the potential energy $c|x|^s/s$ for $c = 1$ and different values of s (Prob. 3.1)

3.2 The solution is similar to that of Prob. 3.1. Letting $H = E < 0$ one finds for the maximum elongation $x_M = [k/(s\,|E|)]^{1/s}$, where the units of k depend on the value of s. The motion is limited to the interval $[-x_M, +x_M]$ and the potential energy is symmetric with respect to the origin. Using (2.47) and exploiting the symmetry yield

$$T = 4\sqrt{\frac{m}{2}}\int_0^{x_M}\frac{\mathrm{d}x}{\sqrt{E-V(x)}} = \sqrt{\frac{8\,m}{|E|}}\int_0^{x_M}[(x_M/x)^s-1]^{-1/2}\,\mathrm{d}x\,.$$

Letting $\xi = x/x_M$ and using (C.109, C.110) yield

$$T = \sqrt{\frac{8\,m}{|E|}}\,x_M\int_0^1\frac{\mathrm{d}\xi}{\sqrt{1/\xi^s-1}} = \sqrt{8\,\pi\,m}\,\frac{\Gamma(1/s+1/2)}{s\,\Gamma(1/s+1)}\left(\frac{k}{s}\right)^{1/s}|E|^{-1/s-1/2}\,.$$

The Coulomb case $s = 1$ yields $T = \sqrt{2\,m}\,\pi\,k\,|E|^{-3/2}$ (in fact, in the Coulomb case and for a closed trajectory the period is always proportional to $|E|^{-3/2}$, compare with (3.81)). Note that in the case considered here the particle crosses the origin because the initial conditions are such that its trajectory is aligned with the x axis. The limit $s \to \infty$ yields $s^{1/s} \to 1$, $c^{1/s} \to 1$ whence, using (C.108, C.110), $\lim_{s\to\infty} T = 0$. The potential energy is shown in Fig. D.6 for some values of s.

3.3 The O reference is chosen as in Sect. 3.13.5, whence $T_{1a} = E = (m_1/m)\,E_B$. From (3.36) one extracts $\mu/s_0 = \tan[(\pi-\chi)/4]$, to find

$$\frac{2\,\mu/s_0}{1-(\mu/s_0)^2} = \frac{2\,\mu\,s_0}{s_0^2-\mu^2} = \tan\left(\frac{\pi-\chi}{2}\right) = \frac{1}{\tan(\chi/2)}\,,$$

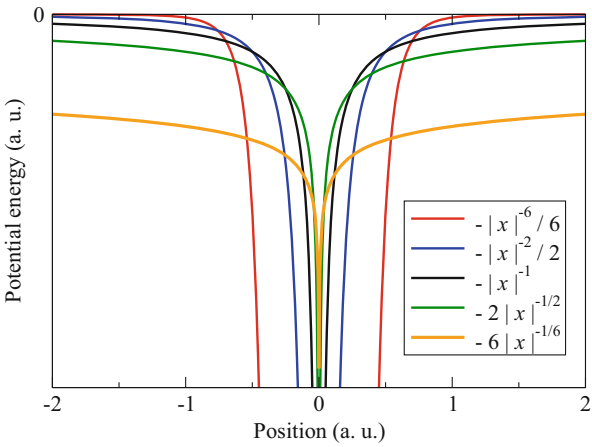


Fig. D.6 Form of the potential energy $-k|x|^{-s}/s$ for $k = 1$ and different values of s (problem 3.2)

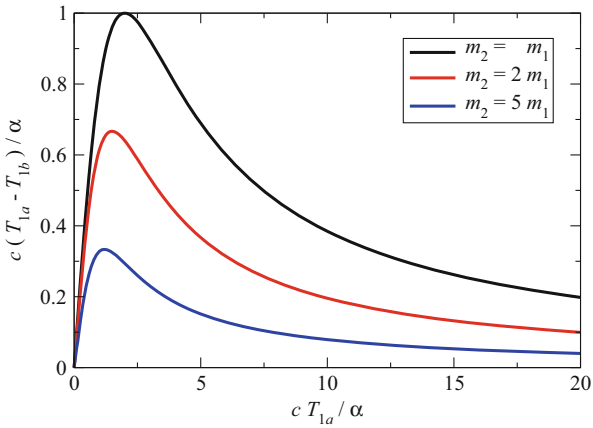


Fig. D.7 Normalized loss of energy $c\,(T_{1a} - T_{1b})/\alpha$ as a function of the normalized initial energy $c\,T_{1a}/\alpha$ (problem 3.3), for different values of the ratio m_1/m_2

where s_0 is given by the second of (3.33). It follows that $s_0^2 - \mu^2 = 2\,\lambda\,s_0$ and $\tan(\chi/2) = \lambda/\mu$. Then, noting that (3.23) contains $\sin^2(\chi/2) = \tan^2(\chi/2)/[1 + \tan^2(\chi/2)]$, and using (3.73), one finds $\sin^2(\chi/2) = 1/(1 + c^2/\lambda^2)$. The expression of λ is taken from the first of (3.32), with α given by (3.75). Inserting the result into (3.23) yields

$$T_{1b} = \frac{\alpha^2\,(1 - m_1/m_2)^2 + c^2\,T_{1a}^2}{\alpha^2\,(1 + m_1/m_2)^2 + c^2\,T_{1a}^2}\;T_{1a}\,, \quad T_{1a} - T_{1b} = \frac{4\,(m_1/m_2)\,T_{1a}}{(1 + m_1/m_2)^2 + (c/\alpha)^2\,T_{1a}^2}\,.$$

Obviously it is $T_{1b} < T_{1a}$. It follows that $T_{1a} - T_{1b}$ is the loss of energy due to the collision. It is also interesting to note that using the normalized energies $c\,T_{1a}/\alpha$ and $c\,T_{1b}/\alpha$ makes the expressions above to depend on the m_1/m_2 ratio only. The loss of energy is drawn in normalized form in Fig. D.7 for different values of m_1/m_2.

Problems of Chap. 4

4.1 Using primes to indicate derivatives, a first integration yields

$$\varphi' = \varphi'(c) - H\,, \qquad H(x) = \int_c^x \frac{\varrho(\xi)}{\varepsilon_0}\,\mathrm{d}\xi\,,$$

where H is integrated by parts:

$$\int_a^x H(\xi)\,\mathrm{d}\xi = x\,H(x) - a\,H(a) - \int_a^x \xi\,\frac{\varrho(\xi)}{\varepsilon_0}\,\mathrm{d}\xi\,.$$

Integrating φ' and using the expression of $\int_a^x H(\xi)\,\mathrm{d}\xi$ yield the solution

$$\varphi = \varphi(a) + \varphi'(c)\,(x-a) - x\int_c^x \frac{\varrho(\xi)}{\varepsilon_0}\,\mathrm{d}\xi + a\int_c^a \frac{\varrho(\xi)}{\varepsilon_0}\,\mathrm{d}\xi + \int_a^x \xi\,\frac{\varrho(\xi)}{\varepsilon_0}\,\mathrm{d}\xi\,.$$

4.2 Letting $c = a$ in the solution to Prob. 4.1 yields at any point x within $[a, b]$ the expression

$$\varphi(x) = \varphi(a) + \varphi'(a)\,(x-a) - x\int_a^x \frac{\varrho(\xi)}{\varepsilon_0}\,\mathrm{d}\xi + \int_a^x \xi\,\frac{\varrho(\xi)}{\varepsilon_0}\,\mathrm{d}\xi\,.$$

For $x > b$ it is $\varrho = 0$ so that the solution of $\varphi'' = 0$ is linear and has the form $\varphi(x) = \varphi(b) + \varphi'(b)\,(x-b)$. The term $\varphi(b)$ in the latter is obtained by letting $x = b$ in the above expression of $\varphi(x)$. One finds $\varphi(b) = \varphi(a) + \varphi'(a)\,(b-a) - b\,M_0 + M_1$, with

$$M_0 = \int_a^b \frac{\varrho(\xi)}{\varepsilon_0}\,\mathrm{d}\xi\,, \qquad M_1 = \int_a^b \xi\,\frac{\varrho(\xi)}{\varepsilon_0}\,\mathrm{d}\xi$$

the first two moments of ϱ/ε_0 (compare with Sect. C.6). The derivative φ' is found from Prob. 4.1 with $c = a$, and reads

$$\varphi'(x) = \varphi'(a) - \int_a^x \frac{\varrho(\xi)}{\varepsilon_0}\,\mathrm{d}\xi\,,$$

whence $\varphi'(b) = \varphi'(a) - M_0$. Using the expressions of $\varphi(b)$, $\varphi'(b)$ thus found yields

$$\varphi(x) = \varphi(a) + \varphi'(a)\,(x-a) - M_0\,x + M_1\,, \qquad x > b\,.$$

4.3 From the findings of Prob. 4.2 one observes that the solution φ is invariant for any charge density $\tilde{\varrho}$ that leaves M_0 and M_1 unchanged. Due to this, if both M_0 and M_1 differ from zero, the new charge density must contain two adjustable parameters in order to fit the values of M_0, M_1 through the expressions introduced in Prob. 4.2.

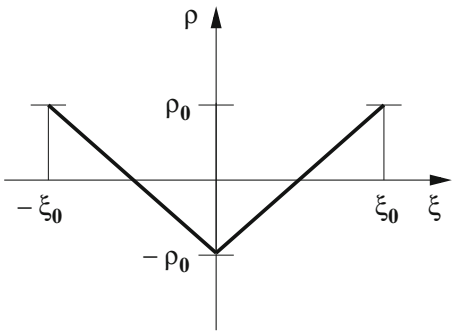


Fig. D.8 Example of charge density such that $M_0 = 0$ and $M_1 = 0$

If only one moment differs from zero, one parameter suffices, while no parameter is necessary if both moments are equal to zero. Figure D.8 gives an example of charge density such that $M_0 = 0$ and $M_1 = 0$.

4.4 The starting point is the solution for $x > b$ found in Prob. 4.2. When the charge density is removed, the new solution reads

$$\varphi(x) = \tilde{\varphi}(a) + \tilde{\varphi}'(a)\,(x - a)\,.$$

For $x > b$ the two solutions become equal to each other by letting $\tilde{\varphi}(a) = \varphi(a) - M_0\,a + M_1$ and $\tilde{\varphi}'(a) = \varphi'(a) - M_0$.

4.5 Considering that the value of h is unknown, the integrals that define the moments (Prob. 4.2) must be extended from $-\infty$ to $+\infty$. One finds $\mu = M_0$, $h = M_1/M_0$. If $h \geq a$, the solution is given by $\varphi = \varphi(a) + \varphi'(a)\,(x - a) - M_0\,x + M_1$ for $x \geq h$, while it is given by $\varphi = \varphi(a) + \varphi'(a)\,(x - a)$ for $x < h$. If $h < a$, the solution is given by $\varphi = \varphi(a) + \varphi'(a)\,(x - a)$ for $x \geq h$, while it is given by $\varphi = \varphi(a) + \varphi'(a)\,(x - a) - M_0\,x + M_1$ for $x < h$. At h the electric potential is continuous, $\varphi(h^+) = \varphi(h^-) = \varphi(a) + \varphi'(a)\,(h - a)$, whereas the electric field is discontinuous, $\varphi'(h^-) - \varphi'(h^+) = M_0$. The case $M_0 \neq 0$, $M_1 = 0$ yields $\mu = M_0$, $h = 0$, while the case $M_0 = 0$, $M_1 \neq 0$ cannot be fulfilled by $\mu\,\delta(x - h)$.

4.6 From (4.60), the force acting on the particle is

$$\mathbf{F} = q\,\mathbf{u} \wedge \mathbf{B} = q \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_x & u_y & u_z \\ 0 & 0 & B \end{bmatrix} = q\,u_y\,B\,\mathbf{i} - q\,u_x\,B\,\mathbf{j}\,,$$

whence Newton's law reads

$$m\,\dot{u}_x = q\,B\,u_y\,, \qquad m\,\dot{u}_y = -q\,B\,u_x\,, \qquad m\,\dot{u}_z = 0\,.$$

Letting u_{x0}, u_{y0}, u_{z0} be the components of the initial velocity, from the third relation above one derives $z(t) = z_0 + u_{z0}\,t$, with $z_0 = z(0)$. Combining the other two relations yields

$$\ddot{u}_x = -(q\,B/m)^2\,u_x = -\omega^2\,u_x\,, \qquad \ddot{u}_y = -(q\,B/m)^2\,u_y = -\omega^2\,u_y\,,$$

$\omega = q\,B/m > 0$, whose solution is

$$u_x = a_x\,\cos(\omega\,t) + b_x\,\sin(\omega\,t)\,, \qquad u_y = a_y\,\cos(\omega\,t) + b_y\,\sin(\omega\,t)\,.$$

The expression of u_x yields $a_x = u_{x0}$; observing that $\dot{u}_{x0} = \omega\,b_x$ and $\dot{u}_{x0} = \omega\,u_{y0}$, one finds $b_x = u_{y0}$. By the same token, $b_y = -u_{x0}$. Now, let the initial conditions be $u_{x0} > 0$ and $u_{y0} = u_{z0} = 0$, whence $z = z_0$, $b_x = a_y = 0$; thus, $u_x = u_{x0}\,\cos(\omega\,t)$ and $u_y = -u_{x0}\,\sin(\omega\,t)$. In conclusion,

$$x = x_0 + (u_{x0}/\omega)\,\sin(\omega\,t)\,, \qquad y = y_0 + (u_{x0}/\omega)\,\cos(\omega\,t)\,.$$

Combining the above yields $u^2 = u_x^2 + u_y^2 = u_{x0}^2$ and

$$(x - x_0)^2 + (y - y_0)^2 = u_{x0}^2/\omega^2\,, \qquad z = z_0\,,$$

that is, the trajectory is a circumference of radius $r = u_{x0}/\omega$ belonging to the plane $z = z_0$.

Problems of Chap. 5

5.1 From $n = n(x_1)$ one finds that $\mathrm{grad}\,n = \mathbf{i}_1\,\mathrm{d}n/\mathrm{d}x_1$ is parallel to x_1 whereas $\mathrm{d}n/\mathrm{d}x_2 = \mathrm{d}n/\mathrm{d}x_3 = 0$. From the eikonal equation (5.57) it follows

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(n\,\frac{\mathrm{d}x_2}{\mathrm{d}s}\right) = 0\,, \qquad \frac{\mathrm{d}}{\mathrm{d}s}\left(n\,\frac{\mathrm{d}x_3}{\mathrm{d}s}\right) = 0\,,$$

whence $n\,\mathrm{d}x_2/\mathrm{d}s = \mathrm{const}$, $n\,\mathrm{d}x_3/\mathrm{d}s = \mathrm{const}$. The ratio of the latter relations yields $\mathrm{d}x_2/\mathrm{d}x_3 = \mathrm{const}$, namely, $x_2 = a\,x_3 + b$, where a, b are constants. This expression is one of the two parametric equations $u(x_1, x_2, x_3) = 0$, $v(x_1, x_2, x_3) = 0$ describing the ray, and shows that the ray belongs to a plane parallel to x_1. By a suitable rotation around x_1, such that $x_2 \to x_2'$, $x_3 \to x_3'$, the plane of the ray is made parallel to the plane $x_1\,x_2'$, so that the third coordinate x_3' is fixed. In the new reference, let ϑ be the angle between the direction of the ray and x_1 at some point P; it is $\mathrm{d}x_1 = \cos\vartheta\,\mathrm{d}s$, $\mathrm{d}x_2' = \sin\vartheta\,\mathrm{d}s$. The eikonal equation in the new reference then provides

$$n\,\frac{\mathrm{d}x_2'}{\mathrm{d}s} = \mathrm{const}\,, \qquad n\,\sin\vartheta = \mathrm{const}\,.$$

5.2 Like in problem 5.1 one considers the case where the refraction index depends only on the x_1 coordinate. Let the medium be made of three regions separated by two planes parallel to each other. The two external regions A and B have a

constant refraction index, n_A and, respectively, $n_B \neq n_A$. The internal region I, whose thickness is s, has a refraction index that varies continuously from n_A to n_B as x_1 varies from the A-I interface to the I-B interface. Applying the solution of problem 5.1 to this case shows that $n_A \sin\vartheta_A = \text{const}$ everywhere in region A, hence the ray is a straight line there; similarly it is $n_B \sin\vartheta_B = \text{const}$ everywhere in region B, with the same constant. It follows

$$n_B \sin\vartheta_B = n_A \sin\vartheta_A \,.$$

Unless $\vartheta_A = 0$, the position of the ray along the x_2' axis at the I-B interface is different from that at the A-I interface; if, however, s is made to vanish, the position becomes the same, thus yielding the *Descartes law of refraction*: the ray crossing an interface between two media is continuous, whereas its slopes on the two sides of the interface fulfill the relation above. The result still holds in the cases where the interface between the two media is not planar, provided its curvature is small enough to make Geometrical Optics applicable.

5.3 Using (5.65) and $\operatorname{grad}\sigma = \mathbf{g}/g$, one finds

$$\frac{\partial\tau}{\partial x_i} = -\frac{\partial\sigma/\partial x_i}{\partial\sigma/\partial\tau} = -\frac{g_i/g}{c - \mathbf{u}\cdot\mathbf{g}/g} = -\frac{\partial\tau}{\partial t}\,\frac{g_i}{c\,g}\,, \qquad \operatorname{grad}\tau = -\frac{\partial\tau}{\partial t}\,\frac{\mathbf{g}}{c\,g}\,.$$

Combining the third relation in the above with $g = c\,(t-\tau)$ yields $\operatorname{grad} g = -c\,\operatorname{grad}\tau = (\mathbf{g}/g)\,\partial\tau/\partial t$. It follows

$$\frac{\partial g}{\partial t} = \frac{\partial g}{\partial\tau}\,\frac{\partial\tau}{\partial t} = -\frac{\partial\tau}{\partial t}\,\mathbf{u}\cdot\frac{\mathbf{g}}{g}\,, \qquad \frac{\partial\mathbf{g}}{\partial t} = -\frac{\partial\mathbf{g}}{\partial\tau}\,\frac{\partial\tau}{\partial t} = -\frac{\partial\tau}{\partial t}\,\mathbf{u}\,,$$

whence

$$\frac{\partial}{\partial t}\left(\frac{\mathbf{g}}{g}\right) = \frac{\mathbf{u}\cdot\mathbf{g}\,\mathbf{g} - \mathbf{g}\cdot\mathbf{g}\,\mathbf{u}}{g^3}\,\frac{\partial\tau}{\partial t}\,.$$

Starting from the above relations and from $\partial\mathbf{u}/\partial t = \dot{\mathbf{u}}\,\partial\tau/\partial t$, and using (5.65) repeatedly, one finds

$$\frac{\partial^2\tau}{\partial t^2} = \frac{\partial}{\partial t}\left(\frac{c}{c - \mathbf{u}\cdot\mathbf{g}/g}\right) = c\,\frac{\partial\,(\mathbf{u}\cdot\mathbf{g}/g)\,/\partial t}{(c - \mathbf{u}\cdot\mathbf{g}/g)^2} = \frac{1}{c}\left(\frac{\partial\tau}{\partial t}\right)^2\,\frac{\partial\,(\mathbf{u}\cdot\mathbf{g}/g)}{\partial t}\,,$$

where

$$\frac{\partial\,(\mathbf{u}\cdot\mathbf{g}/g)}{\partial t} = \frac{\mathbf{g}}{g}\cdot\frac{\partial\mathbf{u}}{\partial t} + \mathbf{u}\cdot\frac{\partial(\mathbf{g}/g)}{\partial t} = \frac{\mathbf{g}}{g}\cdot\dot{\mathbf{u}}\,\frac{\partial\tau}{\partial t} + \mathbf{u}\cdot\frac{\mathbf{u}\cdot\mathbf{g}\,\mathbf{g} - \mathbf{g}\cdot\mathbf{g}\,\mathbf{u}}{g^3}\,\frac{\partial\tau}{\partial t}\,.$$

Combining the above finally yields the result sought,

$$\frac{\partial^2\tau}{\partial t^2} = \frac{1}{c\,g^3}\left(\frac{\partial\tau}{\partial t}\right)^3 \left[(\mathbf{g}\cdot\mathbf{u})^2 - g^2\,u^2 + g^2\,\mathbf{g}\cdot\dot{\mathbf{u}}\right] .$$

5.4 It suffices to use (5.65) in (5.6), (5.7), and to replace μ_0 with $1/(\varepsilon_0\,c^2)$ in the second one. The corresponding expressions of $\mathbf{E}$, $\mathbf{B}$ are found from (4.26),

$$\frac{4\,\pi\,\varepsilon_0}{e}\,\mathbf{E} = -\,\mathrm{grad}\left(\frac{\partial\tau}{\partial t}\,\frac{1}{g}\right) - \frac{1}{c^2}\,\frac{\partial}{\partial t}\left(\frac{\partial\tau}{\partial t}\,\frac{\mathbf{u}}{g}\right) , \qquad \frac{4\,\pi\,\varepsilon_0\,c^2}{e}\,\mathbf{B} = \mathrm{rot}\left(\frac{\partial\tau}{\partial t}\,\frac{\mathbf{u}}{g}\right) .$$

5.5 The expression of the electric field $\mathbf{E}$ derived in Prob. 5.4 is made of two terms. For the first one, one calculates

$$\mathbf{S} = -\,\mathrm{grad}\left(\frac{\partial\tau}{\partial t}\,\frac{1}{g}\right) = -\frac{1}{g}\,\frac{\partial\,\mathrm{grad}\,\tau}{\partial t} + \frac{\partial\tau}{\partial t}\,\frac{\mathrm{grad}\,g}{g^2} ,$$

where the time derivative and the gradient have been exchanged in the first term. The $\mathrm{grad}\,\tau$ and $\mathrm{grad}\,g$ terms are replaced, respectively, with $-(\partial\tau/\partial t)\,\mathbf{g}/(c\,g)$ and $(\partial\tau/\partial t)\,\mathbf{g}/g$ (Prob. 5.3), to yield

$$\mathbf{S} = \frac{\mathbf{g}}{c\,g^2}\,\frac{\partial^2\tau}{\partial t^2} + \frac{1}{c\,g}\,\frac{\partial\tau}{\partial t}\,\frac{\partial(\mathbf{g}/g)}{\partial t} + \frac{\mathbf{g}}{g^3}\left(\frac{\partial\tau}{\partial t}\right)^2 .$$

As for the second term in the expression of the electric field $\mathbf{E}$ derived in Prob. 5.4, one calculates

$$\mathbf{V} = -\frac{1}{c^2}\,\frac{\partial}{\partial t}\left(\frac{\partial\tau}{\partial t}\,\frac{\mathbf{u}}{g}\right) = -\frac{1}{c^2}\,\frac{\mathbf{u}}{g}\,\frac{\partial^2\tau}{\partial t^2} - \frac{1}{c^2}\left(\frac{\partial\tau}{\partial t}\right)^2\left(\frac{\dot{\mathbf{u}}}{g} + \frac{\mathbf{u}\cdot\mathbf{g}\,\mathbf{u}}{g^3}\right) ,$$

where the relations $\partial\mathbf{u}/\partial t = \dot{\mathbf{u}}\,\partial\tau/\partial t$ and $\partial g/\partial t = -(\partial\tau/\partial t)\,\mathbf{u}\cdot\mathbf{g}/g$ have been used (Prob. 5.3). Next, one forms the sum $\mathbf{T} = \mathbf{S} + \mathbf{V} = 4\,\pi\,\varepsilon_0\,\mathbf{E}/e$ and groups some terms, to find

$$\mathbf{T} = \frac{c\,\mathbf{g} - g\,\mathbf{u}}{c^2\,g^2}\,\frac{\partial^2\tau}{\partial t^2} + \frac{1}{c\,g}\,\frac{\partial\tau}{\partial t}\,\frac{\partial(\mathbf{g}/g)}{\partial t} + \left(\frac{\partial\tau}{\partial t}\right)^2\frac{c^2\,\mathbf{g} - g^2\,\dot{\mathbf{u}} - \mathbf{u}\cdot\mathbf{g}\,\mathbf{u}}{c^2\,g^3} .$$

The right-hand side of the above has three summands, $\mathbf{T}_i$, $i = 1, 2, 3$; the first of them, using the expression of $\partial^2\tau/\partial t^2$ from Prob. 5.3, reads

$$\mathbf{T}_1 = \left(\frac{\partial\tau}{\partial t}\right)^3 \frac{(\mathbf{g}\cdot\mathbf{u})^2 - g^2\,u^2 + g^2\,\mathbf{g}\cdot\dot{\mathbf{u}}}{c^3\,g^5}\,(c\,\mathbf{g} - g\,\mathbf{u}) .$$

The second and third summands, using the expression of $\partial(\mathbf{g}/g)/\partial t$ from Prob. 5.3, yield

$$\mathbf{T}_2+\mathbf{T}_3=\left(\frac{\partial\tau}{\partial t}\right)^2\frac{c\,\mathbf{g}\cdot\mathbf{u}\,\mathbf{g}-c\,\mathbf{g}\cdot\mathbf{g}\,\mathbf{u}+c^2\,g\,\mathbf{g}-g^3\,\dot{\mathbf{u}}-g\,\mathbf{g}\cdot\mathbf{u}\,\mathbf{u}}{c^2\,g^4}\,,$$

which can be rearranged by introducing the relation $c\,g=(c\,g-\mathbf{g}\cdot\mathbf{u})\,(\partial\tau/\partial t)$ taken from (5.65)

$$\mathbf{T}_2+\mathbf{T}_3=\left(\frac{\partial\tau}{\partial t}\right)^3\frac{c\,\mathbf{g}\cdot\mathbf{u}\,\mathbf{g}-c\,\mathbf{g}\cdot\mathbf{g}\,\mathbf{u}+c^2\,g\,\mathbf{g}-g^3\,\dot{\mathbf{u}}-g\,\mathbf{g}\cdot\mathbf{u}\,\mathbf{u}}{c^3\,g^5}\,(c\,g-\mathbf{g}\cdot\mathbf{u})\,.$$

From the above one calculates $c^3\,g^5\,(\partial\tau/\partial t)^{-3}\,(\mathbf{T}_1+\mathbf{T}_2+\mathbf{T}_3)$. In so doing, the products $c\,(\mathbf{g}\cdot\mathbf{u})^2\,\mathbf{g}$, $c^2\,g\,(\mathbf{g}\cdot\mathbf{u})\,\mathbf{g}$, and $c\,g^2\,(\mathbf{g}\cdot\mathbf{u})\,\mathbf{u}$ are canceled out by similar products having the opposite sign. The remaining products are regrouped to yield, after eliminating $c^3\,g^2$ from both sides,

$$\frac{g^3}{(\partial\tau/\partial t)^3}\,\mathbf{T}=\left(1-\frac{u^2}{c^2}\right)\left(\mathbf{g}-g\,\frac{\mathbf{u}}{c}\right)+\frac{1}{c^2}\left[\left(\mathbf{g}-g\,\frac{\mathbf{u}}{c}\right)\mathbf{g}\cdot\dot{\mathbf{u}}-\mathbf{g}\cdot\left(\mathbf{g}-g\,\frac{\mathbf{u}}{c}\right)\dot{\mathbf{u}}\right].$$

The term in brackets is given a more compact form by letting $\mathbf{a}=\dot{\mathbf{u}}$, $\mathbf{b}=\mathbf{g}-g\,\mathbf{u}/c$, $\mathbf{c}=\mathbf{g}$ and using the second identity in (A.33); combining the result with $\mathbf{T}=4\,\pi\,\varepsilon_0\,\mathbf{E}/e$ finally yields

$$\frac{4\,\pi\,\varepsilon_0\,g^3}{e\,(\partial\tau/\partial t)^3}\,\mathbf{E}=\left(1-\frac{u^2}{c^2}\right)\left(\mathbf{g}-g\,\frac{\mathbf{u}}{c}\right)+\mathbf{g}\wedge\left[\left(\mathbf{g}-g\,\frac{\mathbf{u}}{c}\right)\wedge\frac{\dot{\mathbf{u}}}{c^2}\right],$$

that coincides with (5.66).

5.6 Vector multiplying by $\mathbf{g}/g$ the equation before last of Prob. 5.5 yields

$$\frac{g^2}{(\partial\tau/\partial t)^3}\,\mathbf{T}\wedge\mathbf{g}=-\mathbf{g}\cdot\left(\mathbf{g}-g\,\frac{\mathbf{u}}{c}\right)\frac{\dot{\mathbf{u}}}{c^2}\wedge\frac{\mathbf{g}}{g}+\left[\left(1-\frac{u^2}{c^2}\right)+\mathbf{g}\cdot\frac{\dot{\mathbf{u}}}{c^2}\right]\mathbf{g}\wedge\frac{\mathbf{u}}{c}\,,$$

where the identity $\mathbf{g}\wedge\mathbf{g}=0$ has been accounted for. Similarly, vector multiplying by $\mathbf{u}/c$ the same equation provides

$$-\frac{g^3}{(\partial\tau/\partial t)^3}\,\mathbf{T}\wedge\frac{\mathbf{u}}{c}=\mathbf{g}\cdot\left(\mathbf{g}-g\,\frac{\mathbf{u}}{c}\right)\frac{\dot{\mathbf{u}}}{c^2}\wedge\frac{\mathbf{u}}{c}-\left[\left(1-\frac{u^2}{c^2}\right)+\mathbf{g}\cdot\frac{\dot{\mathbf{u}}}{c^2}\right]\mathbf{g}\wedge\frac{\mathbf{u}}{c}\,,$$

where the identity $\mathbf{u}\wedge\mathbf{u}=0$ has been accounted for. When the last two equations are added up, the following equation ensues:

$$\frac{g^2}{(\partial\tau/\partial t)^3}\,\mathbf{T}\wedge\left(\mathbf{g}-g\,\frac{\mathbf{u}}{c}\right)=-\mathbf{g}\cdot\left(\mathbf{g}-g\,\frac{\mathbf{u}}{c}\right)\frac{\dot{\mathbf{u}}}{c^2}\wedge\frac{\mathbf{g}}{g}+\mathbf{g}\cdot\left(\mathbf{g}-g\,\frac{\mathbf{u}}{c}\right)\frac{\dot{\mathbf{u}}}{c^2}\wedge\frac{\mathbf{u}}{c}\,.$$

Dividing both sides by g and using $\mathbf{T}=4\,\pi\,\varepsilon_0\,\mathbf{E}/e$ yield the result sought.

5.7 The first component of rot $\mathbf{u}$ reads

$$(\mathrm{rot}\,\mathbf{u})_1 = \frac{\partial u_3}{\partial x_2} - \frac{\partial u_2}{\partial x_3} = \frac{\mathrm{d}u_3}{\mathrm{d}\tau}\,\frac{\partial \tau}{\partial x_2} - \frac{\mathrm{d}u_2}{\mathrm{d}\tau}\,\frac{\partial \tau}{\partial x_3} = (\mathrm{grad}\,\tau \wedge \dot{\mathbf{u}})_1 \,,$$

and the like for the other components. Then, the expression of grad τ is taken from Prob. 5.3, thus yielding

$$\frac{4\,\pi\,\varepsilon_0\,c^2}{e}\,\mathbf{B} = -\left(\frac{\partial \tau}{\partial t}\right)^2 \frac{\mathbf{g}}{c\,g^2} \wedge \dot{\mathbf{u}} - \mathbf{u} \wedge \mathrm{grad}\left(\frac{1}{g}\,\frac{\partial \tau}{\partial t}\right) .$$

The second term at the right-hand side needs further elaboration, which is carried out in Prob. 5.8.

5.8 One notes that the expression of $\mathbf{E}$ derived in Prob. 5.4, and that of $\mathbf{B}$ derived in Prob. 5.7, have a term in common, namely, grad$[(1/g)\,\partial\tau/\partial t]$. Eliminating the latter yields

$$\frac{4\,\pi\,\varepsilon_0\,c^2}{e}\,\mathbf{B} = -\left(\frac{\partial \tau}{\partial t}\right)^2 \frac{\mathbf{g}}{c\,g^2} \wedge \dot{\mathbf{u}} + \mathbf{u} \wedge \left[\frac{4\,\pi\,\varepsilon_0}{e}\,\mathbf{E} + \frac{1}{c^2}\,\frac{\partial}{\partial t}\left(\frac{\partial \tau}{\partial t}\,\frac{\mathbf{u}}{g}\right)\right] .$$

Expanding the time derivative of $(\mathbf{u}/g)\,\partial\tau/\partial t$ yields three terms, one of which is $(1/g)\,(\partial\tau/\partial t)\,\partial\mathbf{u}/\partial t = (1/g)\,(\partial\tau/\partial t)^2\,\dot{\mathbf{u}}$. The other two terms do not contribute to $\mathbf{B}$ because they are proportional to $\mathbf{u}$ (in fact, the $\mathbf{u} \wedge \mathbf{u} = 0$ identity holds). Thus,

$$\frac{4\,\pi\,\varepsilon_0\,c^2}{e}\,\mathbf{B} = \mathbf{u} \wedge \frac{4\,\pi\,\varepsilon_0}{e}\,\mathbf{E} - \frac{1}{c\,g^2}\left(\frac{\partial \tau}{\partial t}\right)^2 \left(\mathbf{g} - g\,\frac{\mathbf{u}}{c}\right) \wedge \dot{\mathbf{u}} .$$

Dividing both sides by $c\,(\partial\tau/\partial t)^3/g^3$ yields

$$\frac{g^3}{(\partial\tau/\partial t)^3}\,\frac{4\,\pi\,\varepsilon_0\,c}{e}\,\mathbf{B} = \frac{g^3}{(\partial\tau/\partial t)^3}\,\frac{\mathbf{u}}{c} \wedge \frac{4\,\pi\,\varepsilon_0}{e}\,\mathbf{E} - \frac{g}{\partial\tau/\partial t}\left(\mathbf{g} - g\,\frac{\mathbf{u}}{c}\right) \wedge \frac{\dot{\mathbf{u}}}{c^2} .$$

Due to (5.65) it is $1/(\partial\tau/\partial t) = 1 - \mathbf{g}\cdot\mathbf{u}/(g\,c)$, while $\mathbf{E}$ is expressible in terms of $\partial\tau/\partial t$, $\mathbf{g}$, $\mathbf{u}$, and $\dot{\mathbf{u}} = \partial\mathbf{u}/\partial\tau$ as shown in Prob. 5.6. Thus, the above equation is the result sought.

5.9 Rearranging the first equation of Prob. 5.6 after multiplying both sides by g yields

$$\frac{4\,\pi\,\varepsilon_0\,g^2}{e\,(\partial\tau/\partial t)^3}\,\mathbf{g} \wedge \mathbf{E} = \frac{4\,\pi\,\varepsilon_0\,g^3}{e\,(\partial\tau/\partial t)^3}\,\frac{\mathbf{u}}{c} \wedge \mathbf{E} - \frac{g}{\partial\tau/\partial t}\left(\mathbf{g} - g\,\frac{\mathbf{u}}{c}\right) \wedge \frac{\dot{\mathbf{u}}}{c^2} ,$$

whose right-hand side is identical to that of the last equation of Prob. 5.8. Equating the left-hand sides then yields

$$\mathbf{B} = \frac{\mathbf{g}}{g} \wedge \frac{\mathbf{E}}{c} ,$$

that coincides with (5.67).

Problems of Chap. 6

6.1 Letting $\vartheta = \beta\, h\, \nu$, with $\beta = 1/(k_B\, T)$, the Boltzmann distribution takes the form $N_n = N_0\, \exp(-n\, \vartheta)$, whence

$$\sum_{n=0}^{\infty} N_n = N_0\, [1 + \exp(-\vartheta) + \exp(-2\, \vartheta) + \ldots] = \frac{N_0}{1 - \exp(-\vartheta)} ,$$

and

$$\sum_{n=0}^{\infty} n\, h\, \nu\, N_n = h\, \nu\, N_0\, [\exp(-\vartheta) + 2\, \exp(-2\, \vartheta) + 3\, \exp(-3\, \vartheta) + \ldots] .$$

Observing that $n\, \exp(-n\, \vartheta) = -\mathrm{d}\exp(-n\, \vartheta)/\mathrm{d}\vartheta$, one finds

$$\sum_{n=0}^{\infty} n\, h\, \nu\, N_n = -h\, \nu\, \frac{\mathrm{d}}{\mathrm{d}\vartheta} \left(\sum_{n=0}^{\infty} N_n - N_0 \right) = h\, \nu\, \frac{N_0\, \exp(-\vartheta)}{[1 - \exp(-\vartheta)]^2} ,$$

whence

$$\mathrm{Av}[E_n] = \frac{\sum_{n=0}^{\infty} n\, h\, \nu\, N_n}{\sum_{n=0}^{\infty} N_n} = \frac{h\, \nu}{\exp(\vartheta) - 1} = \frac{h\, \nu}{\exp(\beta\, h\, \nu) - 1} .$$

Problems of Chap. 7

7.1 From (7.13), the power per unit area emitted by a black body is $\sigma\, T^4$, with $\sigma = 5.67 \times 10^{-12}$ W cm^{-2} K^{-4} the Stefan-Boltzmann constant. Thus, the power emitted by the whole surface of the sun is $P = 4\, \pi\, r^2\, \sigma\, T^4$, with $T = 5{,}780$ K. As the emitted radiation spreads uniformly over a spherical surface whose radius is R, the power per unit area at the outer edge of the earth's atmosphere (*solar constant*) is

$$\frac{\mathrm{d}P_E}{\mathrm{d}A} = \frac{P}{4\, \pi\, R^2} = \frac{1}{216^2}\, \sigma\, T^4 \simeq 136 \quad \text{mW cm}^{-2} .$$

Problems of Chap. 8

8.1 Consider the homogeneous equation associated with (8.76), $g'' + a\,g' + b\,g = 0$, and let $f = g\,h$ and $u = h'$; this splits (8.76) into the system

$$g'' + a\,g' + b\,g = 0\,, \qquad g\,u' + (2\,g' + a\,g)\,u = c\,.$$

If g is known, then u is found by integrating a first-order equation, whence $f = g \int u\,\mathrm{d}\xi$. To find g one lets $A(x) = \int a\,\mathrm{d}\xi$, $g = \exp(-A/2)\,w$, thus transforming the homogeneous equation for g into

$$w'' + q\,w = 0\,, \qquad q = b - a^2/4 - a'/2\,,$$

which is a time-independent Schrödinger equation.

8.2 As both w_1 and w_2 are solutions corresponding to the same eigenvalue, it is $w_1'' + q\,w_1 = 0$ and $w_2'' + q\,w_2 = 0$. Multiply the first one by w_2, the second one by w_1, and subtract; this yields $w_2\,w_1'' = w_1\,w_2''$. The latter is recast as $(w_2\,w_1')' = (w_1\,w_2')'$, namely

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(w_1\,w_2' - w_2\,w_1'\right) = \frac{\mathrm{d}W}{\mathrm{d}x} = 0\,, \qquad W = \text{const},$$

with W the Wronskian. As shown in Sect. A.12, the condition $W = 0$ is necessary for the linear dependence of w_1, w_2; as a consequence, the condition $W \neq 0$ is sufficient for the linear independence of w_1, w_2. The calculation carried out in this problem does not tell whether the constant is zero or not; to discriminate between the two cases, one must add some prescription on the solutions (an example is given in Prob. 8.3).

8.3 The procedure shown here is used also in Sect. 11.4. As both w_1 and w_2 are solutions corresponding to the same eigenvalue, it is $w_1'' + q\,w_1 = 0$ and $w_2'' + q\,w_2 = 0$. Letting $w_2 = v\,w_1$ in the latter yields $v\,w_1'' + 2\,v'\,w_1' + v''\,w_1 + v\,q\,w_1 = 0$, where the first and last summands cancel each other due to the equation fulfilled by w_1. Thus, letting a be any point of I, the relation $2\,v'\,w_1' + v''\,w_1 = 0$ is equivalent to

$$\left(w_1^2\,v'\right)' = 0\,, \qquad w_1^2\,v' = w_1^2(a)\,v'(a)\,, \qquad v = v(a) + w_1^2(a)\,v'(a)\int_a^x \frac{\mathrm{d}\xi}{w_1^2(\xi)}\,,$$

where x belongs to I as well. Now, if it were $v'(a) = 0$, it would be $w_2 = v(a)\,w_1$, that is, the two solutions would be linearly dependent in contrast to the hypothesis. It follows that $v'(a) \neq 0$ and, as a consequence, v is monotonic in I because $w_1(a) \neq 0$ and the integrand $1/w_1^2$ is positive. Given these premises, it is found by inspection that $w_2 = v\,w_1$ fulfills $w_2'' + q\,w_2 = 0$. The Wronskian of the two solutions w_1, w_2 (Sect. A.12) reads $W = w_1\,w_2' - w_2\,w_1' = w_1^2(a)\,v'(a) \neq 0$.

By way of example, in the case of a free particle (Sect. 8.2.1) one of the two solutions of (8.1) can be written as $w_1 = \cos(k\,x)$; then, letting $I \equiv [-\pi/4, +\pi/4]$ and $a = 0$ it follows $v = v(0) + v'(0)\,\tan(k\,x)/k$. Although this is not relevant for the reasoning here, one may note that w_1 is a fundamental solution (Sect. 11.4); in order to make w_2 a fundamental solution as well, one lets $v(0) = 0$, $v'(0) = 1$ to find $w_2 = v\,w_1 = \sin(k\,x)/k$. The two solutions are linearly independent, and $W = 1$.

As a second example, consider the case of the energy well of Sect. 11.5; as shown in (11.47), one of the two solutions in the left region ($x < 0$) is $w_1 = \exp(\alpha\,x)$. Letting $I \equiv (-\infty, 0]$ and $a = 0$ one finds $v = v(0) + v'(0)\,\exp(-\alpha\,x)\,\sinh(\alpha\,x)/\alpha$; thus, $w_2 = v(0)\,\exp(\alpha\,x) + v'(0)\,\sinh(\alpha\,x)/\alpha$. If $v'(0) \neq 0$, such a solution is linearly independent from w_1; however, as $\sinh(\alpha\,x)$ diverges for $x \to -\infty$, the result is not acceptable from a physical point of view: one must then let $v'(0) = 0$ which, however, renders $w_2 = v(0)\,w_1$. Thus, imposing the additional prescription that w_2 should not diverge makes the two solutions linearly dependent, whence $W = 0$.

8.4 Let w_1 be the eigenfunction corresponding to E, namely, w_1 fulfills the equation $w_1''(x) + q(x)\,w_1(x) = 0$. It follows that $w_1''(-x) + q(-x)\,w_1(-x) = 0$ also holds. On the other hand, if V is even, q is also even; as a consequence,

$$w_1''(-x) + q(x)\,w_1(-x) = 0$$

is fulfilled with the same eigenvalue. Being the latter simple, it must be $w_1(-x) = a\,w_1(x)$, with a some nonvanishing constant. As this relation holds for any x, one may replace x with $-x$ to find $w_1(x) = a\,w_1(-x)$; multiplying by a both sides of the latter yields $a^2\,w_1(-x) = a\,w_1(x) = w_1(-x)$, whence $a = \pm 1$ and $w_1(-x) = \pm w_1(x)$. An example of this property is found from the analysis of the linear harmonic oscillator (Sect. 12.2).

Problems of Chap. 9

9.1 Inserting the expression of c_k into the one-dimensional form of (9.26) yields

$$A\,(x - u\,t\,; k_0) = \frac{\sqrt{\sigma/2}}{\pi^{3/4}} \int_{-\infty}^{+\infty} \exp\left[\mathrm{i}\,(x - u\,t)\,(k - k_0) - \sigma^2\,k^2/2\right] \mathrm{d}k\,.$$

Following the same procedure as in Sect. C.8 one finds

$$\mathrm{i}\,(x - u\,t)\,(k - k_0) - \frac{1}{2}\,\sigma^2\,k^2 = -\frac{(x - u\,t)^2}{2\,\sigma^2} - \frac{\sigma^2}{2}\,\left(k - j\,\frac{x - u\,t}{\sigma^2}\right)^2\,,$$

whence

$$A\,(x-u\,t\,;k_0)=\frac{1}{\pi^{1/4}\,\sqrt{\sigma}}\,\exp\left[-\mathrm{i}\,k_0\;(x-u\,t)-\frac{(x-u\,t)^2}{2\,\sigma^2}\right].$$

The particle's localization is determined by

$$|A(x-u\,t)|^2=\frac{1}{\sqrt{\pi}\,\sigma}\,\exp\left[-\frac{(x-u\,t)^2}{\sigma^2}\right].$$

Using again the results of Sect. C.8 yields $||A||=1$.

9.2 Remembering that $|\psi|^2=|A|^2$, the one-dimensional form of (9.23) reads

$$x_0(t)=\int_{-\infty}^{+\infty}x\,|A|^2\,\mathrm{d}x=\int_{-\infty}^{+\infty}(x-u\,t)\,|A|^2\,\mathrm{d}x+u\,t\int_{-\infty}^{+\infty}|A|^2\,\mathrm{d}x\,.$$

Letting $s=x-u\,t$ one finds that the integral of $s\,|A(s)|^2$ vanishes because the integrand is odd and the integration domain is symmetric with respect to the origin. Using the result $||A||=1$ of Prob. 9.1 then yields $x_0(t)=u\,t$.

Problems of Chap. 10

10.1 To determine the time evolution of the expectation value of the wavepacket for a free particle one starts from the general expression (9.5), with $w_{\mathbf{k}}(\mathbf{r})$ and $E_{\mathbf{k}}=\hbar\,\omega_{\mathbf{k}}$ given by (9.22), and $c_{\mathbf{k}}$ given by the second relation in (9.7). The wave function is assumed normalized, $\int_{-\infty}^{+\infty}|\psi(\mathbf{r},t)|^2\,\mathrm{d}^3r=\int_{-\infty}^{+\infty}|c(\mathbf{k})|^2\,\mathrm{d}^3k=1$. Using the first spatial coordinate x_1 and defining $m_{\mathbf{k}}=c_{\mathbf{k}}\,\exp(-\mathrm{i}\,\omega_{\mathbf{k}}\,t)$, the following are of use: $x_1\,w_{\mathbf{k}}=-\mathrm{i}\,\partial w_{\mathbf{k}}/\partial k_1$, $x_1^2\,w_{\mathbf{k}}=-\partial^2 w_{\mathbf{k}}/\partial k_1^2$, and

$$-\int_{-\infty}^{+\infty}m_{\mathbf{k}}\,\frac{\partial w_{\mathbf{k}}}{\partial k_1}\,\mathrm{d}k_1=\int_{-\infty}^{+\infty}w_{\mathbf{k}}\,\frac{\partial m_{\mathbf{k}}}{\partial k_1}\,\mathrm{d}k_1\,,\quad\int_{-\infty}^{+\infty}m_{\mathbf{k}}\,\frac{\partial^2 w_{\mathbf{k}}}{\partial k_1^2}\,\mathrm{d}k_1=\int_{-\infty}^{+\infty}w_{\mathbf{k}}\,\frac{\partial^2 m_{\mathbf{k}}}{\partial k_1^2}\,\mathrm{d}k_1\,,$$

where the last two equalities are obtained by integrating by parts and observing that, due to the normalization condition, $c_{\mathbf{k}}$ and $\partial c_{\mathbf{k}}/\partial k_1$ vanish at infinity. In turn,

$$\mathrm{i}\,\frac{\partial m_{\mathbf{k}}}{\partial k_1}=\left(u_1\,t\,c_{\mathbf{k}}+\mathrm{i}\,\frac{\partial c_{\mathbf{k}}}{\partial k_1}\right)\exp(-\mathrm{i}\,\omega_{\mathbf{k}}\,t)\,,\qquad u_1=\frac{\partial\omega_{\mathbf{k}}}{\partial k_1}=\frac{\hbar\,k_1}{m}\,,$$

$$-\frac{\partial^2 m_{\mathbf{k}}}{\partial k_1^2}=\left[\left(u_1^2\,t^2+\mathrm{i}\,\frac{\hbar}{m}\,t\right)c_{\mathbf{k}}+2\,\mathrm{i}\,u_1\,t\,\frac{\partial c_{\mathbf{k}}}{\partial k_1}-\frac{\partial^2 c_{\mathbf{k}}}{\partial k_1^2}\right]\exp(-\mathrm{i}\,\omega_{\mathbf{k}}\,t)\,.$$

The expectation value $\langle x_1 \rangle = \langle \psi | x_1 | \psi \rangle$ involves an integration over $\mathbf{r}$ to calculate the scalar product, an integration over $\mathbf{k}$ to calculate the integral expression of ψ, and an integration over $\mathbf{k}'$ to calculate the integral expression of ψ^*. Performing the integration over $\mathbf{r}$ first, letting $c'_{\mathbf{k}} = c(\mathbf{k}')$, $\omega'_{\mathbf{k}} = \omega(\mathbf{k}')$, and using (C.56) yields

$$\langle x_1 \rangle = \iiint_{-\infty}^{+\infty} \left(u_1 \, t \, |c_{\mathbf{k}}|^2 + \mathrm{i} \, c_{\mathbf{k}}^* \frac{\partial c_{\mathbf{k}}}{\partial k_1} \right) \mathrm{d}^3 k \, .$$

Letting $c_{\mathbf{k}} = a_{\mathbf{k}} + \mathrm{i} \, b_{\mathbf{k}}$, with $a_{\mathbf{k}}$ and $b_{\mathbf{k}}$ real, and using the asymptotic vanishing of $c_{\mathbf{k}}$, one finds

$$\iiint_{-\infty}^{+\infty} \mathrm{i} \, c_{\mathbf{k}}^* \frac{\partial c_{\mathbf{k}}}{\partial k_1} \, \mathrm{d}^3 k = x_{01} \, , \qquad x_{01} = \iiint_{-\infty}^{+\infty} \left(\frac{\partial a_{\mathbf{k}}}{\partial k_1} \, b_{\mathbf{k}} - \frac{\partial b_{\mathbf{k}}}{\partial k_1} \, a_{\mathbf{k}} \right) \mathrm{d}^3 k \, ,$$

where x_{01} is a real constant. Repeating the calculation for x_2 and x_3, and letting $\mathbf{u} = \mathrm{grad}_{\mathbf{k}} \omega$, $\mathbf{r}_0 = (x_{01}, x_{02}, x_{03})$, finally yields

$$\langle \mathbf{r} \rangle = \mathbf{r}_0 + \iiint_{-\infty}^{+\infty} \mathbf{u} \, t \, |c_{\mathbf{k}}|^2 \, \mathrm{d}^3 k \, , \qquad \frac{\mathrm{d}}{\mathrm{d}t} \langle \mathbf{r} \rangle = \iiint_{-\infty}^{+\infty} \mathbf{u} \, |c_{\mathbf{k}}|^2 \, \mathrm{d}^3 k = \mathrm{const} \, .$$

If $|c_{\mathbf{k}}|^2$ is even with respect to all components of $\mathbf{k}$, the expectation value of $\mathbf{r}$ does not change with respect to the initial value $\mathbf{r}_0$. Otherwise, it moves at constant speed.

10.2 The time evolution of the standard deviation of position is found following the same line and using the same symbols and relations as in Prob. 10.1, starting with

$$\langle x_1^2 \rangle = \iiint_{-\infty}^{+\infty} \left[\left(u_1^2 \, t^2 + \mathrm{i} \, \frac{\hbar}{m} \, t \right) |c_{\mathbf{k}}|^2 + 2\mathrm{i} \, u_1 \, t \, c_{\mathbf{k}}^* \frac{\partial c_{\mathbf{k}}}{\partial k_1} - c_{\mathbf{k}}^* \frac{\partial^2 c_{\mathbf{k}}}{\partial k_1^2} \right] \mathrm{d}^3 k \, .$$

An integration by parts combined with the normalization condition for $c_{\mathbf{k}}$ shows that

$$\iiint_{-\infty}^{+\infty} 2\mathrm{i} \, u_1 \, t \, c_{\mathbf{k}}^* \frac{\partial c_{\mathbf{k}}}{\partial k_1} \, \mathrm{d}^3 k = -\mathrm{i} \, \frac{\hbar}{m} \, t + 2 \, t \iiint_{-\infty}^{+\infty} u_1 \left(\frac{\partial a_{\mathbf{k}}}{\partial k_1} \, b_{\mathbf{k}} - \frac{\partial b_{\mathbf{k}}}{\partial k_1} \, a_{\mathbf{k}} \right) \mathrm{d}^3 k \, ,$$

where the second term at the right-hand side is real, whereas the first one cancels out in the expression of $\langle x_1^2 \rangle$. Finally, another integration by parts yields

$$- \iiint_{-\infty}^{+\infty} c_{\mathbf{k}}^* \frac{\partial^2 c_{\mathbf{k}}}{\partial k_1^2} \, \mathrm{d}^3 k = \iiint_{-\infty}^{+\infty} \left| \frac{\partial c_{\mathbf{k}}}{\partial k_1} \right|^2 \mathrm{d}^3 k \, .$$

In conclusion,

$$\langle x_1^2 \rangle = \iiint_{-\infty}^{+\infty} \left| u_1 \, t \, c_{\mathbf{k}} + \mathrm{i} \, \frac{\partial c_{\mathbf{k}}}{\partial k_1} \right|^2 \mathrm{d}^3 k \, .$$

Repeating the calculation for x_2 and x_3 yields

$$\langle \mathbf{r}\cdot\mathbf{r}\rangle = \int_{-\infty}^{+\infty} |\mathbf{u}\, t\, c_{\mathbf{k}} + \mathrm{i}\,\mathrm{grad}_{\mathbf{k}} c_{\mathbf{k}}|^2 \,\mathrm{d}^3k\,,$$

where the definition of the squared length of a complex vector is found in (A.2) and (A.4). The standard deviation of the wave packet in the $\mathbf{r}$ space is the positive square root of $\langle \mathbf{r}\cdot\mathbf{r}\rangle - \langle \mathbf{r}\rangle\cdot\langle \mathbf{r}\rangle = \sum_{i=1}^{3}(\Delta x_i)^2$, where the expression of $\langle \mathbf{r}\rangle$ was derived in Prob. 10.1. It is easily shown that the standard deviation diverges with t. In fact, the leading term of $\langle x_1^2\rangle$ and, respectively, $\langle x_1\rangle^2$ is

$$\langle x_1^2\rangle \sim t^2 \iiint_{-\infty}^{+\infty} u_1^2\,|c_{\mathbf{k}}|^2\,\mathrm{d}^3k\,, \qquad \langle x_1\rangle^2 \sim t^2 \left(\iiint_{-\infty}^{+\infty} u_1\,|c_{\mathbf{k}}|^2\,\mathrm{d}^3k\right)^2,$$

the first of which is positive, whereas the second one is nonnegative. Letting $f = c_{\mathbf{k}}$, $g = u_1 c_{\mathbf{k}}$ in the Schwartz inequality (8.15) and using the normalization condition of $c_{\mathbf{k}}$ yields

$$\iiint_{-\infty}^{+\infty} u_1^2 |c_{\mathbf{k}}|^2\,\mathrm{d}^3k > \left(\iiint_{-\infty}^{+\infty} u_1 |c_{\mathbf{k}}|^2\,\mathrm{d}^3k\right)^2,$$

where the strict inequality holds because f and g are not proportional to each other. For the leading term it follows that $(\Delta x_1)^2 = \langle x_1^2\rangle - \langle x_1\rangle^2 \sim \text{const}\times t^2$, where the constant is strictly positive. The same reasoning applies to x_2, x_3. In conclusion, the standard deviation Δx_i associated with the ith coordinate diverges in time with the first power of t.

10.3 Still with reference to the wave packet of a free particle used in Probs. 10.1 and 10.2, the time evolution of the expectation value in the $\mathbf{p}$ space is found starting with the first component p_1 of momentum. The corresponding operator is $\hat{p}_1 = -\mathrm{i}\,\hbar\,\partial/\partial x_1$, and the following relations are of use: $\hat{p}_1 w_{\mathbf{k}} = \hbar\, k_1\, w_{\mathbf{k}}$, $\hat{p}_1^2 w_{\mathbf{k}} = \hbar^2\, k_1^2\, w_{\mathbf{k}}$. The expectation value $\langle p_1\rangle = \langle \psi | p_1 | \psi\rangle$ involves an integration over $\mathbf{r}$ to calculate the scalar product, an integration over $\mathbf{k}$ to calculate the integral expression of ψ, and an integration over $\mathbf{k}'$ to calculate the integral expression of ψ^*. Performing the integration over $\mathbf{r}$ first, letting $c'_{\mathbf{k}} = c(\mathbf{k}')$, $\omega'_{\mathbf{k}} = \omega(\mathbf{k}')$, and using (C.56) yields

$$\langle p_1\rangle = \iiint_{-\infty}^{+\infty} \hbar\, k_1\, |c_{\mathbf{k}}|^2\,\mathrm{d}^3k = p_{01}\,.$$

The real constant p_{01} defined above is independent of time. In conclusion, repeating the calculation for p_2 and p_3, and letting $\mathbf{p}_0 = (p_{01}, p_{02}, p_{03})$, the following holds: $\langle \mathbf{p}\rangle = \mathbf{p}_0$. If $|c_{\mathbf{k}}|^2$ is even with respect to all components of $\mathbf{k}$, the expectation value of $\mathbf{p}$ is zero.

10.4 The calculation of $\langle \psi | \hat{p}_1^2 | \psi\rangle$ is carried out following the same line as in Prob. 10.3, leading to

$$\langle p_1^2 \rangle = \iiint_{-\infty}^{+\infty} \hbar^2 \, k_1^2 \, |c_{\mathbf{k}}|^2 \, \mathrm{d}^3 k \, .$$

Repeating the calculation for x_2 and x_3 yields

$$\langle \mathbf{p} \cdot \mathbf{p} \rangle = \iiint_{-\infty}^{+\infty} \hbar \, \mathbf{k} \cdot \hbar \, \mathbf{k} \, |c_{\mathbf{k}}|^2 \, \mathrm{d}^3 k \, .$$

In turn, the standard deviation of the wave packet in the $\mathbf{p}$ space is the positive square root of $\langle \mathbf{p} \cdot \mathbf{p} \rangle - \langle \mathbf{p} \rangle \cdot \langle \mathbf{p} \rangle = \sum_{i=1}^{3} (\Delta p_i)^2$. Letting $f = c_{\mathbf{k}}$, $g = \hbar k_1 c_{\mathbf{k}}$ in the Schwartz inequality (8.15) and using the normalization condition of $c_{\mathbf{k}}$ yields

$$\iiint_{-\infty}^{+\infty} \hbar^2 \, k_1^2 \, |c_{\mathbf{k}}|^2 \, \mathrm{d}^3 k > \left(\iiint_{-\infty}^{+\infty} \hbar \, k_1 \, |c_{\mathbf{k}}|^2 \, \mathrm{d}^3 k \right)^2 ,$$

where the strict inequality holds because f and g are not proportional to each other. It follows that $(\Delta p_1)^2 = \langle p_1^2 \rangle - \langle p_1 \rangle^2$ is strictly positive and constant in time. The same reasoning applies to p_2, p_3. In conclusion, the standard deviation Δp_i associated with the ith component of momentum is constant in time.

10.5 One finds $\langle x \rangle = x_0$, $\mathrm{d}\hbar\,\beta/\mathrm{d}x = \hbar \, k_0$, $\langle p_e \rangle = \hbar \, k_0$,

$$\left\langle \frac{p_e^2}{2\,m} \right\rangle = \frac{\hbar^2 \, k_0^2}{2\,m} \, , \qquad \langle Q \rangle = \frac{\hbar^2}{8\,m\,\sigma^2} \, , \qquad \langle T \rangle = \frac{\hbar^2}{2\,m} \left(k_0^2 + \frac{1}{4\,\sigma^2} \right) .$$

One notes that for a fixed $\langle T \rangle$ all nonnegative values of the "convective" and "thermal" parts that add up to $\langle T \rangle$ are allowed. In the particular case of a free particle, where $\langle T \rangle = \langle E \rangle$, the above shows that different values of the average momentum and "dispersion" may combine to yield the same total energy.

Problems of Chap. 11

11.1 Letting $b^- = \pi \, a \sqrt{2\,m} \, (\sqrt{E} - \sqrt{E - V_0})/\hbar$, $b^+ = \pi \, a \sqrt{2\,m} \, (\sqrt{E} + \sqrt{E - V_0})/\hbar$ and remembering that $\sinh b \simeq b$ when $|b| \ll 1$ yields, with m fixed,

$$R(a \to 0) = \left(\frac{b^-}{b^+} \right)^2 = \frac{(\sqrt{E} - \sqrt{E - V_0})^2}{(\sqrt{E} + \sqrt{E - V_0})^2} \, ,$$

that coincides with the first relation in (11.11). Conversely, when $a > 0$ is fixed and m is let grow one finds

$$R \simeq \exp[2\,(b^- - 2\,b^+)] = \exp\left(-4\,\pi\,a\,\sqrt{2\,m}\,\sqrt{E - V_0}/\hbar\right)\,,$$

namely, $\lim_{m\to\infty} R = 0$, thus recovering the classical limit.

11.2 The maximum of the cotangent's argument $s\,\sqrt{2\,m\,(E - V_0)/\hbar^2}$ is found by letting $E = 0$. It is found

$$\gamma = \frac{s}{\hbar}\,\sqrt{-2\,m\,V_0} \simeq 13.4\,, \qquad \frac{13.4}{\pi} \simeq 4.3\,.$$

As a consequence, the cotangent has four complete branches and one incomplete branch in the interval $V_0 < E < 0$, corresponding to five eigenvalues $E_1, \ldots, E_5$. Using the normalized parameter $0 < \eta = \sqrt{1 - E/V_0} < 1$, the equation to be solved reads

$$\frac{\eta^2 - 1/2}{\eta\,\sqrt{1 - \eta^2}} = \cot\,(\gamma\,\eta)\ .$$

Over the η axis, the 5 branches belong to the intervals $(0, \pi/\gamma)$, $(\pi/\gamma, 2\pi/\gamma)$, $(2\pi/\gamma, 3\pi/\gamma)$, $(3\pi/\gamma, 4\pi/\gamma)$, $(4\pi/\gamma, 1)$.

Problems of Chap. 13

13.1 Letting $Z = 1$ one finds that the lowest total energy of the electron in the hydrogen atom has the value

$$E_1(Z = 1) = -\frac{m_0}{2\,\hbar^2}\left(\frac{q^2}{4\,\pi\,\varepsilon_0}\right)^2\,.$$

As noted in Sect. 13.5.2, the electron is bound as long as $E < 0$. As a consequence, the minimum energy for which it becomes free is $\lim_{n\to\infty} E_n = 0$. The hydrogen atom's ionization energy is thus found to be

$$E_{\rm ion} = 0 - E_1(Z = 1) = |E_1(Z = 1)| = \frac{m_0}{2\,\hbar^2}\left(\frac{q^2}{4\,\pi\,\varepsilon_0}\right)^2\,.$$

Replacing the constants' values of Table D.1 yields $E_{\rm ion} \simeq 2.18 \times 10^{-18}$ J $\simeq 13.6$ eV.

13.2 The time-dependent wave function is in this case $\psi = w(E_{\rm min})\,\exp(-{\rm i}\,E_{\rm min}\,t/\hbar)$, whence $|\psi|^2 = \exp(-2\,r/a)/(\pi\,a^3)$. Taking the Jacobian determinant $J = r^2\,\sin\vartheta$ from (B.3) and using the definitions of Sect. 10.5 one finds

$$\langle r \rangle = \int_0^\infty \int_0^\pi \int_0^{2\pi} r \, \frac{\exp(-2\,r/a)}{\pi\, a^3} \, r^2 \, \sin\vartheta \, \mathrm{d}r \, \mathrm{d}\vartheta \, \mathrm{d}\varphi = \frac{3}{2}\, a \, .$$

From (13.96) one finds $a_1 = a(Z = 1) = 4\,\pi\,\hbar^2\,\varepsilon_0/(m_0\,q^2) \simeq 5.3 \times 10^{-11}$ m $\simeq$ 0.53 Å, where the constants' values are taken from Table D.1. Note that $a_1 = r_1/2$, with r_1 the radius of the ground state derived from the Bohr hypothesis (Sect. 7.4.4). The expectation value of r turns out to be $\langle r \rangle \simeq 0.8$ Å.

13.3 The time-independent Schrödinger equation to be discretized is (11.28); it reads

$$\frac{\mathrm{d}^2 w}{\mathrm{d}x^2} + q\,w = 0\,, \qquad q(x) = \frac{2\,m}{\hbar^2}\,(E - V)\,.$$

The symbols of node, nodal value, element, and cell are those introduced in Sects. 13.6.6 and A.13.2. The elements are taken equal to each other, $h_1 = \ldots = h_{N+1} = h$. Expanding into a series the unknown function w at x_i, using h as increment, and combining the result with $w'' = -q\,w$, provides

$$w_{i+1} = w_i + w_i'\,h - q_i\,w_i\,\frac{h^2}{2} - (q_i\,w_i)'\,\frac{h^3}{6} - (q_i\,w_i)''\,\frac{h^4}{24} \quad \cdots$$

Then, one expresses w_{i-1} using $-h$ as increment, thus yielding a new series whose odd-degree terms have the opposite sign with respect to those of the first series. Adding up the two series, and leaving out the terms with the derivatives of w of the sixth order or higher, yields

$$2\,w_i - w_{i+1} - w_{i-1} \simeq q_i\,w_i\,h^2 + (q_i\,w_i)''\,\frac{h^4}{12}\,.$$

The second derivative of the above, after leaving out the term with the sixth derivative of w and using again $w'' = -q\,w$, reads

$$-2\,q_i\,w_i + q_{i+1}\,w_{i+1} + q_{i-1}\,w_{i-1} \simeq (q_i\,w_i)''\,h^2\,.$$

One then combines the last two equations after multiplying both sides of the second one by $h^2/12$; this results in the elimination of $(q_i\,w_i)''$, whence

$$-\left(1 + \frac{h^2}{12}\,q_{i-1}\right) w_{i-1} + \left(2 - \frac{10\,h^2}{12}\,q_i\right) w_i - \left(1 + \frac{h^2}{12}\,q_{i+1}\right) w_{i+1} = 0\,.$$

Problems of Chap. 14

14.1 From $h^{(0)}_{\mathbf{bg}} = [A/(2\,\pi)^3]/(q_c^2 + q^2)$ and $q = |\mathbf{b} - \mathbf{g}|$ one finds

$$H^{(0)}_{\mathbf{b}}(E_{\mathbf{g}}) = \frac{A^2}{(2\,\pi)^6} \int_0^{\pi} \int_0^{2\,\pi} \frac{1}{(q_c^2 + q^2)^2} \sin\vartheta \, \mathrm{d}\vartheta \, \mathrm{d}\varphi \,,$$

$A = \kappa\, Z\, e^2/\varepsilon_0$. Observing that $\mathbf{b}$ is a fixed vector one can use it as the reference for angle ϑ, so that $q^2 = (\mathbf{b} - \mathbf{g}) \cdot (\mathbf{b} - \mathbf{g}) = b^2 + g^2 - 2\,b\,g\,\cos\vartheta$. From $g = b$ it follows $q^2 = 4\,g^2\,\sin^2(\vartheta/2)$. On the other hand, it is $\sin\vartheta\,\mathrm{d}\vartheta = \mathrm{d}\sin^2(\vartheta/2)$ whence, integrating over φ and letting $\mu = \sin^2(\vartheta/2)$,

$$H^{(0)}_{\mathbf{b}}(E_{\mathbf{g}}) = \frac{A^2}{(2\,\pi)^5} \int_0^1 \frac{\mathrm{d}\mu}{(q_c^2 + 4\,g^2\,\mu)^2} = \frac{A^2/(2\,\pi)^5}{q_c^2\,(q_c^2 + 4\,g^2)} \,.$$

The dependence on $E_{\mathbf{g}}$ is found through the relation $E_{\mathbf{g}} = \hbar^2\, g^2/(2\,m)$.

14.2 The first term in brackets in (14.51) transforms as

$$|a_{\mathbf{k}\,\mathbf{k}'}|^2\, \delta(E_{\mathbf{k}} - E_{\mathbf{k}'} - \hbar\,\omega_{\mathbf{q}}) \;\; \rightarrow \;\; |a_{\mathbf{k}'\,\mathbf{k}}|^2\, \delta\left[-(E_{\mathbf{k}} - E_{\mathbf{k}'} + \hbar\,\omega_{\mathbf{q}})\right] \,,$$

while the second one transforms as

$$|a_{\mathbf{k}'\,\mathbf{k}}|^2\, \delta(E_{\mathbf{k}} - E_{\mathbf{k}'} + \hbar\,\omega_{\mathbf{q}}) \;\; \rightarrow \;\; |a_{\mathbf{k}\,\mathbf{k}'}|^2\, \delta\left[-(E_{\mathbf{k}} - E_{\mathbf{k}'} - \hbar\,\omega_{\mathbf{q}})\right] \,.$$

As δ is invariant upon change of sign of its argument, the negative sign before the parenthesis in the two relations above can be eliminated. This shows that the two summands in (14.51) merely exchange places upon exchange of the indices, so that the whole sum is invariant. In conclusion, (14.51) is invariant upon exchange of the indices $\mathbf{k}$ and $\mathbf{k}'$. One notes in passing that the coefficient $|a_{\mathbf{k}\,\mathbf{k}'}|^2$ is also invariant, due to the hermiticity of $\mathscr{H}$.

14.3 Due to the form of the exponential, the summands are made of the product of three independent factors:

$$\sum_{m=0}^{N_c-1} \exp[\mathrm{i}\,(\mathbf{k} - \mathbf{k}' + \mathbf{k}_d) \cdot \mathbf{l}_m] = \sum_{m_1=0}^{N_1-1} \exp(2\,\pi\,\mathrm{i}\,\gamma_1\,m_1) \sum_{m_2=0}^{N_2-1} \ldots \sum_{m_3=0}^{N_3-1} \exp(2\,\pi\,\mathrm{i}\,\gamma_3\,m_3) \,.$$

Each factor in the above is recast as

$$\sum_{m_i=0}^{N_i-1} \exp(2\,\pi\,\mathrm{i}\,\gamma_i\,m_i) = \exp[\pi\,\mathrm{i}\,\gamma_i\,(N_i - 1)]\, \frac{\sin(\pi\,\gamma_i\,N_i)}{\sin(\pi\,\gamma_i)} \,,$$

whence

$$\left| \sum_{m=0}^{N_c-1} \exp[\mathrm{i}\,(\mathbf{k}-\mathbf{k}'+\mathbf{k}_d)\cdot\mathbf{l}_m] \right|^2 = \prod_{s=1}^{3} \frac{\sin^2(\pi\,\gamma_s N_s)}{\sin^2(\pi\,\gamma_s)}\,.$$

14.4 From the solution to Prob. 14.3 one takes the relation

$$\eta = \sum_{m_i=0}^{N_i-1} \exp(2\,\pi\,\mathrm{i}\,\gamma_i\,m_i) = \exp[\pi\,\mathrm{i}\,\gamma_i\,(N_i-1)]\,\frac{\sin(\pi\,\gamma_i\,N_i)}{\sin(\pi\,\gamma_i)}\,.$$

It has already been shown in Prob. 14.3 that when N_i is large, the modulus of η is significant only when γ_i is integer. Now assume, first, that γ_i differs slightly from an even number, say, $\gamma_i = 2\,n+\varepsilon$, where n is an integer and $|\varepsilon| \ll 1$. It follows

$$\eta = \eta_e = \exp\,[\mathrm{i}\,\pi\,\varepsilon(N_i-1)]\,\frac{\sin(\pi\,\varepsilon\,N_i)}{\sin(\pi\,\varepsilon)}\,,$$

whence $\lim_{\varepsilon\to 0}\eta_e = N_i$. Similarly, if γ_i differs slightly from an odd number, say, $\gamma_i = 2\,n+1+\varepsilon$, it follows

$$\eta = \eta_o = \exp\,[\mathrm{i}\,\pi\,(\varepsilon+1)\,(N_i-1)]\,\frac{\sin[\pi\,(\varepsilon+1)\,N_i]}{\sin[\pi\,(\varepsilon+1)]}\,,$$

where the exponential may be written as $(-1)^{N_i-1}\,\exp[\mathrm{i}\,\pi\,\varepsilon\,(N_i-1)]$. In turn, after expanding its numerator and denominator, the fraction in the definition of η_o reads $(-1)^{N_i-1}\,\sin(\pi\,\varepsilon\,N_i)/\sin(\pi\,\varepsilon)$. The product of the last two expressions equals the right-hand side of the definition of η_e, so that $\lim_{\varepsilon\to 0} f_o = N_i$.

14.5 In this case $u_{\mathbf{k}}$ is constant; remembering that its square modulus is the inverse of a volume, it is sensible to let $u_{\mathbf{k}} = 1/\sqrt{\Omega}$, whence

$$Y_{\mathbf{k}\,\mathbf{k}'}(\mathbf{g}_\gamma, V=0) = \frac{1}{\Omega}\int_{\Omega_0} \exp(\mathrm{i}\,\mathbf{g}_\gamma\cdot\mathbf{r})\,\mathrm{d}^3 r\,.$$

The evaluation of the integral above is the same as that occurring in the theory of the electromagnetic field (Sect. C.15), the only difference being that the reference is not orthogonal:

$$\mathbf{r} = \mu_1\mathbf{a}_1 + \mu_2\mathbf{a}_2 + \mu_3\mathbf{a}_3\,,$$

with μ_i real numbers and $\mathrm{d}^3 r = \tau_l\,\mathrm{d}^3\mu$ (Sect. 17.3). As the integral does not depend on the form chosen for the cell, one may use for Ω_0 a prismatic cell whose sides coincide with $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, respectively, so that one vertex coincides with the origin. As a consequence, the limits of μ_i for the Ω_0 cell are 0 and 1. Then, remembering that $\mathbf{a}_i\cdot\mathbf{b}_j = \delta_{ij}$, one finds that $Y_{\mathbf{k}\,\mathbf{k}'}(\mathbf{g}_\gamma, V=0)$ is given by the factor $\tau_l/\Omega = 1/N_c$ multiplied by three integrals of the form

$$Y_i = \int_0^1 \exp(2\,\pi\,\mathrm{i}\,\gamma_i\,\mu_i)\,\mathrm{d}\mu_i = \exp(\pi\,\mathrm{i}\,\gamma_i)\,\frac{\sin(\pi\,\gamma_i)}{\pi\,\gamma_i}\,,$$

$i = 1, 2, 3$. As γ_i is an integer, the above function vanishes for every γ_i with the exception of $\gamma_i = 0$. In conclusion,

$$\begin{aligned} &G_{\mathbf{k}\mathbf{k}'}(\mathbf{g}_\gamma, V = 0) = 1/N_c^2\,, \qquad &&\mathbf{g}_\gamma = 0 \\ &G_{\mathbf{k}\mathbf{k}'}(\mathbf{g}_\gamma, V = 0) = 0\,, &&\mathbf{g}_\gamma \neq 0\,. \end{aligned}$$

Problems of Chap. 15

15.1 After selecting a number $0 < \eta < 1/2$, define E^+ and E^- such that $P(E^+) = \eta$, $P(E^-) = 1 - \eta$. It follows

$$E^+ - E^- = 2\,k_B\,T\,\log\frac{1-\eta}{\eta}\,.$$

Letting, e.g., $\eta = 0.1$ one finds $E(P = 0.1) - E(P = 0.9) = 2\,k_B\,T\log 9 \simeq 4.39\,k_B\,T$. Similarly, letting $\eta = 0.01$ one finds $E(P = 0.01) - E(P = 0.99) = 2\,k_B\,T\log 99 \simeq 9.19\,k_B\,T$. At $T = 300$ K it is $k_B\,T \simeq 25.8$ meV. From the above results one finds $E(P = 0.1) - E(P = 0.9) \simeq 113$ meV and $E(P = 0.01) - E(P = 0.99) \simeq 237$ meV, respectively.

15.2 From the qualitative standpoint one may discuss the problem by remembering that, from the findings of Sect. C.4, $P(E) - 1/2$ is odd with respect to $E - E_F$; as a consequence, a decrease in the occupation probability produced, by a temperature change, at $E = E_F - \Delta E$, corresponds to a probability increase by the same amount at $E = E_F + \Delta E$. If it were $g(E) = \text{const}$, a number of particles would relocate from one side of E_F to the other. However, $g(E)$ is not constant in general; it follows that the number of states available on the two sides of E_F is not the same and, when temperature changes, E_F must shift in order to compensate for this.

From the quantitative standpoint, and for the conditions specified by the problem, one may let $E_1 = 0$ and $E_U \to \infty$ in the first relation of (15.48). This yields [70]

$$N_S = \int_0^\infty g(E)\,P(E)\,\mathrm{d}E = -\int_0^\infty G(E)\,\frac{\mathrm{d}P}{\mathrm{d}E}\,\mathrm{d}E\,, \qquad G(E) = \int_0^E g(E')\,\mathrm{d}E'\,.$$

Function G is now expanded into a Taylor series around E_F to the second order, namely, $G(E) \simeq G(E_F) + g(E_F)\,(E - E_F) + (1/2)\,g'(E_F)\,(E - E_F)^2$. Inserting the zero-order term into the integral provides $G(E_F)\,P(E = 0) \simeq G(E_F)$. Then, using again the findings of Sect. C.4, the integral of the first-order term yields

$$\frac{g(E_F)}{k_B\,T}\int_0^\infty (E-E_F)\,P\,(1-P)\,\mathrm{d}E = 0\,,$$

because the integrand is odd[1] with respect to $E-E_F$. For the integral of the second-order term one finds, letting $E-E_F = k_B\,T\,\eta$ and $p(\eta) = 1/[\exp(\eta)+1]$,

$$\frac{g'(E_F)}{2\,k_B\,T}\int_0^\infty (E-E_F)^2\,P\,(1-P)\,\mathrm{d}E = -g'(E_F)\,(k_B\,T)^2\int_{\eta_F}^\infty \frac{1}{2}\,\eta^2\,p'(\eta)\,\mathrm{d}\eta\,,$$

with $\eta_F = -E_F/(k_B\,T)$. In the same order of approximation as for $P(E=0) \simeq 1$ one replaces η_F with $-\infty$; then, observing that the integrand is even with respect to $\eta = 0$, integrating by parts, and using the first relation in (C.141), one finds

$$-\int_{-\infty}^\infty \frac{1}{2}\,\eta^2\,p'(\eta)\,\mathrm{d}\eta = -\int_0^\infty \eta^2\,p'(\eta)\,\mathrm{d}\eta = 2\int_0^\infty \eta\,p(\eta)\,\mathrm{d}\eta = \frac{\pi^2}{6}\,.$$

Collecting the results yields $N_S \simeq G(E_F) + (\pi^2/6)\,g'(E_F)\,(k_B\,T)^2$. As expected, the relation between E_F and T depends on the form of the density of states g: if it were $g' = 0$, E_F would not depend on temperature. Note that the integral involving the third derivative of G also vanishes, still because the integrand is odd; it follows that the first term left out in this approximation is of the fourth order in $E-E_F$. In the $T \to 0$ limit the above relation becomes $N_S \simeq G(E_{F0})$, with $E_{F0} = E_F(T=0)$; equating the two expressions of N_S and using the definition of G provides

$$\int_{E_{F0}}^{E_F} g(E)\,\mathrm{d}E \simeq -\frac{\pi^2}{6}\,g'(E_F)\,(k_B\,T)^2\,.$$

If T is close to zero the above becomes

$$E_F \simeq E_{F0} - \frac{\pi^2}{6}\,\frac{g'(E_{F0})}{g(E_{F0})}\,(k_B\,T)^2\,,$$

showing that if $g'(E_{F0}) > 0$ (which is the typical case), E_F decreases when temperature increases.

15.3 Object of the combinatorial problem is placing N_r particles, not subjected to the exclusion principle, into g_r states. The particles are not distinguishable. Since the exclusion principle does not apply, both cases $g_r \geq N_r$ and $g_r < N_r$ are possible. They are considered below, with the aid of tables whose columns are labeled by numbers that refer to the particles; the entries of each line are letters that indicate

[1] In principle, the statement would be correct if the integral was started from $-\infty$. It is approximately true due to the condition $P(E=0) \simeq 1$.

the states. The left table shows the placement of two particles, labeled 1 and 2, into three states, labeled A, B, and C; in the right table, four particles, labeled from 1 to 4, are placed into the same three states.

1	2
A	A
A	B
A	C
B	B
B	C
C	C

1	2	3	4
A	A	A	A
A	A	A	B
·	·	·	·
·	·	·	·
B	C	C	C
C	C	C	C

The state labels may be repeated within the same line because the exclusion principle does not apply; e.g., the first line of the left table shows that in this placement both particles 1 and 2 belong to state A. In each line the state labels are ordered in ascending order because, due to indistinguishability, permutations of these labels do not introduce new placements. It is found by inspection that the number of placements (lines) of each table equals the number C_n^k of k-combinations with repetitions of n objects, where $n = g_r$ is the number of states and $k = N_r$ is the number of particles. In particular it is $C_n^1 = n$.

To find C_n^k it is necessary to determine its relation with C_n^{k-1}. To this purpose, one starts from the total number of states $k\,C_n^k$ appearing in the table, and observes that each state appears the same number of times; thus, the number of appearances of single state, say, A, is $k\,C_n^k/n$. Now, one selects in the table only the placements where state A appears, and suppresses A once in each placement; this yields the $(k-1)$-combinations of n elements, whose number is C_n^{k-1}. The number of appearances of A in the latter is, by the same token, $(k-1)\,C_n^{k-1}/n$; as a consequence, the number of appearances of A in the original table (before suppression) can be calculated in an alternative way as

$$C_n^{k-1} + \frac{k-1}{n}\,C_n^{k-1} = \frac{n+k-1}{n}\,C_n^{k-1}\,.$$

Equating the latter to $k\,C_n^k/n$ provides $C_n^k = [(n+k-1)/k]\,C_n^{k-1}$; iterating the above as $C_n^{k-1} = [(n+k-2)/(k-1)]\,C_n^{k-2}$, down to $C_n^2 = [(n+1)/2]\,C_n^1$ and $C_n^1 = n$, and multiplying term-by-term the relations thus found, eventually yield

$$C_n^k = \frac{(n+k-1)\,(n+k-2)\cdots(n+1)\,n}{k\,(k-1)\cdots 1} = \binom{n+k-1}{k} = \binom{g_r+N_r-1}{N_r}\,.$$

Problems of Chap. 18

18.1 The derivation of (18.21) follows the same line as that of (15.49). The starting point is determining the number of way of placing the electrons into the available states, subjected to the constraints of exclusion and indistinguishability. The states of the bands are the same as for an undoped semiconductor; for such states the number of placements $W = W_1\, W_2\, \ldots\, W_r\, \ldots$ still applies, where W_r is the number of ways in which N_r particles can be placed into the g_r states of the rth energy interval; its value is given by (15.43). In a doped semiconductor one must then consider the additional states associated with the dopants. Considering an n-doped semiconductor where a constant concentration N_D of donors is present, the total number of donor atoms is $K_D = \Omega\, N_D$, with Ω the volume of the crystal. In each donor atom, the states able to accommodate the electron are those not involved in the formation of the covalent bonds with the semiconductor atoms; thus, if the available orbitals are f_D, the available states are $d_D = 2 f_D$ due to spin.

Given these premises, to the placements W considered above one must add the placements of a number of electrons, say, k_D, into the donor atoms; the latter counting, however, must be carried out differently: in each group of d_D states it is possible in fact to accommodate one electron at most, while in each group of g_r states of the bands it is possible to accommodate more than one electron. Let W_D be the number of ways to place k_D electrons into the K_D groups of states, with the provision that at most one electron can be placed into each group ($k_D \le K_D$), and that a given group has d_D states able to accommodate the electron. If it were $d_D = 1$, the outcome would be $W'_D = K_D!/[k_D!\,(K_D - k_D)!]$; in fact, the counting would be identical to that leading to (15.49). As $d_D > 1$, there are d_D ways to place the first electron into one of the dopant atoms; for each of the above placements, there are d_D to place the second electron into another atom, and so on. In conclusion one finds

$$W_D = d_D^{k_D}\, W'_D = d_D^{k_D}\, \frac{K_D!}{k_D!\,(K_D - k_D)!}\,,$$

and the total number of possible placements is $W_D\, W$. Given the constraints (compare with (15.38), (15.38))

$$F_E = E_S - k_D\, E_D - \sum_i N_i\, E_i\,, \qquad F_N = N - k_D - \sum_i N_i\,,$$

one maximizes (compare with (15.40))

$$F(N_1, N_2, \ldots, k_D, \alpha, \beta) = \log W + \log W_D + \alpha\, F_N + \beta\, F_E\,.$$

The treatment of W is identical to that of Sect. 15.8.1; that of W_D yields $\mathrm{d} \log W_D / \mathrm{d}k_D = \alpha + \beta\, E_D$, whence

$$\alpha + \beta\, E_D = \frac{\mathrm{d}}{\mathrm{d}k_D}\left[d_D^{k_D}\,\frac{K_D!}{k_D!\,(K_D-k_D)!}\right] \simeq \log d_D + \log\left(\frac{K_D}{k_D}-1\right).$$

Taking the exponential of both sides yields (18.21).

Problems of Chap. 19

19.1 Using (19.115) and adding up the two expressions in (19.118) one finds

$$\mu_p = q\left(\frac{m_{hh}^{1/2}}{m_{hh}^{3/2}+m_{hl}^{3/2}} + \frac{m_{hl}^{1/2}}{m_{hh}^{3/2}+m_{hl}^{3/2}}\right)\tau_p = \frac{q\,\tau_p}{\overline{m}_h}.$$

Using the values taken from Table 17.3 yields

$$\frac{m_0}{\overline{m}_h} = \frac{0.5^{1/2}}{0.5^{3/2}+0.16^{3/2}} + \frac{0.16^{1/2}}{0.5^{3/2}+0.16^{3/2}},$$

whence $\overline{m}_h \simeq 0.377\, m_0$. As for a_p, using the common value of the relaxation time in (19.122) yields

$$a_p = \frac{q\,\tau_p}{\mu_p^2}\left(\frac{\mu_{ph}}{m_{hh}} + \frac{\mu_{pl}}{m_{hl}}\right).$$

Replacing the expressions (19.118) of μ_{ph}, μ_{pl},

$$a_p = \frac{m_{hh}^{3/2}+m_{hl}^{3/2}}{m_{hh}^{1/2}\,m_{hl}^{1/2}\left(m_{hh}^{1/2}+m_{hl}^{1/2}\right)} = \frac{0.5^{3/2}+0.16^{3/2}}{0.5^{1/2}\,0.16^{1/2}\left(0.5^{1/2}+0.16^{1/2}\right)} \simeq 1.33.$$

Problems of Chap. 21

21.1 The equilibrium condition is considered. In the n region of the device, at a sufficiently large distance from the junction, due to the spatial uniformity of the material the charge neutrality condition (18.23) holds. Using the form (18.24) of the latter, one finds

$$N_C\,\Phi_{1/2}(\xi_e) = N_V\,\Phi_{1/2}(\xi_h) + \frac{N_D}{d_D\,\exp(\xi_e-\xi_D)+1},$$

where $-\xi_D = (E_C - E_D)/(k_B\,T) > 0$ is the normalized distance between the edge of the conduction band and the ground level of the donor atoms. In turn, ξ_e and

ξ_h are given by (18.56) and (18.57), respectively. It follows that when the above equation is coupled with the relation $\xi_e + \xi_h = -E_G/(k_B\,T)$ to form a system in the two unknowns ξ_e, ξ_h, the solution determines the position of $E_F + q\,\varphi(+\infty)$ with respect to the band edges E_C and E_V in the n region, at a given temperature $T > 0$ and donor concentration N_D.

Similarly, in the p region (18.38), (18.39) hold, namely,

$$N_C\,\Phi_{1/2}(\xi_e) + \frac{N_A}{(1/d_A)\,\exp(\xi_h - \xi_A) + 1} = N_V\,\Phi_{1/2}(\xi_h)\,,$$

where $-\xi_A = (E_A - E_V)/(k_B\,T) > 0$ is the normalized distance between the ground level of the acceptor atoms and the edge of the valence band. The above, along with the relation $\xi_e + \xi_h = -E_G/(k_B\,T)$, again forms a system in the two unknowns ξ_e, ξ_h, whose solution determines this time the position of $E_F + q\,\varphi(-\infty)$ with respect to the band edges E_C and E_V in the p region, at a given temperature $T > 0$ and acceptor concentration N_A. The built-in potential is then found from

$$q\,\psi_0 = [E_F + q\,\varphi(+\infty)] - [E_F + q\,\varphi(-\infty)]\ .$$

Due to the presence of the Fermi integrals, the charge-neutrality equations cannot be solved analytically as they stand. The approximating expressions for the Fermi integrals shown in Sect. C.13 ease the calculation somewhat.

21.2 The relations to be used are (21.54), (21.58), and (21.60). If $k_p = 0$, $k_n > 0$, one finds $Y_p = 0$, $Y_n = 1 - \exp[-m(b)]$. On the other hand, it is in this case $m(b) = \int_a^b k_n\,\mathrm{d}x > 0$, whence $Y_n < 1$. If, instead, $k_n = 0$, $k_p > 0$, one finds $Y_n = 0$, $Y_p = 1 - \exp[m(b)]$ with $m(b) = -\int_a^b k_p\,\mathrm{d}x < 0$, whence $Y_p < 1$. In conclusion, the condition for avalanche never occurs.

21.3 The dopant concentrations are such that the nondegeneracy and complete-ionization conditions hold. The product $N_A\,N_D$ is the same as in (21.127), so that $\psi_0 \simeq 0.65$ V. Next, one applies (21.44), with $\psi = \psi_0 - V = 2.75$ V and $1/N_D + 1/N_A \simeq 10^{-14}$ cm^3. In the calculation, one may profit by (21.131), where $1/N_D + 1/N_A = 1.1 \times 10^{-15}$ cm^3 and $\psi = \psi_0 = 0.65$ V; in fact, it suffices to replace the new values to find that the term in brackets of (21.131) increases by the factor $(2.75/0.65)\times(10/1.1) \simeq 38.46$ whence, extracting the square root, $l = l_n + l_p \simeq 6.2$ μm. From the second relation in (21.39) it then follows $l_n = l/(1 + N_D/N_A) \simeq l$ and $l_p = l_n\,N_D/N_A \simeq 6.2$ nm.

21.4 In silicon at room temperature it is $n_i \simeq 10^{10}$ cm^{-3}, whence $l\,n_i/\tau_g \simeq 6.2 \times 10^{10}$ cm^{-2} s^{-1}. It follows $q\,A_e\,l\,n_i/\tau_g \simeq 10$ nA.

21.5 Introducing the values indicated in the text into $\tau_0 = A_e\,C_0\,\psi_0/I_G$, one finds $\tau_0 = 143$ ms.

21.6 The linearization of (21.69) starting from $V_D(0) = V_0 + R I_G = -2.6 + 10^5 \times 50 \times 10^{-9} \simeq -2.6$ V yields

$$V_i \simeq V_D(0) + \left(\frac{\mathrm{d}V_D}{\mathrm{d}t}\right)_{t=0} T_i = V_0 + R I_G + \psi_0 \left(1 - \frac{V_0 + R I_G}{\psi_0}\right)^{1/2} \frac{T_i}{\tau_0},$$

namely,

$$V_i \simeq -2.6 + 0.65 \left(1 + \frac{2.6}{0.65}\right)^{1/2} \frac{15}{143} \simeq -2.45 \quad \mathrm{V}.$$

21.7 From (21.70) one obtains

$$G = \frac{(-2.45 + 2.6)\ \mathrm{V}}{10^5\ \Omega \times 50 \times 10^{-9}\ \mathrm{A}} = 30.$$

21.8 Defining $\xi = (1 - V_D/\psi_0)^{1/2}$ and $\alpha = [1 - (V_0 + R I_G)/\psi_0]^{1/2}$ transforms (21.71) into

$$-\frac{\alpha\, \mathrm{d}t}{R A_e C_0} = \frac{\mathrm{d}\xi}{\xi - \alpha} - \frac{\mathrm{d}\xi}{\xi + \alpha} = \mathrm{d}\log\frac{\xi - \alpha}{\xi + \alpha}.$$

Letting $\xi_i = (1 - V_i/\psi_0)^{1/2}$, $\tau = 2 R A_e C_0/\alpha$ in the above equation and fixing the origin of t at the beginning of the sampling time yield

$$\frac{1 - V_D/\psi_0}{1 - (V_0 + R I_G)/\psi_0} = \left[\frac{\xi_i \cosh(t/\tau) + \alpha \sinh(t/\tau)}{\xi_i \sinh(t/\tau) + \alpha \cosh(t/\tau)}\right]^2.$$

The solution fulfills the initial condition $V_D = V_i$; the limit for $t \to \infty$ is $V_D = V_0 + R I_G$. Using the values of Probs. 21.5 and 21.6 yields $\tau \simeq 0.98$ ms; this shows that, in practice, the duration of the sampling time can be limited to a few milliseconds.

21.9 Defining $\xi = (1 - V_D/\psi_0)^{1/3}$ and $\alpha = [1 - (V_0 + R I_G)/\psi_0]^{1/3}$ transforms (21.71) into

$$-\frac{\alpha\, \mathrm{d}t}{R A_e C_0} = \mathrm{d}\log(\xi - \alpha) - \frac{1}{2}\,\mathrm{d}\log(\xi^2 + \alpha\,\xi + \alpha^2) + \sqrt{3}\,\mathrm{d}\arctan\left(\frac{2\,\xi + \alpha}{\sqrt{3}\,\alpha}\right),$$

whose integration provides the inverse relation $t = t(V_D)$.

21.10 It is $1 - (V_0 + R I_G)/\psi_0 \simeq 5$, whence

$$V_D \simeq \psi_0 \left[1 - \left(\sqrt{5} - t/(2\,\tau_0)\right)^2\right], \qquad m = 2,$$

with $V_D = V_i = -2.45$ V for $t = 15$ ms, and

$$V_D \simeq \psi_0 \left[1 - \left(25^{1/3} - t/(1.5\,\tau_0)\right)^{3/2}\right], \qquad m = 3\,,$$

with $V_D = V_i = -2.48$ V for $t = 15$ ms. The corresponding curves are shown in Fig. 21.16.

21.11 It is $\alpha^2 = 1 - (V_0 + R\,I_G)/\psi_0 \simeq 5$, whence, from Prob. 21.8,

$$\frac{V_D}{\psi_0} \simeq 1 - 5\left[\frac{\xi_i\,\cosh(t/\tau) + \alpha\,\sinh(t/\tau)}{\xi_i\,\sinh(t/\tau) + \alpha\,\cosh(t/\tau)}\right]^2, \quad m = 2\,, \quad \xi_i^2 = 1 - V_i/\psi_0 \simeq 4.77\,.$$

As for the $m = 3$ case, one takes the result of Prob. 21.9 and draws the curve representing the inverse function. The corresponding curves are shown in Fig. 21.16.

21.12 Multiply by $\exp(-s/L_n)$ both sides of (21.79) and subtract the result from (21.78); rearranging the terms yields

$$A_n = \frac{n_{p0}\,F - \sum_i C_i\,\exp(-k_i\,s) + \exp(-s/L_n)\,\sum_i C_i}{2\,\sinh(s/L_n)}\,.$$

Similarly, multiply by $\exp(s/L_n)$ both sides of (21.79) and subtract the result from (21.78), thus yielding

$$B_n = -\frac{n_{p0}\,F - \sum_i C_i\,\exp(-k_i\,s) + \exp(s/L_n)\,\sum_i C_i}{2\,\sinh(s/L_n)}\,.$$

21.13 With reference to Fig. D.9, and using x as independent variable, one finds for the area between $p(x)$ and its asymptotic value:

$$\int_0^\infty (p - p_{n0})\,\mathrm{d}x = A_p \int_0^\infty \exp(-x/L_p)\,\mathrm{d}x = L_p\,A_p\,.$$

The same area is obviously found using p as independent variable; thus,

$$\frac{1}{A_p}\int_{p_{n0}}^{p_{n0}+A_p} x(p)\,\mathrm{d}p = L_p\,.$$

Note that the equation of the tangent to $p(x)$ at $x = 0$ is $y = p_{n0} + A_p\,(1 - x/L_p)$; thus, the tangent intersects the asymptotic value at $x = L_p$. This property of the exponential function is exploited, e.g., for measuring the minority-carrier lifetimes (Sect. 25.2).

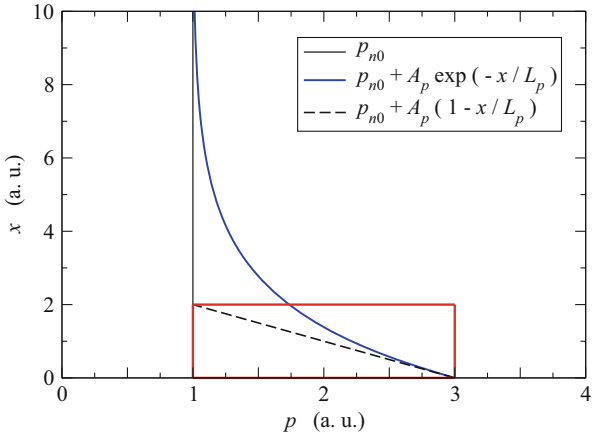


Fig. D.9 The *blue line* is the inverse function of (21.141), drawn with $p_{n0} = 1$, $A_p = 2$, $L_p = 2$ in arbitrary units. The tangent to $p(x)$ at $x = 0$ (*dashed line*) intersects the asymptotic value p_{n0} at $x = L_p$. The area of the rectangle marked in red is equal to the area between $p(x)$ and the asymptotic value

Problems of Chap. 22

22.1 The differential capacitance of the MOS structure, extracted from (22.19) and (22.20), reads

$$\frac{C}{C_{\rm ox}} = \frac{1}{1 + C_{\rm ox}/C_{\rm sc}}\,, \qquad C_{\rm sc} = \pm\frac{\sqrt{2}\,\varepsilon_{\rm sc}}{L_A}\,\frac{{\rm d}F}{{\rm d}u_s} > 0\,,$$

where the plus (minus) sign holds for $u_s > 0$ ($u_s < 0$). From (22.3) and (22.5) one finds ${\rm d}F^2/{\rm d}u_s = A(u_s)$; on the other hand, it is ${\rm d}F^2/{\rm d}u_s = 2F\ {\rm d}F/{\rm d}u_s$, whence

$$C_{\rm sc} = \pm\frac{\varepsilon_{\rm sc}}{\sqrt{2}\,L_A}\,\frac{A}{F}\,, \qquad \frac{C_{\rm ox}}{C_{\rm sc}} = \pm\frac{F}{rA}\,, \qquad r = \frac{\varepsilon_{\rm sc}\,t_{\rm ox}}{\varepsilon_{\rm ox}\,\sqrt{2}\,L_A}\,.$$

Then, the $C(V_G)$ relation is found by eliminating u_s from

$$u'_G = u_s \pm 2\,r\,F\,, \qquad \frac{C}{C_{\rm ox}} = \frac{1}{1 \pm F/(rA)}\,.$$

In particular, from (22.26) one finds $C(V'_G = 0) = C_{\rm ox}/[1 + 1/(\sqrt{2}\,r)]$.

22.2 Using the normalized form (22.132) of the electron and hole concentrations, the Poisson equation reads (compare with (22.133) and (22.134))

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = \frac{1}{L_D^2} A\,, \qquad A = \exp(u - \chi_n) - 1 + \frac{n_i^2}{N_D^2}\left[1 - \exp(\chi_p - u)\right],$$

with L_D the electrons' Debye length (21.9). When the channel is well formed, the term $\exp(u - \chi_n)$, proportional to the electron concentration, is negligible; also, $-1 + (n_i/N_D)^2 \simeq -1$. Using the gradual-channel approximation and the relation $(n_i/N_D)^2 = \exp(2\,u_F)$, the above reduces to (compare with (22.135))

$$L_D^2 \frac{\partial^2 u}{\partial x^2} = -1 - \frac{n_i^2}{N_D^2} \exp(\chi_p - u) < 0\,.$$

Following the same reasoning as in Sect. 22.11.1 yields (compare with (22.137))

$$\left(\frac{q\,E_{sx}}{k_B\,T}\right)^2 = \frac{2}{L_D^2} F^2\,, \qquad F^2 = \exp(\chi_p + 2\,u_F)\,\left[\exp(-u_s) - 1\right] - u_s\,.$$

The exponential term in the above becomes dominant when $\chi_p + 2\,u_F - u_s \geq 0$, whence the threshold condition is given by $\chi_p + 2\,u_F = u_s$. Remembering (22.131), the latter becomes $u_s = u_p + u_F$, which is the normalized form of (22.102).

22.3 With reference to Table 22.3, the current sought corresponds to the part of the $V_{\text{out}}(V_{\text{in}})$ curve between points B and C. Letting $V_{GS} = V_{\text{in}}^{\text{sat}}$ in (22.94) and combining the result with (22.128) yield

$$\tilde{I}_{\text{sat}} = \frac{1}{2}\,\beta_n\,\left(V_{\text{in}}^{\text{sat}} - V_{Tn}\right)^2 = \frac{1}{2}\,\frac{\left(V_{SS} - V_{Tn} - V_{Tp}\right)^2}{\left(1/\sqrt{\beta_n} + 1/\sqrt{\beta_p}\right)^2}\,.$$

22.4 Between A and B the n-channel transistor works in the saturation region, whence $I_{Dn} = (\beta_n/2)\,(V_{\text{in}} - V_{Tn})^2$. The p-channel transistor, instead, works in the linear region, whence $\beta_p\,[(V_{SS} - V_{\text{in}} - V_{Tp})\,(V_{SS} - V_{\text{out}}) - (V_{SS} - V_{\text{out}})^2/2]$. Equating the two current and solving for V_{out} yield

$$V_{\text{out}} = V_{\text{in}} + V_{Tp} + \sqrt{(V_{SS} - V_{Tp} - V_{\text{in}})^2 - (\beta_n/\beta_p)\,(V_{\text{in}} - V_{Tn})^2}\,.$$

The negative sign before the square root must be discarded because it must be $V_{\text{out}} = V_{SS}$ when $V_{\text{in}} = V_{Tn}$. Between C and D the p-channel transistor works in the saturation region, whence $I_{Sp} = (\beta_p/2)\,(V_{SS} - V_{\text{in}} - V_{Tp})^2$. The n-channel transistor, instead, works in the linear region, whence $I_{Dn} = \beta_n\,[(V_{\text{in}} - V_{Tn})\,V_{\text{out}} - V_{\text{out}}^2/2]$. Proceeding as above yields

$$V_{\text{out}} = V_{\text{in}} - V_{Tn} - \sqrt{(V_{\text{in}} - V_{Tn})^2 - (\beta_p/\beta_n)\,(V_{SS} - V_{\text{in}} - V_{Tp})^2}\,.$$

The positive sign before the square root must be discarded because it must be $V_{\text{out}} = 0$ when $V_{\text{in}} = V_{SS} - V_{Tp}$.

22.5 The relations describing the CMOS inverter become particularly simple if $V_{Tp} = V_{Tn} = V_T$ and $\beta_p = \beta_n = \beta$. The former condition can be achieved by a suitable channel implant, the latter by designing the gates' geometry in such a way that the equality $\mu_p\,(W/L)_p = \mu_n\,(W/L)_n$ holds. When the two conditions apply, one finds from (22.128) and from Prob. 22.3:

$$V_{\text{in}}^{\text{sat}} = \frac{1}{2}\,V_{SS}\,, \qquad \tilde{I}_{\text{sat}} = \frac{1}{8}\,\beta\,\left(V_{SS} - 2\,V_T\right)^2\,.$$

In turn, the $V_{\text{out}}(V_{\text{in}})$ relation between points A and B found in Prob. 22.4 becomes

$$V_{\text{out}} = V_{\text{in}} + V_T + \sqrt{V_{SS}^2 - 2\,V_{SS}\,(V_T + V_{\text{in}}) + 4\,V_T\,V_{\text{in}}}\,.$$

Similarly, the relation between points C and D becomes

$$V_{\text{out}} = V_{\text{in}} - V_T - \sqrt{-V_{SS}^2 + 2\,V_{SS}\,(V_T + V_{\text{in}}) - 4\,V_T\,V_{\text{in}}}\,.$$

Shifting the origin by letting $V'_{\text{in}} = V_{\text{in}} - V_{SS}/2$ and $V'_{\text{out}} = V_{\text{out}} - V_{SS}/2$ transforms the $V_{\text{out}}(V_{\text{in}})$ relation corresponding to the interval between points A and B into

$$V'_{\text{out}} = V'_{\text{in}} + V_T + 2\,\sqrt{(V_T - V_{SS}/2)\,V'_{\text{in}}}\,.$$

Note that both factors under the square root are negative. Still in the new variables, the $V_{\text{out}}(V_{\text{in}})$ relation between C and D becomes

$$V'_{\text{out}} = -\left[(-V'_{\text{in}}) + V_T + 2\,\sqrt{(V_T - V_{SS}/2)\,(-V'_{\text{in}})}\right]\,,$$

showing that these two portions of the curve are antisymmetric. The proof for the remaining parts of the curve is trivial (Prob. 22.6).

22.6 Let $0 \le V_{\text{in}} \le V_{Tn}$: the current is zero at A and in the region to the left of A; the corresponding output voltage is V_{SS}. As V_{in} increases from V_{Tn} to $V_{\text{in}}^{\text{sat}}$, the current grows from 0 to $\tilde{I}_{\text{sat}}$ and V_{out} decreases from V_{SS} (point A) to $V_{\text{in}}^{\text{sat}} + V_{Tp}$ (point B); the expression of this part of the curve is worked out in Prob. 22.4. Similarly, let $V_{SS} - V_{Tp} \le V_{\text{in}} \le V_{SS}$: the current is zero at D and in the region to the right of D; the corresponding output voltage is zero. As V_{in} decreases from $V_{SS} - V_{Tp}$ to $V_{\text{in}}^{\text{sat}}$, the current grows from 0 to $\tilde{I}_{\text{sat}}$ and V_{out} increases from zero (point D) to $V_{\text{in}}^{\text{sat}} - V_{Tn}$ (point C); the expression is worked out in Prob. 22.4 as well. Finally, the part of the curve between points B and C is vertical with $V_{\text{in}} = V_{\text{in}}^{\text{sat}}$, while V_{out} ranges between $V_{\text{in}}^{\text{sat}} + V_{Tp}$ (point B) and $V_{\text{in}}^{\text{sat}} - V_{Tn}$ (point C).

22.7 Assume that the threshold voltage V_{Tn} fulfills the inequality $V_{SS} > V_{Tn}$, where V_{SS} is the bias applied to the resistor's upper contact (Fig. 22.27). Then, let $0 \le V_i \le V_{Tn}$, so that the transistor is in the *off* condition. It follows $I_R = (V_{SS} - V_{\text{out}})/R = I_{Dn} = 0$, whence $V_{\text{out}} = V_{SS}$. As long as $0 \le V_{\text{in}} \le V_{Tn}$, the output voltage is

the highest and no power is dissipated by the inverter. Let now the input voltage be increased by a small amount starting from $V_{\rm in} = V_{Tn}$, to become $V_{\rm in} = V_{Tn} + \delta V_{\rm in}$ with $\delta V_{\rm in} > 0$. It follows $V_{GSn} - V_{Tn} = \delta V_{\rm in}$, which makes the transistor to turn on, $I_R = I_{Dn} > 0$. In turn, the current flow makes $V_{\rm out}$ to decrease and become $V_{\rm out} = V_{SS} - \delta V_{\rm out}$, with $\delta V_{\rm out} > 0$. One finds for the current $R\,I_R = V_{SS} - V_{\rm out} = \delta V_{\rm out}$. From $V_{DSn} = V_{SS} - \delta V_{\rm out}$ one obtains

$$(V_{GSn} - V_{Tn}) - V_{DSn} = \delta V_{\rm in} - (V_{SS} - \delta V_{\rm out}) = (\delta V_{\rm in} + \delta V_{\rm out}) - V_{SS}\,.$$

As $\delta V_{\rm in} + \delta V_{\rm out}$ can be made as small as we please, it follows $(V_{GSn} - V_{Tn}) - V_{DSn} < 0$, namely, the transistor is in the saturation region, whence

$$\frac{V_{SS} - V_{\rm out}}{R} = \frac{1}{2}\,\beta_n\,(V_{\rm in} - V_{Tn})^2\ .$$

If the input voltage is further increased, the transistor will eventually turn from the saturation to the linear region. This happens for $V_{DSn} = V_{GSn} - V_{Tn}$, namely $V_{\rm out} = V_{\rm in} - V_{Tn}$. Replacing in the above provides an algebraic equation whose solution yields

$$\tilde{V}_{\rm in} = V_{Tn} + \frac{\sqrt{1 + 2\,R\,\beta_n\,V_{SS}} - 1}{R\,\beta_n}\,.$$

The minus sign before the square root must be discarded because it would make $\tilde{V}_{\rm in} < V_{Tn}$. If $\tilde{V}_{\rm in} < V_{SS}$, then there is an interval of input voltages $\tilde{V}_{\rm in} \le V_{\rm in} \le V_{SS}$ where the transistor is in the linear region, namely,

$$\frac{V_{SS} - V_{\rm out}}{R} = \beta_n \left[(V_{\rm in} - V_{Tn})\ V_{\rm out} - \frac{1}{2}\,V_{\rm out}^2 \right].$$

The above is applicable up to the maximum input voltage $V_{\rm in} = V_{SS}$. Using the normalized quantities $x = V_{\rm in}/V_{SS}$, $y = V_{\rm out}/V_{SS}$, $a = V_{Tn}/V_{SS}$, $b = R\,\beta_n\,V_{SS}$, and $\tilde{x} = \tilde{V}_{\rm in}/V_{SS} = a + (\sqrt{1 + 2\,b} - 1)/b$, yields

$$y = 1 \quad \text{for} \quad 0 \le x \le a\,, \qquad y = 1 - \frac{1}{2}\,b\,(x - a)^2 \quad \text{for} \quad a \le x \le \tilde{x} < 1\,,$$

and

$$y = x - a + 1/b - \sqrt{(x - a + 1/b)^2 - 2/b}\,, \qquad \tilde{x} \le x \le 1\,.$$

In the latter, the plus sign before the square root has been discarded because $V_{\rm out}$ vanishes as R increases, hence $\lim_{b\to\infty} y = 0$. One finds that $V_{\rm out}$ reaches its minimum $V_{\rm out}^{\rm min}$ for $V_{\rm in} = V_{\rm in}^{\rm max} = V_{SS}$; correspondingly, the current and dissipated power reach their maximum values $I_{\rm max} = (V_{SS} - V_{\rm out}^{\rm min})/R$, $P_{\rm max} = V_{SS}\,I_{\rm max}$.

22.8 Remembering that the relative dielectric constant of SiO_2 is $\varepsilon_{ox}^{rel} \simeq 3.9$, it follows

$$\beta_n = \frac{W}{L}\,\mu_n\,C_{ox} \simeq 993\,\frac{\text{cm}^2}{\text{V s}} \times \frac{3.9 \times 8.85 \times 10^{-14}}{5 \times 10^{-7}}\,\frac{\text{F}}{\text{cm}^2} \simeq 6.85 \times 10^{-4}\,\frac{\text{A}}{\text{V}^2}\,,$$

$$b = 3.5 \times 10^3\ \Omega \times 6.85 \times 10^{-4}\,\frac{\text{A}}{\text{V}^2} \times 5\ \text{V} \simeq 12\,, \qquad \tilde{x} = \frac{1}{10} + \frac{1}{3} \simeq 0.433\,,$$

$$\tilde{V}_{in} = 0.433 \times 5\ \text{V} \simeq 2.17\ \text{V}\,, \qquad V_{out}(\tilde{V}_{in}) = 2.17\ \text{V} - 0.5\ \text{V} \simeq 1.67\ \text{V}\,.$$

The curve is shown in Fig. D.10, using the normalized variables $x = V_{in}/V_{SS}$, $y = V_{out}/V_{SS}$. The dynamic range is smaller than V_{SS}; in fact, it is $V_{out}(0) = V_{SS}$ whereas $V_{out}(V_{SS}) = V_{out}^{min} > 0$ (compare with the comments in Sect. 22.11.3).

22.9 Assume that the threshold voltage V_{Tp} fulfills the inequality $V_{SS} > V_{Tp}$ where V_{SS} is the bias applied to the transistor's source (Fig. 22.28). Then, let $0 \le V_{SS} - V_{in} \le V_{Tp}$, so that the transistor is in the *off* condition. It follows $I_R = V_{out}/R = I_{Sp} = 0$, whence $V_{out} = 0$. As long as $0 \le V_{SS} - V_{in} \le V_{Tp}$, the output voltage is the lowest and no power is dissipated by the inverter. Let now the input voltage be decreased by a small amount starting from $V_{SS} - V_{in} = V_{Tp}$, to become $V_{in} = V_{SS} - V_{Tp} - \delta V_{in}$ with $\delta V_{in} > 0$. It follows $V_{SGp} - V_{Tp} = \delta V_{in}$, which makes the transistor to turn on, $I_R = I_{Sp} > 0$. In turn, the current flow makes V_{out} to increase and become $V_{out} = \delta V_{out}$, with $\delta V_{out} > 0$. One finds for the current $R\,I_R = \delta V_{out}$. From $V_{SDp} = V_{SS} - \delta V_{out}$ one obtains

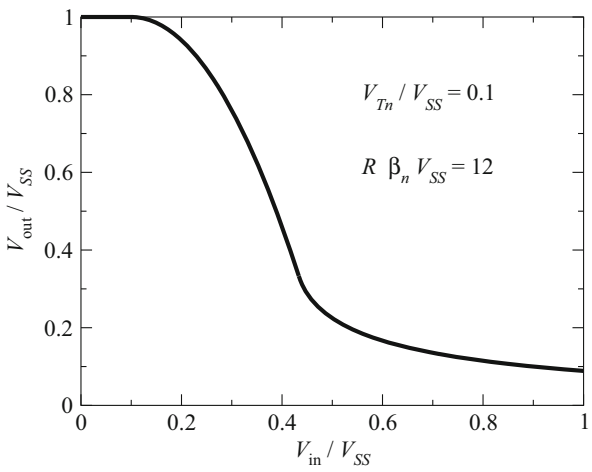


Fig. D.10 Input-output curve of the inverter made of an *n*-channel MOSFET with a resistive load (Probs. 22.7 and 22.8)

$$\left(V_{SGp} - V_{Tp}\right) - V_{SDp} = (\delta V_{\rm in} + \delta V_{\rm out}) - V_{SS}\,.$$

As $\delta V_{\rm in} + \delta V_{\rm out}$ can be made as small as we please, it follows $(V_{SGp} - V_{Tp}) - V_{SDp} < 0$, namely, the transistor is in the saturation region, whence

$$\frac{V_{\rm out}}{R} = \frac{1}{2}\,\beta_p\,\left(V_{SS} - V_{\rm in} - V_{Tp}\right)^2\,.$$

If the input voltage is further decreased, the transistor will eventually turn from the saturation to the linear region. This happens for $V_{SDp} = V_{SGp} - V_{Tp}$, namely $V_{\rm out} = V_{\rm in} + V_{Tp}$. Replacing in the above provides an algebraic equation whose solution yields

$$\tilde{V}_{\rm in} = V_{SS} - V_{Tp} - \frac{\sqrt{1 + 2\,R\,\beta_p\,V_{SS}} - 1}{R\,\beta_p}\,.$$

The plus sign before the square root must be discarded because it would make $V_{\rm out}(\tilde{V}_{\rm in}) > V_{SS}$. If $\tilde{V}_{\rm in} > 0$, then there is an interval of input voltages $0 \le V_{\rm in} \le \tilde{V}_{\rm in}$ where the transistor is in the linear region, namely,

$$\frac{V_{\rm out}}{R} = \beta_p\,\left[\left(V_{SS} - V_{\rm in} - V_{Tp}\right)\,\left(V_{SS} - V_{\rm out}\right) - \frac{1}{2}\,\left(V_{SS} - V_{\rm out}\right)^2\right]\,.$$

The above is applicable down to the minimum input voltage $V_{\rm in} = 0$. Using the normalized quantities $x = V_{\rm in}/V_{SS}$, $y = V_{\rm out}/V_{SS}$, $c = V_{Tp}/V_{SS}$, $d = R\,\beta_p\,V_{SS}$, and $\tilde{x} = \tilde{V}_{\rm in}/V_{SS} = 1 - c - (\sqrt{1 + 2\,d} - 1)/d$ yields

$$y = 0 \quad \text{for} \quad 1 - c \le x \le 1\,, \qquad y = \frac{1}{2}\,d\,(1 - x - c)^2 \quad \text{for} \quad 0 < \tilde{x} \le x < 1 - c\,,$$

and

$$y = x + c - 1/d + \sqrt{(1 - x - c + 1/d)^2 - 2/d}\,.$$

In the latter, the minus sign before the square root has been discarded because $V_{\rm out}$ tends to V_{SS} as R increases, hence $\lim_{d\to\infty} y = 1$. The expressions of Prob. 22.7 transform into those found here by replacing x, $\tilde{x}$, y, a, b with $1 - x$, $1 - \tilde{x}$, $1 - y$, c, d, respectively. One finds that $V_{\rm out}$ reaches its maximum $V_{\rm out}^{\rm max}$ for $V_{\rm in} = V_{\rm in}^{\rm min} = 0$; correspondingly, the current and dissipated power reach their maximum values $I_{\rm max} = V_{\rm out}^{\rm max}/R$, $P_{\rm max} = V_{SS}\,I_{\rm max}$.

22.10 Remembering that the relative dielectric constant of SiO_2 is $\varepsilon_{\rm ox}^{\rm rel} \simeq 3.9$, it follows

$$\beta_p = \frac{W}{L}\,\mu_p\,C_{\rm ox} \simeq 662\,\frac{\rm cm^2}{\rm V\,s} \times \frac{3.9 \times 8.85 \times 10^{-14}}{5 \times 10^{-7}}\,\frac{\rm F}{\rm cm^2} \simeq 4.57 \times 10^{-4}\,\frac{\rm A}{\rm V^2}\,,$$

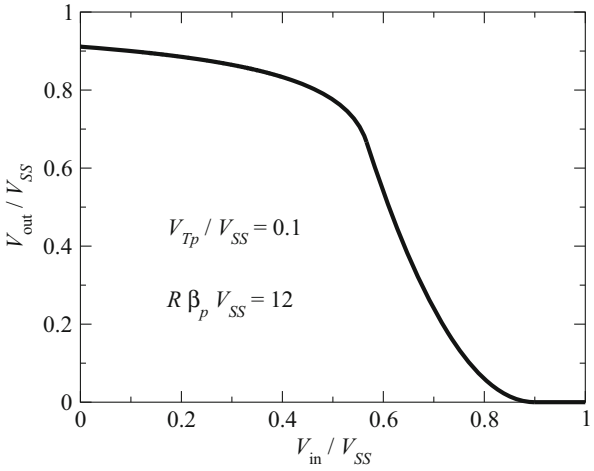


Fig. D.11 Input-output curve of the inverter made of an *p*-channel MOSFET with a resistive load (Probs. 22.9 and 22.10)

$$d = 5.25 \times 10^3\ \Omega \times 4.57 \times 10^{-4}\ \frac{\mathrm{A}}{\mathrm{V}^2} \times 5\ \mathrm{V} \simeq 12\,, \qquad \tilde{x} = \frac{9}{10} - \frac{1}{3} \simeq 0.567\,,$$

$$\tilde{V}_{\mathrm{in}} = 0.567 \times 5\ \mathrm{V} \simeq 2.83\ \mathrm{V}\,, \qquad V_{\mathrm{out}}(\tilde{V}_{\mathrm{in}}) = 2.83\ \mathrm{V} + 0.5\ \mathrm{V} \simeq 3.33\ \mathrm{V}\,.$$

The curve is shown in Fig. D.11, using the normalized variables $x = V_{\mathrm{in}}/V_{SS}$, $y = V_{\mathrm{out}}/V_{SS}$. The dynamic range is smaller than V_{SS}; in fact, it is $V_{\mathrm{out}}(V_{SS}) = 0$ whereas $V_{\mathrm{out}}(0) = V_{\mathrm{out}}^{\mathrm{max}} < V_{SS}$ (compare with the comments in Sect. 22.11.3).

22.11 As mentioned in the text, the first part of I_D derives from the first term (that is, the unity) at the right-hand side of (22.149). It reads

$$I'_D - I''_D = \beta\ (k_B\,T/q)^2\ \left[\left(u'_{GB}\,u_s - u_s^2/2\right) - (2/3)\,(1/g)\,u_s^{3/2}\right]_{u_s(0)}^{u_s(L)}\,.$$

To calculate the other part of the current, with reference to the second term at the right-hand side of (22.148), one preliminarily finds the zeros of the denominator, that read

$$u_s^{\pm} = u'_{GB} + \frac{1}{2\,r^2}\left(1 \pm \sqrt{1 + 4\,r^2\,u'_{GB}}\right)\,,$$

where the positive or negative signs hold together. The range of values of u_s is $u_s < u_s^-$. Using the zeros, the second term at the right-hand side of (22.148) becomes

$$\frac{\mathrm{d}\chi_n}{\mathrm{d}u_s} - 1 = \frac{1}{u_s^+ - u_s} + \frac{1}{u_s^- - u_s}\,, \qquad u_s^+ > u_s^- > u_s\,.$$

The above form is not suitable because it diverges when $u_s \to u_s^-$. In fact, the zero is canceled by a zero belonging to the other factor in (22.149); it is then convenient to cancel the zeros beforehand by recasting the integrand as

$$Y(u_s) = \left[\left(u'_{GB} - u_s\right) - (1/r)\,\sqrt{u_s}\right]\left(\frac{\mathrm{d}\chi_n}{\mathrm{d}u_s} - 1\right) = 2\,\frac{r\,\left(u'_{GB} - u_s\right) + 1/(2\,r)}{r\,\left(u'_{GB} - u_s\right) + \sqrt{u_s}}\,.$$

If the expression of u_s^- is replaced into the denominator at the right-hand side of the above, the result is strictly positive, whence no divergence occurs. The integrand is further simplified to yield $Y(u_s) = 2 - 2\,[\sqrt{u_s} - 1/(2\,r)]/[r\,(u'_{GB} - u_s) + \sqrt{u_s}]$; multiplying the latter by $\mathrm{d}u_s$ and using the new variable $w = \sqrt{u_s} - 1/(2\,r)$ yield

$$Y(w) = \left(w + \frac{1}{2\,r}\right)\left(4 + \frac{2/r}{k + w} - \frac{2/r}{k - w}\right)\mathrm{d}w\,, \qquad k^2 = u'_{GB} + \frac{1}{4\,r^2}\,,$$

whose integration is elementary.

22.12 The working conditions of the device are such that the surface potential is near saturation at each point of the channel. This allows one to calculate (22.149) by letting $u_s(0) \simeq u_s^-$ and assuming that u_s varies little along the channel. The integral in (22.149) is thus evaluated as the product of the integration domain $u_s(L) - u_s(0)$ times the integrand calculated at $u_s(0)$. The integration domain is in turn extracted from (22.147) by recasting it as

$$r^2\,\left(u'_{GB} - u_s\right)^2 - u_s - \exp(u_s - 2\,u_F - \chi_n) = 0\,,$$

where the exponential is small because $\chi_n > u_s - 2\,u_F$ due to the saturation condition. If the exponential was missing, the solutions of the above would be the zeros already found in Prob. 22.11. Thus, the solutions here are determined by replacing $1+4\,r^2\,u'_{GB}$ with $1+4\,r^2\,u'_{GB}+4\,r^2\,\exp(u_s-2\,u_F-\chi_n)$ under the square root in Prob. 22.11. Discarding the solution corresponding to the positive sign, expanding the square root to first order, and letting $u_s = u_s^-$ in the exponential yield

$$u_s \simeq u'_{GB} + \frac{1}{2\,r^2}\left[1 - \sqrt{1 + 4\,r^2\,u'_{GB}}\left(1 + \frac{2\,r^2\,\exp(u_s^- - 2\,u_F - \chi_n)}{1 + 4\,r^2\,u'_{GB}}\right)\right],$$

whence

$$u_s(L) - u_s(0) \simeq \frac{\exp[-\chi_n(0)] - \exp[-\chi_n(L)]}{(1 + 4\,r^2\,u'_{GB})^{1/2}\;\exp(2\,u_F - u_s^-)}\,,$$

with $\chi_n(0) = q\,V_{SB}/(k_B\,T)$ and $\chi(L) = q\,V_{DB}/(k_B\,T)$. The analysis is completed by calculating the integrand of (22.149) after letting $u_s = u_s(0) = u_s^-$. The first part of the integrand derives from the first term (that is, the unity) at the right-hand side of (22.149); it is found by inspection that this part vanishes for $u_s = u_s^-$. Thus, the remaining part is $Y(u_s^-)$ (Prob. 22.11); observing that $\sqrt{u_s^-} = r\,(u'_{GB} - u_s^-)$, one finds $Y(u_s^-) = 1 + 1/[2\,r^2\,(u'_{GB} - u_s^-)]$, whence (22.150) ensues.

22.13 Consider the integration along the channel carried out in (22.73); if the integration were interrupted at some point y internal to the channel, the outcome would be

$$y\,I = W \int_{\varphi_n(0)}^{\varphi_n^*} \mu_e(\varphi_n)\,Q_i(\varphi_n)\;\mathrm{d}\varphi_n\,,$$

with φ_n^* the electron quasi-Fermi potential at y. Similarly, (22.149) is replaced with

$$\frac{y}{L}\,I_D = \beta\left(\frac{k_B\,T}{q}\right)^2 \int_{u_s(0)}^{u_s^*} \left[\left(u'_{GB} - u_s\right) - (1/r)\,\sqrt{u_s}\right] \frac{\mathrm{d}\chi_n}{\mathrm{d}u_s}\;\mathrm{d}u_s\,,$$

where I_D is calculated beforehand from (22.149). Calculating the above integral with u_s^* ranging from $u_s(0)$ to $u_s(L)$ yields a set of values; multiplying each of them by L/I_D yields the corresponding position y, thus providing the inverse function $y = y(u_s^*)$. For each u_s^* one finds the corresponding χ_n from (22.147) which, combined with the former, provides the other inverse function $y = y(\chi_n)$.

22.14 From the data it follows $L_n = 100\ \mu\mathrm{m}$. As $\tau_{i+1} < \tau_i$ and, in particular, $\tau_1 = \tau_0/9$, the slowest decay rate corresponds to the term with $i = 0$, whereas the terms with $i > 0$ are less important. Using $a = 10\ \mu\mathrm{m}$ yields $\tau_0 \simeq \tau_n/100$, $\tau_0^\star \simeq \tau_0$.

Problems of Chap. 23

23.1 The maximum initial profile is $N(x = 0, t = 0) = 2Q/\sqrt{\pi c_1}$. Remembering (23.26), at the end of the diffusion process the profile has become $N(x, t = t_P) = 2Q\,[\pi(c_1 + c_2)]^{-1/2} \times \exp\left[-x^2/(c_1 + c_2)\right]$, whence $N(x = 0, t = t_P) = 2Q\,[\pi(c_1 + c_2)]^{-1/2}$, with t_P the process duration and $c_2 = 4D\,t_P$. From $N(x = 0, t = t_P) = (2/3)\,N(x = 0, t = 0)$ it follows $1/\sqrt{c_1 + c_2} = 2/(3\,\sqrt{c_1})$, $c_2 = (5/4)\,c_1$ and, finally, $t_P = (5/16)\,(c_1/D) = 5{,}000$ s.

23.2 The initial and final profiles are $N_i(x) = 2\,Q\,\exp(-x^2/c_1)/(\pi c_1)^{1/2}$ and, from (23.26), $N_f(x) = 2\,Q\,\exp[-x^2/(c_1 + c_2)]/[\pi(c_1 + c_2)]^{1/2}$. Letting $N_f = N_i$ and defining $r = [(c_1 + c_2)/c_1]^{1/2}$ and $a^2 = (c_1 + c_2)\,c_1/c_2$ yield $r = \exp(x^2/a^2)$, whence $\bar{x} = 10^4 \times a\,(\log r)^{1/2} \simeq 2.68\ \mu\mathrm{m}$.

23.3 Converting to seconds one finds $t_P = 14{,}400$ s, whence $c_2 = 4\,D\,t_P = 14.4 \times 10^{-6}$ cm^2. Considering that $N(x, t = 0) = 2\,Q\,\exp(-x^2/c_1)/(\pi c_1)^{1/2}$ and, from (23.26), $N(x, t = t_P) = 2\,Q\,\exp[-x^2/(c_1 + c_2)]/[\pi\,(c_1 + c_2)]^{1/2}$, the ratio sought is $N(x = 0, t = t_P)/N(x = 0, t = 0) = [c_1/(c_1 + c_2)]^{1/2} = 0.385$.

23.4 Due to (23.26), the final profile is $N_f = 2\,Q\,\exp[-x^2/(c_1+c_2)]/[\pi(c_1+c_2)]^{1/2}$. From $N(\bar{x}) = N_f(0)$ one has $\exp(-\bar{x}^2/c_1) = [c_1/(c_1 + c_2)]^{1/2}$. As a consequence, the position sought is $\bar{x} = [(c_1/2)\,\log(1 + c_2/c_1)]^{1/2} = 0.83\ \mu$m.

23.5 Due to (23.26) the final profile is $N_f = 2\,Q\,\exp[-x^2/(c_1+c_2)]/[\pi(c_1+c_2)]^{1/2}$. Let $\alpha = 1.1$. From the condition $N_f(0) = N(x = \alpha\,\sqrt{c_1})$ one derives $1/[\pi(c_1 + c_2)]^{1/2} = \exp(-\alpha^2)/(\pi c_1)^{1/2}$, whence $c_1 = c_2/[\exp(2\alpha^2) - 1] = 0.976 \times 10^{-9}$ cm^2.

23.6 Letting $c_2 = 4\,D\,t = 2.4 \times 10^{-8}$ cm^2 and $r = c_2/(c_1 + c_2) = 4/7$, one eliminates Q at $x = x_0$ to find $N_1/N_2 = \sqrt{1 + r}\,\exp[-r\,x_0^2/(c_1 + 2\,c_2)]$, whence $x_0 = 10^4 \times [(1/r)\,(c_1 + 2\,c_2)\,\log(\sqrt{1 + r}\,N_2/N_1)]^{1/2} = 4.06\ \mu$m.

23.7 Remembering (C.71) one has $Q = \int_0^\infty N_S\,\mathrm{erfc}(x/\sqrt{c})\,\mathrm{d}x = N_S\,\sqrt{c/\pi}$, whence $c = 4\,D\,t = \lambda^2$, $t = \lambda^2/(4\,D) = 2.99 \times 10^{16}$ nm^2 s^{-1} $\simeq 5$ min.

23.8 As indicated in Sect. 23.5.2, the initial profile must preliminarily be mirrored onto the negative axis with $N_0 = 2\,Q\,(h + x)/h^2$ for $-h \le x \le 0$ and with $N_0 = 0$ for $x < -h$. Then, the profile is obtained from (23.13) as the portion of $N(x) = (2\,Q/h^2)\,(I^- + I^+)$ calculated for $x \ge 0$, where

$$I^- = \int_{-h}^{0} (h + \xi)\,\Delta(x - \xi, t)\,\mathrm{d}\xi\,, \qquad I^+ = \int_{0}^{+h} (h - \xi)\,\Delta(x - \xi, t)\,\mathrm{d}\xi$$

and $\Delta(x-\xi, t)$ is given by (C.76). Letting $\eta = (\xi - x)/(4\,a)^{1/2}$, $\mu = x/h$, $b = 4\,a/h^2$, and using (C.66) yield

$$N = (Q/h)\,\left[(\mu + 1)\,Y^- + (\mu - 1)\,Y^+ - (b/\pi)^{1/2}\,(Z^- + Z^+)\right]\,,$$

with $Y^\mp = \mathrm{erfc}(\mu/\sqrt{b}) - \mathrm{erfc}[(\mu \pm 1)/\sqrt{b}\,]$ and $Z^\mp = \exp(-\mu^2/b) - \exp[-(\mu \pm 1)^2/b]$. When $t \to 0^+$ it is $b \to 0$. This makes the $Z^\mp$ terms to vanish, while the terms containing $Y^\mp$ render the initial condition N_0. When $t > 0$ the dose is found by integrating $h\,N$ over μ from 0 to $+\infty$. A somewhat lengthy calculation based on the expressions of Appendix C.7 shows that the integral of $(b/\pi)^{1/2}\,(Z^- + Z^+)$ vanishes whereas that of $(\mu+1)\,Y^-+(\mu-1)\,Y^+$ yields unity. As expected, the result of the dose calculation is Q. The normalized profile $h\,N/Q$ is shown in Fig. D.12 as a function of the normalized coordinate μ at different values of the parameter $b = b(t)$.

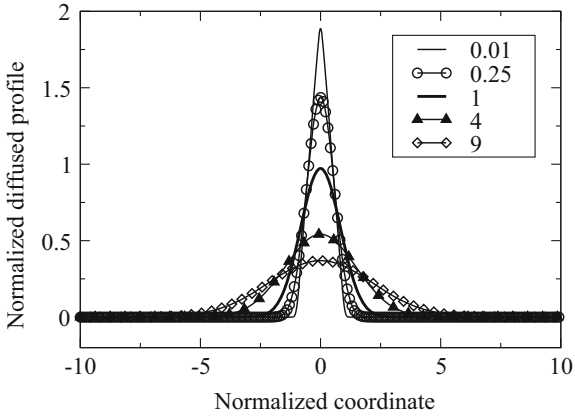


Fig. D.12 Normalized profiles $h\,N/Q$ resulting from the drive-in diffusion of problem 23.8. The coordinate is $\mu = x/h$. Each profile corresponds to the value of $b(t)$ shown in the legend. The parameter is defined by $b = 4\,a/h^2$, while $a = a(t)$ is defined by the second of (23.10). As explained in Sect. 23.5.2, only the profile's portion on the right of the origin must be considered

Problems of Chap. 24

24.1 The relation between time t and oxide thickness s is given by (24.11), $s^2/c_p + s/c_l = t'$ with $t' = t + s_i^2/c_p + s_i/c_l$. Solving for s and discarding the negative solution yield $s = [(c_p^2/c_l^2 + 4c_p t')^{1/2} - c_p/c_l]/2$, with $c_p/c_l = 0.05\ \mu$m. It follows $t' = t + s_i^2/c_p + s_i/c_l \simeq 136/60 + 0.995 + 0.237 = 3.50$ h and $4c_p t' = 0.620\ \mu\text{m}^2$. The total oxide thickness and the thickness of silicon that is consumed in the second process are, respectively, $s = [(0.05^2 + 4c_p t')^{1/2} - 0.05]/2 \simeq 0.369\ \mu$m and $h = 0.44\,(s - s_i) \simeq 70$ nm.

24.2 The gradient sought is found by remembering that, from (24.10), it is $c_p = 2w\,k_0\,N_G\,D_O$ and $c_l = w\,k_0\,N_G\,v_r$, whence $c_l/c_p = v_r/(2\,D_O) = 20\ \mu\text{m}^{-1}$, $v_r/D_O = 4 \times 10^5\ \text{cm}^{-1}$. On the other hand, from (24.7, 24.6), $-D_O\ \text{grad}\,N = -D_O\,\text{d}N/\text{d}x = v_r N_I$ whence $\text{d}N/\text{d}x = -N_O\,v_r/D_O = -4 \times 10^{17}\ \text{cm}^{-4}$.

24.3 Converting the units one finds $c_p = 738\ \text{nm}^2\text{min}^{-1}$, $c_l = 14.8\ \text{nm min}^{-1}$. Letting h be the thickness of silicon consumed one has $h = 0.44\,(s - s_i)$, $s = s_i + h/0.44$ whence, from (24.11), $t = (s^2 - s_i^2)/c_p + (s - s_i)/c_l = 150$ min.

24.4 Converting the units yields $c_p = 2{,}000\ \text{nm}^2\ \text{min}^{-1}$, $c_l = 50\ \text{nm min}^{-1}$. From (24.10) it follows $v_r/D_O = 2c_l/c_p = 0.05$, $1 + s\,v_r/D_O = 2$ whence, using (24.8), $N_I = N_O/(1 + s\,v_r/D_O) = 15 \times 10^{11}\ \text{cm}^{-3}$.

24.5 Converting the units yields $c_p = 11.4\ \text{nm}^2\ \text{s}^{-1}$, $c_l = 0.237\ \text{nm s}^{-1}$. From (24.10) the duration of the first process is found as $t_1 = (s_f^2 - s_i^2)/c_p + (s_f - s_i)/c_l$. Similarly, that of the second process is $t_2 = (s^2 - s_f^2)/c_p + (s - s_f)/c_l$. As the coefficients c_p, c_l are the same one adds the expressions of t_1 and t_2 to each other and lets $s = s_i + \Delta s_1 + \Delta s_2$. This yields $t_1 + t_2 = (s^2 - s_i^2)/c_p + (s - s_i)/c_l = 202$ s.

24.6 Converting the units yields $c_l = 16.67$ nm min^{-1}. Using the definitions (24.10) of c_p and c_l one finds $r = D_O/v_r = c_p/(2\,c_l)$, whence $t_P = [(s^2 - s_i^2)/(2r) + s - s_i]/c_l = 13.9$ min.

24.7 Letting $t = t_P$ and multiplying the first of (24.10) by c_p yield $s^2 + bs + c = 0$ with $b = c_p/c_l$, $c = -s_i^2 - (c_p/c_l)\,s_i - c_p t$. Here s_i and $c_p t_P$ are given while $c_p/c_l = 2D_O/v_r = 2.25 \times 10^{-6}$ cm. Solving for s and discarding the negative root provide the final thickness $s = [(b^2 - 4c)^{1/2} - b]/2 = 76.1$ nm.

24.8 From the relation $\Delta P = s\,\pi r^2\,p_{\mathrm{Si}}$, where ΔP, s are the weight and thickness of the epitaxial layer, and r the wafer's radius, one finds $2\,r = 2\,\sqrt{\Delta P/(\pi s p_{\mathrm{Si}})} \simeq 20.4$ cm $\simeq 8$ in.

24.9 The surface concentration N_S of $SiCl_4$ is found from the relations $s/t = c_l = w\,F_2 = w\,v_r\,N_S$, whence $N_S = s/(w\,v_r\,t) = 1 \times 10^{16}$ cm^{-3}.

24.10 Using $1/w = 5 \times 10^{22}$ cm^{-3} in the relations (24.22, 24.23) yields $t = s/c_l = s/(wF_I) = 2$ min.

24.11 Letting t_P be the duration of the process one has, from (24.22, 24.23), $c_l = s/t_P = w\,v_r\,N_I$, whence, using 1 μm $= 10^{-4}$ cm, $v_r = s/(w\,N_I\,t_P) = 200$ cm min^{-1}.

24.12 From (24.21) and the second of (24.23) one finds $b = (v_r + v_G)/(v_G\,w\,N_G) = (N_G/N_S)\,(w\,N_G)^{-1} = a/(w\,N_G)$, whence $N_G = a/(w\,b) = 2 \times (5 \times 10^{22}/4.87 \times 10^5) = 2.05 \times 10^{17}$ cm^{-3}.

Problems of Chap. 25

25.1 As the problem is independent of time it is $\mathbf{E} = -\operatorname{grad}\varphi$. At any point along the boundary of D, let $\mathbf{n}$ be the unit vector normal to the boundary and pointing inwards, and $\mathbf{t}$ the tangent unit vector such that $\mathbf{k} = \mathbf{t} \wedge \mathbf{n}$, with $\mathbf{k}$ the unit vector of the z axis. Due to the steady-state condition and the hypothesis that one type of carriers prevail, it is $\operatorname{div}\mathbf{J} = 0$, with $\mathbf{J}$ the two-dimensional current field whose components J_x, J_y are defined by (25.34). As the domain is simply connected, there exist a function $\psi(x, y)$ such that $\mathbf{J} = \operatorname{rot}(\psi\,\mathbf{k})$; thus, letting $\mathbf{i}$, $\mathbf{j}$ be the unit vectors of the x and y axes, it is

$$\mathbf{J} = \frac{\partial\psi}{\partial y}\,\mathbf{i} - \frac{\partial\psi}{\partial x}\,\mathbf{j}\,, \qquad \mathbf{k} \wedge \mathbf{J} = \frac{\partial\psi}{\partial y}\,\mathbf{j} + \frac{\partial\psi}{\partial x}\,\mathbf{i} = \operatorname{grad}\psi\,.$$

Taking the divergence of the second equation above yields

$$-\nabla^2\psi = \mathbf{k} \cdot \operatorname{rot}\mathbf{J} = \frac{\partial J_y}{\partial x} - \frac{\partial J_x}{\partial y}\,.$$

Then, one differentiates the first of (25.33) with respect to y, the second one with respect to x, and subtracts the results using $\operatorname{div}\mathbf{J} = 0$; this provides

$$\sigma\left(\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y}\right) = \frac{\partial J_y}{\partial x} - \frac{\partial J_x}{\partial y}\,,$$

where the term in parenthesis is the component of $\operatorname{rot}\mathbf{E}$ along the z axis. On the other hand, it is $\operatorname{rot}\mathbf{E} = 0$ due to the steady-state condition. Combining this with the two equations above yields $\nabla^2\psi = 0$.

25.2 The triad of unit vectors $\mathbf{k} = \mathbf{t}\wedge\mathbf{n}$ is considered here at any point belonging to the interior of D or to its boundary; thus, $\mathbf{t}$ is not necessarily parallel to a boundary. Starting from $\mathbf{E} = E_t\,\mathbf{t} + E_n\,\mathbf{n}$, letting $\theta = \mu_H\,B$, and observing that $\mathbf{J}\wedge\mathbf{k}\cdot\mathbf{t} = \mathbf{J}\cdot\mathbf{k}\wedge\mathbf{t} = \mathbf{J}\cdot\mathbf{n}$ and $\mathbf{J}\wedge\mathbf{k}\cdot\mathbf{n} = \mathbf{J}\cdot\mathbf{k}\wedge\mathbf{n} = -\mathbf{J}\cdot\mathbf{t}$, the two-dimensional current field (25.33) is recast as

$$\sigma\,E_t = J_t + \theta\,J_n\,, \qquad \sigma\,E_n = J_n - \theta\,J_t\,.$$

Suffixes t and n indicate the component along $\mathbf{t}$ or $\mathbf{n}$. Another relation is found from $\mathbf{k}\wedge\mathbf{J} = \operatorname{grad}\psi$, whence

$$\mathbf{k}\wedge\mathbf{J} = \frac{\partial\psi}{\partial t}\,\mathbf{t} + \frac{\partial\psi}{\partial n}\,\mathbf{n}\,, \qquad J_t = \frac{\partial\psi}{\partial n}\,, \qquad J_n = -\frac{\partial\psi}{\partial t}\,,$$

where $\partial\psi/\partial t$ and $\partial\psi/\partial n$ indicate the derivative in the direction of $\mathbf{t}$ or $\mathbf{n}$. The boundary conditions for ψ are obtained from the above. Along a contact it is $E_t = 0$, whence $J_t = -\theta\,J_n$; integrating the above along the ith contact,

$$h\int_{s_i}\frac{\partial\psi}{\partial n}\,\mathrm{d}s' = -\theta\,I_i\,, \qquad \frac{\partial\psi}{\partial n} = -\frac{\theta\,I_i}{h\,s_i}\,,$$

where s_i, h are the contact's length and thickness, s' is the curvilinear abscissa along the contact, and I_i the current through it. As $\mathbf{n}$ points inwards, the direction of I_i is taken inwards as well. The second form of the above equation is due to the fact that due to the smallness of the contact, J_n is uniformly distributed along it; such a relation provides a nonhomogeneous Neumann condition for ψ along the contacts.

Along an insulating boundary it is $J_n = 0$, whence $\partial\psi/\partial t = 0$, $E_n = -\theta\,E_t$, and $J_t = \sigma\,E_t$. This shows that $\psi = \text{const}$ along an insulating boundary, corresponding to a Dirichlet condition, yet unknown.[2] Let ψ_{i-1}, ψ_i be the values of ψ along two insulating boundaries that are separated by the ith contact. Integrating $-\partial\psi/\partial t$ along the contact yields

[2]Note that the boundary conditions of ψ are dual to those of the electric potential φ: the latter is an unknown constant along the contacts, whereas a mixed Neumann condition $E_n = -\theta\,E_t$ holds along the insulating boundaries, to be compared with $J_t = -\theta\,J_n$.

$$\psi_{i-1} - \psi_i = -\int_{s_i} \frac{\partial \psi}{\partial t}\, \mathrm{d}s' = \int_{s_i} J_n\, \mathrm{d}s' = \frac{I_i}{h}\,.$$

From $\operatorname{div}\mathbf{J} = 0$ it also follows $\sum_{i=1}^{n} I_i = 0$ and $\sum_{i=1}^{n} (\psi_{i-1} - \psi_i) = 0$, with n the number of contacts and $\psi_0 = \psi_n$. The differential equation $\nabla^2 \psi = 0$, supplemented with the boundary conditions shown above, provides ψ apart from an additive constant. Prescribing the value of ψ along one of the insulating boundaries, e.g., ψ_0, fixes the Dirichlet conditions along the other insulating boundaries. In this way, the problem is completely defined.

25.3 From the definitions of q_n, k_n, and λ_n given in Sect. 25.5 one finds $|\lambda_n\, q_n| \le 4\,\theta\, I/(h\, H^2\, k_n^2) = 4\,\theta\, I/[h\,\pi^2\,(2\,n+1)^2]$. Adding up over n and combining the result with the relation $\sum_{n=0}^{\infty}(2\,n+1)^{-2} = \pi^2/8$ [59] shows that the series (25.40) is uniformly and absolutely convergent. This makes derivation and integration term-by-term admissible. Differentiating (25.40) term-by-term with respect to y, letting $y = L/2$, and using $k_n H = (2\,n+1)\,\pi$ yield

$$\left[\frac{\partial Q(x,y)}{\partial y}\right]_{y=L/2} = -\frac{\theta\, I}{h\, H}\,\frac{4}{\pi}\,\sum_{n=0}^{\infty}(-1)^n\,\frac{\cos(k_n\, x)}{2\,n+1}\,.$$

The series above converges, but not absolutely; in order to prove that it is in fact the derivative of Q one observes that, due to the boundary condition (25.39), the right-hand side of the above equation must equal $-\theta\, I/(h\, H)$. This is indeed true, thanks to one of the possible series expansions of π valid in the interval $0 \le x/H \le \pi/2$ [59]. As $\cos[(2\,n+1)\,x/H]$ is even, the expansion holds in the interval $-\pi/2 \le x/H \le 0$ as well.

25.4 The calculation is similar to that of Sect. 25.5.2. Taking by way of example the side $x = H/2$, let y_c be the distance of the probes from the x axis (Fig. 25.5). The voltage drop sought is $V_c = \varphi(H/2, y_c) - \varphi(H/2, -y_c)$. Combining the second of (25.34) with $\psi = P + Q$, with (25.40) and (25.41), and with the results of Prob. 25.1, yields

$$-\sigma\, E_y = \frac{\mathrm{d}P}{\mathrm{d}x} + \frac{\partial Q}{\partial x} + \theta\,\frac{\partial Q}{\partial y}\,.$$

Using the above, the voltage V_c reads

$$V_c = -\int_{-y_c}^{+y_c} E_y(H/2, y)\, \mathrm{d}y = \frac{1}{\sigma}\,\int_{-y_c}^{+y_c}\frac{\mathrm{d}P}{\mathrm{d}x}\,\mathrm{d}y = \frac{\psi_1 - \psi_0}{\sigma\, H}\,2\,y_c = \frac{2\,y_c}{\sigma\, h\, H}\,I\,.$$

The term $\partial Q/\partial x$ does not contribute to the integral because Q is odd with respect to y; in turn, the term $\partial Q/\partial y$ does not contribute because Q fulfills the homogeneous Dirichlet condition at $x = H/2$. The ratio $G_c = I/V_c$ is the conductance measured at the probes, whence $\sigma\, h = 2\,y_c G_c/H$; replacing the latter into the first relation of (25.45) yields

$$V_{H0} = -\frac{H}{2\,y_c}\,\mu_H\,B\,V_c\,.$$

The last relation shows that, as anticipated in Sect. 25.5.1, the device thickness h does not explicitly appear in the final calculation. The same result would be found if the product $\sigma\,h$ was measured by the van der Pauw method outlined in Sect. 25.7.

25.5 A fast calculation of $S(y)$ is accomplished by taking the interval $0 \le y \le L/2$, introducing the ratio $\omega = H/L$ and the normalized coordinate $\eta = y/L$; letting $\alpha_n = k_n\,y = (2\,n+1)\,\pi\,\eta/\omega$, with $0 \le \eta \le 1/2$, and $\beta_n = k_n\,L/2 = (2\,n+1)\,\pi/(2\,\omega)$, one finds

$$\frac{V_H(\eta)}{V_{H0}} = 1 - \frac{8}{\pi^2}\sum_{n=0}^{\infty}\frac{\exp(\alpha_n-\beta_n)}{(2n+1)^2}\times\frac{1+\exp(-2\,\alpha_n)}{1+\exp(-2\,\beta_n)}\,,$$

with $-1/2 \le \alpha_n - \beta_n \le 0$. The values in the remaining part of the domain are found from $S(-y) = S(y)$. The above result shows in passing that the local Hall voltage does not separately depend on the device's width and length, but only on the ratio H/L.

References

1. M. Abramowitz, I.A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1970)
2. N.D. Arora, J.R. Hauser, D.J. Roulston, Electron and hole mobilities in silicon as a function of concentration and temperature. IEEE Trans. Electron Dev. **ED-29**(2), 292–295 (1982)
3. N.W. Ashcroft, N.D. Mermin, *Solid State Physics* (Saunders, Philadelphia, 1976)
4. G. Baccarani, *Dispositivi* MOS (Pàtron, Bologna, 1982) (in Italian)
5. G. Baccarani, M. Rudan, G. Spadini, Analytical IGFET model including drift and diffusion currents. Solid-State Electron Dev. **2**, 62–68 (1978)
6. G. Baccarani, M.R. Wordeman, R.H. Dennard, Generalized acaling theory and its application to a 1/4 micrometer MOSFET design. IEEE Trans. Electron Dev. **31**(4), 452–462 (1984)
7. J. Bak, D.J. Newman, *Complex Analysis* (Springer, New York, 1997)
8. H.P. Baltes, R.S. Popoviĉ, Integrated semiconductor magnetic field sensors. Proc. IEEE **74**(11), 1107–1132 (1986)
9. J. Bardeen, W.H. Brattain, The transistor, a semiconductor triode. Phys. Rev. **74**, 230 (1948)
10. R. Becker, *Electromagnetic Fields and Interactions* (Dover, New York, 1982)
11. G. Birkhoff, Tres observaciones sobre el algebra lineal. Univ. Nac. Tucuman Rev. Ser. A **5**, 147–150 (1946) (in Spanish)
12. J.S. Blakemore, *Semiconductor Statistics* (Dover, New York, 1987)
13. D. Bohm, *Quantum Theory* (Dover, New York, 1989)
14. D. Bohm, J.P. Vigier, Model of the causal interpretation of quantum theory in terms of a fluid with irregular fluctuations. Phys. Rev. **96**(1), 208–216 (1954)
15. M. Born, E. Wolf, *Principles of Optics*, 6th edn. (Pergamon Press, London, 1980)
16. A. Bravais, Mémoire sur les systèmes formés par les points distribués régulièrement sur un plan ou dans l'espace. J. Ecole Polytech. **19**(1), 128 (1850) (in French. English version: Memoir 1, Crystallographic Society of America, 1949)

17. R. Brunetti, P. Golinelli, L. Reggiani, M. Rudan, Hot-carrier thermal conductivity for hydrodynamic analyses, in *Proc. of the 26th Solid State Device Research Conference (ESSDERC)*, ed. by G. Baccarani, M. Rudan, pp. 829–832. Edition Frontiers (1996)
18. R. Brunetti, M.C. Vecchi, M. Rudan, Monte Carlo analysis of anisotropy in the transport relaxation times for the hydrodynamic model, in *Fourth Int. Workshop on Computational Electronics (IWCE)*, ed. by C. Gardner, Phoenix (1995)
19. F. Buscemi, E. Piccinini, R. Brunetti, M. Rudan, High-order solution scheme for transport in low-D devices, in *Simulation of Semiconductor Processes and Devices 2014 (SISPAD)*, pp. 161–164, ed. by S. Odanaka, N. Mori (IEEE, Yokohama, September 2014)
20. F. Buscemi, M. Rudan, E. Piccinini, R. Brunetti, A 5th-order method for 1D-device solution, in *2014 Int. Workshop on Computational Electronics (IWCE)*, pp. 1–4, ed. by P. Dollfus (IEEE, Paris, June 2014)
21. C.Y. Chang, S.M. Sze, ULSI *Technology* (McGraw-Hill, New York, 1996)
22. P.C. Chow, Computer solutions to the Schrodinger equation. Am. J. Phys. **40**, 730–734 (1972). http://doi.org/10.1119/1.1986627
23. A.G. Chynoweth, Ionization rates for electrons and holes in silicon. Phys. Rev. **109**(5), 1537 (1958)
24. C. Cohen-Tannoudji, B. Diu, F. Laloë, *Quantum Mechanics* (Wiley, New York, 1977)
25. L. Colalongo, M. Valdinoci, G. Baccarani, A. Pecora, I. Policicchio, G. Fortunato, F. Plais, P. Legagneux, C. Reita, D. Pribat, Analysis and characterization of polycrystalline silicon thin-film transistors, in *Proc. of the 25th Solid State Device Research Conference (ESSDERC)*, pp. 75–78, ed. by H.C. de Graaff, H. van Kranenburg (The Hague, September 1995). Editions Frontieres
26. L. Colalongo, M. Valdinoci, A. Pellegrini, M. Rudan, Dynamic modeling of amorphous- and polycrystalline-silicon devices. IEEE Trans. Electron Dev. **ED-45**, 826–833 (1998)
27. L. Colalongo, M. Valdinoci, M. Rudan, A physically-based analytical model for *a*-Si devices including drift and diffusion currents, in *Simulation of Semiconductor Processes and Devices 1999 (SISPAD)*, pp. 179–182, ed. by K. Taniguchi, N. Nakayama (IEEE, Kyoto, September 1999)
28. L. Colalongo, M. Valdinoci, M. Rudan, G. Baccarani, Charge-sheet analytical model for amorphous silicon TFTs, in *Proc. of the 29th Solid State Device Research Conference (ESSDERC)*, pp. 244–245, ed. by H.E. Maes, R.P. Mertens, G. Declerck, H. Grünbacher, Leuven, September (1999). Edition Frontiers
29. A.H. Compton, A quantum theory of the scattering of x-rays by light elements. Phys. Rev. **21**(5), 483–502 (1923)
30. L.N. Cooper, Bound electron pairs in a degenerate Fermi gas. Phys. Rev. **104**, 1189–1190 (1956)
31. C.R. Crowell, The Richardson constant for thermionic emission in Schottky barrier diodes. Solid-State Electron. **8**(4), 395–399 (1965)
32. C.G. Darwin, R.H. Fowler, Fluctuations in an assembly in statistical equilibrium. Proc. Cambridge Phil. Soc. **21**(391), 730 (1923)
33. S. Datta, *Quantum Transport: Atom to Transistor* (Cambridge University Press, Cambridge, 2006)
34. L. de Broglie, La mécanique ondulatoire et la structure atomique de la matière et du rayonnement. J. Phys. Radium **8**(5), 225–241 (1927) (in French)
35. L. de Broglie, Interpretation of quantum mechanics by the double solution theory, in *Annales de la Fondation Louis de Broglie*, vol. 12(4), pp. 1–23 (Fondation Louis de Broglie, 1987) (translated from the original 1972 version, in French)
36. N.G. de Bruijn, *Asymptotic Methods in Analysis* (Dover, New York, 1981)
37. E. De Castro, *Fondamenti di Elettronica — Fisica elettronica ed elementi di teoria dei dispositivi* (UTET, Torino, 1975) (in Italian)

38. B.E. Deal, A.S. Grove, General relationship for the thermal oxidation of silicon. J. Appl. Phys. **36**, 3770 (1965)
39. P. Debye, E. Hückel, The theory of electrolytes. I. Lowering of freezing point and related phenomena. Physikalische Zeitschrift **24**, 185–206 (1923)
40. P. Dienes, *The Taylor Series* (Dover, New York, 1957)
41. P.A.M. Dirac, The quantum theory of the emission and absorption of radiation. Proc. R. Soc. Lond. A **114**, 243–265 (1927)
42. P.A.M. Dirac, The quantum theory of the electron. Proc. R. Soc. Lond. A **117**, 610–624 (1928)
43. P.A.M. Dirac, *The Principles of Quantum Mechanics*, 4th edn. (Oxford University Press, Oxford, 1992)
44. H.M. Edwards, *Riemann's Zeta Function* (Dover, New York, 2001)
45. A. Einstein, Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt. Annalen der Physik **17**(6), 132–148 (1905) (in German). English translation: D. ter Haar, The Old Quantum Theory, Pergamon Press, pp. 91–107, 1967
46. A. Einstein, On the movement of small particles suspended in a stationary liquid demanded by the molecular-kinetic theory of heat, in *Investigations on the Theory of the Brownian Movement*, chapter 1 (Dover, New York, 1956)
47. E. Fermi, *Thermodynamics* (Dover, New York, 1956)
48. A. Fick. Über Diffusion. Ann. der Physik **94**, 59–86 (1855) (in German). Phil. Mag. **10**, 30 (1855) (in English)
49. G. Floquet, Sur les équations différentielles linéaires à coefficients périodiques. Annales Scientifiques de l'Ec. Norm. Sup. **12**(2), 7–88 (1883) (in French)
50. A. Forghieri, R. Guerrieri, P. Ciampolini, A. Gnudi, M. Rudan, G. Baccarani, A new discretization strategy of the semiconductor equations comprising momentum and energy balance. IEEE Trans. CAD ICAS **7**(2), 231–242 (1988)
51. G. Fortunato, P. Migliorato, Model for the above-threshold characteristics and threshold voltage in polycrystalline silicon transistors. J. Appl. Phys. **68**(5), 2463–2467 (1990)
52. D.H. Frisch, L. Wilets, Development of the Maxwell-Lorentz equations from special relativity and Gauss's Law. Am. J. Phys. **24**, 574–579 (1956)
53. L.J. Giacoletto, Diode and transistor equivalent circuits for transient operation. IEEE J. Solid-State Circuits **4**(2), 80–83 (1969)
54. A. Gnudi, F. Odeh, M. Rudan, Investigation of non-local transport phenomena in small semiconductor devices. Eur. Trans. Telecomm. Relat. Technol. **1**(3), 307–312 (77–82) (1990)
55. I.I. Gol'dman, V.D. Krivchenkov, *Problems in Quantum Mechanics* (Pergamon Press, London, 1961)
56. H. Goldstein, C. Poole, J. Safko, *Classical Mechanics*, 3rd edn. (Addison Wesley, San Francisco, 2002)
57. P. Golinelli, R. Brunetti, L. Varani, L. Reggiani, M. Rudan, Monte Carlo calculation of hot-carrier thermal conductivity in semiconductors, in *Proc. of the Ninth Intl. Conf. on Hot Carriers in Semiconductors (HCIS-IX)*, pp. 405–408, Chicago, 1995, ed. by K. Hess, J.P. Leburton, U. Ravaioli (Plenum Press, New York, 1995)
58. G.A. Gottwald, M. Oliver, Boltzmann's dilemma — an introduction to statistical mechanics via the Kac ring. SIAM Rev. **51**, 613–635 (2009)
59. I.S. Gradshteyn, I.M. Ryzhik, *Table of Integrals, Series, and Products* (Academic Press, New York, 1980)
60. T. Grasser, R. Korsik, C. Jungemann, H. Kosina, S. Selberherr, A non-parabolic six moments model for the simulation of sub-100nm semiconductor devices. J. Comp. Electron. **3**, 183–187 (2004)
61. A.S. Grove, *Physics and Technology of Semiconductor Devices* (Wiley, New York, 1967)
62. G. Hardy, J.E. Littlewood, G. Pólya, *Inequalities*, 2nd edn. (Cambridge University Press, Cambridge, 1952)

63. W. Heisenberg, Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. Zeitschrift für Physik **43**, 172–198 (1927) (in German). English translation: J.A. Wheeler, H. Zurek, *Quantum Theory and Measurement*, pp. 62–84 (Princeton Univ. Press, Princeton, 1983)
64. S.-M. Hong, A.-T. Pham, C. Jungemann, Deterministic solvers for the Boltzmann transport equation, in *Computational Microelectronics*, ed. by S. Selberherr (Springer, Wien-New York, 2011)
65. K. Huang, *Statistical Mechanics*, 2nd edn. (Wiley, New York, 1987)
66. E.L. Ince, *Ordinary Differential Equations* (Dover, New York, 1956)
67. L. Infeld, On a new treatment of some eingenvalue problems. Phys. Rev. **59**, 737–747 (1941)
68. ITRS, *International Technology Roadmap for Semiconductors*. www.itrs2.net
69. J.D. Jackson, *Classical Electrodynamics*, 2nd edn. (Wiley, New York, 1975)
70. C. Jacoboni, *Theory of Electron Transport in Semiconductors*, 1st edn. (Springer, New York, 2010)
71. C. Jacoboni, R. Brunetti, P. Bordone, *Theory of Transport Properties of Semiconductor Nanostructures*, vol. 4 of *Electronics Materials*, ed. by E. Schöll, chapter 3 *"Monte Carlo Simulation of Semiconductor Transport"*, 1st edn., pp. 59–101 (Chapman and Hall, London, 1998)
72. C. Jacoboni, P. Lugli, The Monte Carlo method for semiconductor device simulation, in *Computational Microelectronics*, ed. by S. Selberherr (Springer, Wien-New York, 1989)
73. C. Jacoboni, L. Reggiani, The Monte Carlo method for the solution of charge transport in semiconductors with applications to covalent materials. Rev. Mod. Phys. **55**, 645–705 (1983)
74. W. Jones, N.H. March, *Theoretical Solid State Physics* (Dover, New York, 1973)
75. C. Jungemann, M. Bollhöfer, B. Meinerzhagen, Convergence of the legendre polynomial expansion of the Boltzmann equation for nanoscale devices, in *Proc. of the 35th Solid State Device Research Conference (ESSDERC)*, pp. 341–344, ed. by G. Ghibaudo, T. Skotnicki, S. Cristoloveanu, M. Brillouët, Grenoble (September 2005)
76. M. Kac, Some remarks on the use of probability in classical statistical mechanics. Acad. Roy. Belg. Bull. Cl. Sci. (5) **42**, 356–361 (1956).
77. E.H. Kennard, Zur quantenmechanik einfacher bewegungstypen. Zeitschrift für Physik **44**, 326 (1927) (in German)
78. P. Kiréev, *La Physique des Semiconducteurs* (MIR, Moscou, 1975)
79. C. Kittel, *Introduction to Solid State Physics*, 7th edn. (Wiley, New York, 1953)
80. D.B.M. Klaassen, A unified mobility model for device simulation — I. Model equations and concentration dependence. Solid-State Electr. **35**(7), 953–959 (1992)
81. A.N. Kolmogorov, S.V. Fomin, *Introductory Real Analysis* (Dover, New York, 1975)
82. C. Lanczos, *The Variational Principles in Mechanics*, 4th edn. (Dover, New York, 1970)
83. L. Landau, E. Lifchitz, *Physique statistique* (MIR, Moscou, 1967) (in French)
84. L. Landau, E. Lifchitz, *Mécanique* (MIR, Moscou, 1969) (in French)
85. L. Landau, E. Lifchitz, *Théorie des Champs* (MIR, Moscou, 1970) (in French)
86. A. Landé, *New Foundations of Quantum Mechanics* (Cambridge University Press, Cambridge, 1965)
87. A. Landé, Solution of the Gibbs entropy paradox. Phil. Sci. **32**(2), 192–193 (1965)
88. R.A. Levy (ed.), *Microelectronics Materials and Processes*, vol. E-164 of NATO ASI (Kluwer, Boston, 1986)
89. M.J. Lighthill, *Fourier Analysis and Generalized Functions* (Cambridge University Press, Cambridge, 1962)
90. H.J. Lippmann, F. Kuhrt, Der Geometrieeinfluss auf den Hall Effekt bei rechteckig ein Halbleiterplatten. Z. f. Naturf. **13a**(2), 474–483 (1958)
91. C. Lombardi, S. Manzini, A. Saporito, M. Vanzi, A physically based mobility model for numerical simulation of nonplanar devices. IEEE Trans. CAD **7**(11), 1164–1171 (1988)
92. A.M. Lyapounov, Problème Général de la Stabilité du Mouvement. Ann. Fac. Sc. Univ. Toulouse **9**(2), 203–475 (1907) (in French)

93. W. Maly, *Atlas of* IC *Technologies: An Introduction to* VLSI *Processes* (The Benjamin/Cummings Publishing, Menlo Park, 1987)
94. M. Marcus, H. Minc, *A Survey of Matrix Theory and Matrix Inequalities* (Dover, New York, 1992)
95. E. Merzbacher, *Quantum Mechanics* (Wiley, New York, 1970)
96. A. Messiah, *Mécanique Quantique* (Dunod, Paris, 1969) (in French). English edition: *Quantum Mechanics* (Dover, New York, 1999)
97. R.S. Muller, T.I. Kamins, *Device Electronics for Integrated Circuits*, 2nd edn. (Wiley, New York, 1986)
98. M. Muskat, E. Hutchisson, Symmetry of the transmission coefficients for the passage of particles through potential barriers. Proc. Natl. Acad. Sci. USA **23**, 197–201 (1937)
99. D.A. Neamen, *Semiconductor Physics and Devices* (Irwin, Homewood, 1992)
100. B. Numerov, Publ. Obs. Central Astrophys. **2**, 188 (1933) (in Russian)
101. W. Pauli, The connection between spin and statistics. Phys. Rev. **58**, 716–722 (1940)
102. R. Peierls, *Bird of Passage: Recollections of a Physicist*. Princeton Legacy Library (Princeton University Press, Princeton, 1985)
103. A. Pellegrini, L. Colalongo, M. Valdinoci, M. Rudan, AC analysis of amorphous silicon devices. IEEE Trans. CAD **15**(11), 1324–1331 (1996)
104. M. Planck, On an improvement of Wien's equation for the spectrum, in *The Old Quantum Theory*, p. 79 (Pergamon Press, London, 1967)
105. S. Reggiani, M. Rudan, E. Gnani, G. Baccarani, Investigation about the high-temperature impact-ionization coefficient in silicon, in *Proc. of the 34th Solid State Device Research Conference (ESSDERC)*, pp. 245–248, ed. by R.P. Mertens, C.L. Claeys (IEEE, Leuven, 2004)
106. S. Reggiani, M.C. Vecchi, A. Greiner, M. Rudan, Modeling hole surface- and bulk-mobility in the frame of a spherical-harmonics solution of the BTE, in *Simulation of Semiconductor Processes and Devices 1998 (SISPAD)*, pp. 316–319, ed. by K. De Meyer, S. Biesemans (Springer, Wien, Austria, 1998)
107. F. Reif, *Fundamentals of Statistical and Thermal Physics* (McGraw-Hill, New York, 1985)
108. M. Rudan, G. Baccarani, On the structure and closure condition of the hydrodynamic model. VLSI Des. (Special Issue, J. Jerome, Ed.) **3**(2), 115–129 (1995)
109. M. Rudan, E. Gnani, S. Reggiani, G. Baccarani, The density-gradient correction as a disguised pilot wave of de Broglie, in *Simulation of Semiconductor Processes and Devices 2004 (SISPAD)*, pp. 13–16, ed. by G. Wachutka, G. Schrag (Springer, Munich, 2004)
110. M. Rudan, A. Gnudi, E. Gnani, S. Reggiani, G. Baccarani, Improving the accuracy of the Schrödinger-Poisson solution in CNWs and CNTs, in *Simulation of Semiconductor Processes and Devices 2010 (SISPAD)*, pp. 307–310, ed. by G. Baccarani, M. Rudan (IEEE, Bologna, 2010)
111. M. Rudan, A. Gnudi, W. Quade, Process and device modeling for microelectronics, chapter 2, in *"A Generalized Approach to the Hydrodynamic Model of Semiconductor Equations"*, pp. 109–154, ed. by G. Baccarani (Elsevier, Amsterdam, 1993)
112. M. Rudan, M. Lorenzini, R. Brunetti, *Theory of Transport Properties of Semiconductor Nanostructures*, vol. 4 of *Electronics Materials, E. Schöll, Ed.*, chapter 2 *"Hydrodynamic Simulation of Semiconductor Devices"*, 1st edn., pp. 27–57 (Chapman and Hall, London, 1998)
113. M. Rudan, F. Odeh, Multi-dimensional discretization scheme for the hydrodynamic model of semiconductor devices. COMPEL **5**(3), 149–183 (1986)
114. M. Rudan, F. Odeh, J. White, Numerical solution of the hydrodynamic model for a one-dimensional semiconductor device. COMPEL **6**(3), 151–170 (1987)
115. M. Rudan, S. Reggiani, E. Gnani, G. Baccarani, C. Corvasce, D. Barlini, M. Ciappa, W. Fichtner, M. Denison, N. Jensen, G. Groos, M. Stecher, Theory and experimental validation of a new analytical model for the position-dependent hall voltage in devices with arbitrary aspect ratio. IEEE Trans. Electron Dev. **53**(2), 314–322 (2006)

116. M. Rudan, M.C. Vecchi, D. Ventura, *Mathematical Problems in Semiconductor Physics*, chapter *"The Hydrodynamic Model in Semiconductors — Coefficient Calculation for the Conduction Band of Silicon"*, pp. 186–214. Number 340 in Pitman Research Notes in Mathematical Series, ed. by P. Marcati, P.A. Markowich, R. Natalini (Longman, Harlow, 1995)
117. D.L. Scharfetter, H.K. Gummel, Large-signal analysis of a silicon Read diode oscillator. IEEE Trans. Electron Dev. **ED-16**(1), 64–77 (1969)
118. E. Schrödinger, Quantisierung als eigenwertproblem (erste mitteilung). Annalen der Physik **384**(4), 361–376 (1926) (in German)
119. E. Schrödinger, *Statistical Thermodynamics* (Dover, New York, 1989)
120. M. Schwartz, *Principles of Electrodynamics* (Dover, New York, 1987)
121. W. Shockley, The theory of *p-n* junctions in semiconductors and *p-n* junction transistors. Bell Syst. Tech. J. **28**, 435 (1949)
122. W. Shockley, *Electrons and Holes in Semiconductors* (Van Nostrand Book Co., New York, 1950)
123. W. Shockley, H.J. Queisser, Detailed balance limit of efficiency of *p-n* junction solar cells. J. Appl. Phys. **32**, 510 (1961)
124. J.C. Slater, *Quantum Theory of Matter* (McGraw-Hill, New York, 1968)
125. J.W. Slotboom, H.C. DeGraaff, Measurement of bandgap narrowing in silicon bipolar transistors. Solid-State Electron. **19**(2), 857–862 (1976)
126. J.W. Slotboom, H.C. DeGraaff, Bandgap narrowing in silicon bipolar transistors. IEEE Trans. Electron Dev. **28**(8), 1123–1125 (1977)
127. J.C.F. Sturm, Mémoire sur les équations différentielles linéaires du deuxième ordre. J. de Mathématiques Pures et Appliquées **1**, 106–186 (1836) (in French)
128. S.M. Sze, *Physics of Semiconductor Devices* (Wiley, New York, 1981)
129. S.M. Sze, *Semiconductor Devices — Physics and Technology* (Wiley, New York, 1985)
130. S.M. Sze, VLSI *Technology* (McGraw-Hill, New York, 1988)
131. S. Takagi, A. Toriumi, M. Iwase, H. Tango, On the Universality of inversion layer mobility in Si MOSFET's: Part I — effects of substrate impurity concentration. IEEE Trans. Electron Dev. **41**(12), 2357–2362 (1994)
132. S. Taschini, R. Brunetti, M. Rudan, Particle and energy fluxes in semiconductors: full-band hydrodynamic equations and the thermodynamic limit. Phys. Rev. B **60**(19), 13582–13591 (1999)
133. D. ter Haar, On a heuristic point of view about the creation and conversion of light, in *The Old Quantum Theory*, pp. 91–107 (Pergamon Press, London, 1967)
134. R. Thoma, A. Emunds, B. Meinerzhagen, H.-J. Peifer, W.L. Engl, Hydrodynamic equations for semiconductors with nonparabolic band structure. IEEE Trans. Electron Dev. **38**(6), 1343–1353 (1991)
135. R.C. Tolman, Note on the derivation from the principle of relativity of the fifth fundamental equation of the Maxwell-Lorentz theory. Phil. Mag. S. 6 **21**(123), 296–301 (1911)
136. R.C. Tolman, *Statistical Mechanics* (Dover, NewYork, 1979)
137. M. Valdinoci, A. Gnudi, M. Rudan, G. Fortunato, Analysis of amorphous silicon devices, in *Int. Workshop on Numerical Modeling of Processes and Devices for Integrated Circuits (NUPAD V)*, pp. 19–22, Honolulu (June 1994)
138. M. Valdinoci, D. Ventura, M.C. Vecchi, M. Rudan, G. Baccarani, F. Illien, A. Stricker, L. Zullino, Impact-ionization in silicon at large operating temperature, in *Simulation of Semiconductor Processes and Devices 1999 (SISPAD)*, pp. 27–30, ed. by K. Taniguchi, N. Nakayama (IEEE, Kyoto, September 1999)
139. L.J. van der Pauw, A method of measuring specific resistivity and Hall effect of discs of arbitrary shape. Philips Res. Rep. **13**(1), 1–9 (1958)
140. M.C. Vecchi, M. Rudan, Modeling electron and hole transport with full-band structure effects by means of the spherical-harmonics expansion ot the BTE. IEEE Trans. Electron Dev. **45**(1), 230–238 (1998)

141. G.H. Wannier, *Statistical Physics* (Dover, New York, 1996)
142. J. Ward Brown, R.V. Churchill, *Complex Variables and Applications* (McGraw-Hill, New York, 1996)
143. R. Weinstock, *Calculus of Variations* (Dover, New York, 1974)
144. H. Weyl, Quantenmechanik und gruppentheorie. Zeitschrift für Physik **46**(1–2), 1–46 (1927) (in German)
145. J.A. Wheeler, W.H. Zurek, *Quantum Theory and Measurement*, pp. 62–84 (Princeton Univ. Press, Princeton, 1983)
146. N. Wiener, *The Fourier Integral and Certain of Its Applications* (Dover, New York, 1958)
147. D.M. Young, *Iterative Solution of Large Linear Systems* (Academic Press, New York, 1971)

Index

M. Rudan, *Physics of Semiconductor Devices*, DOI 10.1007/978-3-319-63154-7

F

M

N