Quantum Mechanics: Genesis and Achievements

Alexander Komech

Quantum Mechanics: Genesis and Achievements

Springer

Alexander Komech
Faculty of Mathematics
University of Vienna
Vienna, Austria

Resume The intent of this book is a concise introduction to the field of nonrelativistic and relativistic quantum mechanics. A historical account of the Heisenberg and the Schrödinger theories is given, and traditional applications to the hydrogen atom are outlined. All details of calculations are given and supplemented with related materials from the Maxwell Electrodynamics and Special Relativity.
The book can be used as a source for a two-semester lecture course on Nonrelativistic Quantum Mechanics, as well as for a one-semester course on either the Relativistic Quantum Mechanics, Classical Electrodynamics, or the Classical Field Theory.

ISBN 978-94-007-5541-3 ISBN 978-94-007-5542-0 (eBook)
DOI 10.1007/978-94-007-5542-0
Springer Dordrecht Heidelberg New York London

Library of Congress Control Number: 2012951292

© Springer Science+Business Media Dordrecht 2013
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

To the memory of
Vera and Alexey Komech,
Esther Braginskaya, and Mark Vishik

Preface

This book is based on a three-semester course of lectures delivered by the author at the Faculty of Mathematics of the Vienna University. For preliminary versions of the manuscript, see [113, 114].

The main emphasis of this book is on nonrelativistic and relativistic quantum mechanics with standard applications to the hydrogen atom. Our main intention is to present the quantum mechanics in a comprehensive manner, accessible for a mathematician. The exposition is formalized (whenever possible) on the basis of coupled Maxwell–Schrödinger and Maxwell–Dirac equations. This intention agrees with Hilbert's 6th Problem ("Axiomatize Theoretical Physics"), and Heisenberg's nonlinear programme [83, 84].

Our exposition starts with a chronological analysis of crucial empirical observations and their theoretical systematization, explaining *inter alia* the motivation behind the Heisenberg and Schrödinger equations. The introduction of quantum observables stems from the agreement with corresponding classical observables for short wavelength solutions, Hamilton–Jacobi's theorem being taken into account. Also, the relation between the quantum observables and the Noether symmetry theory is discussed. The Lagrangian formalism is used as a fundamental unifying principle to lay the basis for introduction of the coupled Maxwell–Schrödinger and Maxwell–Dirac equations.

Of course, the equations and observables could be (and formally should be) accepted as axioms. On the other hand, it is crucially important to know the experimental and mathematical facts behind quantum formalism to embed it in the whole of physics.

Moreover, the modern form of quantum theory seems to be far from its completeness, like geometry in pre-Euclidean era. It is therefore particularly important to understand the degree of confidence to individual constituents of the quantum formalism. This is why we pay so much attention to the origin and motivation of the formalism.

The hydrogen spectrum and the atom radiation are calculated with all detail. Parallels between quantum and classical description are traced everywhere to motivate the introduction of quantum differential cross section, magnetic moment, etc.

The scattering problems are solved by application of the perturbation procedure to the coupled Maxwell–Schrödinger equations. We point out some deficiency in the perturbation procedure, which should be fixed with a nonperturbative approach, however this correction is still an open problem.

The introduction of the electron spin is discussed in detail from experimental and theoretical point of view. We calculate the Landé formula for the *gyromagnetic ratio* via the spin-orbital interaction of Russell–Saunders, which explains the Einstein–de Haas experiment and the anomalous Zeeman effect. Further we prove the relativistic covariance of the Dirac equation, obtain the corresponding intrinsic spin momentum, and the corresponding nonrelativistic approximations. Finally, we calculate the hydrogen spectrum via the Dirac equation.

We make explicit invoked assumptions and approximations, and discuss a plausible treatment of some logical leaps in the theory. However, we did not try to establish new rigorous results. Generally, our exposition is not mathematically rigorous. For example, we do not distinguish between Hermitian symmetric and the selfadjoint operators, even though the spectral resolution is used repeatedly.

In appendices (Chaps. 12 and 13), we explain related details from Classical Electrodynamics and Special Relativity, Geometrical Optics, the Hamilton–Jacobi theorem, an updated version of the Noether theory of currents, and the limiting amplitude principle.

In Chap. 14, we collect classical calculations lying in the base of the 'old quantum mechanics'.

Further Reading Our main goal is to give a concise explanation of mathematical principles of Quantum Mechanics. More technical details and a systematic comparison with experimental data can be found in [7, 11, 12, 20, 23, 34, 63, 75, 81, 130, 131, 145, 160, 171, 179, 191]. The books [46, 79, 89, 143] and [93, 185], respectively, explain basic concepts of Quantum Mechanics and Classical Electrodynamics. A suitable introduction to the mathematical theory of the Schrödinger equation is contained in [10].

We develop the methods of quantum mechanics for the hydrogen atom which can be extended to other one-electron atoms (lithium, sodium, potassium atoms, etc), and do not touch multi-electron problems of quantum chemistry [34, 41, 178, 187]. We also do not touch the Stability of Matter [31, 134], the Quantum Electrodynamics and Quantum Field Theory [13, 14, 33, 64, 77, 85, 137, 138, 158, 159, 163, 189, 195, 196].

Acknowledgements Over a period of years the author has been pleased to have many stimulating discussions with A. Shnirelman and H. Spohn. The author is grateful to E. Kopylova for reading and checking the manuscript, and to C. Adam for his help in writing preliminary versions of the Exercises (Sects. 14.1–14.6). The author is indebted to the Faculty of Mathematics of Vienna University, and the Institute for the Information Transmission Problems of the Russian Academy of Sciences for providing congenial facilities for the work.

The author is also grateful to the Max Planck Institute for Mathematical Sciences (Leipzig), the München Technical University, and the Team ONDES/POEMS (INRIA, Rocquencourt) for their hospitality.

The work was supported in part by the Department of Mechanics and Mathematics of the Moscow State University, by the Alexander von Humboldt Foundation, by the Austrian Science Fund (FWF) (projects Nos. P19138-N13 and P22198-N13), and by grants from the Deutsche Forschungsgemeinschaft and the Russian Fund for Basic Research.

Leipzig–Moscow–München–Rocquencourt–Vienna A. Komech

Contents

Chapter 1
Genesis of Quantum Mechanics

Abstract Quantum mechanics rests on the centuries-long development of the theory of structure and stability of matter. The development of quantum mechanics is supposed to have started with the concept of atom, and further progressed in parallel with the development of chemistry, mechanics, optics and electrodynamics. As long ago as at the end of the 19th century, the atoms were already considered as vibrating systems composed of positive and negative charged particles. However, the pure electromagnetic theory of atom stability turned out to be inconsistent, and the key ideas came from the thermodynamics of radiation.

The fundamental principles of quantum mechanics can be perceived only from systematic analysis of empirical observations and their theoretic classification. We begin with early discoveries in chemistry and in spectral observations, Thomson's discovery of the electron, Lorentz's electron theory, and Abraham's mass-energy identification.

Further we dwell upon Kirchhoff's laws of radiation, Rayleigh–Jeans's and Planck's theories of black-body radiation, and the 'old' quantum mechanics of Niels Bohr. Next we show how these achievements emerged in the development of Heisenberg's matrix theory of quantum mechanics, de Broglie's wave–particle dualism and the generalizations thereof by Schrödinger.

Finally, we describe the quasiclassical asymptotics ('geometrical optics') for short-wavelength solutions of the Schrödinger equation, which provides a 'bridge' between the quantum and classical description of matter.

1.1 Atoms and Spectra

Historically, the first information about the structure of matter was obtained from chemical reactions, melting processes, and scattering of radiation by matter. The latter phenomenon occurs commonly in the theory of the rainbow and from the colors of burning carbon.

1.1.1 Concept of Atom

The explicit statement that matter is composed of indivisible particles called atoms (from the Greek word 'uncuttable') can be traced back to the Greek philosophers

A. Komech, *Quantum Mechanics: Genesis and Achievements*,
DOI 10.1007/978-94-007-5542-0_1,

Leucippus and Democritus of Abdera (taught in the latter part of the fifth century B.C.), who taught that all matter is made up of atoms and empty space. On the one hand, the concept of an atom reflects the divisibility of a homogeneous substance, and on the other, the existence of a minimal portion of the matter that retains its peculiarity. In both these aspects, the concept of an atom resembles that of fabric. Moreover, the words 'matter' and 'fabric' coincide in several languages reflecting this similarity.

However, this concept has become widely accepted only after the appearance of Dalton's law of multiple proportions (1803).

1.1.2 Early Spectra Observations

The first spectral analysis of sunlight was made in 1666 by Newton, who is reputed to have discovered that a prism could disperse white light into its constituent (rainbow) colors and a further prism recombine the colors to restore the white light. More systematic observations were done by Melville in 1752 on burning substances in flame.

The wave nature of light was suggested by Hooke in 1660's and Huygens in 1678. Different colors were shown to have different wavelengths by Young in 1801 using diffraction gratings. The first man-made diffraction grating was made around 1785 by an American astronomer David Rittenhouse, even though its underlying principle was discovered by Gregory in 1667s by sending sunlight through bird feather and observing the produced diffraction pattern.

In 1814, Fraunhofer compared light from the Sun with radiation from sodium chloride in flame and noticed dark lines in the Sun's spectrum. With the aid of diffraction gratings, Fraunhofer eventually observed 574 mysterious dark spectral lines.

The spectrum of hydrogen, which turned out to be crucial in providing first insight into atomic structure over half a century later, was observed by Ångström in 1853. David Alter described the spectrum of hydrogen and of other gases in 1855.

The first systematic study of radiation and spectra was made by Bunsen and Kirchhoff in the early 1860s, who determined the spectra of more than 20 elements. Their investigation made evident that the spectra are element-specific. This fact was the starting point for the *spectral analysis* of chemical substances.

Wavelengths of more than 1000 spectral lines in the solar spectrum were measured with very high precision by Ångström in 1868.

1.1.3 Resonance Nature of Spectra

In 1849, Foucault examined the light spectrum produced by the voltaic arc between carbon poles. He saw a bright double yellow line at exactly the same wavelength as Fraunhofer's dark D-line in the solar spectrum. Proceeding further, Foucault sent the sun's light through the arc and then through a prism to have observed that the D-lines in the spectrum were even darker than usual. After applying other sources, he pointed out the arc emits the D-line, but absorbs it when it is emitted by a different radiator [65].

Foucault discovered that if a brighter radiator giving a continuous spectrum was placed behind the arc, so that the spectrum of this radiator was seen through the arc, then the D-line appeared dark, just as in the solar spectrum [192, p. 369].

Similarly, Kirchhoff and Bunsen identified some of the Fraunhofer dark lines with the spectral lines of oxygen. This means that the dark lines are the result of absorption of the sunlight by the oxygen atoms in the Earth atmosphere, as explained by Kirchhoff and Bunsen in 1859.

The agreement of the emitted and absorbed frequencies, as discovered by Foucault, Kirchhoff and Bunsen, suggested the resonance nature of the absorption and the oscillatory model of atoms: in 1852 Stokes suggested that atomic spectra are produced by periodic vibrations of constituent parts of atoms and molecules.

1.1.4 Combination Principle

Johann Balmer spent many years on deciphering Ångström's numerical data on the hydrogen spectra. His efforts culminated in 1885 in the astonishing discovery of empirical formula for the wavelength of some hydrogen spectral lines,

$$\omega_{mn} = \omega_m - \omega_n, \tag{1.1}$$

with $m \geq n = 2$ and the *spectral terms*

$$\omega_n = -B/n^2, \quad n = 1, 2, \ldots, \qquad B \approx 104 \times 10^{30}\ \mathrm{s}^{-1}. \tag{1.2}$$

In 1888, Rydberg suggested the general *Combination Principle* (1.1) with all $m, n = 1, 2, \ldots$ for the hydrogen spectral lines. This suggestion was experimentally confirmed by Paschen (1908), Lyman (1909), and Brackett (1914), who discovered the hydrogen spectral lines (1.1) corresponding to $n = 3$, $n = 1$, and $n = 4$, respectively. The combination principle, as further extended to all elements by Ritz in 1908, is one of the underlying principles for *Heisenberg's matrix theory of quantum mechanics*; it was justified later by the *Schrödinger theory*.

1.2 Electrodynamics

The next step toward understanding the structure of matter was provided by electromagnetic phenomena: magnetization of iron needles, electric discharge, the Zeeman splitting of spectral lines in a magnetic field, etc.

1.2.1 Ampère Theory of Magnetization

The magnetization of iron needles was known in China since 2000 BC. In 1820, Ersted discovered the magnetic field generated by electric current. In the same year, Ampère, Biot, Savart and Laplace examined the relation between current and magnetic field to discover what we now call the *Ampère law* and the *Biot–Savart–Laplace law*. These investigations suggested to Ampère in 1820 the presence of atomic and molecular circular currents; with this basis the theory was capable of explaining the magnetization of iron needles due to reorientation of microscopic currents under the influence of the Earth's magnetic field.

1.2.2 Maxwell Electrodynamics

In 1855–1861, Maxwell put forward the equations for the electric and magnetic fields $\mathbf{E}(t,\mathbf{x})$, $\mathbf{B}(t,\mathbf{x})$ as generated by density $\rho(t,\mathbf{x})$ and current $\mathbf{j}(t,\mathbf{x})$. The equations result from a mathematical synthesis of the empirical Coulomb's, Biot–Savart–Laplace's, and Faraday's laws.

The Maxwell Equations

Written in *unrationalized Gaussian units* (cgs), the Maxwell equations are as follows:

$$\begin{cases} \operatorname{div}\mathbf{E}(t,\mathbf{x}) = 4\pi\rho(t,\mathbf{x}), & \operatorname{curl}\mathbf{E}(t,\mathbf{x}) = -\frac{1}{c}\dot{\mathbf{B}}(t,\mathbf{x}), \\ \operatorname{div}\mathbf{B}(t,\mathbf{x}) = 0, & \operatorname{curl}\mathbf{B}(t,\mathbf{x}) = \frac{4\pi}{c}\mathbf{j}(t,\mathbf{x}) + \frac{1}{c}\dot{\mathbf{E}}(t,\mathbf{x}); \end{cases} \tag{1.3}$$

here c is the speed of light in vacuum, and the dot indicates differentiation with respects to time. The Maxwell equations imply the *continuity equation for charge*

$$\dot{\rho}(t,\mathbf{x}) + \operatorname{div}\mathbf{j}(t,\mathbf{x}) = 0 \tag{1.4}$$

which is equivalent to the charge conservation law. Of course, this equivalence cannot be considered as the proof of the (experimentally known) charge conservation, since it was used in derivation of the Maxwell equations.

The Lorentz Force

In accordance with Ampère's law (1820), the force acting on a current element in a magnetic field is as follows (in unrationalized Gaussian units):

$$\mathbf{F} = q\frac{\mathbf{v}}{c} \times \mathbf{B}(\mathbf{x}); \tag{1.5}$$

here q is the charge of a particle, $\mathbf{v}$ is its velocity, and $\mathbf{x}$ is its position. Taking into account the action of electric field, this gives the *Lorentz equation* for a charged particle

$$m\ddot{\mathbf{x}}(t) = q\left[\mathbf{E}\big(t, \mathbf{x}(t)\big) + \frac{\dot{\mathbf{x}}(t)}{c} \times \mathbf{B}\big(t, \mathbf{x}(t)\big)\right]; \tag{1.6}$$

here m stands for the electron mass, and $\mathbf{x}(t)$, for its trajectory.

1.2.3 Cathode Rays: Thomson's Discovery of Electron

Cathode rays indicate the existence of electric flux in the vacuum. These rays were discovered first by Geissler and Plucker around 1855 in a *vacuum tube* (glass bulb with two electrodes) with better vacuum conditions than those used by Faraday in 1836–1838.

Plucker produced "…a bright stream-like glow between the electrodes, which was much brighter than any achieved in previous experiments. He found that the glow responded to a magnetic field: it could be moved by a powerful magnet. The discovery indicated that the stream crossing the vacuum was composed of charged particles rather than rays.

The next scientist to conduct important research using vacuum tubes was Hittorf in 1869. A student of Plucker, Hittorf further improved the method for creating a vacuum within glass tubes of his own design. He observed that the luminescent glow increased dramatically as the pressure within the tube continued to decrease. He also placed tiny obstacles inside the tube in the path between the two electrodes. When a current was applied, the glow would be partially obscured by these obstacles, casting shadows. This further confirmed the idea that the glow was caused by a particle emission…" [139].

The Geissler vacuum tube was considerably improved by Crookes in 1875, whose works supported the previous observations on cathode rays. Moreover, he "…installed tiny vanes within his tubes. As the current was applied, the vanes would turn slightly, as if they were blown by a gust of wind…" [139].

The rays were emitted from the cathode, and their deflection in a magnetic field demonstrated their negative charge. French physicist Jean Baptiste Perrin showed in 1895 that the rays deposit negative electric charge on a charge collector placed inside cathode-ray tube [152].

"…Some physicists, like Goldstein, Hertz, and Lenard, thought that this phenomenon is like the light, due to vibrations of the 'ether' or even that it is the light of short wavelength. It is easily understood that such rays may have a rectilinear path, excite phosphorescence, and affect photographic plates. Others, like Crookes and J. J. Thomson, thought that these rays are formed by the negatively charged matter moving with great velocity, and on this hypothesis their mechanical properties, as well as the manner in which they become curved in a magnetic field, are readily explicable…" [152].

The discussion was cut short after the famous series of experiments conducted by Thomson in 1893–1897 [182, 183], who succeeded in deflecting cathode rays by an electrostatic field, thus showing that they consist of negatively charged particles. He also systematized all previous observations and demonstrated the particle-like behavior of the cathode rays, which is in agreement with the Lorentz equation (1.6). Since this equation is capable of yielding only the ratio q/m, one cannot separately obtain q and m from observations of interactions between particles and the Maxwell field.

Cathode rays were shown by Thomson as being identical with a beam of negatively charged particles (*electrons*). He also introduced the name *electron* for these particles (the term electron had been used earlier, and Thomson preferred at first to call it simply a ‘corpuscle’ [69, p. 311]). This study has led to the first measurement of the ratio

$$|e|/\mathrm{m} = 1.76 \times 10^7 \text{ emu/g}, \tag{1.7}$$

where $e < 0$ is the electron charge, and m is its mass. This ratio proved to be fairly close to its present value, which is $1.75882012(15) \times 10^7$ emu/g.

Magnetic deflection of the cathode rays was also observed by Kauffmann in 1901–1906 [98], who also evaluated the ratio e/m; the resulting value has occurred to be quite close to that obtained by Thomson.

On the other hand, subsequent development suggested that cathode rays should be looked upon not as a beam of particles but rather as *waves*, whose diffraction was observed in 1924–1927 by Davisson and Germer [40].

The contradiction between these two wave/particle points of view was removed by geometrical optics for short-wavelength solutions of the Schrödinger equation (see Sect. 3.2.2).

1.2.4 Elementary Electric Charge

Faraday Laws of Electrolysis

In 1834, Faraday estimated $F \approx 96{,}500$ C/mol for the absolute value of charge carried by ions *of valency one* in a mole of a chemical substance (the modern value is $F = 96{,}485.3365(21)$ C/mol). The value $N_A \approx 7.15 \times 10^{23}\ \mathrm{mol}^{-1}$ for the Avogadro’s number was estimated in 1909 by Perrin on the basis of Einstein’s theory of Brownian motion [144, 165] (the modern value is $N_A = 6.022141 \times 10^{23}$/mol). This gives the value of the electron’s charge as the ratio of the Faraday constant F to the Avogadro number N_A; using the values obtained by Faraday and Perrin, this gives

$$e = -\frac{F}{N_A} \approx -1.34 \times 10^{-19}\ \mathrm{C} = -1.34 \times 10^{-20}\ \mathrm{emu} \tag{1.8}$$

(the modern value is $e = -1.6021766 \times 10^{-20}$ emu). As a result, the ratio (1.7) gives the corresponding value of the electron mass:

$$\mathrm{m} \approx \frac{e}{1.76 \times 10^{7}\ \mathrm{emu/g}} \approx 0.76 \times 10^{-27}\ \mathrm{g} \tag{1.9}$$

(the modern value is $\mathrm{m} = 9.10938188 \times 10^{-28}$ g). On the other hand, the mass of a hydrogen atom is $M_H = 1/N_A$. Hence, the electron mass constitutes a very small part of the atomic mass: using the Faraday value of F and the Thomson ratio (1.7), it is found that

$$\frac{M_H}{\mathrm{m}} = \frac{1}{\mathrm{m} \cdot N_A} = \frac{1.76 \times 10^{7}}{F} \approx 1820 \pm 10, \tag{1.10}$$

which is quite close to the modern value of 1837.

Millikan's 'Oil Drop' Experiment

In 1909–1914, Millikan, celebrated for his 'oil-drop' method for finding the charge on the electron, has measured elementary electric charges by taking into account the balance between the downward gravitational force with the upward buoyant and electric forces on tiny charged droplets of oil suspended between two metal electrodes [141]. By repeating the experiment for several droplets, Millikan was able to confirm that the charges of droplets differ by *integer* multiples of some fundamental value, and calculated it to be $1.5924(17) \times 10^{-19}$ C; this value agrees (within 1% accuracy) with the currently accepted value

$$|e| = 1.602176487(40) \times 10^{-19}\ \mathrm{C}. \tag{1.11}$$

The error in Millikan's result was caused by systematic errors in the calculation of air viscosity.

1.2.5 The Zeeman Effect

In 1895, Zeeman [193] discovered the splitting of spectral lines of atoms and molecules in a magnetic field. This reflected the fact that the atomic and molecular spectra are produced by moving charged particles—a further argument to support Stokes's ideas on the oscillatory nature of the spectra.

1.2.6 Lorentz's Theory of Electrons

Observation of cathode rays is further evidence for the presence of electric charges in atoms; this was already known after Faraday's discovery of the laws of electrolysis. The presence of electric currents in atoms was suggested earlier by Ampère's theory of magnetization.

These facts supported the view on the nature of the electromagnetic theory of matter as arising from collection of charged particles interacting with the electric and magnetic fields, $\mathbf{E}(t,\mathbf{x})$ and $\mathbf{B}(t,\mathbf{x})$, which are governed by the Maxwell equations (1.3). In particular, Lorentz's book [136] summarizes almost all of the work on the 'electron theory' of matter up to 1878. The theory is based on the Maxwell electrodynamics and on some specific assumptions about the atomic structure.

Namely, Lorentz used the classical model of atom as a damped oscillator subject to a Maxwell field (cf. (1.6)):

$$\mathrm{m}\big[\ddot{\mathbf{x}}(t)+\gamma\dot{\mathbf{x}}(t)+\omega_0^2\mathbf{x}(t)\big]=e\Bigg[\mathbf{E}(t,\mathbf{x}(t)+\frac{\dot{\mathbf{x}}(t)}{c}\times\mathbf{B}(t,\mathbf{x}(t)\Bigg]. \tag{1.12}$$

The friction term $\gamma\dot{\mathbf{x}}$ was introduced to secure stability under external perturbation $\mathbf{E}=\mathbf{E}_0e^{-i\omega t}$ for all frequencies $\omega\in\mathbb{R}$. The friction models a radiation of energy to infinity.

The Lorentz theory has partially explained many fundamental phenomena, *viz.*, polarization and dispersion (Lecture 14.4), the normal Zeeman effect (Lecture 14.5), etc.

1.2.7 Abraham: Mass–Energy Identification

The Thomson experiments and Lorentz's theory suggested for a purely electromagnetic description of the matter. In particular, stability and dynamics of the electron were to be explained by the Maxwell electrodynamics.

In 1902, Abraham analyzed the dynamics of moving electron [1, 2]. The main idea was that the comoving field carries the energy and momentum, as defined by (12.94) and (12.103) on $\Omega=\mathbb{R}^3$, as follows:

$$\begin{aligned}\mathcal{E}(\mathbf{v})&=\frac{1}{8\pi}\int\big[\mathbf{E}_\mathbf{v}^2(\mathbf{x}-\mathbf{v}t)+\mathbf{B}_\mathbf{v}^2(\mathbf{x}-\mathbf{v}t)\big]\,d\mathbf{x},\\ \mathbf{G}(\mathbf{v})&=\frac{1}{4\pi c}\int\mathbf{E}_\mathbf{v}(\mathbf{x}-\mathbf{v}t)\times\mathbf{B}_\mathbf{v}(\mathbf{x}-\mathbf{v}t)\,d\mathbf{x};\end{aligned} \tag{1.13}$$

here $\mathbf{E}_\mathbf{v}$ and $\mathbf{B}_\mathbf{v}$ are the comoving fields, and $\mathbf{v}$ is electron velocity.

The first important observation was that the field energy (1.13) for the point electron is infinite, since the Coulombic fields $\mathbf{E}_\mathbf{v}(\mathbf{x})$ and $\mathbf{B}_\mathbf{v}(\mathbf{x})$ are of order $|\mathbf{x}|^{-2}$ as $\mathbf{x}\to 0$. This explains why did Abraham introduce the model of the *extended electron* with the charge uniformly distributed over the sphere $|\mathbf{x}|=\varepsilon>0$. Then the energy is given by the first integral of type (1.13) over $|\mathbf{x}|>\varepsilon$, which is finite of order $1/\varepsilon$.

The principal discovery of Abraham was the relation

$$\mathbf{G}(v)=\frac{\mathrm{m}_e\mathbf{v}}{\sqrt{1-\beta^2}}\quad\text{with }\mathrm{m}_e=\frac{4}{3}\frac{\mathcal{E}(0)}{c^2}, \tag{1.14}$$

where $\beta := |\mathbf{v}|/c$, see [7, vol. I, (65.9)] and Sect. 12.9. Hence, the total momentum of the moving electron is as follows:

$$\mathbf{P}(v) = \mathrm{m}\mathbf{v} + \frac{\mathrm{m}_e \mathbf{v}}{\sqrt{1-\beta^2}}. \tag{1.15}$$

This gives the total mass of the electron:

$$M = \mathrm{m} + \frac{\mathrm{m}_e}{\sqrt{1-\beta^2}}; \tag{1.16}$$

here m is the 'bare' mass of the electron, and $\frac{\mathrm{m}_e}{\sqrt{1-\beta^2}}$ is its 'electromagnetic mass'.

On the other hand, from the Kauffmann experiments (1901–1906) it follows that the total electron momentum increases with velocity proportionally to $(\sqrt{1-\beta^2})^{-1}$. This is why Abraham suggested to completely reduce the bare mass of the electron setting $\mathrm{m} = 0$ ("Voilà l'atome dématérialisé, … la matière disparaît…" [90, pp. 63, 87, 88]). In this case, (1.14) becomes

$$Mc^2 = \frac{4}{3}\mathcal{E}(0). \tag{1.17}$$

Feynman has also suggested that 'almost all' mass of the electron is due to its electromagnetic energy [63].

This identification of the mass with the energy, as suggested by Abraham, was the first step to the wave theory of matter. Such a theory was later developed by de Broglie and Schrödinger.

Einstein had corrected relation (1.17) and extended it to any type of energy and mass as follows:

$$\mathcal{E} = Mc^2. \tag{1.18}$$

This formula holds for selfconsistent relativistic models [53]. The Abraham puzzling factor $4/3$ can be explained as being due to the fact that the classical model is not self-consistent because of electrostatic repulsion. The theoretical and practical importance of this formula can hardly be overrated. In particular, this formula provided the main stimulus for practical use of nuclear energy.

1.3 Thermodynamics

Thomson's discovery of the electron seems to confirm the Lorentz electron theory. However, the Lorentz basic equation (1.12) contains only one frequency ω_0, and so is incapable of explaining Combination Principle (1.1) and predicting spectral terms (1.2).

This problem was solved for the first time by Bohr and Debye in 1913 in the old Bohr quantum mechanics (Sect. 1.4), and more satisfactory, by Heisenberg and Schrödinger in 1925–1926.

The crucial role in the development of the old quantum mechanics was played by Planck's investigation (started in 1894) of the thermodynamic equilibrium in the radiation-matter system. Planck's studies were inspired by the *radiation law*, which was discovered by Kirchhoff in 1859 [99].

1.3.1 Equilibrium Radiation

The Kirchhoff discovery is based on the coincidence of the absorption and emission spectrum of the carbon (1849), as observed by Foucault, and the similar coincidence for sodium, as verified by Bunsen and Kirchhoff (1859). In short,

the emissivity is proportional to the absorptivity

This statement is supported by the following experiments:

(i) Let us take two cups (black and white) at normal temperature (20 °C). The absorption under an external illumination of the black cup obviously exceeds that of the white one. Now let us heat the cups in an oven up to 200–300 °C. Then the thermal radiation of the black cup is more intense than that of the white one—this can be felt by hand.
(ii) A small aperture in a platinum cavity is black at normal temperature (20 °C); it shines brightly at high temperatures (1600 °C).

Taking into account these and other observations, Kirchhoff suggested that the thermodynamic equilibrium of a substance at any fixed temperature T should be looked upon as a *statistical equilibrium* between the absorbed and emitted light: the spectrum, which is the set of all frequencies ω, and the intensity $I(\omega) = I_T(\omega)$ of the absorbed and emitted light should coincide.

On the other hand, the equilibrium radiation field is substance-dependent. This suggests a new question on thermodynamic equilibrium between different substances: the surrounding radiation field should correspond to each of the substances. Hence, it is natural to suppose the existence of a universal equilibrium radiation field corresponding to any fixed temperature.

In this radiation field, each substance emits (and absorbs) its characteristic components of light, and reflects other components. In the equilibrium, the sum of the emitted and reflected light coincides with the equilibrium radiation. This means that the resulting radiation of all substances is identical, i.e. of the same color. This phenomenon was well known from experiments on melting metals as early as in the sixth millennium BC: it is impossible to distinguish visually the heated iron ore from the burning carbon in thermodynamic equilibrium during fusion in the blast furnace. This coincidence of colors of the iron ore and carbon was used as an indication of the start of the melting.

This coincidence occurs *inside* the oven and bears evidence on the thermodynamic equilibrium in the absence of *external illumination*; the latter allows one to distinguish different substances exactly because the light source is not in the thermodynamic equilibrium with the substances, and the emission does not balance the absorption. For example, the emitted light is negligible under a bright illumination by the sun, electricity, etc. at room temperatures (20 °C). On the other hand, different objects cannot be distinguished if the source is at the same temperature as the substances (as in the dark room).

The universal equilibrium spectral density is referred to as the 'black-body radiation', since by definition the reflection of the 'black-body' is zero, and hence in this case its emission (radiation) would equal the equilibrium density. For example, the radiation spectrum of the burning carbon is a good approximation to the equilibrium spectral density.

It is worth pointing out, however, that the universal equilibrium spectral density is independent of the radiating body; hence the reference to the 'black-body' is only a convention. Also, perfect black bodies do not exist in nature. In practice, the 'black-body' is modeled by a small aperture in a heated cavity, inasmuch as the aperture absorbs all incident waves.

A precise formulation of Kirchhoff's law is as follows:

For every temperature $T > 0$, *there is a specific spectral intensity* $I(\omega) = I_T(\omega)$ *of light* (*per unit volume*) *which is in statistical equilibrium with any substance of temperature* T.

Kirchhoff deduced this law from the *second law of thermodynamics*. Namely, let us consider the light inside two cavities made of different materials, which is in thermodynamic equilibrium with the walls at the same temperature. Let us bring the cavities in the contact through a common aperture with an optical filter, allowing the light exchange at a fixed frequency. Given different spectral densities at this frequency, this results in energy exchange between the cavities to cause one cavity to heat and the other one to cool down, no work being done [7].

1.3.2 Experiments and Theory

The experimental measurements of the equilibrium spectral density $I(\omega)$ were performed by Tyndall (1865), Stefan (1879), Crova (1880), Langley (1886), Weber (1887), and Paschen and Wanner (1895–1899). In 1899, Lummer and Pringscheim, and Kurlbaum and Rubens made very precise measurements involving $\omega = 2\pi c/\lambda$, which corresponds to wavelengths λ ranging between 0.4 µm (dark violet H-line of ionized calcium) and 50 µm (far-infrared spectra), where $\mu\text{m} = 10^{-4}$ cm and temperatures ranged between -188 °C and 1600 °C.

The results of empirical observations were summarized by Wien in 1896: for *large frequencies* ω corresponding to (small) wavelengths λ between 0.4 µm and 18 µm, Wien's formula reads as follows:

$$I(\omega) \sim \omega^3 e^{-\frac{\gamma\omega}{T}}, \quad \gamma \approx 0.76 \times 10^{-11}\,\mathrm{K\,s},\ \omega \gg 1. \tag{1.19}$$

In 1879, Stefan has experimentally discovered the law of total radiated energy density:

$$\int_0^\infty I(\omega)\, d\omega = CT^4 \tag{1.20}$$

(the radiation of heat per unit of time and area of a body of absolute temperature T is proportional to T^4). Boltzmann deduced this *Stefan–Boltzmann* law in 1884 from the second law of thermodynamics [7].

Next step was made by Wien who set forth the *scaling law*

$$I(\omega) = \omega^3 f\left(\frac{\omega}{T}\right), \tag{1.21}$$

which was suggested by the displacement of the maximum value depending on T. This law was based on experimental evidence: at high temperatures the spectrum of burning carbon shifts to smaller wavelength. Wien theoretically justified this formula in 1893 by further analyzing the second law of thermodynamics. Clearly, (1.21) supports the law (1.20).

In 1894, Rayleigh and Jeans deduced the formula

$$I(\omega) = \frac{\omega^2 kT}{\pi^2 c^3} \tag{1.22}$$

where

$$k = R/N_A = 1.38 \times 10^{-16}\ \mathrm{erg/K} \tag{1.23}$$

is the *Boltzmann constant*, $R = 8.3141 \times 10^7\ \frac{\mathrm{erg}}{\mathrm{mol\,K}}$ is the universal gas constant, and $N_A = 6.022 \times 10^{23}\ \mathrm{mol}^{-1}$ is the Avogadro number. The underlying analysis resides in the classical electrodynamics and in the Boltzmann's *equipartition principle* (1876) for the oscillators of the Maxwell field.

Unfortunately, for $\omega \gg 1$ this formula cannot be valid, since it gives infinite total radiation (1.20). On the other hand, the Rayleigh–Jeans formula is supported by experimental evidence for small frequencies ω.

Conversely, the Wien distribution agrees well with experimental data for large ω/T, but it contradicts the experimental evidence for small ω/T.

In 1900, Planck discovered his famous formula

$$I(\omega) = \frac{\hbar\omega^3}{\pi^2 c^3} \frac{e^{-\frac{\hbar\omega}{kT}}}{1 - e^{-\frac{\hbar\omega}{kT}}} \tag{1.24}$$

where $\hbar$ is the *Planck constant*,

$$\hbar \approx 1.05 \times 10^{-27}\ \mathrm{erg\,s}. \tag{1.25}$$

Formula (1.24) accords remarkably well with experimental observations in the entire observed range of frequencies. In particular, it agrees with the Wien formula (1.19) (for large frequencies), and with the Rayleigh–Jeans formula (1.22) (for small frequencies). Both formulas (1.22) and (1.24) agree with the Wien's displacement law (1.21).

Planck's law leads to the revolutionary *quantization* (*discretization*) of energy of the oscillators of the Maxwell field. The Planck's ideas served as a key tool in the creation of quantum mechanics by Einstein, Bohr, Heisenberg and Schrödinger.

1.3.3 The Rayleigh–Jeans Law: Ultraviolet Divergence

The derivation of the Rayleigh–Jeans formula (1.22) depends on the Maxwell electrodynamics and the Boltzmann thermodynamics. In the free space (when $\rho(t,\mathbf{x}) = 0$ and $\mathbf{j}(t,\mathbf{x}) = 0$), the Maxwell equations (1.3) reduce to the system

$$\begin{cases} \operatorname{div}\mathbf{E}(t,\mathbf{x}) = 0, & \operatorname{curl}\mathbf{E}(t,\mathbf{x}) = -\dfrac{1}{c}\dot{\mathbf{B}}(t,\mathbf{x}), \\ \operatorname{div}\mathbf{B}(t,\mathbf{x}) = 0, & \operatorname{curl}\mathbf{B}(t,\mathbf{x}) = \dfrac{1}{c}\dot{\mathbf{E}}(t,\mathbf{x}). \end{cases} \tag{1.26}$$

The spectral density $I(\omega)$ is *formally* defined, for $\omega > 0$, by

$$I(\omega) := \frac{1}{8\pi} M\big[\tilde{\mathbf{E}}^2_\omega(t,\mathbf{x}) + \tilde{\mathbf{B}}^2_\omega(t,\mathbf{x})\big] \quad \text{(in unrationalized Gaussian units)} \tag{1.27}$$

(see (7.76) for the correct definition). Here M means the (mathematical) expectation, $\tilde{\mathbf{E}}_\omega(t,\mathbf{x})$ and $\tilde{\mathbf{B}}_\omega(t,\mathbf{x})$ are, respectively, the Fourier components of the equilibrium electric and magnetic fields at temperature T. In other words,

$$\begin{aligned} \tilde{\mathbf{E}}_\omega(t,\mathbf{x}) &= e^{-i\omega t}\mathbf{E}(\omega,\mathbf{x}) + e^{i\omega t}\mathbf{E}(-\omega,\mathbf{x}), \\ \tilde{\mathbf{B}}_\omega(t,\mathbf{x}) &= e^{-i\omega t}\mathbf{B}(\omega,\mathbf{x}) + e^{i\omega t}\mathbf{B}(-\omega,\mathbf{x}), \end{aligned} \tag{1.28}$$

where $\mathbf{E}(\omega,\mathbf{x})$ and $\mathbf{B}(\omega,\mathbf{x})$ are the Fourier amplitudes of the fields,

$$\mathbf{E}(t,\mathbf{x}) = \int_{\mathbb{R}} e^{-i\omega t}\mathbf{E}(\omega,\mathbf{x})\,d\omega, \qquad \mathbf{B}(t,\mathbf{x}) = \int_{\mathbb{R}} e^{-i\omega t}\mathbf{B}(\omega,\mathbf{x})\,d\omega. \tag{1.29}$$

Consider the lattice

$$\Gamma = \big\{\mathbf{k} = \big(\mathbf{k}^1,\mathbf{k}^2,\mathbf{k}^3\big) \in \pi\mathbb{Z}^3/L : \mathbf{k}^1,\mathbf{k}^2,\mathbf{k}^3 \geq 0\big\}.$$

I. We will represent the free Maxwell field in the cubic cavity $[0, L]^3$ with metal walls as a collection of independent harmonic oscillators. Both the tangential component of $\mathbf{E}(t,\mathbf{x})$ and the normal component of $\mathbf{B}(t,\mathbf{x})$ vanish on the conducting boundary. The corresponding eigenmodes are harmonics of the form

$$
\begin{aligned}
\mathbf{E}_{\mathbf{k}\alpha}(\mathbf{x}) = \big(&\mathbf{e}^1_{\mathbf{k}\alpha} \cos \mathbf{k}_1\mathbf{x}_1 \sin \mathbf{k}_2\mathbf{x}_2 \sin \mathbf{k}_3\mathbf{x}_3, \mathbf{e}^2_{\mathbf{k}\alpha} \sin \mathbf{k}_1\mathbf{x}_1 \cos \mathbf{k}_2\mathbf{x}_2 \sin \mathbf{k}_3\mathbf{x}_3,\\
&\mathbf{e}^3_{\mathbf{k}\alpha} \sin \mathbf{k}_1\mathbf{x}_1 \sin \mathbf{k}_2\mathbf{x}_2 \cos \mathbf{k}_3\mathbf{x}_3\big)\\
\mathbf{B}_{\mathbf{k}\alpha}(\mathbf{x}) = \big(&\mathbf{b}^1_{\mathbf{k}\alpha} \sin \mathbf{k}_1\mathbf{x}_1 \cos \mathbf{k}_2\mathbf{x}_2 \cos \mathbf{k}_3\mathbf{x}_3, \mathbf{b}^2_{\mathbf{k}\alpha} \cos \mathbf{k}_1\mathbf{x}_1 \sin \mathbf{k}_2\mathbf{x}_2 \cos \mathbf{k}_3\mathbf{x}_3,\\
&\mathbf{b}^3_{\mathbf{k}\alpha} \cos \mathbf{k}_1\mathbf{x}_1 \cos \mathbf{k}_2\mathbf{x}_2 \sin \mathbf{k}_3\mathbf{x}_3\big).
\end{aligned}
$$

To satisfy the Maxwell equations $\operatorname{div} \mathbf{E}_{\mathbf{k}\alpha}(\mathbf{x}) = 0$ and $\operatorname{div} \mathbf{B}_{\mathbf{k}\alpha}(\mathbf{x}) = 0$, we choose two orthogonal real vectors $\mathbf{e}_{\mathbf{k}\alpha} := (\mathbf{e}^1_{\mathbf{k}\alpha}, \mathbf{e}^2_{\mathbf{k}\alpha}, \mathbf{e}^3_{\mathbf{k}\alpha})$ and $\mathbf{b}_{\mathbf{k}\alpha} := (\mathbf{b}^1_{\mathbf{k}\alpha}, \mathbf{b}^2_{\mathbf{k}\alpha}, \mathbf{b}^3_{\mathbf{k}\alpha})$ so as to have $\mathbf{e}_{\mathbf{k}\alpha} \cdot \mathbf{k} = \mathbf{b}_{\mathbf{k}\alpha} \cdot \mathbf{k} = 0$ for $\alpha = 1, 2$. The Maxwell field can be expanded into a Fourier series as follows:

$$
\begin{aligned}
\mathbf{E}(t, \mathbf{x}) &= \sum_{\alpha} \sum_{\mathbf{k}\in\Gamma} \sqrt{\frac{8\pi\omega(\mathbf{k})}{L^3}} \mathbf{E}_{\mathbf{k}\alpha}(\mathbf{x}) q_{\mathbf{k}\alpha}(t),\\
\mathbf{B}(t, \mathbf{x}) &= \sum_{\alpha} \sum_{\mathbf{k}\in\Gamma} \sqrt{\frac{8\pi\omega(\mathbf{k})}{L^3}} \mathbf{B}_{\mathbf{k}\alpha}(\mathbf{x}) p_{\mathbf{k}\alpha}(t);
\end{aligned}
\tag{1.30}
$$

here $\omega(\mathbf{k}) = |\mathbf{k}|c$. The normalization factors guarantee that the variables $q_{\mathbf{k}\alpha}$, $p_{\mathbf{k}\alpha}$ are canonically conjugate. Finally, the Maxwell equations (1.26) for the curl reduce to

$$
\begin{aligned}
\mathbf{k} \times \mathbf{e}_{\mathbf{k}\alpha} q_{\mathbf{k}\alpha}(t) &= -\frac{1}{c} \mathbf{b}_{\mathbf{k}\alpha} \dot{p}_{\mathbf{k}\alpha}(t),\\
\mathbf{k} \times \mathbf{b}_{\mathbf{k}\alpha} p_{\mathbf{k}\alpha}(t) &= -\frac{1}{c} \mathbf{e}_{\mathbf{k}\alpha} \dot{q}_{\mathbf{k}\alpha}(t), \quad \mathbf{k} \in \Gamma,\ \alpha = 1, 2.
\end{aligned}
\tag{1.31}
$$

Now we choose arbitrary mutually orthogonal 'polarization' unit vectors $\mathbf{e}_{\mathbf{k}\alpha} \perp \mathbf{k} \in \mathbb{R}^3$, $\alpha = 1, 2$, and set $\mathbf{b}_{\mathbf{k}\alpha} := \frac{\mathbf{k}}{|\mathbf{k}|} \times \mathbf{e}_{\mathbf{k}\alpha}$. Then $\frac{\mathbf{k}}{|\mathbf{k}|} \times \mathbf{b}_{\mathbf{k}\alpha} = -\mathbf{e}_{\mathbf{k}\alpha}$, and so (1.31) reads as the Hamilton equations for harmonic oscillator

$$
q_{\mathbf{k}\alpha}(t) = -\frac{1}{\omega(\mathbf{k})} \dot{p}_{\mathbf{k}\alpha}(t), \qquad p_{\mathbf{k}\alpha}(t) = \frac{1}{\omega(\mathbf{k})} \dot{q}_{\mathbf{k}\alpha}(t), \quad \mathbf{k} \in \Gamma,\ \alpha = 1, 2. \tag{1.32}
$$

This means that the free Maxwell field is equivalent to the infinite collection of independent harmonic oscillators. Finally, the total electromagnetic energy in the finite volume $[0, L]^3$ is as follows:

$$
\mathcal{E} = \frac{1}{8\pi} \int_{[0,L]^3} \big[\mathbf{E}^2(t, \mathbf{x}) + \mathbf{B}^2(t, \mathbf{x})\big]\, dx = \frac{1}{2} \sum_{\alpha=1,2} \sum_{\mathbf{k}\in\Gamma} \omega(\mathbf{k}) \big[q^2_{\mathbf{k}\alpha}(t) + p^2_{\mathbf{k}\alpha}(t)\big],
\tag{1.33}
$$

because of orthogonality of the eigenmodes $\mathbf{E}_{\mathbf{k}\alpha}(\mathbf{x})$ and $\mathbf{B}_{\mathbf{k}\alpha}(\mathbf{x})$, and due to the choice of the unit vectors $\mathbf{e}_{\mathbf{k}\alpha}$ and $\mathbf{b}_{\mathbf{k}\alpha}$ and of the normalization factors in (1.30).

The system of oscillators (1.32) is Hamiltonian with the Hamilton function (1.33).

II. The Boltzmann *equipartition principle* assigns the *mean energy*

$$
\bar{E} = kT \tag{1.34}
$$

to *every independent degree of freedom*. Here, the bar denotes the mathematical expectation. In particular, contribution of each oscillator (1.32) to the total energy (1.33) is as follows:

$$\frac{1}{2}\omega(\mathbf{k})\big[\overline{q^2_{\mathbf{k}\alpha}(t)}+\overline{p^2_{\mathbf{k}\alpha}(t)}\big]=kT, \quad \mathbf{k}\in\Gamma,\ \alpha=1,2. \tag{1.35}$$

Hence, the expectation of total energy (1.33) is infinite, even in the finite volume $[0,L]^3$!

III. Formula (1.22) follows from the geometry of the lattice Γ if we make $L\to\infty$. Namely, for each $\mathbf{k}\in\Gamma$, there are two independent oscillators corresponding to the two possible polarizations $\mathbf{e}_{\mathbf{k}\alpha}\perp\mathbf{k}$, the frequency of both oscillators being $|\mathbf{k}|c$; the number of wave vectors $\mathbf{k}\in\Gamma$ in a spherical layer $r\le|\mathbf{k}|\le r+dr$ is asymptotically

$$N(r)\,dr\sim\frac{1}{8}\frac{4\pi r^2\,dr}{(\pi/L)^3}, \quad L\to\infty. \tag{1.36}$$

Hence the mean total energy of the oscillators with frequencies from $\omega=cr$ to $\omega+d\omega=c(r+dr)$ is

$$L^3 I(\omega)\,d\omega\sim 2\bar{E}N(r)\,dr, \quad L\to\infty. \tag{1.37}$$

As a result,

$$I(\omega)\sim\frac{\pi\omega^2}{(\pi c)^3}\bar{E}, \quad L\to\infty. \tag{1.38}$$

Now making $L\to\infty$, this gives (1.22), since $\bar{E}=kT$.

As we noted above, the Rayleigh–Jeans law gives the absurd *infinite total energy* per unit volume,

$$\int_0^\infty I(\omega)\,d\omega=\infty; \tag{1.39}$$

this is known as the *ultraviolet divergence*.

Remark 1.1 The equipartition principle (1.34) means that the *mean total energy* of the Maxwell field is infinite, and hence, for a finite L, the *mean energy density* is also infinite. The divergence in (1.39) means that this infinity of the mean energy density retains in the limit $L\to\infty$.

1.3.4 Planck's Law: Quantization of Energy

Thus, by 1896, the situation was the following: the Rayleigh–Jeans distribution was satisfactory for small frequencies, while Wien's law works well for large frequencies. Next step would be to interpolate between these two asymptotics with the main goal to avoid the ultraviolet divergence (1.39), providing the fast decay of the mean energy $\bar{E}$ as $\omega \to \infty$.

The problem was resolved by Planck in 1900 [154, 155]. Planck's derivation of the spectral formula (1.24) depended on some delicate thermodynamic arguments [172, Sect. 20]. In 1907, Einstein re-interpreted Planck's arguments in the framework of the statistical mechanics, and discovered that the same formula can be obtained by a *discretization of the energy distribution* [57] (see also [177]).

Namely, the assignment (1.34) agrees with the universal Boltzmann–Gibbs statistical distribution for the energy $E \ge 0$ (*canonical ensemble*),

$$p(E) = \frac{1}{Z} e^{-\frac{E}{kT}}, \quad Z := \int_0^\infty e^{-\frac{E}{kT}}\, dE = kT, \tag{1.40}$$

since

$$\bar{E} := \int_0^\infty E p(E)\, dE = -\left[\frac{d}{d\beta} \log \int_0^\infty e^{-\beta E}\, dE \right]_{\beta = 1/kT} = kT. \tag{1.41}$$

Analysing Planck's thermodynamic arguments, Einstein suggested to substitute the continuous Boltzmann distribution (1.40) by its *discrete version*

$$p_\varepsilon(E) = \frac{1}{Z} \sum_{n=0}^\infty e^{-\frac{E}{kT}} \delta(E - n\varepsilon), \qquad Z := \int \sum_{n=0}^\infty e^{-\frac{E}{kT}} \delta(E - n\varepsilon)\, dE, \tag{1.42}$$

with an energy step $\varepsilon > 0$. This means that only the discrete energies $E = n\varepsilon$ with integer $n = 0, 1, 2, \ldots$ are allowed. Now (1.41) changes to

$$\bar{E} = \frac{1}{Z} \sum_{n=0}^\infty n\varepsilon e^{-\frac{n\varepsilon}{kT}}, \tag{1.43}$$

which is the *Riemann sum* corresponding to the integral (1.40). Calculating as in (1.41), Einstein has arrived at the Planck's expression for $\bar{E}$:

$$\begin{aligned} \bar{E} &:= \int E p_\varepsilon(E)\, dE = -\left[\frac{d}{d\beta} \log \sum_{n=0}^\infty e^{-\beta n \varepsilon} \right]_{\beta = 1/kT} \\ &= \left[\frac{d}{d\beta} \log\bigl(1 - e^{-\beta\varepsilon}\bigr) \right]_{\beta = 1/kT} = \frac{\varepsilon e^{-\frac{\varepsilon}{kT}}}{1 - e^{-\frac{\varepsilon}{kT}}} \end{aligned} \tag{1.44}$$

(instead of (1.34)). To reconcile with the Wien's scaling law (1.21), one should set

$$\varepsilon = \hbar\omega \tag{1.45}$$

with a constant $\hbar > 0$. Then $\bar{E}$ remarkably decays for large ω and tends to (1.34) as $\varepsilon \to 0$.

Now, (1.38) with $\bar{E} = \bar{E}(\omega)$ implies the Planck's famous formula (1.24). Finally, we should set

$$\hbar = k\gamma \approx 1.05 \times 10^{-27} \text{ erg s} \tag{1.46}$$

to reconcile the spectral law (1.24) with the empirical Wien formula (1.19) for large ω/T. Moreover, (1.24) also agrees with the Rayleigh–Jeans spectral law (1.22) for small ω/T.

Formula (1.45) was probably suggested by a comparison of the empirical Wien's law (1.19) with the Maxwell–Boltzmann–Gibbs probability distribution

$$p(E) \sim e^{-\frac{E}{kT}}, \tag{1.47}$$

where E stands for the energy of a state. This comparison formally suggests the relation

$$\frac{\gamma\omega}{T} = \frac{E}{kT}, \tag{1.48}$$

which implies the formula of type (1.45)

$$E = \hbar\omega \tag{1.49}$$

with $\hbar$ given by (1.46).

1.3.5 Photoelectric Effect: Einstein's Rules

A new treatment of Planck's *energy quanta* (1.45) was suggested by Einstein in 1905 in his theory of the *photoelectric effect* [58] aimed at explaining the Lenard experimental observations (1902), see Sect. 8.4. Einstein suggested the law

$$K = \hbar\omega - W \tag{1.50}$$

for the (maximum) kinetic energy of the emitted electrons. He regarded this identity as the energy conservation law, where $\hbar\omega$ is the energy of the *absorbed photon* and the constant W is the *work function*, which is the minimum energy needed to remove an electron from a solid. Einstein was awarded the *Nobel Prize* (1921) for his theory of the photoelectric effect.

1.3.6 Einstein and Debye: Specific Heat of Solids

In 1906, Einstein has applied Planck's ideas to the specific heat problem [59], regarding a solid as a collection of independent oscillators with fixed frequency ω. He employed formulas (1.44), (1.45) to calculate the internal energy $U(T, V) = N\bar{E}$

and the specific heat $C_V = \partial_T U$. The results obtained were in a satisfactory agreement with experimental evidence for sufficiently large T. However, for small T, the discrepancy was quite pronounced.

In 1912, Debye went further, representing the Hamiltonian of a solid as the sum of energies of independent harmonic oscillators ('phonons') and applied formulas (1.44), (1.45) to each oscillator [42, 172]. His formula for specific heat provides a much better agreement with observations for small T than Einstein's formula.

1.4 'Old Quantum Mechanics'

Following the Lorentz electron theory (1895), atoms were modeled by classical oscillators (1.12). The theory gives the perfect description for the *normal Zeemann splitting* of the spectral lines (see Sect. 14.5). In 1902, Zeeman and Lorentz were awarded the Nobel Prize for these discoveries. However, the Lorentz theory cannot predict the unperturbed spectral lines $\omega_{mn} = \omega_m - \omega_n$, and the origin of the spectral terms ω_n remained mysterious. Moreover, the model (1.12) is unstable due to the friction term.

1.4.1 Planetary Model and Inconsistency of Classical Physics

The first 'planetary' model of an atom was introduced by Perrin in 1901: a positively charged particle is surrounded by a finite number of negative electrons revolving like 'small planets' around the positive charge (nucleus) [152]. Perrin suggested that the angular velocities of the electrons motion correspond to the frequencies of the radiation.

In 1904, Thomson proposed the *plum pudding model*: the atom is composed of negative corpuscles (electrons) surrounded by the 'soup' of positive charges to balance the electrons's negative charges.

This model allowed to justify the Lorentz equation (1.12). Namely, the term $\omega_0^2\mathbf{x}(t)$ exactly corresponds to the Coulomb force acting on the electron immersed in the soup of positive constant charge density. This follows by the Gauss theorem.

The Thomson model was rejected after the famous Rutherford experiments (1911) on the scattering of α-particles which showed the presence of a positively charged 'point' nucleus. Then the planetary model became again popular.

However, the circling electrons move with the centripetal acceleration, thereby losing their energy in accordance with the classical electrodynamics and eventually falling on the nucleus (the corresponding rigorous result was established only in 1997, see [112]). Therefore, the classical planetary model contradicts the stability of atoms. This being so, the classical physics had proved to be inconsistent in explaining the structure and stability of the matter. The problem required new robust ideas.

1.4.2 Bohr's Postulates

A revolutionary 'Columbus's' solution was put forward by Bohr in 1913 [15]:

I. Neglecting objections from the classical electrodynamics, Bohr *postulated* (i) the existence of *quantized stationary orbits* $|E_n\rangle$, with energies E_n, where the electrons *do not radiate*, and (ii) transitions between the orbits,

$$|E_m\rangle \mapsto |E_n\rangle. \tag{1.51}$$

II. Bohr identified the frequency of electromagnetic wave, which is radiated (or absorbed) in this transition:

$$\omega_{mn} = \frac{E_m - E_n}{\hbar}. \tag{1.52}$$

The above formula, written in the form

$$\hbar\omega_{mn} = E_m - E_n, \tag{1.53}$$

expresses the energy conservation, inasmuch as $\hbar\omega_{mn}$ is the energy of the radiated (or absorbed) photon in accordance with the Planck formula (1.45) and the Einstein formula (1.50). Furthermore, a comparison with the Rydberg–Ritz's Combination Principle (1.1) implies that

$$E_n = \hbar\omega_n + \text{const}. \tag{1.54}$$

1.4.3 Debye Quantum Rules

Bohr's postulates left open the question on the determination of energy levels E_n. In 1913, Debye introduced the *quantum condition* for the determining the energies E_n of *periodic orbits* of the electron in the atom:

$$S = 2\pi\hbar n, \quad n = 1, 2, 3, \ldots, \tag{1.55}$$

where $S := \oint p\,dq$ is the action integral over the period. The condition was suggested by the Planck energy distribution (1.42) with the discrete energy levels

$$E_n = n\varepsilon = n\hbar\omega. \tag{1.56}$$

Namely, Planck postulated the discreteness of energies for the oscillating eigenmodes (1.30) of the Maxwell field. In this case, (1.56) is equivalent to (1.55), since the action is $S = 2\pi E/\omega$ for the harmonic oscillators (1.32). The quantum condition was also motivated by the Ehrenfest ideas on adiabatic invariants [147].

The most convincing confirmation of (1.55) came from the Balmer formula (1.2), written as

$$E_n = -\frac{b}{n^2}, \quad n = 1, 2, \ldots, \tag{1.57}$$

where $b = \hbar B$. We will show that (1.55) implies (1.57) with

$$b = \frac{\mathrm{m}e^4}{2\hbar^2}, \tag{1.58}$$

where m and e are the rest mass and the charge of the electron in *unrationalized Gaussian units* (cgs), and $\hbar$ is the Plank constant (1.25); their modern values (in these units) are as follows:

$$e = -4.8032068 \times 10^{-10}\ \text{esu}, \qquad \mathrm{m} = 9.10938188 \times 10^{-28}\ \text{g}. \tag{1.59}$$

Calculating (1.58), we obtain $b \approx 21.79 \times 10^{-12}$ erg $= 21.79 \times 10^{-19}$ J. Dividing by the electron charge $|e| = 1.602 \times 10^{-19}$ C (SI), it follows from (1.57) that $E_1/e = b/|e| \approx 13.60$ V, which perfectly agrees with the experimentally known ionization energy 13.60 e V of the hydrogen.

We still need to derive (1.57) from (1.55). Let us consider the circular electron orbit of radius R. The energy conservation and the Newton equation read, respectively,

$$E = \frac{mv^2}{2} - \frac{e^2}{R}, \quad mv^2 = \frac{e^2}{R}. \tag{1.60}$$

Hence, $E = -\frac{e^2}{2R} = -\frac{mv^2}{2}$, and so $v = \sqrt{-\frac{2E}{m}}$ and $R = -\frac{e^2}{2E}$. Further, (1.55) now becomes $mv2\pi R = 2\pi\hbar n$; that is,

$$m\sqrt{-\frac{2E}{m}}\left[-\frac{e^2}{2E}\right] = \hbar n, \tag{1.61}$$

which is equivalent to (1.57), (1.58). The same result also holds for the elliptic orbits (see Sect. 14.2).

This confirmation of the quantum condition (1.55) was the great triumph of the Bohr–Debye quantum theory. Moreover, it buried the hope that pure electromagnetic theory of matter is possible, since the Planck constant with dimension of action cannot be obtained from fundamental parameters of classical electrodynamics (speed of light, the electron mass, and charge).

In 1916, Sommerfeld and Wilson extended the Debye quantum condition to periodic trajectories of multidimensional systems (cf. (14.10)):

$$\oint p_k\, dq_k = 2\pi\hbar n_k, \quad n_k = 1, 2, 3, \ldots \tag{1.62}$$

for every degree of freedom q_k.

1.4.4 Boltzmann Distribution in Old Quantum Mechanics

In 1917, Einstein derived Planck's formula (1.24) from the Bohr postulates, avoiding the energy discretization (1.42) introduced by Planck. In this derivation, the key role is played by the Boltzmann distribution for the energies of stationary orbits. This distribution was introduced by Boltzmann in 1887 for a gas in thermodynamic equilibrium at temperature $T > 0$; Boltzmann himself attributed this formula to Maxwell.

Suppose that the Boltzmann characteristic energy kT is less than the ionization energy, which is about 13.6 e V for the hydrogen atom. Taking into account $1\,\mathrm{eV} \approx 1.602 \times 10^{-12}$ erg, it follows that $kT \ll 13.6 \times 1.6 \times 10^{-12}$, which is equivalent to

$$T \ll \frac{13.602 \times 1.6 \times 10^{-12}}{1.38 \times 10^{-16}} \approx 13.6 \times 11600\,\mathrm{K} \approx 157 \times 10^3\,\mathrm{K}. \tag{1.63}$$

Below such temperatures, the interaction between the gas and the heat bath can be considered as a small perturbation of atoms: the atoms stay in their own quantum stationary states, sometimes undergoing the Bohr quantum transitions. In the statistical equilibrium, the *mean number* N of atoms and the *mean energy* $E(T)$ of the gas should be constant:

$$N = \sum N_m(T) = \mathrm{const}, \qquad E(T) = \sum E_m N_m(T) = \mathrm{const}, \tag{1.64}$$

where $N_m(T)$ is the *mean number* of atoms on the Bohr orbit $|E_m\rangle$. Therefore, the equilibrium distribution of atoms is given by the Boltzmann formula

$$\frac{N_m(T)}{N} \approx \frac{1}{Z} e^{-\frac{E_m}{kT}} \tag{1.65}$$

where Z is the normalization factor

$$Z = \sum e^{-\frac{E_m}{kT}} \tag{1.66}$$

ensuring the first identity of (1.64). For sufficiently large N formula (1.65) follows from Boltzmann’s classical arguments (1877) on the maximum probability (or entropy) [17], [172, (29.10)].

1.4.5 Einstein Theory of Radiation: Counting Photons

In 1917, Einstein deduced the Planck formula (1.24) by equating the *photons emission rate* to the *rate of their absorption*. The Einstein theory relies on Bohr’s postulates on transitions between quantum stationary states and the Boltzmann distribution (1.65). Additionally, Einstein postulated the existence of ‘photons’, as well as certain relations for the rate of Bohr’s transitions.

Namely, let us consider gas of atoms at temperature T in statistical equilibrium with the electromagnetic radiation. Let $|E_m\rangle$ and $|E_n\rangle$ be two stationary orbits of electrons with the energies $E_m < E_n$. Denote by $\omega_{nm} := (E_n - E_m)/\hbar$ the frequency of the absorbed 'photon', and by N_m and N_n, the number of the atoms in the state $|E_m\rangle$ and $|E_n\rangle$, respectively. The Boltzmann distribution (1.65) gives the relation

$$N_n/N_m = e^{-\frac{E_n}{kT}}/e^{-\frac{E_m}{kT}}. \tag{1.67}$$

Einstein had put forward the following new postulates:

I. The rate of absorption, $|E_m\rangle \to |E_n\rangle$, is proportional to $I(\omega_{nm})$:

$$Z_{mn} = N_m B_{mn} I(\omega_{nm}), \tag{1.68}$$

where Z_{mn} is the number of transitions per unit time.

II. The rate of emission, $|E_n\rangle \to |E_m\rangle$, consists of two terms:

$$Z_{nm} = N_n\big[A_m + B_{nm} I(\omega_{nm})\big]. \tag{1.69}$$

Here Z_{mn} is the number of transitions per unit time, A_m is the rate of 'spontaneous emission' (since it does not depend on the photons density $I(\omega_{mn})$), $B_{nm}I(\omega_{mn})$ is the rate of the 'induced emission'. Similarly, Z_{mn} in (1.68) is the rate of the 'induced absorption'.

III. The Rayleigh–Jeans law (1.22) holds asymptotically for small $\dfrac{\omega_{mn}}{T}$; i.e.,

$$I(\omega_{nm}) \sim \frac{\omega_{mn}^2 kT}{\pi^2 c^3} \quad \text{as } \frac{\omega_{mn}}{T} \to 0 \tag{1.70}$$

Remark 1.2 The proportionality to N_m and N_n in (1.68) and (1.69) 'stems' from the mutual independence of atoms. Similarly, the proportionality to spectral distribution $I(\omega_{nm})$ in (1.68) and (1.69) is motivated by regarding the distribution as the density of photons, and by the mutual independence of photons in their interaction with atoms.

In statistical equilibrium, $Z_{mn} = Z_{nm}$, that is,

$$e^{-\frac{E_m}{kT}} B_{mn} I(\omega_{mn}) = e^{-\frac{E_n}{kT}}\big[A_{nm} + B_{nm} I(\omega_{mn})\big], \tag{1.71}$$

according to (1.67). Solving this equation, we obtain

$$I(\omega_{mn}) = \frac{A_{nm}}{B_{mn} e^{\frac{\hbar\omega_{mn}}{kT}} - B_{nm}}. \tag{1.72}$$

To explain why $B_{mn} = B_{nm}$, Einstein applies the Taylor expansion in (1.72), getting

$$I(\omega_{mn}) = \frac{A_{nm}}{B_{mn} - B_{nm} + B_{mn}\dfrac{\hbar\omega_{mn}}{kT}} \tag{1.73}$$

for small $\frac{\omega_{mn}}{T}$. This relation yields (1.70) as $T \to \infty$ only if $B_{mn} = B_{nm}$. Hence,

$$\frac{A_{nm}}{B_{nm}} = \frac{\hbar\omega^3}{\pi^2 c^3}, \tag{1.74}$$

assuming additionally that coefficients A_{nm} and B_{nm} are independent of T. Now (1.72) implies the Planck formula (1.24).

In 1926, Dirac reproduced this formula in the framework of the Schrödinger wave mechanics [43] (see Sect. 7.3).

1.4.6 Bohr's Correspondence Principle

In 1913, Bohr introduced his *correspondence principle* and then reformulated it in several different forms until 1924. The principle was inspired by the belief that the quantum predictions turn into the corresponding classical ones as $\hbar \to 0$. For example, Einstein's discretization (1.42) with the step (1.45) turns into the Boltzmann equipartition law (1.34).

One of Bohr's formulations, dated 1920, reads: *the quantum theory should reproduce classical physics in the limit of large quantum numbers*, see [16].

Example The discrete energy levels (1.57) of the hydrogen become quasicontinuous for large n and tend to zero. This corresponds to the fact that Kepler's bounded orbits (ellipses) exist only for negative energy $E < 0$. Moreover, both quantum and classical orbits do not exist for positive energies.

In the Schrödinger's wave mechanics, the nonexistence of quantum orbits with a positive energy follows from Kato's theorem on the absence of embedded eigenvalues. Thus, Kato's theorem can be considered as a particular realization of Bohr's Correspondence Principle.

The correspondence principle was applied by Ladenburg [127] to determine the Einstein coefficient of absorption B_{mn}, by Kramers [121, 122], and Kramers and Heisenberg [123], to the dispersion theory (see the Kramers–Kronig formula (8.83)), and by many others. The correspondence principle is used, in particular,

(i) For substitution of difference quotient $\Delta F/\Delta J$ with $\Delta J = n\hbar$ instead of classical derivatives of the type dF/dJ.

Example Bohr's postulate (1.52) reads $\omega = \Delta E/\Delta J$, where $J = S/2\pi$, by (1.55). This corresponds to the classical formula $\omega = dE/dJ$, see (2.7) below;

(ii) For substitution of expressions of the type $J(J+1)$ instead of J^2, as in the formulas (14.66), (14.68) for the effective gyromagnetic ratio. The classical formula (14.66) has been deduced by Landé in 1921 from classical vector model for spin-orbital interaction to explain the Einstein–de Haas experiment and the anomalous Zeemann effect (Sect. 10.3).

Bohr applied the correspondence principle to multi-electron atoms to explain the atom radiation and the periodic law, basing his arguments on the geometric configuration of the orbits of different electrons in an atom, the screening effects, etc. The most prominent application was Heisenberg's discovery of the Matrix Quantum Mechanics. The detailed exposition of the correspondence principle and its applications can be found in the Pauli's survey [147] and in [188].

Chapter 2
Heisenberg's Matrix Mechanics

Abstract The quantum selection rule and its generalizations are capable of predicting energies of the stationary orbits; however they should be obtained in a more general framework of a universal theory, which could provide the intensities of spectral lines, scattering cross sections, etc.

Such a dynamical theory has been discovered first by Heisenberg in 1925 by developing Bohr's correspondence principle. The Heisenberg 'matrix mechanics' serves as a tool for extending the quantum selection rule (1.55) to arbitrary quantum systems, independently of the periodicity of trajectories of the corresponding classical models. The stationary energies appear to be the eigenvalues of the matrix Hamiltonian.

All equations and predictions of the Heisenberg theory turn into the classical one as $\hbar \to 0$; this agrees with the Bohr Correspondence Principle.

Heisenberg's theory, as was developed immediately by Born, Jordan, Pauli and others, is capable of producing the Hydrogen spectra, the selection rules and intensities of spectral lines, the quantization of the Maxwell field, etc. Up to now, Heisenberg's theory serves as the ground for the quantum electrodynamics and for modern quantum field theory.

2.1 Heisenberg's Matrix Formalism

Heisenberg suggested a novel revolutionary treatment of classical kinematics by combining Bohr's postulates (1.52) with the quantum selection rule (1.55) and the Bohr's Correspondence Principle.

2.1.1 Classical Oscillator

In 1925, Heisenberg [82] applied the correspondence principle to the one-dimensional oscillator

$$\dot{x}(t) = p(t), \qquad \dot{p}(t) = f\big(x(t)\big), \qquad f(x) = -V'(x) \tag{2.1}$$

A. Komech, *Quantum Mechanics: Genesis and Achievements*,
DOI 10.1007/978-94-007-5542-0_2,

introducing a non commutative algebra of 'quantum' observables and the corresponding dynamical equations. This is the Hamilton equation with the Hamiltonian function

$$H(x,p)=\frac{p^2}{2}+V(x); \tag{2.2}$$

the energy is conserved,

$$\frac{p^2(t)}{2}+V\big(x(t)\big)=E, \quad t\in\mathbb{R}. \tag{2.3}$$

Let us assume that the potential is confining; i.e.,

$$V(x)\to\infty, \quad |x|\to\infty.$$

Then each solution is periodic in time, the period being given by

$$T=\oint dt=\int_\gamma \frac{dx}{\sqrt{2(E-V(x))}}, \tag{2.4}$$

where γ is the trajectory $H(x,p)=E$, and $dt=dx/dp$. This period can be infinite when E is a critical value of $V(x)$. It can be written as

$$T=\frac{dS}{dE}, \tag{2.5}$$

where $S=\oint p\,dx$ is the action integral over the period; i.e.,

$$S=\int_\gamma \sqrt{2\big(E-V(x)\big)}\,dx. \tag{2.6}$$

Equivalently, (2.5) can be written as

$$\omega=\frac{dE}{dJ}, \quad J:=S/2\pi, \tag{2.7}$$

where $\omega=2\pi/T$ is the angular frequency. The corresponding Fourier series for the trajectories read

$$x(t)=\sum_{\nu\in\mathbb{Z}}x_\nu e^{i\nu\omega t}, \qquad p(t)=\sum_{\nu\in\mathbb{Z}}p_\nu e^{i\nu\omega t}, \tag{2.8}$$

where the frequencies can be expressed as follows:

$$\nu\omega=\lim_{\varepsilon\to 0}\frac{E(J+\nu\varepsilon)-E(J)}{\varepsilon}, \quad \nu\in\mathbb{Z}. \tag{2.9}$$

Thus the classical framework is capable of yielding only the frequencies $\nu\omega$.

2.1.2 *Quantum Oscillator*

Main features of the program of Heisenberg are as follows:

I. The position and momentum of an electron in the atom are not observable, since the Bohr postulates treat 'stationary orbits' rather than the electron motion. Respectively, these classical dynamical variables should be *reinterpreted* by taking into account the Bohr's correspondence principle.
II. The *form* of all classical relations between the dynamical variables (the form of dynamical equations, the expressions for the Hamiltonian, for the angular momentum, for the dipole moment, etc.) *should be kept.*

Correspondence Principles and Quantum Observables

The quantum selection rule (1.55) gives $J = n\hbar$ for $n = 1, 2, 3, \ldots$. On the other hand, by the Bohr postulate (1.52),

$$\omega_{mn} = \frac{E(m\hbar) - E(n\hbar)}{\hbar}. \tag{2.10}$$

The key Heisenberg's observation was the parallelism between (2.10) and the relation (2.9), which can be written approximately as follows:

$$(m - n)\omega \approx \frac{E(m\hbar) - E(n\hbar)}{\hbar}, \quad |m - n| \ll m. \tag{2.11}$$

Comparing (2.11) and (2.10), Heisenberg suggested, by the Bohr's correspondence principle, that ω_{mn} should be quantum analogues of the classical 'obertones' $(m - n)\omega$. This suggestion is confirmed by the Rydberg–Ritz Combination Principle (1.1), as written in the form

$$\omega_{mn} = \omega_{mk} + \omega_{kn} \tag{2.12}$$

which agrees with the classical 'combination principle'

$$\omega(m - n) = \omega(m - k) + \omega(k - n). \tag{2.13}$$

Respectively, the classical Fourier coefficients x_{m-n} and p_{m-n} should be substituted by new *quantum amplitudes* $\hat{x}_{mn}$ and $\hat{p}_{mn}$, and the Correspondence Principle should be treated as the asymptotics

$$\begin{aligned} \hat{x}_{mn} e^{i\omega_{mn}t} &\approx x_{m-n} e^{i\omega(m-n)t}, \\ \hat{p}_{mn} e^{i\omega_{mn}t} &\approx p_{m-n} e^{i\omega(m-n)t}, \quad |m - n| \ll m. \end{aligned} \tag{2.14}$$

Hence, the quantum analogues $\hat{x}(t)$ and $\hat{p}(t)$ of the classical observable $x(t)$ and $p(t)$ should be constructed in terms of collections of all quantum amplitudes

$$\begin{aligned} \hat{x}(t) &= \left\{\hat{x}_{mn} e^{i\omega_{mn}t} : m, n = 1, 2, \ldots\right\}, \\ \hat{p}(t) &= \left\{\hat{p}_{mn} e^{i\omega_{mn}t} : m, n = 1, 2, \ldots\right\}, \end{aligned} \tag{2.15}$$

representing the system at time t. The form of the observables was intended to give an appropriate description of Bohr's transitions (1.51), introducing the unknown frequencies ω_{mn} and the corresponding amplitudes $\hat{x}_{mn}$ and $\hat{p}_{mn}$, which are provisionally responsible for the transitions.

Matrix Algebra

Heisenberg's main intention was to keep all the classical relations for new 'quantum observables' (2.15), taking the correspondence (2.14) as the 'Ariadne's thread'. First, Heisenberg kept the form of dynamical equations (2.1), postulating

$$\partial_t \hat{x}(t) = \hat{p}(t), \qquad \partial_t \hat{p}(t) = f\big(\hat{x}(t)\big) \tag{2.16}$$

where the derivatives and the first linear equation are well defined for the quantum observables (2.15). On the other hand, the meaning of the nonlinear function $f(\hat{x})$ should be reinterpreted for general nonlinear functions f, at least for the polynomial functions.

For example, let us discuss the case of $f(x) = -\lambda x^2$ considered by Heisenberg. The multiplication of the Fourier series (2.8) gives

$$x^2(t) = \sum_{\mu,\nu\in\mathbb{Z}} x_\mu x_\nu e^{i(\mu+\nu)\omega t}. \tag{2.17}$$

Equivalently, the Fourier component of $x^2(t)$ with frequency $\omega(m-n)$ is given by

$$\big(x^2\big)_{m-n} e^{i\omega(m-n)t} = \sum_{k\in\mathbb{Z}} x_{m-k} e^{i\omega(m-k)t} x_{k-n} e^{i\omega(k-n)t}. \tag{2.18}$$

Now the correspondence principle (2.14) suggests the definition

$$\big(\hat{x}^2\big)_{mn} e^{i\omega_{mn}t} := \sum_{k\in\mathbb{Z}} \hat{x}_{mk}(t) e^{i\omega_{mk}t} \hat{x}_{kn}(t) e^{i\omega_{kn}t}, \tag{2.19}$$

which agrees with the Rydberg–Ritz Combination Principle (2.12). In other words,

$$\big(\hat{x}^2(t)\big)_{mn} := \sum_{k\in\mathbb{Z}} \hat{x}_{mk}(t)\hat{x}_{kn}(t), \tag{2.20}$$

where $\hat{x}_{mk}(t) := \hat{x}_{mk} e^{i\omega_{mk}t}$ and $\hat{x}_{kn}(t) := \hat{x}_{kn} e^{i\omega_{kn}t}$. The rule (2.20) was recognized as the matrix multiplication by Born, who remembered the lectures delivered by Jakob Rosanes at the Breslau University when reading the Heisenberg manuscript. Definition (2.20) has become the cornerstone of the *matrix mechanics* of Heisenberg—it means that a quantum observable $\hat{x}(t)$ is the *matrix of an operator* with the *matrix entries* $x_{mn}(t)$.

It is worth noting that the classical observables (2.8) correspond to particular case of the *Töplitz matrices*

$$\hat{x}(t) = \big(x_{m-n} e^{i\omega(m-n)t}\big), \qquad \hat{p}(t) = \big(p_{m-n} e^{i\omega(m-n)t}\big) \tag{2.21}$$

which are Hermitian, since the functions $x(t)$ and $p(t)$ are real. Respectively, the corresponding quantum observables (2.15) are postulated to be Hermitian.

Now *the quantum equations* (2.16) are well-defined at least for any polynomial function f.

Dynamical Equations

Dynamical equations for $\hat{x}(t)$ and $\hat{p}(t)$ follow by differentiating (2.15): using the key Bohr's relation (1.52), we obtain

$$\partial_t \hat{x}(t) = i\frac{E_m - E_n}{\hbar} x_{mn} e^{i\omega_{mn}t}, \qquad \partial_t \hat{p}(t) = i\frac{E_m - E_n}{\hbar} p_{mn} e^{i\omega_{mn}t}. \tag{2.22}$$

In the matrix form, the equations are known as the *Heisenberg equations*,

$$i\hbar\partial_t \hat{x}(t) = \big[\hat{x}(t), \hat{E}\big], \qquad i\hbar\partial_t \hat{p}(t) = \big[\hat{p}(t), \hat{E}\big], \tag{2.23}$$

in which $\hat{E}$ is the diagonal matrix $E_{mn} = E_m \delta_{mn}$. The next crucial step was the identification

$$\hat{E} = H\big(\hat{x}(t), \hat{p}(t)\big). \tag{2.24}$$

Now the Heisenberg equations (2.23) take the form

$$i\hbar\partial_t \hat{x}(t) = \big[\hat{x}(t), H\big(\hat{x}(t), \hat{p}(t)\big)\big], \qquad i\hbar\partial_t \hat{p}(t) = \big[\hat{p}(t), H\big(\hat{x}(t), \hat{p}(t)\big)\big]. \tag{2.25}$$

From these equations it follows that the dynamics of *any polynomial observable* $M(\hat{x}(t), \hat{p}(t))$ should be described by the similar equation

$$i\hbar\partial_t M\big(\hat{x}(t), \hat{p}(t)\big) = \big[M\big(\hat{x}(t), \hat{p}(t)\big), H\big(\hat{x}(t), \hat{p}(t)\big)\big] \tag{2.26}$$

(by the Jacobi identity for commutators). In particular,

$$i\hbar\partial_t H\big(\hat{x}(t), \hat{p}(t)\big) = \big[H\big(\hat{x}(t), \hat{p}(t)\big), H\big(\hat{x}(t), \hat{p}(t)\big)\big] = 0, \tag{2.27}$$

which means the *conservation of quantum energy* and justifies identification (2.24).

Commutation Relations

Comparing (2.25) with the postulate (2.16), it follows that

$$i\hbar\hat{p}(t) = \big[\hat{x}(t), H(\hat{x}, \hat{p})\big], \qquad i\hbar f\big(\hat{x}(t)\big) = \big[\hat{p}(t), H(\hat{x}, \hat{p})\big]. \tag{2.28}$$

Applying to a particular Hamilton function $H = p^2/2 + x$, which corresponds to $f(x) = -1$, this establishes

$$i\hbar\dot{\hat{p}}(t) = \big[\hat{x}(t), \hat{p}^2(t)/2\big], \qquad -i\hbar = \big[\hat{p}(t), \hat{x}(t)\big]. \tag{2.29}$$

Here the first identity follows from the second one; the latter is known to be the fundamental *commutation relation* obtained first in the Born and Jordan's paper [18]:

$$\big[\hat{x}(t), \hat{p}(t)\big] = i\hbar. \tag{2.30}$$

Example 2.1 The examples of linear operators in $C_0^\infty(\mathbb{R})$ satisfying (2.30) are as follows:

$$\hat{x}\psi(x) = x\psi(x), \qquad \hat{p}\psi(x) = -i\hbar\frac{d}{dx}\psi(x), \quad \psi \in C_0^\infty(\mathbb{R}). \tag{2.31}$$

Exercise 2.2 Prove the formulas

$$\big[\hat{x}(t), \hat{p}^N(t)\big] = i\hbar N\hat{p}^{N-1}(t), \qquad \big[\hat{p}(t), \hat{x}^N(t)\big] = -i\hbar N\hat{x}^{N-1}(t) \tag{2.32}$$

for any $N = 1, 2, \ldots$. **Hint:** Use (2.30).

2.2 Early Applications of Heisenberg Theory

Heisenberg's theory was immediately applied in [19, 148] to atoms with the energy-matrix

$$\hat{E} = \frac{1}{2\mathrm{m}}\hat{p}^2 + V(\hat{x}), \tag{2.33}$$

where $V(x)$ is a nucleus potential defined for $x \in \mathbb{R}^3$, and $\hat{x}$, $\hat{p}$ are subject to commutation relations

$$[\hat{x}_j, \hat{p}_k] = i\hbar\delta_{jk}, \qquad [\hat{x}_j, \hat{x}_k] = [\hat{p}_j, \hat{p}_k] = 0, \quad j, k = 1, 2, 3, \tag{2.34}$$

which 'follow' similarly to (2.29). Next triumph of Heisenberg's theory was the quantization of a Maxwell field by Dirac [44], who automatically implied the Planck's spectral law (1.24). Up to now, Heisenberg approach's has great value in quantum field theory (see [13, 33, 77, 138, 158, 159, 163, 189, 195, 196]).

2.2.1 *Eigenvalue Problem*

The main advantage of Heisenberg's theory was the identification of the spectral terms E_n with eigenvalues of the Hermitian operator $\hat{E}$. Namely, the matrix $\hat{E}$ in (2.23) is diagonal by definition, and E_n are its eigenvalues. To obtain E_n, one should find such dynamical variables $\hat{x}$ and $\hat{p}$ that the commutation relations (2.34) hold, and at the same time, the corresponding matrix (2.24) should be diagonal.

However, the dynamical variables $\hat{x}$ and $\hat{p}$ are not specified uniquely by (2.34). For example, we can choose operators (2.31) and make the transformation $\hat{x} \mapsto T^{-1}\hat{x}T$, $\hat{p} \mapsto T^{-1}\hat{p}T$, under which the commutation relations (2.34) are invariant. Then operator (2.24) transforms as follows: $\hat{E} \mapsto T^{-1}\hat{E}T$. The matrix $\hat{E}$ from (2.24) is Hermitian, like $\hat{p}$ and $\hat{x}$. Hence, one could expect to reduce $\hat{E}$ to the diagonal form:

$$T^{-1}\hat{E}T = \begin{pmatrix} \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & E_2 & 0 & 0 & 0 & 0 & \cdots \\ \cdots & 0 & E_1 & 0 & 0 & 0 & \cdots \\ \cdots & 0 & 0 & E_0 & 0 & 0 & \cdots \\ \cdots & 0 & 0 & 0 & E_{-1} & 0 & \cdots \\ \cdots & 0 & 0 & 0 & 0 & E_{-2} & \cdots \\ \cdots & 0 & 0 & 0 & 0 & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}. \tag{2.35}$$

The existence of the transformation T was known at that time (1925) for bounded Hermitian operators with discrete and continuous spectrum in Hilbert spaces [86, 88]. The authors [19] suggested the existence also for unbounded operators—this was proved later by J. von Neumann [142].

Reduction (2.35) is obviously equivalent to the eigenvalue problem. In fact, (2.35) reads

$$\sum_n \hat{E}_{mn} T_{nk} = T_{mk} E_k, \quad m, k = 0, \pm 1, \ldots. \tag{2.36}$$

In other words, the columns $e_k = (T_{mk} : m = 0, \pm 1, \ldots)$ of the matrix T are the eigenvectors of $\hat{E}$:

$$\hat{E} e_k = E_k e_k, \quad k = 0, \pm 1, \ldots. \tag{2.37}$$

The eigenvectors can be normalized: for the discrete eigenvalues,

$$\langle e_k, e_l \rangle = \delta_{kl}, \tag{2.38}$$

where $\langle \cdot, \cdot \rangle$ stands for the inner product in the Hilbert space l^2, and similarly for the eigenfunctions of continuous spectrum. Then the matrix T is orthogonal.

For the Hydrogen atom with the Coulomb potential $V(x)$, the spectrum E_n was obtained in this way by Pauli [148], who had obtained the Balmer formula (1.2) with $\omega_n = E_n/\hbar$.

Example: Quantization of Harmonic Oscillator

The solution of the eigenvalue problem (2.37) for one-dimensional harmonic oscillator was a cornerstown for further development of quantum mechanics and quantum field theory. In this case, the Hamilton function reads: $H=\frac{1}{2}p^2+\frac{1}{2}\omega^2x^2$, and

$$\hat{E}=\frac{1}{2}\hat{p}^2+\frac{1}{2}\omega^2\hat{x}^2,\quad [\hat{x},\hat{p}]=i\hbar. \tag{2.39}$$

The following proposition will be proved in Sect. 14.8.

Proposition 2.1 *Eigenvalues of $\hat{E}$ are given by $E_n=\hbar\omega(n+\frac{1}{2})$, $n=0,1,\dots$.*

2.2.2 Intensity of Spectral Lines

The key observation in the celebrated 'three-man paper' [19] was that the matrix element $\hat{x}_{nn'}=\langle e_n,\hat{x}e_{n'}\rangle$ of the transformed matrix $T^*\hat{x}T$ should be responsible for intensity of the corresponding spectral line $\omega_{nn'}=\omega_n-\omega_{n'}$:

$$I_{nn'}=\frac{2e^2}{3c^3}\hat{x}_{nn'}^2\omega_{nn'}^4. \tag{2.40}$$

This expression is suggested by the Hertz formula (12.127) for radiation of the dipole with the moment

$$\mathbf{p}=e\hat{x}_{nn'}. \tag{2.41}$$

This identification was motivated by the following facts:

(i) $\mathbf{p}=e\mathbf{x}$ in the classical theory;
(ii) Formula (12.127) for radiation of a harmonic dipole $\mathbf{p}(t)=ex_{m-n}\cos\omega(m-n)t$ was used previously by Heisenberg and Kramers for calculation of the dispersion formula [123];
(iii) The correspondence principle suggests that the Fourier coefficient $x_{n-n'}$ and the frequency $\omega(n-n')$ of the classical dipole should be substituted by the matrix element $\hat{x}_{nn'}$ and the frequency $\omega_{nn'}$ respectively.

2.2.3 The Normal Zeeman Effect

Formula (2.40) was applied in [19] for the derivation of selection rules and intensity of spectral lines in the normal Zeeman effect (see (9.14)). Calculation of the corresponding intensities (2.40) relies on the following arguments:

(i) The commutation relations

$$[\hat{M}_1,\hat{M}_2]=-i\hbar\hat{M}_3 \tag{2.42}$$

hold for the components of quantum angular momentum $\hat{M}:=\hat{p}\times\hat{x}$ (cf. (6.39)).

(ii) Operators $\hat{E}$, $\hat{M}^2$ and $\hat{M}_3$ commute and hence can be simultaneously diagonalized (cf. Lemma 6.10(ii)). Therefore, the eigenvectors of $\hat{E}$ can be numbered by the corresponding eigenvalues E, m and m_3 of the operators $\hat{E}$, $\sqrt{\hat{M}^2}$ and $\hat{M}_3$, respectively.

The selection rules mean that the matrix elements $\hat{x}_{nn'} = \langle e_n, \hat{x}e_{n'}\rangle$ and intensities (2.40) vanish if $m_3' \neq m_3, m_3 \pm \hbar$. Here, the eigenvector e_n corresponds to the triple (E, m, m_3), and $e_{n'}$, to the triple (E', m', m_3'). In terms of Bohr's postulate, *the transitions* $(E, m, m_3) \to (E', m', m_3')$ *are forbidden if* $m_3' \neq m_3, m_3 \pm \hbar$. In short, the selection rule $m_3 \to m_3', m_3' \pm \hbar$ holds.

2.2.4 Quantization of Maxwell Field and Planck's Law

Dirac [44] was first to quantize a Maxwell field by representing it as the system of independent harmonic oscillators (1.32). The system is Hamiltonian with the Hamilton function (1.33) and the canonical variables $q_{\mathbf{k}\alpha}$, $p_{\mathbf{k}\alpha}$. Therefore, the commutation relations for the corresponding Heisenberg matrices should be

$$\left\{\begin{array}{l} \left[\hat{q}_{\mathbf{k}\alpha}(t), \hat{p}_{\mathbf{k}'\alpha'}(t)\right] = i\hbar\delta_{\mathbf{k}\alpha,\mathbf{k}'\alpha'}, \\ \left[\hat{q}_{\mathbf{k}\alpha}(t), \hat{q}_{\mathbf{k}'\alpha'}(t)\right] = \left[\hat{p}_{\mathbf{k}\alpha}(t), \hat{p}_{\mathbf{k}'\alpha'}(t)\right] = 0, \end{array}\right| \quad \mathbf{k}, \mathbf{k}' \in \Gamma, \ \alpha, \alpha' = 1, 2, \tag{2.43}$$

which is suggested by (2.34). Further, (1.33) implies that the corresponding energy operator reads:

$$\hat{\mathcal{E}} = \sum_{\alpha=1,2} \sum_{\mathbf{k}\in\Gamma} \hat{\mathcal{E}}_{\mathbf{k}\alpha} \tag{2.44}$$

where $\hat{\mathcal{E}}_{\mathbf{k}\alpha} = \frac{\omega(\mathbf{k})}{2}[\hat{q}^2_{\mathbf{k}\alpha} + \hat{p}^2_{\mathbf{k}\alpha}]$. Finally, Proposition 2.1 with $\omega = 1$ implies that eigenvalues of $\hat{\mathcal{E}}_{\mathbf{k}\alpha}$ equal $E_{\mathbf{k}\alpha n} = \hbar\omega(\mathbf{k})(n + \frac{1}{2})$ with $n = 0, 1, \ldots$. Therefore, eigenvalues of $\hat{\mathcal{E}}$ are given by finite sums of $E_{\mathbf{k}\alpha, n(\mathbf{k}\alpha)}$, since $\hat{\mathcal{E}}_{\mathbf{k}\alpha}$ all commute, and hence can be simultaneously diagonalized. Hence, the Boltzmann distribution (1.65) implies that the eigenvalues of $\hat{\mathcal{E}}_{\mathbf{k}\alpha}$ are independent for different $\mathbf{k}\alpha$.

Application of the Boltzmann distribution (1.65) to the quantized Maxwell field immediately implies Planck's low (1.24). First, the probabilities (1.65) of the eigenvalues $E_{\mathbf{k}\alpha n}$ are given by

$$p(E_{\mathbf{k}\alpha n}) = \frac{1}{Z_{\mathbf{k}\alpha}} e^{-\frac{E_{\mathbf{k}\alpha n}}{kT}}, \quad Z_{\mathbf{k}\alpha} := \sum_n e^{-\frac{E_{\mathbf{k}\alpha n}}{kT}}. \tag{2.45}$$

This means that we have a discretization of the Einstein type (1.42) for the energy distribution of each field oscillator. Further arguments mainly repeat those of (1.43)–(1.44) and (1.36)–(1.37). In fact, the mean value of $E_{\mathbf{k}\alpha}$ is as follows:

$$\bar{E}_{\mathbf{k}\alpha} = \sum_n E_{\mathbf{k}\alpha n} p(E_{\mathbf{k}\alpha n}) = -\left[\frac{d}{d\beta} \log \sum_{n=0}^{\infty} e^{-\beta\hbar\omega(\mathbf{k})(n+\frac{1}{2})}\right]_{\beta=1/kT}$$

$$= -\left[\frac{d}{d\beta} \log \frac{e^{-\beta\hbar\omega(\mathbf{k})/2}}{1 - e^{-\beta\hbar\omega(\mathbf{k})}}\right]_{\beta=1/kT} = \hbar\omega(\mathbf{k})/2 + \frac{\hbar\omega(\mathbf{k}) e^{-\frac{\hbar\omega(\mathbf{k})}{kT}}}{1 - e^{-\frac{\hbar\omega(\mathbf{k})}{kT}}}; \quad (2.46)$$

this coincides with (1.44) up to the additional term $\hbar\omega(\mathbf{k})/2$, which makes the total energy infinite after summation in $\mathbf{k}$. We will drop this term, since the energy is defined up to an additive constant and should vanish at $T = 0$. Then we obtain (1.44):

$$\bar{E}_{\mathbf{k}\alpha} = \bar{E}\big(\omega(\mathbf{k})\big) := \frac{\hbar\omega(\mathbf{k}) e^{-\frac{\hbar\omega(\mathbf{k})}{kT}}}{1 - e^{-\frac{\hbar\omega(\mathbf{k})}{kT}}}. \quad (2.47)$$

Therefore, repeating the arguments of (1.36)–(1.37), we arrive at (1.38) and (1.24).

Chapter 3
Schrödinger's Wave Mechanics

Abstract In 1926, Schrödinger proposed an alternative wave theory of quantization by developing de Broglie's wave particle duality. The Schrödinger theory also regards the Hamiltonian as an operator in a Hilbert space, and the stationary energies, as the corresponding eigenvalues.

The Schrödinger wave equation alone cannot be justified experimentally. One should complete the equation with the corresponding quantum observables: energy, momentum, angular momentum, charge and current, etc. A fundamental requirement for the introduction of the quantum observables is their agreement with the corresponding classical observables; this follows from the 'quasiclassical asymptotics'.

Schrödinger showed that his theory is equivalent to the Heisenberg's matrix theory. Moreover, both theories turn into the classical one as $\hbar \to 0$: Heisenberg's theory implies this correspondence directly, while for the Schrödinger theory, this follows from the quasiclassical asymptotics.

3.1 Wave-Particle Duality: de Broglie and Schrödinger

The wave description of matter was suggested for the first time by Abraham's theory of electromagnetic mass (1902), see Sect. 1.2.7.

The wave treatment of the cathode rays by Hertz, Lenard, and others was restored since 1923 in de Broglie's and Schrödinger's wave theory of matter relying on the Einstein's theory of special relativity and Planck's theory of radiation. The wave nature of cathode rays was confirmed by the Davisson and Germer experiments on the electron diffraction by crystals (1924–1927), see Sect. 5.2.2.

3.1.1 De Broglie Waves of Matter

The Planck's radiation law and Einstein's theory of the photo-electric effect (1905) suggested the corpuscular treatment of the light as a collection of particles (photons) restoring the Newton corpuscular theory of light. Thus the light is thought of as an electromagnetic wave in the Maxwell electrodynamics and as a collection of

A. Komech, *Quantum Mechanics: Genesis and Achievements*,
DOI 10.1007/978-94-007-5542-0_3,

particles in the photoelectric effect. This 'wave-particle duality' of the light was broadly discussed since 1905, and in particular, at the First Solvay Congress in Brussels (1911), one of main topics of which was the Einstein theory of the specific heat of solids.

One of scientific secretaries of the congress was Marcel de Broglie, and he was involved in preparing the proceedings of the Congress for publication. His young brother, Louis de Broglie, fervently studied the manuscripts, and decided "dedicate all my life for the study of the true nature of the mysterious quants introduced by Planck ten years earlier in theoretical physics" [27, p. 458].

In 1923, Louis de Broglie extended the wave-particle duality from the light to the matter [24–26]. Namely, he introduced the waves of matter, assigning the wave function $\psi(x,t) = Ae^{i(\mathbf{kx}-\omega t)}$ to homogeneous beam of particles with fixed momentum $\mathbf{p}$ and energy E. He found the corresponding wave vector $\mathbf{k}$ and the frequency ω by applying the Plank relation (1.45) and Einstein's Special Relativity.

Assume for a moment that the speed of light $c = 1$. Then a *homogeneous beam* of *free relativistic particles* is characterized by energy-momentum vector $(E, \mathbf{p})$ satisfying the relativistic relation (13.11)

$$E^2 = \mathbf{p}^2 + \mathrm{m}^2 \tag{3.1}$$

with *rest mass* $\mathrm{m} \geq 0$. Louis de Broglie postulated a correspondence between a homogeneous beam and a plane wave

$$\textbf{free particles of energy } E \textbf{ and momentum } \mathbf{p} \ \Leftrightarrow \ \psi(t, \mathbf{x}) = Ae^{i(\mathbf{kx}-\omega t)}, \tag{3.2}$$

where the vector $(\omega, \mathbf{k})$ is a function of $(E, \mathbf{p})$; i.e.,

$$(\omega, \mathbf{k}) = F(E, \mathbf{p}). \tag{3.3}$$

Second, de Broglie applied the Einstein's special theory of relativity to show that the vector $(\omega, \mathbf{k})$ is proportional to $(E, \mathbf{p})$,

$$(E, \mathbf{p}) = \hbar(\omega, \mathbf{k}). \tag{3.4}$$

Proof by Special Relativity Arguments The wave function (3.2) must be Lorentz invariant, hence the Lorentz group transforms both vectors, $(\omega, \mathbf{k})$ and $(E, \mathbf{p})$, identically, as the *covariant vectors*. Further, the *relativistic covariance* requires that the function F commutes with the Lorentz transforms (see Sect. 12.3),

$$FL(E, \mathbf{p}) = LF(E, \mathbf{p}), \quad L \in \mathrm{SO}(3,1). \tag{3.5}$$

On the other hand, relation (3.1) implies that there exists $L_0 \in \mathrm{SO}(3,1)$ such that

$$L_0(E, \mathbf{p}) = (\mathrm{m}, 0, 0, 0). \tag{3.6}$$

Then (3.5) with $L = L_0$ gives that

$$F(E, \mathbf{p}) = L_0^{-1} F(\mathrm{m}, 0, 0, 0). \tag{3.7}$$

Finally, it follows from (3.5) with $(E, \mathbf{p}) = (\mathrm{m}, 0, 0, 0)$ that

$$FL(\mathrm{m}, 0, 0, 0) = LF(\mathrm{m}, 0, 0, 0), \quad L \in \mathrm{SO}(3, 1). \tag{3.8}$$

Taking rotations $L \in \mathrm{SO}(3) \subset \mathrm{SO}(3, 1)$, this establishes

$$F(\mathrm{m}, 0, 0, 0) = LF(\mathrm{m}, 0, 0, 0), \quad L \in \mathrm{SO}(3), \tag{3.9}$$

since the vector $(\mathrm{m}, 0, 0, 0)$ is invariant under the rotations. Hence, the vector $F(\mathrm{m}, 0, 0, 0)$ is also invariant under all rotations; i.e.,

$$F(\mathrm{m}, 0, 0, 0) = \big(f(\mathrm{m}), 0, 0, 0\big). \tag{3.10}$$

We can define $\hbar_1$ by

$$\mathrm{m} := \hbar_1 f(\mathrm{m}), \tag{3.11}$$

since $f(\mathrm{m}) \neq 0$. Now (3.7) reads

$$\begin{aligned} F(E, \mathbf{p}) &= L_0^{-1} F(\mathrm{m}, 0, 0, 0) = L_0^{-1}\big(f(\mathrm{m}), 0, 0, 0\big) \\ &= \frac{1}{\hbar_1} L_0^{-1}(\mathrm{m}, 0, 0, 0) = \frac{(E, \mathbf{p})}{\hbar_1}, \end{aligned} \tag{3.12}$$

since $L_0^{-1}(\mathrm{m}, 0, 0, 0) = (E, \mathbf{p})$ by (3.6). Hence

$$(\omega, \mathbf{k}) = \frac{(E, \mathbf{p})}{\hbar_1}, \tag{3.13}$$

by (3.3). Finally, the Planck formula (1.45) suggests that

$$\hbar_1 = \hbar. \tag{3.14}$$

Then (3.13) becomes (3.4). □

In all calculations (3.1)–(3.12) it has been assumed that the speed of light is $c = 1$. To restore the value of c, is suffices to substitute everywhere E/c, ω/c and mc^2 for E, ω and m, respectively.

De Broglie's relation (3.4) was confirmed experimentally by Davisson and Germer's observations on the diffraction of cathode rays [40]. In 1929, de Broglie was awarded the Nobel Prize for his discovery of (3.4).

3.1.2 De Broglie Wavelength and Dispersion Relations

Identity (3.4) plays a crucial role in the entire quantum theory. In particular, it implies the famous de Broglie formula for the 'particle wavelength' $\lambda = 2\pi/|\mathbf{k}|$:

$$\lambda = 2\pi\hbar/|\mathbf{p}|. \tag{3.15}$$

It also gives the relativistic *dispersion relation* (i.e., the relation between the wave vector and frequency)

$$\hbar^2\omega^2/c^2 = \hbar^2\mathbf{k}^2 + \mathrm{m}^2c^2, \tag{3.16}$$

which follows from expression for Hamiltonian of relativistic particle (see (13.11)):

$$E^2/c^2 = \mathbf{p}^2 + \mathrm{m}^2c^2. \tag{3.17}$$

For $m = 0$, we have $|\mathbf{p}| = E/c = \hbar\omega/c$, and hence formula (3.4) gives

$$\lambda = 2\pi c/\omega, \tag{3.18}$$

as it holds for the light. Hence, the particles of light ('the photons') should have zero *rest mass*.

For small values of $|\mathbf{p}| \ll \mathrm{m}c$, the following *non-relativistic approximation* holds,

$$E = \sqrt{\mathbf{p}^2c^2 + \mathrm{m}^2c^4} \approx \mathrm{m}c^2 + \frac{\mathbf{p}^2}{2\mathrm{m}}. \tag{3.19}$$

Dropping here the 'unessential' additive constant $\mathrm{m}c^2$, we arrive at the non-relativistic dispersion relation

$$\hbar\omega = \frac{\hbar^2\mathbf{k}^2}{2\mathrm{m}}. \tag{3.20}$$

Exercise 3.1 Calculate the wavelength of the electron beam in a television tube with voltage $U \approx 10$ KV. **Hints:**

(i) Electron energy at rest is $E_0 := mc^2 \approx 0.51$ MeV; hence the electron energy after acceleration $E = E_0 + eU \approx 0.52$ MeV, provided that the electron was initially at rest. Now (3.17) gives

$$\mathbf{p}^2c^2 = E^2 - E_0^2 \approx \left(52^2 - 51^2\right) \times 10^8 \ (\mathrm{e\,V})^2 \approx 10^{10} \ (\mathrm{e\,V})^2 \tag{3.21}$$

(ii) Substituting $\mathrm{e\,V} \approx 1.6 \times 10^{-12}$ erg, we obtain

$$|\mathbf{p}|c \approx 10^5 \times 1.6 \times 10^{-12} = 1.6 \times 10^{-7}. \tag{3.22}$$

Taking $c \approx 3 \times 10^{10}$ cm/s, it follows that $|\mathbf{p}| \approx 0.5 \times 10^{-17}$. Finally, (3.15) with (1.25) gives

$$\lambda \approx \frac{6 \times 10^{-27}}{0.5 \times 10^{-17}} \approx 10^{-9} \text{ cm} = 0.1 \text{ Å} \tag{3.23}$$

where $\text{Å} = 10^{-8}$ cm is *one Ångström*, which is about the size of the hydrogen atom.

For a comparison, the wavelengths of visible light is as follows:

$\lambda = 3800$ Å for violet light,
$\lambda = 4200$ Å for indigo light,
$\lambda = 4400$ Å for blue light,
$\lambda = 5000$ Å for cyan light,
$\lambda = 5200$ Å for green light,
$\lambda = 5650$ Å for yellow light,
$\lambda = 5900$ Å for orange light,
$\lambda = 6250$ Å for red light.

Respectively, scattering of visible light by atoms is negligible, while for the electron beam in television tube, the scattering is significant. This difference explains high efficiency of electron microscopes.

3.1.3 Canonical Quantization I: Free Particles

The de Broglie plane wave should be implemented into a more general context of wave fields and wave propagation. The corresponding wave equations are provided by the dispersion relations (3.16) and (3.20): they imply the *free* Klein-Gordon resp. Schrödinger equations for the plane wave $\psi(t,\mathbf{x}) = e^{i(\mathbf{kx}-\omega t)} = e^{i(\mathbf{px}-Et)/\hbar}$ describing the free particles:

$$\frac{1}{c^2}[i\hbar\partial_t]^2\psi(t,\mathbf{x}) = \big[(-i\hbar\boldsymbol{\nabla})^2 + \mathrm{m}^2c^2\big]\psi(t,\mathbf{x}), \tag{3.24}$$

$$i\hbar\partial_t\psi(t,\mathbf{x}) = \frac{1}{2\mathrm{m}}[-i\hbar\boldsymbol{\nabla}]^2\psi(t,\mathbf{x}) \tag{3.25}$$

where $\boldsymbol{\nabla} = (\nabla_1, \nabla_2, \nabla_3)$, $\nabla_k := \frac{\partial}{\partial \mathbf{x}^k}$. Formally, the equations follow from the energy-momentum relations (3.17), (3.19) by substitution (cf. (2.31))

$$E \mapsto \hat{E} := i\hbar\partial_t, \qquad \mathbf{p} \mapsto \hat{\mathbf{p}} := -i\hbar\boldsymbol{\nabla} \tag{3.26}$$

which is called “canonical quantization”.

3.1.4 Canonical Quantization II: Bound Particles

In 1925–1926, Klein, Gordon and Schrödinger extended the wave equations (3.24) and (3.25) to the case of a bound electron interacting with an external Maxwell field

$$\mathbf{E}(t,\mathbf{x}) = -\boldsymbol{\nabla}\phi(t,\mathbf{x}) - \dot{\mathbf{A}}(t,\mathbf{x}), \qquad \mathbf{B}(t,\mathbf{x}) = \operatorname{curl}\mathbf{A}(t,\mathbf{x}), \tag{3.27}$$

where $\phi(t,\mathbf{x})$ and $\mathbf{A}(t,\mathbf{x})$ are, respectively, the corresponding external scalar and vector Maxwell potentials. For example, in the hydrogen atom, $\phi(t,\mathbf{x}) = -e/|x|$ is the Coulombic potential of the nucleus and $\mathbf{A}(t,\mathbf{x}) \approx 0$, provided we neglect the magnetic field generated by the nucleus motion and magnetic moment.

For a relativistic electron in an external Maxwell field, energy E is given by (12.93):

$$\big[E - e\phi(t,\mathbf{x})\big]^2/c^2 = \Big[\mathbf{p} - \frac{e}{c}\mathbf{A}(t,\mathbf{x})\Big]^2 + \mathrm{m}^2c^2 \tag{3.28}$$

where e is the charge of the electron. Applying the canonical quantization (3.26), Klein, Gordon and Schrödinger obtained the wave equation

$$\frac{1}{c^2}\big[i\hbar\partial_t - e\phi(t,\mathbf{x})\big]^2\psi(t,\mathbf{x}) = \Big[-i\hbar\nabla - \frac{e}{c}\mathbf{A}(t,\mathbf{x})\Big]^2\psi(t,\mathbf{x}) + \mathrm{m}^2c^2\psi(t,\mathbf{x}). \tag{3.29}$$

For a nonrelativistic electron in an external Maxwell field, the energy E is given by (12.90):

$$E = \mathcal{H}(t,\mathbf{x},\mathbf{p}) = \frac{1}{2\mathrm{m}}\Big[\mathbf{p} - \frac{e}{c}\mathbf{A}(t,\mathbf{x})\Big]^2 + e\phi(t,\mathbf{x}). \tag{3.30}$$

This means that Hamiltonian form of the Lorentz equation (1.6) reads

$$\begin{cases} \dot{\mathbf{x}}(t) = \nabla_{\mathbf{p}}\mathcal{H}(t,\mathbf{x},\mathbf{p}) = \dfrac{1}{\mathrm{m}}\Big(\mathbf{p} - \dfrac{e}{c}\mathbf{A}(t,\mathbf{x})\Big), \\ \dot{\mathbf{p}}(t) = -\nabla_{\mathbf{x}}\mathcal{H}(t,\mathbf{x},\mathbf{p}) = -e\nabla\phi(t,\mathbf{x}) + \dfrac{e}{\mathrm{m}c}\nabla\Big[\mathbf{A}(t,\mathbf{x})\cdot\Big(\mathbf{p} - \dfrac{e}{c}\mathbf{A}(t,\mathbf{x})\Big)\Big]. \end{cases} \tag{3.31}$$

Exercise 3.2 Prove the equivalence of (1.6) and (3.31). **Hint:** Calculate $\ddot{\mathbf{x}}_1(t)$ differentiating the first equation of (3.31), and use the second equation together with relations (3.27).

Applying the canonical quantization (3.26) to (3.30), Schrödinger derived the corresponding nonrelativistic wave equation [161]

$$i\hbar\partial_t\psi(t,\mathbf{x}) = \frac{1}{2\mathrm{m}}\Big[-i\hbar\nabla - \frac{e}{c}\mathbf{A}(t,\mathbf{x})\Big]^2\psi(t,\mathbf{x}) + e\phi(t,\mathbf{x})\psi(t,\mathbf{x}), \tag{3.32}$$

which can be written as follows:

$$\begin{aligned} &i\hbar\partial_t\psi(t,\mathbf{x}) = H(t)\psi(t,\mathbf{x}), \\ &\quad H(t) = \mathcal{H}(t,\mathbf{x},-i\hbar\nabla) = \frac{1}{2\mathrm{m}}\Big[-i\hbar\nabla - \frac{e}{c}\mathbf{A}(t,\mathbf{x})\Big]^2 + e\phi(t,\mathbf{x}). \end{aligned} \tag{3.33}$$

Physical Interpretation of the Wave Function relies on the Schrödinger identification for the electric charge and current densities

$$\rho(t,\mathbf{x}) := e\big|\psi(t,\mathbf{x})\big|^2, \qquad \mathbf{j}(t,\mathbf{x}) := \frac{e}{\mathrm{m}}\mathrm{Re}\bigg(\overline{\psi(t,\mathbf{x})}\Big[-i\hbar\nabla - \frac{e}{c}\mathbf{A}(t,\mathbf{x})\Big]\psi(t,\mathbf{x})\bigg). \tag{3.34}$$

These expressions are justified by the agreement with the classical case for short wavelength 'packets' (see Sect. 3.3 below); they are related by the continuity equation for charge (1.4), which holds for any solution of the Schrödinger equation.

Respectively, the normalization

$$\int \big|\psi(t,\mathbf{x})\big|^2 d\mathbf{x} = 1 \tag{3.35}$$

should hold for wave functions describing *one electron*. Modern experiments give the following values (in *unrationalized Gaussian units*):

$$\begin{aligned} e &= -4.8032068\times 10^{-10}\ \text{esu}, & \mathrm{m} &= 9.10938188\times 10^{-28}\ \text{g},\\ \hbar &= 1.05457266\times 10^{-27}\ \text{erg s}, & c &= 29979245800\ \text{cm/s}. \end{aligned} \tag{3.36}$$

3.1.5 Quantum Stationary States: Eigenvalue Problem

For *static* external Maxwell potentials $\phi_{\text{ext}}(\mathbf{x}), \mathbf{A}_{\text{ext}}(\mathbf{x})$, the Schrödinger equation (3.33) reads

$$\begin{aligned} &i\hbar\partial_t\psi(t,\mathbf{x}) = H\psi(t,\mathbf{x}),\\ &H = \mathcal{H}(\mathbf{x}, -i\hbar\nabla) = \frac{1}{2\mathrm{m}}\Big[-i\hbar\nabla - \frac{e}{c}\mathbf{A}_{\text{ext}}(\mathbf{x})\Big]^2 + e\phi_{\text{ext}}(\mathbf{x}). \end{aligned} \tag{3.37}$$

In this case, Schrödinger identified the *quantum stationary states* of the energy E with the solutions of type

$$\psi(t,\mathbf{x}) = \psi(\mathbf{x})e^{-iEt/\hbar}. \tag{3.38}$$

The Schrödinger equation (3.37) becomes then the *eigenvalue problem*

$$E\psi(\mathbf{x}) = H\psi(\mathbf{x}), \tag{3.39}$$

which is equivalent to (2.37). This eigenvalue problem agrees with the quantum rules (1.55) and (1.62) asymptotically for large 'quantum numbers' (Bohr–Sommerfeld quantum conditions, see [5, 78]).

Moreover, for the case of *free particles*, identification (3.38) agrees with the de Broglie wave function $\psi(t,\mathbf{x}) = e^{i(\mathbf{px}-Et)/\hbar} = e^{i\mathbf{px}/\hbar}e^{-iEt/\hbar}$. For *bound particles* in an external potential, identification (3.38) reflects the fact that the space is 'twisted' by the external field, while the time remains 'free'.

So the reasons for the Schrödinger identification (3.38) seem to be rather algebraic. On the other hand, we suggest that the identification also has a deep dynamical background in the framework of the nonlinear coupled Maxwell–Schrödinger equations (see (4.26) below) with static external Maxwell potentials $\phi_{\text{ext}}(\mathbf{x})$, $\mathbf{A}_{\text{ext}}(\mathbf{x})$:

$$\begin{cases} \left[i\hbar\partial_t - e\big(\phi(t,\mathbf{x}) + \phi_{\text{ext}}(\mathbf{x})\big)\right]\psi(t,\mathbf{x}) \\ \quad = \dfrac{1}{2\mathrm{m}}\left[-i\hbar\nabla - \dfrac{e}{c}\big(\mathbf{A}(t,\mathbf{x}) + \mathbf{A}_{\text{ext}}(\mathbf{x})\big)\right]^2 \psi(t,\mathbf{x}), \\ \dfrac{1}{4\pi}\Box\phi(t,\mathbf{x}) = e\big|\psi(t,\mathbf{x})\big|^2, \\ \dfrac{1}{4\pi}\Box\mathbf{A}(t,\mathbf{x}) = \dfrac{e}{\mathrm{m}c}\mathrm{Re}\left[\overline{\psi(t,\mathbf{x})}\left[-i\hbar\nabla - \dfrac{e}{c}\big(\mathbf{A}(t,\mathbf{x}) + \mathbf{A}_{\text{ext}}(\mathbf{x})\big)\right]\psi(t,\mathbf{x})\right]. \end{cases}$$

This being so, we suggest that the solutions

$$\psi(t,\mathbf{x}) = \psi_\omega(\mathbf{x})e^{-i\omega t}, \qquad \phi(t,\mathbf{x}) = \phi_\omega(\mathbf{x}), \qquad \mathbf{A}(t,\mathbf{x}) = \mathbf{A}_\omega(\mathbf{x}), \tag{3.40}$$

with $\omega \in \mathbb{R}$, form a global attractor for the coupled equations. In other words, for every finite energy solution $\psi(t,\mathbf{x})$, $\phi(t,\mathbf{x})$, $\mathbf{A}(t,\mathbf{x})$, the following long-time asymptotics hold,

$$\big(\psi(t,\mathbf{x}), \phi(t,\mathbf{x}), \mathbf{A}(t,\mathbf{x})\big) \sim \big(\psi_\pm(\mathbf{x})e^{-i\omega_\pm t}, \phi_\pm(\mathbf{x}), \mathbf{A}_\pm(\mathbf{x})\big), \quad t \to \pm\infty, \tag{3.41}$$

which give the dynamical interpretation to the Bohr's transitions (1.51). First results on global attraction of type (3.41) were obtained recently for some model nonlinear Hamiltonian wave equations [100–112] (see the survey [106]). However, for the coupled Maxwell–Schrödinger equations, the proof of transitions (3.41) is still an open problem.

Note that attraction (3.41) *does not hold* for the linear Schrödinger equation (3.37) due to the principle of superposition.

3.1.6 Stationary Perturbation Theory

In practice it is very important to have a simple formula for the correction ΔE to the eigenvalue E of the perturbed Schrödinger operator $H + \varepsilon H'$ with $|\varepsilon| \ll 1$; i.e.,

$$(E + \Delta E)\psi'(\mathbf{x}) = \big(H + \varepsilon H'\big)\psi'(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^3. \tag{3.42}$$

Let us consider the case of a **simple eigenvalue** E. In this case, the formula becomes (see [11, 160])

$$\Delta E \sim \big\langle \psi, \varepsilon H'\psi \big\rangle, \quad \varepsilon \to 0, \tag{3.43}$$

where ψ is the corresponding unperturbed eigenfunction *satisfying the normalization condition* (3.35), and $\langle\cdot,\cdot\rangle$ is the Hermitian inner product in $L^2(\mathbb{R}^3)$.

Let us prove the formula for the case of Hermitian operators H and H' in a finite dimensional complex space. The corresponding eigenvalues and eigenfunctions being smooth functions of ε, and so we have

$$\Delta E \sim \mathcal{O}(\varepsilon) \quad \text{and} \quad \psi' = \psi + \varepsilon\phi + \mathcal{O}\big(\varepsilon^2\big), \quad \varepsilon \to 0. \tag{3.44}$$

Substituting this into (3.42), this establishes

$$(E + \Delta E)\big(\psi + \varepsilon\phi + \mathcal{O}\big(\varepsilon^2\big)\big) = \big(H + \varepsilon H'\big)\big(\psi + \varepsilon\phi + \mathcal{O}\big(\varepsilon^2\big)\big). \tag{3.45}$$

The first order terms give

$$E\varepsilon\phi + \Delta E\psi = H\varepsilon\phi + \varepsilon H'\psi + \mathcal{O}\big(\varepsilon^2\big). \tag{3.46}$$

We can assume the normalization condition $\|\psi'\| = 1$; hence $\langle\psi, \phi\rangle = 0$. Therefore, multiplying (3.45) by ψ, it follows that

$$\Delta E\langle\psi, \psi\rangle = \big\langle\varepsilon H'\psi + \mathcal{O}\big(\varepsilon^2\big), \psi\big\rangle, \tag{3.47}$$

since $\langle H\varepsilon\phi, \psi\rangle = \langle\varepsilon\phi, H\psi\rangle = \varepsilon E\langle\phi, \psi\rangle = 0$. Formula (3.43) herewith follows.

3.2 Quasiclassical Asymptotics

One of the *mathematical justifications* of the canonical quantization (3.26) relies on the WKB (Wentzel–Kramers–Brillouin, 1926) *quasiclassical asymptotics* [62]: the Planck constant $\hbar$ is very small, and so the 'short wavelength solutions' to the Schrödinger equation (3.32) propagate along the 'rays', which are trajectories of the corresponding classical dynamics (1.6) (or (3.31)). The asymptotics also play a crucial role in the justification of definitions for *quantum observables*: in the next section we will show that the quantum observables turn into the corresponding classical ones in the limit $\hbar \to 0$. Both these asymptotics are particular appearances of Bohr's Correspondence Principle.

3.2.1 Geometrical Optics

Straightline light propagation is well known from conventional observations of the shadow. Similar property is also well known in acoustics and water waves with small wavelength. However, its explanation by the wave interference was discovered for the first time by Fresnel (1815) by calculating the diffraction pattern from the border of a screen. Fresnel's approach was developed later in short wavelength expansions for diffraction problems by Rayleigh (1877) and Kirchhoff (1882).

A systematic theory of the short wavelength asymptotic expansions in the diffraction theory was developed by Debye in 1909. Substitution of the short wavelength expansions into wave equations leads to the *eikonal equation* for the phase function. The German term 'Eikonal' (from Greek $\varepsilon\iota\kappa\omega\nu =$ image) was introduced in 1895 by H. Bruns; the reason behind this is because the equation determines the *rays* forming images. By the Hamilton–Jacobi Theorem 13.7, the rays are trajectories of the corresponding Hamilton system. The rays are straight lines in the case of homogeneous media; in a nonhomogeneous case they satisfy the Snell's law of refraction.

3.2.2 Application to Schrödinger Equation

In the case of the Schrödinger equation, the Debye expansion takes the form

$$\psi(t,\mathbf{x}) \sim \left(\sum_{k=0}^{\infty} \hbar^k a_k(t,\mathbf{x})\right) e^{iS(t,\mathbf{x})/\hbar}, \quad \hbar \ll 1. \tag{3.48}$$

In particular, substituting this expansion into the Schrödinger equation (3.33), and equating the leading terms with $\hbar^0 = 1$, we arrive at the eikonal *Hamilton–Jacobi equation*

$$-\dot{S}(t,\mathbf{x}) = \mathcal{H}\big(\mathbf{x}, \nabla S(t,\mathbf{x}), t\big). \tag{3.49}$$

Its solution is the *action functional* given by integral (13.15) over classical trajectories $(\mathbf{x}(t), \mathbf{p}(t))$ of the Hamilton system (3.31) according to the Hamilton–Jacobi Theorem 13.7 (see also [4, 73, 129]). In particular,

$$\mathbf{p}(t) = \nabla S\big(t, \mathbf{x}(t)\big), \tag{3.50}$$

by (13.38). The amplitudes $a_k(t,\mathbf{x})$ are solutions to the *transport differential equations* (13.36) along trajectories of classical dynamics (3.31) (see Sect. 13.2).

Integral representation (13.15) and the transport equations imply that the support of solution (3.48) is a *tube* of type (13.32), which is the union of the *rays* that are trajectories $\mathbf{x}(t)$. More precisely, for any $N > 0$,

$$\psi(t,\mathbf{x}) \sim \mathcal{O}\big(\hbar^N\big) \quad \text{for } \big|\mathbf{x} - \mathbf{x}(t)\big| \geq 2\varepsilon,\ |t| < T_\varepsilon \tag{3.51}$$

if $\psi(0,\mathbf{x}) = 0$ for $|\mathbf{x} - \mathbf{x}(0)| \geq \varepsilon$ (see Corollary 13.10). In other words,

short wavelength solutions of the Schrödinger equation
propagate along trajectories of the Lorentz equation

This 'wave-particle' correspondence is called the *geometrical optics*, suggesting that the cathode rays, which propagate along the trajectories of the Lorentz equation, as observed by J.J. Thomson and W. Kauffmann, might be the short wavelength solutions of the Schrödinger equation (3.32), see Fig. 3.1.

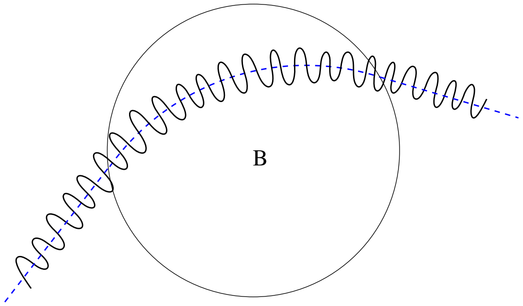


Fig. 3.1 Cathode rays (*wavy line*) refracted by the magnetic field. *Dashed line* is the corresponding classical trajectory (Color figure online)

The geometrical optics partially justifies identification (3.34) of electric charge and current densities. In fact, substituting the asymptotics (3.48) into (3.34), this gives

$$\begin{aligned}\mathbf{j}\big(t,\mathbf{x}(t)\big) &\sim \frac{e}{m}\mathrm{Re}\bigg(\overline{\psi}\Big[\nabla S-\frac{e}{c}\mathbf{A}\Big]\psi\bigg)=\frac{e}{m}\mathrm{Re}\bigg(\overline{\psi}\Big[\mathbf{p}-\frac{e}{c}\mathbf{A}\Big]\psi\bigg)\\ &= e\mathrm{Re}\,(\overline{\psi}\dot{\mathbf{x}}\psi)=e|\psi|^2\dot{\mathbf{x}}(t)=\rho\big(t,\mathbf{x}(t)\big)\dot{\mathbf{x}}(t),\end{aligned}\tag{3.52}$$

where we have used the first equation of (3.31).

3.2.3 Generalizations

The Schrödinger equation was obtained by the canonical quantization (3.26) of the Hamiltonian equations (3.31) for a nonrelativistic particle in a Maxwell field. For a relativistic particle with the Hamiltonian (3.28), the canonical quantization (3.26) leads to the Klein–Gordon equation (3.29). Extension of the geometrical optics to the Klein–Gordon equation is also possible by the same Hamilton–Jacobi theory.

The geometrical optics was extended to the Dirac equation by Pauli [149], and to general hyperbolic systems, by Lax [132]. Further developments culminated in Maslov–Hörmander's theory of the Fourier integral operators: the survey can be found in [78].

3.2.4 Hamilton Optical-Mechanical Analogy

The parallelism between the wave propagation and classical mechanics is known since Hamilton's 'optical-mechanical analogy' (1834), see [4, 129]. The analogy is the *formal* identification of the mechanical trajectories with the optical rays using the Fermat principle for the rays and the Hamilton least action principle for mechanical trajectories. Hamilton considered the action functional (13.15) as a wave, which propagates along the classical trajectories, and applied Huygens's principle for characterization of the wave front. In this way Hamilton has obtained relation (3.50) and equation (3.49). Schrödinger's Nobel lecture (1933) entirely concerns with this optical-mechanical analogy, which was one of primary reasons for the introduction of equation (3.32). Apparently, the key role of the action functional and of the Hamilton–Jacobi equation were suggested by the Debye and Sommerfeld–Wilson *quantum rules* (1.55).

3.2.5 Conclusions

The quasiclassical asymptotics justifies the canonical quantization (1.6)$\mapsto$(3.32) in the limit $\hbar \to 0$. On the other hand, the Planck constant $\hbar$ takes a *finite nonzero value*. Consequently, the quasiclassical asymptotics alone are not sufficient for the introduction of the Schrödinger equation, and hence one needs the Planck and de Broglie arguments to fix the finite value of the constant $\hbar$.

3.3 Quantum Observables and Conservation Laws

We should finish up with the physical interpretation of the Schrödinger equation (3.33)

$$\begin{aligned} i\hbar\partial_t\psi(t,\mathbf{x}) &= H(t)\psi(t,\mathbf{x}), \\ H(t) = \mathcal{H}(x,-i\hbar\nabla,t) &= \frac{1}{2\mathrm{m}}\left[-i\hbar\nabla-\frac{e}{c}\mathbf{A}(t,\mathbf{x})\right]^2+e\phi(t,\mathbf{x}) \end{aligned} \tag{3.53}$$

by introducing suitable quantum observables: energy, momentum, angular momentum and charge. The definitions will be justified by the agreement with the classical observables for one-particle short wavelength solutions (3.48) under the normalization (3.35). Namely, we show that the quantum observables turn into the corresponding classical ones as $\hbar \to 0$.

We will be concerned with the solutions $\psi(t,\mathbf{x})$, which are sufficiently smooth and which decay sufficiently fast as $|\mathbf{x}| \to \infty$, thereby providing the correctness of all differentiations and partial integrations below. Given $\psi_1, \psi_2 \in L^2(\mathbb{R}^3)$, let

$$\langle\psi_1,\psi_2\rangle = \int \psi_1(\mathbf{x})\overline{\psi_2(\mathbf{x})}\,dx \tag{3.54}$$

be the Hermitian inner product in the complex Hilbert space $L^2(\mathbb{R}^3)$. Note that the Schrödinger operator is Hermitian symmetric; i.e.,

$$\big\langle\psi_1, H(t)\psi_2\big\rangle = \big\langle H(t)\psi_1,\psi_2\big\rangle, \qquad \psi_1,\psi_2 \in C_0^\infty\big(\mathbb{R}^3,\mathbb{C}\big). \tag{3.55}$$

We will assume the one-particle normalization condition (3.35).

3.3.1 Quantum Observables

First we define the quantum observables.

I. Energy For quantum stationary state (3.39), the energy is identified with E; this is confirmed by (2.37) and by the excellent agreement of Heisenberg's theory and experimental observations, and in particular, by Pauli's calculation of the Balmer formula [148]. Equation (3.39) with the normalization (3.35) gives the following expression for the energy:

$$E(t) = \langle \psi(t), H\psi(t) \rangle. \tag{3.56}$$

The same expression is suggested for the *mean energy* at time t under normalization (3.35).

II. Momentum and Angular Momentum Similarly, the *mean momentum* and the *mean angular momentum* are defined by

$$\mathbf{p}(t) := \langle \psi(t,\mathbf{x}), \hat{\mathbf{p}}\psi(t,\mathbf{x}) \rangle, \tag{3.57}$$

where $\hat{\mathbf{p}} := -i\hbar\nabla$ as in (3.26), and

$$\mathbf{L}(t) := \langle \psi(t,\mathbf{x}), \hat{\mathbf{L}}\psi(t,\mathbf{x}) \rangle, \tag{3.58}$$

$\hat{\mathbf{L}} := \hat{\mathbf{x}} \times \hat{\mathbf{p}}$, where $\hat{\mathbf{x}}$ is the operator of multiplication by $\mathbf{x}$.

III. Charge Finally, the charge and current densities are defined by (3.34). Respectively, the *mean electric charge* is defined by

$$Q(t) := e\int |\psi(t,\mathbf{x})|^2 \, d\mathbf{x}. \tag{3.59}$$

3.3.2 Conservation Laws

Now we state the corresponding conservation laws.

I. Energy The energy (3.56) is conserved, provided that the Maxwell potentials are independent of time; i.e.,

$$\phi(t,\mathbf{x}) = \phi(\mathbf{x}), \qquad \mathbf{A}(t,\mathbf{x}) = \mathbf{A}(\mathbf{x}). \tag{3.60}$$

Lemma 3.3 *Under conditions* (3.60),

$$E(t) = \text{const}, \quad t \in \mathbb{R}. \tag{3.61}$$

Proof Differentiating, we obtain

$$\begin{aligned}\dot{E}(t) &= \left[\langle\dot{\psi}(t), H\psi(t)\rangle + \langle\psi(t), H\dot{\psi}(t)\rangle\right]\\ &= \left[-\frac{i}{\hbar}\langle H\psi(t), H\psi(t)\rangle + \frac{i}{\hbar}\langle\psi(t), HH\psi(t)\rangle\right] = 0, \end{aligned} \tag{3.62}$$

by (3.37), since the Schrödinger operator $H := H(t)$ is symmetric. □

II. Momentum and Angular Momentum The momentum (3.57) is conserved, provided that the Maxwell potentials $\phi(t, \mathbf{x}), \mathbf{A}(t, \mathbf{x})$ are translation invariant, i.e.,

$$\phi(t, \mathbf{x} + s\mathbf{e}_n) = \phi(t, \mathbf{x}) \quad \text{and} \quad \mathbf{A}(t, \mathbf{x} + s\mathbf{e}_n) = R_n(\theta)\mathbf{A}(t, \mathbf{x}), \quad s \in \mathbb{R}, \tag{3.63}$$

where $\mathbf{e}_n$ is the unit vector in the direction $\mathbf{x}^n$.

Lemma 3.4 *Under condition* (3.63), *the corresponding component* $\mathbf{p}_n$ *of the momentum is conserved*:

$$\mathbf{p}_n(t) = \text{const}, \quad t \in \mathbb{R}. \tag{3.64}$$

Proof Differentiating, we obtain, by (3.53),

$$\begin{aligned}\dot{\mathbf{p}}_n(t) &= \langle\dot{\psi}(t, \mathbf{x}), \hat{\mathbf{p}}_n\psi(t, \mathbf{x})\rangle + \langle\psi(t, \mathbf{x}), \hat{\mathbf{p}}_n\dot{\psi}(t, \mathbf{x})\rangle\\ &= \left[-\frac{i}{\hbar}\langle H(t)\psi(t, \mathbf{x}), \hat{\mathbf{p}}_n\psi(t, \mathbf{x})\rangle + \frac{i}{\hbar}\langle\psi(t, \mathbf{x}), \hat{\mathbf{p}}_n H(t)\psi(t, \mathbf{x})\rangle\right]\\ &= -\frac{i}{\hbar}\langle\psi(t, \mathbf{x}), \left[H(t), \hat{\mathbf{p}}_n\right]\psi(t, \mathbf{x})\rangle = 0 \end{aligned} \tag{3.65}$$

since the Schrödinger operator $H(t)$ is symmetric and since the commutator $[H(t), \hat{\mathbf{p}}_n] := H(t)\hat{\mathbf{p}}_n - \hat{\mathbf{p}}_n H(t) = 0$. These operators commute, in as much as the coefficients of the differential operator $H(t)$ do not depend on $\mathbf{x}^n$. □

The angular momentum (3.58) is conserved, provided that the Maxwell potentials $\phi(t, \mathbf{x}), \mathbf{A}(t, \mathbf{x})$ are rotationally invariant; i.e.,

$$\phi\big(t, R_n(\theta)\mathbf{x}\big) = \phi(t, \mathbf{x}) \quad \text{and} \quad \mathbf{A}\big(t, R_n(\theta)\mathbf{x}\big) = R_n(\theta)\mathbf{A}(t, \mathbf{x}), \quad \theta \in [0, 2\pi], \tag{3.66}$$

where $R_n(\theta)$ stands for the rotations $R_n(\theta)$ around the axis $\mathbf{x}^n$.

Lemma 3.5 *Under condition* (3.66), *the corresponding component* $\mathbf{L}_n(t)$ *of the angular momentum is conserved*:

$$\mathbf{L}_n(t) = \text{const}, \quad t \in \mathbb{R}. \tag{3.67}$$

The proof is quite similar to (3.65), depending on the symmetry of the Schrödinger operator $H(t)$ and on its commutation with $\hat{\mathbf{L}}_n \sim (\mathbf{x} \times \nabla)_n$. Indeed,

$(\mathbf{x} \times \nabla)_n = \frac{\partial}{\partial \theta}$ is a generator of the rotations $R_n(\theta)$. Hence, the Schrödinger operator $H(t)$ commutes with $(\mathbf{x} \times \nabla)_n$, inasmuch as $H(t)$ commutes with the rotations.

Exercise 3.6 Verify that condition (3.66) with $n = 3$ holds for the static uniform magnetic field $\mathbf{B}(\mathbf{x}) = (0, 0, B)$ with the corresponding potentials (see (3.27))

$$\phi(\mathbf{x}) = 0, \qquad \mathbf{A}(\mathbf{x}) = \frac{1}{2}\mathbf{B} \times \mathbf{x} = \frac{B}{2}(-\mathbf{x}_2, \mathbf{x}_1, 0). \tag{3.68}$$

Exercise 3.7 Verify that the axial symmetry condition (3.66) implies the corresponding axial symmetry of the Maxwell field:

$$\begin{aligned} &\mathbf{E}\big(t, R_n(\theta)\mathbf{x}\big) = R_n(\theta)\mathbf{E}(t, \mathbf{x}) \quad \text{and} \\ &\mathbf{B}\big(t, R_n(\theta)\mathbf{x}\big) = R_n(\theta)\mathbf{B}(t, \mathbf{x}), \quad \theta \in [0, 2\pi]. \end{aligned} \tag{3.69}$$

Hint: Use the formulas (3.27).

III. Charge The charge (3.59) is conserved due to the Hermitian symmetry (3.55).

Lemma 3.8 *The charge is conserved,*

$$Q(t) = e \int \big|\psi(t, \mathbf{x})\big|^2 d\mathbf{x} = \text{const}, \quad t \in \mathbb{R}. \tag{3.70}$$

Proof This follows by the differentiation:

$$\dot{Q} = e\big[\langle \dot{\psi}, \psi \rangle + \langle \psi, \dot{\psi} \rangle\big] = e\left[-\frac{i}{\hbar}\big\langle H(t)\psi, \psi\big\rangle + \frac{i}{\hbar}\big\langle \psi, H(t)\psi\big\rangle\right] = 0 \tag{3.71}$$

due to the Hermitian symmetry of the Schrödinger operator $H(t)$. □

Remark 3.9 All conservation laws (3.61), (3.64), (3.67), are particular cases of the general Noether Theorem 13.18, as applied to the corresponding symmetry group: the time translations for (3.61), the space translations for (3.64), the space rotations for (3.67), and the phase rotation for (3.70), see Exercise 13.21. The Noether Theorem additionally provides the *fluxes* of energy, momentum, angular momentum and charge satisfying the corresponding *continuity equations*. The detailed calculations will be given in Sect. 13.4.

3.3.3 Correspondence Principle

Let us show that the quantum observables turn into the corresponding classical ones as $\hbar \to 0$.

I. Energy The definition (3.56) can be justified by the correspondence with the classical Hamiltonian (3.30) for quasiclassical one-particle solutions (3.48) with the normalization (3.48), provided that the potentials $\phi(t,\mathbf{x})$ and $\mathbf{A}(t,\mathbf{x})$ are vary slowly. First, we rewrite (3.56) as follows:

$$\begin{aligned} E(t) &:= \langle\psi(t), H(t)\psi(t)\rangle = \left\langle \psi(t,\mathbf{x}), \left(\frac{1}{2\mathrm{m}}\left[\hat{\mathbf{p}} - \frac{e}{c}\mathbf{A}(t,\mathbf{x})\right]^2 + e\phi(t,\mathbf{x})\right)\psi(t,\mathbf{x})\right\rangle \\ &= \frac{1}{2\mathrm{m}}\left\langle \left[\hat{\mathbf{p}} - \frac{e}{c}\mathbf{A}(t,\mathbf{x})\right]\psi(t,\mathbf{x}), \left[\hat{\mathbf{p}} - \frac{e}{c}\mathbf{A}(t,\mathbf{x})\right]\psi(t,\mathbf{x})\right\rangle \\ &\quad + \langle\psi(t,\mathbf{x}), e\phi(t,\mathbf{x})\psi(t,\mathbf{x})\rangle; \end{aligned} \tag{3.72}$$

here, the last identity follows by partial integration. Further, for the quasiclassical solutions (3.48), we have, asymptotically,

$$-i\hbar\nabla\psi(t,\mathbf{x}) \sim \nabla S(t,\mathbf{x})\psi(t,\mathbf{x}), \quad \hbar \ll 1. \tag{3.73}$$

Therefore, (3.50) and (3.51) with small ε imply, for small t, that

$$\hat{\mathbf{p}}\psi(t,\mathbf{x}) \sim \mathbf{p}(t)\psi(t,\mathbf{x}), \quad \hbar \ll 1. \tag{3.74}$$

Finally, substituting (3.74) into (3.72) and using by (3.35) and (3.51) with small ε, we arrive at the asymptotics

$$\begin{aligned} E(t) &\sim \int \left(\frac{1}{2\mathrm{m}}\left[\mathbf{p}(t) - \frac{e}{c}\mathbf{A}(t,\mathbf{x})\right]^2 |\psi(t,\mathbf{x})|^2 + e\phi(t,\mathbf{x})|\psi(t,\mathbf{x})|^2\right) d\mathbf{x} \\ &\approx \frac{1}{2\mathrm{m}}\left[\mathbf{p}(t) - \frac{e}{c}\mathbf{A}\big(t,\mathbf{x}(t)\big)\right]^2 + e\phi\big(t,\mathbf{x}(t)\big), \quad \hbar \ll 1, \end{aligned} \tag{3.75}$$

since the potentials vary slowly.

II. Momentum and Angular Momentum The definitions (3.57) and (3.58) of the momentum and angular momentum can be justified similarly.

III. Charge and Current The charge definition (3.59) is justified by the energy expression (3.72), where the last term is the integral of the product $\rho(t,\mathbf{x})\phi(t,\mathbf{x})$ with $\rho(t,\mathbf{x}) := e|\psi(t,\mathbf{x})|^2$.

Moreover, the last argument justifies the identification (3.34) for the charge density. Then, for the current density, the identification is justified by the quasiclassical asymptotics (3.52).

3.4 Charge Continuity Equation

Let us verify that the continuity equation (1.4) holds for the charge and current densities (3.34) corresponding to every solution of the Schrödinger equation (3.32).

First we represent the equation and formulas in a real form. Namely, let us identify the complex numbers $\psi = \psi_1 + i\psi_2 \in \mathbb{C}$ with the corresponding real vectors $\binom{\psi_1}{\psi_2} \in \mathbb{R}^2$, so that the multiplication by i will be in correspondence with the application of the skew-symmetric matrix $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Then $\psi(x,t) = \psi_1(x,t) + i\psi_2(x,t)$ is identified with the $\mathbb{R}^2$-valued function $\Psi(x,t) := \binom{\psi_1(x,t)}{\psi_2(x,t)}$, and the Schrödinger equation (3.53) takes the form

$$\hbar J\dot{\Psi} = \mathbb{H}(t)\Psi, \quad \mathbb{H}(t) = \frac{1}{2\mathrm{m}}\Big[-J\hbar\nabla - \frac{e}{c}\mathbf{A}(t,\mathbf{x})\Big]^2 + e\phi(t,\mathbf{x}). \tag{3.76}$$

In the real form, the charge and current densities (3.34) read as follows:

$$\rho(t,\mathbf{x}) = e\big|\Psi(t,\mathbf{x})\big|^2, \qquad \mathbf{j}^k(t,\mathbf{x}) = \frac{e}{\mathrm{m}}\Psi(t,\mathbf{x}) \cdot \Big[-J\hbar\nabla_k - \frac{e}{c}\mathbf{A}^k(t,\mathbf{x})\Big]\Psi(t,\mathbf{x}). \tag{3.77}$$

Now we can prove (1.4):

$$\dot{\rho}(t,\mathbf{x}) + \operatorname{div}\mathbf{j}(t,\mathbf{x}) = 0. \tag{3.78}$$

Indeed, differentiating $\rho(t,\mathbf{x}) = e\Psi(t,\mathbf{x}) \cdot \Psi(t,\mathbf{x})$, and using the Schrödinger equation (3.76), it is found that

$$\hbar\partial_t\rho = 2\hbar e\Psi \cdot \partial_t\Psi = -2e\Psi \cdot \big[J\mathbb{H}(t)\Psi\big]. \tag{3.79}$$

Let us first consider the particular case $\phi = 0$, $\mathbf{A} = 0$. Then $\mathbb{H}(t) = -\frac{\hbar^2}{2m}\Delta$, and (3.78) follows because

$$\begin{aligned}\hbar\partial_t\rho &= -2e\Psi \cdot \big[J\mathbb{H}(t)\Psi\big] = \frac{e\hbar^2}{m}\psi \cdot [J\Delta\Psi] = \frac{e\hbar^2}{m}\sum_{k=1}^{3}\nabla_k\big(\Psi \cdot [J\nabla_k\Psi]\big)\\ &= -\hbar\operatorname{div}\mathbf{j}.\end{aligned} \tag{3.80}$$

Consider the general case. Using (3.79) and (3.76), we have

$$\begin{aligned}\hbar\partial_t\rho &= -2e\Psi \cdot \Big[J\frac{1}{2\mathrm{m}}\Big[-J\hbar\nabla - \frac{e}{c}\mathbf{A}(t,\mathbf{x})\Big]^2 + e\phi(t,\mathbf{x})\Big]\Psi\\ &= \frac{e}{\mathrm{m}}\Psi \cdot \Big[J\Big[J\hbar\nabla + \frac{e}{c}\mathbf{A}(t,\mathbf{x})\Big]^2\Psi\Big],\end{aligned} \tag{3.81}$$

since ϕ is real and since $\Psi \cdot [J\Psi] = 0$. Further, similarly to (3.80), we obtain

$$\begin{aligned}\hbar\partial_t\rho &= \frac{e}{\mathrm{m}}\sum_{k=1}^{3}\Psi\cdot\Big[J\Big[(J\hbar\nabla_k)^2 + J\hbar\nabla_k\frac{e}{c}\mathbf{A}^k(t,\mathbf{x}) + \frac{e}{c}\mathbf{A}^k(t,\mathbf{x})J\hbar\nabla_k\Big]\Psi\Big]\\ &= \frac{e}{\mathrm{m}}\sum_{k=1}^{3}\hbar\nabla_k\Big(\Psi\cdot\Big[\Big[-J\hbar\nabla_k - \frac{e}{c}\mathbf{A}^k(t,\mathbf{x})\Big]\Psi\Big]\Big)\\ &\quad -\frac{e}{\mathrm{m}}\sum_{k=1}^{3}\hbar\nabla_k\Psi\cdot\Big[\Big[-J\hbar\nabla_k - \frac{e}{c}\mathbf{A}^k(t,\mathbf{x})\Big]\Psi\Big]\\ &\quad -\frac{e}{\mathrm{m}}\sum_{k=1}^{3}\Psi\cdot\Big[\frac{e}{c}\mathbf{A}^k(t,\mathbf{x})\hbar\nabla_k\Psi\Big].\end{aligned}$$

It remains to observe that the last two lines cancel, since the potentials $\mathbf{A}^k$ are real.

3.5 Equivalence of Heisenberg's and Schrödinger's Theories

In 1926, Schrödinger had shown the equivalence of his 'Wave Mechanics' to the Heisenberg's 'Matrix Mechanics' (see [162]).

3.5.1 Heisenberg Observables

All Schrödinger's observables (3.72), (3.57), (3.58) and (3.59) are quadratic forms on the phase space $\mathcal{E} := L^2(\mathbb{R}^3)$,

$$\begin{aligned} E(t) &= \big\langle\psi(t), \hat{E}(t)\psi(t)\big\rangle, \qquad & \mathbf{p}(t) &= \big\langle\psi(t), \hat{\mathbf{p}}\psi(t)\big\rangle,\\ \mathbf{L}(t) &= \big\langle\psi(t), \hat{\mathbf{L}}\psi(t)\big\rangle, \qquad & Q(t) &= \big\langle\psi(t), \hat{Q}\psi(t)\big\rangle,\end{aligned} \tag{3.82}$$

where $\langle\psi_1, \psi_2\rangle$ denotes the Hermitian scalar product (3.54), and $\hat{E}(t)$, $\hat{\mathbf{p}}$, $\hat{\mathbf{L}}$, $\hat{Q}$ are the corresponding operators,

$$\hat{E}(t) = H(t), \qquad \hat{\mathbf{p}} = -i\hbar\nabla, \qquad \hat{\mathbf{L}} = \hat{\mathbf{x}}\times\hat{\mathbf{p}}, \qquad \hat{Q} = I. \tag{3.83}$$

Here $\hat{\mathbf{x}}^k\psi = \mathbf{x}^k\psi(\mathbf{x})$ is the quantum observable of the kth coordinate. Operators (3.83) are called the *quantum observables* corresponding to the energy, momentum, angular momentum and charge, respectively.

Exercise 3.10 Verify the commutation relations

$$\big[\hat{\mathbf{p}}_k, \hat{\mathbf{x}}^n\big] = -i\hbar\delta_k^n, \qquad [\hat{\mathbf{p}}_k, \hat{\mathbf{p}}_n] = \big[\hat{\mathbf{x}}^k, \hat{\mathbf{x}}^n\big] = 0, \tag{3.84}$$

where $[\cdot,\cdot]$ denotes the *commutator* $[A, B] := AB - BA$.

We note that

$$H(t) = \mathcal{H}(\hat{\mathbf{x}}, \hat{\mathbf{p}}, t) := \frac{1}{2\mathrm{m}}\left(\hat{\mathbf{p}} - \frac{e}{c}\mathbf{A}(t, \hat{\mathbf{x}})\right)^2 + e\phi(t, \hat{\mathbf{x}}). \tag{3.85}$$

Exercise 3.11 Verify the formula

$$\hat{\mathbf{L}}_n = -i\hbar\frac{\partial}{\partial\varphi_n}, \tag{3.86}$$

where φ_n is the angle of rotation around the vector e_n in a positive direction. **Hint:** Consider $n = 3$ and choose the polar coordinate in the plane $\mathbf{x}^1, \mathbf{x}^2$.

Exercise 3.12 Verify that operators $\hat{\mathbf{p}}_n, \hat{\mathbf{L}}_n$ are symmetric (similarly to $\hat{E}(t) := H(t)$; see (3.55)).

Formulas (3.82) motivate the following

Definition 3.13

(i) A **quantum observable** is a linear symmetric operator $\hat{\mathbf{M}}$ in the Hilbert space $\mathcal{E}$.
(ii) Quadratic form $\mathbf{M}(\psi) := \langle\psi, \hat{\mathbf{M}}\psi\rangle$ is the **mean value** of the observable at the state $\psi \in \mathcal{E}$.

Note that if ψ is a normalized eigenvector for $\hat{\mathbf{M}}$ (i.e., $\hat{\mathbf{M}}\psi = \lambda\psi$ and $\|\psi\| = 1$), then the corresponding mean value is the eigenvalue: $\mathbf{M}(\psi) = \lambda$.

3.5.2 Heisenberg's Picture

Let us introduce the dynamical group $U(t)$ of the Schrödinger equation (3.37) with static potentials.

Definition 3.14 For $t \in \mathbb{R}$, the operator $U(t)$ is defined by $U(t)\psi_0 = \psi(t, \cdot)$, where $\psi(t, \cdot)$ is the solution of the Schrödinger equation (3.37) with initial state ψ_0.

The Schrödinger equation implies that

$$\dot{U}(t) = -\frac{i}{\hbar}HU(t) = -\frac{i}{\hbar}U(t)H, \tag{3.87}$$

and so formally $U(t) = e^{-\frac{i}{\hbar}Ht}$. The charge conservation (3.70) means that $U(t)$ is a unitary operator.

Definition 3.15 The **Heisenberg representation** of a quantum observable $\hat{\mathbf{M}}$ is the operator function

$$\hat{\mathbf{M}}(t) := U(-t)\hat{\mathbf{M}}U(t), \quad t \in \mathbb{R}. \tag{3.88}$$

Using this definition and (3.84), we have the commutation relations

$$\big[\hat{\mathbf{p}}_k(t), \hat{\mathbf{x}}^n(t)\big] = -i\hbar\delta_k^n, \qquad \big[\hat{\mathbf{p}}_k(t), \hat{\mathbf{p}}_n(t)\big] = \big[\hat{\mathbf{x}}^k(t), \hat{\mathbf{x}}^n(t)\big] = 0. \tag{3.89}$$

Furthermore, this definition implies that

$$\mathbf{M}\big(\psi(t)\big) := \big\langle \psi(t), \hat{\mathbf{M}}\psi(t)\big\rangle = \big\langle \psi(0), \hat{\mathbf{M}}(t)\psi(0)\big\rangle, \quad t \in \mathbb{R}, \tag{3.90}$$

for any solution $\psi(t) := \psi(t, \cdot)$ of the Schrödinger equation (3.37).

Exercise 3.16 Verify (3.90). **Hint:** $U(t)$ is a unitary operator.

Exercise 3.17 Verify the identities

$$\hat{E}(t) = \mathcal{H}\big(\hat{\mathbf{x}}(t), \hat{\mathbf{p}}(t), t\big), \qquad \hat{\mathbf{L}}(t) = \hat{\mathbf{x}}(t) \times \hat{\mathbf{p}}(t). \tag{3.91}$$

Exercise 3.18 (cf. (2.32)) Prove the following identities: for $N = 0, 1, 2, \ldots$

$$\begin{aligned} \big[\hat{\mathbf{p}}_k(t), \big(\hat{\mathbf{x}}^n(t)\big)^N\big] &= -i\hbar\delta_k^n N\big(\hat{\mathbf{x}}^n(t)\big)^{N-1}, \\ \big[\hat{\mathbf{x}}^n(t), \big(\hat{\mathbf{p}}_k(t)\big)^N\big] &= i\hbar\delta_k^n N\big(\hat{\mathbf{p}}_k(t)\big)^{N-1}. \end{aligned} \tag{3.92}$$

3.5.3 Heisenberg's Equation

The relations (3.87) imply the **Heisenberg equation** (2.25)

$$i\hbar\dot{\hat{\mathbf{M}}}(t) = \big[\hat{\mathbf{M}}(t), H\big]. \tag{3.93}$$

Exercise 3.19 Verify (3.93).

Using equation (3.93), we obtain the following lemma.

Lemma 3.20 *A quantum observable* $\hat{\mathbf{M}}$ *is conserved* (*i.e.*, $\hat{\mathbf{M}}(t) \equiv \hat{\mathbf{M}}$), *provided* $[\hat{\mathbf{M}}, H] = 0$.

Proof By (3.87),

$$\begin{aligned} \frac{d}{dt}\hat{\mathbf{M}}(t) &= \frac{d}{dt}U(-t)\hat{\mathbf{M}}U(t) = -\dot{U}(-t)\hat{\mathbf{M}}U(t) + U(-t)\hat{\mathbf{M}}\dot{U}(t) \\ &= -\frac{i}{\hbar}\big(-U(-t)H\hat{\mathbf{M}}U(t) + U(-t)\hat{\mathbf{M}}HU(t)\big) \\ &= -\frac{i}{\hbar}U(-t)[\hat{\mathbf{M}}, H]U(t) = 0. \quad \square \end{aligned} \tag{3.94}$$

Note that our proofs of the energy, momentum, angular momentum and charge conservations (3.61), (3.64), (3.67), (3.70) depend on the commutation of the corresponding observables (3.83) with the Hamilton operator.

Example 3.21 The Heisenberg equations for the observables $\hat{\mathbf{x}}(t)$ and $\hat{\mathbf{p}}(t)$ are as follows:

$$\begin{cases} \dot{\hat{\mathbf{x}}}(t) = \dfrac{1}{\mathrm{m}}\Big(\hat{\mathbf{p}}(t) - \dfrac{e}{c}\mathbf{A}\big(t, \hat{\mathbf{x}}(t)\big)\Big), \\ \dot{\hat{\mathbf{p}}}(t) = -e\nabla\phi\big(t, \hat{\mathbf{x}}(t)\big) + \dfrac{e}{2\mathrm{m}c}\Big\{\big[\nabla\mathbf{A}\big(t, \hat{\mathbf{x}}(t)\big)\big] \cdot \Big(\hat{\mathbf{p}}(t) - \dfrac{e}{c}\mathbf{A}\big(t, \hat{\mathbf{x}}(t)\big)\Big) \\ \qquad + \Big(\hat{\mathbf{p}}(t) - \dfrac{e}{c}\mathbf{A}\big(t, \hat{\mathbf{x}}(t)\big)\Big) \cdot \big[\nabla\mathbf{A}\big(t, \hat{\mathbf{x}}(t)\big)\big]\Big\}. \end{cases} \tag{3.95}$$

Exercise 3.22 Prove (3.95). **Hint:** Use (3.93), (3.85), and (3.92) with $N = 1, 2$.

Note that the Heisenberg equations (3.95) are not identical to (3.31), though (3.95) becomes (3.31) after substituting $x(t)$ and $p(t)$ for $\hat{x}(t)$ and $\hat{p}(t)$, respectively. On the other hand, equations (2.16) are formally identical with (2.1). This difference reflects distinct algebraic structures of the Hamiltonian functionals (3.85) and (2.2).

3.5.4 Correspondence Between Heisenberg's and Schrödinger's Pictures

Relations (3.88), (3.90) establish the equivalence of Schrödinger's and Heisenberg's theories. Namely

H (i) In the Heisenberg picture, a quantum observable $\hat{\mathbf{M}}(t)$ changes in time according to the Heisenberg equation (3.93), while the state ψ is unchanged; and
  (ii) the *mean value* of $\hat{\mathbf{M}}(t)$ is defined by the right hand side of (3.90).

S (i) In the Schrödinger picture, a quantum observable $\hat{\mathbf{M}}$ is unchanged, while the state $\psi(t)$ changes in time according to the Schrödinger equation (3.37); and
  (ii) the *mean value* of $\hat{\mathbf{M}}$ is defined by the middle term of (3.90).

Both the mean values coincide, and the quantum observables are related by (3.88).

Chapter 4
Lagrangian Formalism

Abstract The Lagrangian form of the Schrödinger equation is useful in the identification of the quantum energy, momentum and angular momentum, which are the conserved quantities suggested by the Noether's theory of invariants.

On the other hand, the charge and current densities in classical electrodynamics correspond to derivatives of the Lagrangian density with respect to the Maxwell potentials. Respectively, to identify correctly the quantum charge and current, one should introduce the Lagrangian density depending on the wave function ψ *and* on the Maxwell potentials. This Lagrangian density should correspond to the *coupled Maxwell–Schrödinger* equations.

4.1 Hamiltonian and Lagrangian Formalism

The Schrödinger equation can be represented in the Hamiltonian and Lagrangian forms.

4.1.1 Hamiltonian Formalism

In view of (3.55), the Schrödinger operator $H(t)$ is symmetric in the complex Hilbert space $L^2(\mathbb{R}^3, \mathbb{C})$. Correspondingly, it follows from (3.76) that the corresponding real version $\mathbb{H}(t)$ is a symmetric operator in the real Hilbert space $L^2(\mathbb{R}^3, \mathbb{R}^2)$. Hence,

$$\mathbb{H}(t)\Psi = D\mathcal{H}(\Psi, t), \quad \mathcal{H}(\Psi, t) := \frac{1}{2}\big\langle \Psi, \mathbb{H}(t)\Psi \big\rangle_{\mathbb{R}} = \frac{1}{2}\big\langle \psi, H(t)\psi \big\rangle, \tag{4.1}$$

where $\langle \cdot, \cdot \rangle_{\mathbb{R}}$ denotes the inner product on $L^2(\mathbb{R}^3, \mathbb{R}^2)$. Also, $D\mathcal{H}(\Psi, t)$ is the Gâteaux derivative defined as the distribution

$$\big\langle D\mathcal{H}(\Psi, t), \Phi \big\rangle_{\mathbb{R}} := \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \mathcal{H}(\Psi + \varepsilon\Phi, t), \quad \Phi \in C_0^\infty\big(\mathbb{R}^3, \mathbb{R}^2\big). \tag{4.2}$$

Now the Schrödinger equation (3.76) reads

$$\hbar\dot{\Psi} = -J D\mathcal{H}(\Psi, t); \tag{4.3}$$

A. Komech, *Quantum Mechanics: Genesis and Achievements*,
DOI 10.1007/978-94-007-5542-0_4,

this formally coincides with the standard Hamiltonian system (13.8) if we identify ψ_1 with q and ψ_2 with p. Equivalently, we can identify ψ_1 with p and ψ_2 with $-q$.

4.1.2 Lagrangian Formalism

Using the Legendre transform, the Hamiltonian equation (4.3) can be written in the Lagrangian form

$$\frac{d}{dt}D_{\dot\Psi}L\big(\Psi(t),\dot\Psi(t)\big)=D_\Psi L\big(\Psi(t),\dot\Psi(t)\big) \tag{4.4}$$

with the Lagrangian functional $L(\Psi)=\hbar\langle\psi_2,\dot\psi_1\rangle_{\mathbb{R}}-\boldsymbol{\mathcal{H}}(\Psi)$ (cf. (13.7)). Equivalently, we can take the Lagrangian functional $-\hbar\langle\psi_1,\dot\psi_2\rangle_{\mathbb{R}}-\boldsymbol{\mathcal{H}}(\Psi)$. Therefore, we can also take the sum

$$L(\Psi,\dot\Psi)=\hbar\big[\langle\psi_2,\dot\psi_1\rangle_{\mathbb{R}}-\langle\psi_1,\dot\psi_2\rangle_{\mathbb{R}}\big]-2\boldsymbol{\mathcal{H}}(\Psi) \tag{4.5}$$

for the Lagrangian, since the Lagrange equation (4.4) is linear with respect to the Lagrangian functional. Integration by parts implies that the sum admits the integral representation

$$L(\Psi,\dot\Psi)=\int_{\mathbb{R}^3}\mathcal{L}\big(x,\Psi(x),\nabla\Psi(x)\big)\,d\mathbf{x}, \tag{4.6}$$

where $\mathbf{x}=(\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3)\in\mathbb{R}^3$, $x=(t,\mathbf{x})\in\mathbb{R}^4$, and the corresponding *Lagrangian density* $\mathcal{L}(x,\Psi,\nabla\Psi)$ is given by

$$\begin{aligned}&\mathcal{L}(x,\Psi,\nabla\Psi)=\big[J\hbar\nabla_0\Psi-e\phi(t,\mathbf{x})\Psi\big]\cdot\Psi-\frac{1}{2\mathrm{m}}\left|-J\hbar\boldsymbol{\nabla}\Psi-\frac{e}{c}\mathbf{A}(t,\mathbf{x})\Psi\right|^2,\\ &\nabla_0:=\partial_t,\end{aligned} \tag{4.7}$$

since $[J\hbar\nabla_0\Psi]\cdot\Psi=\hbar[\psi_2\dot\psi_1-\psi_1\dot\psi_2]$, where '$\cdot$' is the real inner product on $\mathbb{R}^2$.

4.1.3 Variational Principle and Euler–Lagrange Equations

Let Ω be any open bounded region of $\mathbb{R}^4$, and let $S_\Omega(\Psi)$ be the *action functional* defined, for $\Psi\in C^1(\overline{\Omega},\mathbb{R}^2)$, by

$$S_\Omega(\Psi):=\int_\Omega\mathcal{L}\big(x,\Psi(x),\nabla\Psi(x)\big)\,dx. \tag{4.8}$$

Denote by $DS_\Omega(\Psi)$ the variational Gâteaux derivative, defined as the distribution in Ω

$$\big\langle DS_\Omega(\Psi),\Phi\big\rangle:=\frac{d}{d\varepsilon}\bigg|_{\varepsilon=0}S_\Omega(\Psi+\varepsilon\Phi),\quad \Phi\in C_0^\infty\big(\Omega,\mathbb{R}^2\big). \tag{4.9}$$

Let us show that the Schrödinger equation (3.76) in Ω can be represented in the four-dimensional variational Euler–Lagrange form,

$$D S_\Omega(\Psi) = \mathcal{L}_\Psi\big(x, \Psi(x), \nabla\Psi(x)\big) - \sum_{\alpha=0}^{3} \nabla_\alpha \mathcal{L}_{\nabla_\alpha\Psi}\big(x, \Psi(x), \nabla\Psi(x)\big) = 0,$$
$$x \in \Omega, \tag{4.10}$$

where all the derivatives are understood in the sense of distributions.

For regular solutions $\Psi \in C^2(\mathbb{R}^4)$, the first identity in (4.10) follows by standard integration by parts on the right of (4.9), since the test function $\Phi(x)$ is arbitrary. In the general case, (4.9) implies the first identity in (4.10) in the sense of distributions.

To prove the second identity, we write the Schrödinger equation (3.76) as follows:

$$\mathbb{L}\Psi(x) := \big[J\hbar\nabla_0 - e\phi(x)\big]\Psi(x) - \frac{1}{2\mathrm{m}}\left[-J\hbar\boldsymbol{\nabla} - \frac{e}{c}\mathbf{A}(x)\right]^2 \Psi(x) = 0,$$
$$x \in \mathbb{R}^4, \tag{4.11}$$

where $\mathbb{L}$ is a Hermitean symmetric operator. Integrating by parts, we obtain

$$S_\Omega(\Psi + \varepsilon\Phi) = S_\Omega(\Psi) + 2\varepsilon\langle\Psi, \mathbb{L}\Phi\rangle + \varepsilon^2\langle\Phi, \mathbb{L}\Phi\rangle. \tag{4.12}$$

Hence,

$$D S_\Omega(\Psi) = 2\mathbb{L}\Psi$$

in the sense of distributions. Therefore, the second identity of (4.10) is equivalent to the Schrödinger equation (4.11).

Remark 4.1 The general Noether Theorem 13.18 can be applied to the Schrödinger equation (3.32) in the Lagrangian form (4.10). It implies the conservation laws (3.61), (3.64), (3.67) and (3.70) (see Theorems 13.22, 13.23, and Corollary (13.26)). We gave direct independent proofs in Sect. 3.3 for the convenience of the reader.

Definition 4.2 For the general Lagrangian density $\mathcal{L}(x, \Psi, \nabla\Psi)$ of a field $\Psi(x) \in \mathbb{R}^N$, the *canonically conjugate fields* $\Pi^\nu(x)$ are defined by

$$\Pi^\nu = \mathcal{L}_{\nabla_\nu\Psi}(x, \Psi, \nabla\Psi), \quad \nu = 0, \ldots, 3. \tag{4.13}$$

Example 4.3 For the Schrödinger Lagrangian density (4.7), we get

$$\begin{aligned} \Pi^0(x) &= -J\hbar\Psi(x), \\ \Pi^k(x) &= -\frac{J\hbar}{\mathrm{m}}\left[-J\hbar\boldsymbol{\nabla}_k - \frac{e}{c}\mathbf{A}^k(x)\right]\Psi(x), \quad k = 1, 2, 3. \end{aligned} \tag{4.14}$$

In this notation, the Euler–Lagrange equations (4.10) read:

$$\nabla_\nu \Pi^\nu(x) = \mathcal{L}_\Psi\big(x, \Psi(x), \nabla\Psi(x)\big), \quad x \in \mathbb{R}^4, \tag{4.15}$$

which is equivalent to the Schrödinger equation (3.76). Here and in what follows, we will use Einstein's convention

$$\nabla_\nu \Pi^\nu(x) := \sum_{\nu=0}^{3} \nabla_\nu \Pi^\nu(x), \qquad \nabla_k \Pi^k(x) := \sum_{k=1}^{3} \nabla_k \Pi^k(x).$$

4.2 Maxwell–Schrödinger Equations

The simultaneous evolution of the full system of the electron wave function and Maxwell field is determined by the coupled *semiclassical* Schrödinger–Maxwell equations (3.33), (12.50). The coupling is provided by identification (3.77) of the charge and current densities in (12.50) via the Schrödinger wave function. The coupled system is nonlinear, and so the approximate solutions could be calculated through Born's approximation and perturbation theory.

4.2.1 Lagrangian Density

Let us introduce a Lagrangian density for the coupled system. For the Schrödinger equation (3.33) with given external Maxwell potentials, the Lagrangian density is given by (4.7):

$$\mathcal{L}_S(x, \psi, \nabla\psi) = \big[i\hbar\partial_t - e\phi(x)\big]\psi \cdot \psi - \frac{1}{2\mathrm{m}}\left|-i\hbar\boldsymbol{\nabla}\psi - \frac{e}{c}\mathbf{A}(x)\psi\right|^2, \tag{4.16}$$

where complex values $\psi \in \mathbb{C}$ are identified with the corresponding real vectors $(\operatorname{Re}\psi, \operatorname{Im}\psi) \in \mathbb{R}^2$, and '$\cdot$' denotes the inner product on $\mathbb{R}^2$.

On the other hand, for a Maxwell field in the presence of known charge and current densities, the Lagrangian density is given by (12.56),

$$\mathcal{L}_M(x, \mathcal{A}, \nabla\mathcal{A}) = -\frac{1}{16\pi}\mathcal{F}^{\mu\nu}\mathcal{F}_{\mu\nu} - \mathcal{J}^\nu(x)\mathcal{A}_\nu, \quad (x, \mathcal{A}, \nabla\mathcal{A}) \in \mathbb{R}^4 \times \mathbb{R}^4 \times \mathbb{R}^{16}, \tag{4.17}$$

where $\mathcal{F}^{\mu\nu} := \partial^\mu \mathcal{A}^\nu - \partial^\nu \mathcal{A}^\mu$ and $\mathcal{F}_{\mu\nu} := \partial_\mu \mathcal{A}_\nu - \partial_\nu \mathcal{A}_\mu$. We should combine the densities in a natural way. To do so we choose the combined density as follows:

$$\begin{aligned}\mathcal{L}(x, \psi, \nabla\psi, \mathcal{A}_\mu, \nabla\mathcal{A}_\mu) = [i\hbar\partial_t - e\phi]\psi \cdot \psi - \frac{1}{2\mathrm{m}}\left|-i\hbar\boldsymbol{\nabla}\psi - \frac{e}{c}\mathbf{A}\psi\right|^2 \\ - \frac{1}{16\pi}\mathcal{F}^{\mu\nu}\mathcal{F}_{\mu\nu};\end{aligned} \tag{4.18}$$

here $\phi = \mathcal{A}_0$ and $\mathbf{A}_k = -\mathcal{A}_k$ in accordance with (12.49). The choice is motivated by the following arguments:

I. The variation (4.10) of the density (4.18) in $\psi(x)$ provides the Schrödinger equation (3.32).
II. The introduction of the Lagrangian density (4.18) is equivalent to the identification of the *interaction terms* in the Lagrangian densities (4.16) and (4.17). Similar identification holds in the classical electrodynamics (see Sect. 12.6).

These arguments suggest the Lagrangian density (4.18) for coupled Maxwell–Schrödinger fields. However, a proper justification depends on the subsequent development and experimentally observations, as for every axiomatic theory.

The Euler–Lagrange equations (4.15), corresponding to (4.18), are as follows:

$$\begin{cases} \big[i\hbar\partial_t - e\phi(t,\mathbf{x})\big]\psi(t,\mathbf{x}) = \dfrac{1}{2\mathrm{m}}\left[-i\hbar\boldsymbol{\nabla} - \dfrac{e}{c}\mathbf{A}(t,\mathbf{x})\right]^2\psi(t,\mathbf{x}), \\ \dfrac{1}{4\pi}\nabla_\mu\mathcal{F}^{\mu\nu}(t,\mathbf{x}) = \mathcal{J}^\nu(t,\mathbf{x}), \quad \nu = 0,\ldots,3; \end{cases} \tag{4.19}$$

here $\nabla_0 := \partial_{x_0}$ and $x_0 := ct$. The first (Schrödinger) equation corresponds to the variation in ψ. The second (Maxwell) equation corresponds to the variations in $\mathcal{A}_\mu = (\phi, -\mathbf{A})$: the left hand side is the (negative) of the variation of the last term in the Lagrangian density (4.18), as in the proof of Proposition 12.15, while the 4-currents $\mathcal{J}^\nu$ on the right is the variation (13.99) of the interaction terms containing both ψ and $\mathcal{A}$:

$$\mathcal{J}^\nu(x) = \begin{pmatrix} e|\psi(t,\mathbf{x})|^2, & \nu = 0 \\ \frac{e}{\mathrm{m}c}\psi(t,\mathbf{x})\cdot[-J\hbar\boldsymbol{\nabla}_\nu - \frac{e}{c}\mathbf{A}^\nu(t,\mathbf{x})]\psi(t,\mathbf{x}), & \nu = 1,2,3 \end{pmatrix}. \tag{4.20}$$

Obviously, this 4-current is identical to $(\rho, \frac{\mathbf{j}}{c})$, where ρ and $\mathbf{j}$ are given by (3.34). This identity is a particular case of relations (13.103), which hold by Lemma 13.28, because the Lagrangian density (4.18) is gauge invariant.

The system (4.19) describes the dynamics of the wave field ψ in its 'own' induced Maxwell potentials $\phi(t,\mathbf{x}), \mathbf{A}(t,\mathbf{x})$ generated by the charges and currents of the wave field. Now let us introduce external potentials $\phi_{\text{ext}}(t,\mathbf{x})$, $\mathbf{A}_{\text{ext}}(t,\mathbf{x})$ of a Maxwell field generated by some external sources. The introduction is formalized by the Lagrangian density

$$\begin{aligned}\mathcal{L} = {} & \big[i\hbar\partial_t - e\big(\phi(x) + \phi_{\text{ext}}(x)\big)\big]\psi\cdot\psi - \frac{1}{2\mathrm{m}}\left|-i\hbar\boldsymbol{\nabla}\psi - \frac{e}{c}\big(\mathbf{A}(x) + \mathbf{A}_{\text{ext}}(x)\big)\psi\right|^2 \\ & - \frac{1}{16\pi}\mathcal{F}^{\mu\nu}\mathcal{F}_{\mu\nu}. \end{aligned} \tag{4.21}$$

The corresponding equations read

$$\begin{cases} \big[i\hbar\partial_t - e\big(\phi(t,\mathbf{x})+\phi_{\text{ext}}(t,\mathbf{x})\big)\big]\psi(t,\mathbf{x}) \\ \quad = \dfrac{1}{2\text{m}}\Big[-i\hbar\nabla - \dfrac{e}{c}\big(\mathbf{A}(t,\mathbf{x})+\mathbf{A}_{\text{ext}}(t,\mathbf{x})\big)\Big]^2\psi(t,\mathbf{x}), \\ \dfrac{1}{4\pi}\nabla_\mu\mathcal{F}^{\mu\nu}(t,\mathbf{x}) = \mathcal{J}^\nu(t,\mathbf{x}), \quad \nu=0,\dots,3, \end{cases} \tag{4.22}$$

where

$$\mathcal{J}^\nu(x) = \begin{pmatrix} \rho = e|\psi(t,\mathbf{x})|^2, & \nu=0 \\ \frac{\mathbf{j}^\nu}{c} = \frac{e}{\text{m}c}\psi(t,\mathbf{x})\cdot[-i\hbar\nabla_\nu - \frac{e}{c}(\mathbf{A}^\nu(t,\mathbf{x})+\mathbf{A}^\nu_{\text{ext}}(t,\mathbf{x}))]\psi(t,\mathbf{x}), & \nu=1,2,3 \end{pmatrix}. \tag{4.23}$$

4.2.2 Gauge Invariance and Charge Continuity

Lagrangian densities (4.18) and (4.21) are invariant with respect to the phase rotations $(\psi,\mathcal{A})\mapsto(e^{i\theta}\psi,\mathcal{A})$. Therefore, the *continuity equation for charge and current* holds

$$\dot\rho(t,\mathbf{x}) + \operatorname{div}\mathbf{j}(t,\mathbf{x}) = 0, \quad (t,\mathbf{x})\in\mathbb{R}^4; \tag{4.24}$$

this follows from Corollary 13.26 and (13.97). This result can be also verified by direct calculation (see Sect. 3.4).

4.2.3 Gauge Transformation

More general *gauge transformations* (12.46) do not change Maxwell fields $E(t,\mathbf{x})$, $B(t,\mathbf{x})$ for any real function $\chi(t,\mathbf{x})\in C^1(\mathbb{R}^4)$. Therefore, it would be natural to expect that the solutions to the coupled equations (4.19), (4.22) also do not vary too much under this transformation. More precisely, we can complete the transformation of potentials (12.46) with the corresponding transformation of the wave function:

$$\left|\begin{array}{l} \phi(t,\mathbf{x})\mapsto\phi(t,\mathbf{x})+\dfrac{1}{c}\dot\chi(t,\mathbf{x}), \qquad \mathbf{A}(t,\mathbf{x})\mapsto\mathbf{A}(t,\mathbf{x})-\nabla\chi(t,\mathbf{x}), \\ \psi(t,\mathbf{x})\mapsto e^{-i\frac{e}{c\hbar}\chi(t,\mathbf{x})}\psi(t,\mathbf{x}). \end{array}\right. \tag{4.25}$$

It is easily checked that the new functions also provide a solution of Eqs. (4.19) and (4.22), respectively. Moreover, transformations (4.25) do not change the Lagrangian densities (4.18) and (4.21) and the electric charge and current densities $\rho(t,\mathbf{x})$ and $\mathbf{j}(t,\mathbf{x})$ of (4.23).

An appropriate gauge transformation (4.25) provides the *Lorentz gauge* condition (12.45); then the coupled equations (4.22) reads (cf. (12.47)):

$$\left\{\begin{array}{l} \big[i\hbar\partial_t - e\big(\phi(t,\mathbf{x}) + \phi_{\text{ext}}(t,\mathbf{x})\big)\big]\psi(t,\mathbf{x}) \\ \quad = \dfrac{1}{2\mathrm{m}}\Big[-i\hbar\nabla - \dfrac{e}{c}\big(\mathbf{A}(t,\mathbf{x}) + \mathbf{A}_{\text{ext}}(t,\mathbf{x})\big)\Big]^2\psi(t,\mathbf{x}), \\ \dfrac{1}{4\pi}\Box\phi(t,\mathbf{x}) = e\big|\psi(t,\mathbf{x})\big|^2, \\ \dfrac{1}{4\pi}\Box\mathbf{A}(t,\mathbf{x}) = \dfrac{e}{\mathrm{m}c}\psi(t,\mathbf{x})\cdot\Big[-i\hbar\nabla - \dfrac{e}{c}\big(\mathbf{A}(t,\mathbf{x}) + \mathbf{A}_{\text{ext}}(t,\mathbf{x})\big)\Big]\psi(t,\mathbf{x}); \end{array}\right. \tag{4.26}$$

here, $\Box := \frac{1}{c^2}\partial_t^2 - \Delta$.

4.2.4 Perturbation Theory

The coupled Maxwell–Schrödinger equations (4.22) can be handled perturbatively in a wide class of the problems since the *Sommerfeld fine structure constant* is very small. This is a fundamental feature of quantum mechanics and quantum electrodynamics.

The Born Approximation

Assume that the *induced Maxwell field* $\phi(t,\mathbf{x}), \mathbf{A}(t,\mathbf{x})$ in (4.22) is small as compared to the external field. Then it can be neglected in the interaction terms of the equations, and so we can consider the approximate *linear equations*

$$\left\{\begin{array}{l} \big[i\hbar\partial_t - e\phi_{\text{ext}}(t,\mathbf{x})\big]\psi(t,\mathbf{x}) = \dfrac{1}{2\mathrm{m}}\Big[-i\hbar\nabla - \dfrac{e}{c}\mathbf{A}_{\text{ext}}(t,\mathbf{x})\Big]^2\psi(t,\mathbf{x}), \\ \dfrac{1}{4\pi}\Box\phi(t,\mathbf{x}) = e\big|\psi(t,\mathbf{x})\big|^2, \\ \dfrac{1}{4\pi}\Box\mathbf{A}(t,\mathbf{x}) = \dfrac{e}{\mathrm{m}c}\psi(t,\mathbf{x})\cdot\Big[-i\hbar\nabla - \dfrac{e}{c}\mathbf{A}_{\text{ext}}(t,\mathbf{x})\Big]\psi(t,\mathbf{x}), \end{array}\right. \tag{4.27}$$

which are called the *Born approximation* to (4.22).

Dimension Analysis: The Sommerfeld Fine-Structure Constant

This process of approximations can be iterated substituting the Born approximation into the interaction terms etc. To make it systematically, let us rewrite the system (4.22) as

$$\begin{cases} \left[i\partial_{ct} - \frac{e^2}{\hbar c}\big(\tilde{\phi}(t,\mathbf{x}) + \tilde{\phi}_{\text{ext}}(t,\mathbf{x})\big)\right]\psi(t,\mathbf{x}) \\ \quad = \frac{\hbar}{2\text{m}c}\left[-i\boldsymbol{\nabla} - \frac{e^2}{\hbar c}\big(\tilde{\mathbf{A}}(t,\mathbf{x}) + \tilde{\mathbf{A}}_{\text{ext}}(t,\mathbf{x})\big)\right]^2\psi(t,\mathbf{x}) \\ \frac{1}{4\pi}\Box\tilde{\phi}(t,\mathbf{x}) = \left|\psi(t,\mathbf{x})\right|^2, \\ \frac{1}{4\pi}\Box\tilde{\mathbf{A}}(t,\mathbf{x}) = \frac{\hbar}{\text{m}c}\psi(t,\mathbf{x})\cdot\left[-i\boldsymbol{\nabla} - \frac{e^2}{\hbar c}\big(\tilde{\mathbf{A}}(t,\mathbf{x}) + \tilde{\mathbf{A}}_{\text{ext}}(t,\mathbf{x})\big)\right]\psi(t,\mathbf{x}), \end{cases} \tag{4.28}$$

where $\tilde{\phi}(t,\mathbf{x}) := \phi(t,\mathbf{x})/e$ etc. Using (3.36), one calculate the *fine structure constant*

$$\alpha := \frac{e^2}{\hbar c} \approx \frac{1}{137{,}036}; \tag{4.29}$$

it was introduced by Sommerfeld in 1916. The crucial observation is that the fraction is *dimensionless* and is quite small! This suggests to isolate the interaction terms and treat them as a 'small' perturbation. Consequently, system (4.28) becomes

$$i\dot{Y}(t) = \mathcal{H}Y(t) + \alpha Q\big(Y(t)\big), \quad Y(t) := \big(\tilde{\psi}(\cdot,t), \tilde{\phi}(\cdot,t), \tilde{\mathbf{A}}(\cdot,t)\big). \tag{4.30}$$

Here, $\mathcal{H}$ is the linear operator, which generates the approximation of type (4.27), and $Q(Y)$ is the polynomial operator containing the polynomial terms (in Y) of degree two and three, with zero linear part. The solution of Eq. (4.30) can be expanded in the power series

$$Y(t) = Y_0(t) + \alpha Y_1(t) + \alpha^2 Y_2(t) + \cdots, \tag{4.31}$$

in which each term Y_k can be computed from the corresponding linear equation. The expansion is known as the Dirac nonstationary perturbation theory [43] and is widely used in many problems, though justification of the convergence for the series is still an open problem.

4.3 Klein–Gordon Equation

Let us extend the Lagrangian formalism to the Klein–Gordon equation

$$L_{KG}\psi(x) := \left[i\hbar\nabla_0 - \frac{e}{c}\phi(x)\right]^2\psi(x) - \left[-i\hbar\nabla - \frac{e}{c}\mathbf{A}(x)\right]^2\psi(x) - \text{m}^2c^2\psi(x) = 0, \tag{4.32}$$

where $x_0 = ct$, and respectively, $\nabla_0 = \frac{1}{c}\partial_t$. The equation arises from the relativistic Lorentz equations (12.76) by the canonical quantization (3.26) of the corresponding relativistic energy-momentum relation (12.93). In the real vector form,

$$\begin{aligned}\mathbb{L}_{KG}\psi(x) :&= \Big[J\hbar\nabla_0 - \frac{e}{c}\phi(x)\Big]^2\Psi(x) - \Big[-J\hbar\boldsymbol{\nabla} - \frac{e}{c}\mathbf{A}(x)\Big]^2\Psi(x) + \mathrm{m}^2c^2\Psi(x)\\ &= 0.\end{aligned} \tag{4.33}$$

This equation can be written in the variational form (4.10) with the corresponding Lagrangian density

$$\begin{aligned}\mathcal{L}(x,\Psi,\nabla\Psi) = \frac{1}{2}\Big|\Big[J\hbar\nabla_0\Psi - \frac{e}{c}\phi(x)\Psi\Big]\Big|^2 - \frac{1}{2}\Big| - J\hbar\boldsymbol{\nabla}\Psi - \frac{e}{c}\mathbf{A}(x)\Psi\Big|^2\\ - \frac{1}{2}\mathrm{m}^2c^2|\Psi|^2,\end{aligned} \tag{4.34}$$

where the factor $1/2$ is introduced for convenience of calculations.

Exercise 4.4 Prove that the corresponding Euler–Lagrange equations (4.10) coincide with the Klein–Gordon equation (4.33). **Hint:**

$$\begin{aligned}\mathcal{L}_\Psi = -\frac{e}{c}\phi(x)\Big[J\hbar\nabla_0 - \frac{e}{c}\phi(x)\Big]\Psi(x) + \frac{e}{c}\mathbf{A}^k(x)\Big[-J\hbar\boldsymbol{\nabla}_k - \frac{e}{c}\mathbf{A}^k(x)\Big]\Psi(x)\\ - \mathrm{m}^2c^2\Psi.\end{aligned}$$

The canonically conjugate fields are given by

$$\begin{aligned}\Pi^0(x) &= -J\hbar\Big[J\hbar\nabla_0 - \frac{e}{c}\phi(x)\Big]\Psi(x),\\ \Pi^k(x) &= -J\hbar\Big[-J\hbar\boldsymbol{\nabla}_k - \frac{e}{c}\mathbf{A}^k(x)\Big]\Psi(x).\end{aligned} \tag{4.35}$$

Exercise 4.5 Write the energy, momentum, angular momentum and charge conservation for the Klein–Gordon equation (4.32). **Hint:** Use general formulas of Sect. 13.4.

The corresponding Hamiltonian functional is given by the Legendre transform

$$\begin{aligned}\boldsymbol{\mathcal{H}}\big(\Pi^0,\Psi,t\big) &= \big\langle\Pi^0,\nabla_0\psi\big\rangle - \int\mathcal{L}(x,\Psi,\nabla\Psi)\,d\mathbf{x}\\ &= \frac{1}{2}\Big\langle J\hbar\nabla_0\Psi - \frac{e}{c}\phi(x)\Psi,\ J\hbar\nabla_0\Psi + \frac{e}{c}\phi(x)\Psi\Big\rangle\\ &\quad + \frac{1}{2}\int\Big[\Big| - J\hbar\boldsymbol{\nabla}\Psi - \frac{e}{c}\mathbf{A}(x)\Psi\Big|^2 + \mathrm{m}^2c^2|\Psi|^2\Big]d\mathbf{x}\end{aligned}$$

$$= \frac{1}{2\hbar^2}\Big\langle \Pi^0, \Pi^0 + 2\hbar\frac{e}{c}\phi(x)J\Psi\Big\rangle$$
$$+\frac{1}{2}\int\Big[\Big|-J\hbar\nabla\Psi - \frac{e}{c}\mathbf{A}(x)\Psi\Big|^2 + \mathrm{m}^2c^2|\Psi|^2\Big]d\mathbf{x}. \quad (4.36)$$

Exercise 4.6 Verify that the Klein–Gordon equation (4.33) is equivalent to the Hamilton system

$$\nabla_0\Psi = D_{\Pi^0}\boldsymbol{\mathcal{H}}, \qquad \nabla_0\Pi^0 = -D_\Psi\boldsymbol{\mathcal{H}}, \quad (4.37)$$

where D_{Π^0} and D_Ψ are the corresponding variational derivatives.

Remark 4.7 The middle line of (4.36) can be written as

$$\boldsymbol{\mathcal{H}}\big(\Pi^0, \Psi, t\big) = \frac{\hbar^2}{2}\langle\nabla_0\Psi, \nabla_0\Psi\rangle - \frac{e^2}{2c^2}\big\langle\phi(x)\Psi, \phi(x)\Psi\big\rangle$$
$$+\frac{1}{2}\int\Big[\Big|-J\hbar\nabla\Psi - \frac{e}{c}\mathbf{A}(x)\Psi\Big|^2 + \mathrm{m}^2c^2|\Psi|^2\Big]d\mathbf{x}. \quad (4.38)$$

Hence, the energy (4.36) can be non-positive for large potentials $\phi(x)$. This is one of issues of the known ‘Klein paradox’ [72].

Chapter 5
Wave-Particle Duality

Abstract The wave-particle duality was one of main issues for the Schrödinger theory. Now we will dwell upon this question in framework of the Schrödinger formalism calculating the corresponding density of particles. However, particles cannot be assigned any fixed position and momentum, as there are restrictions due to the Heisenberg *Uncertainty Principle*.

The wave nature of electron beam was confirmed experimentally and is perfectly explained by the Schrödinger wave theory.

However, the problem of wave-particle duality acquires new appearance as 'reduction of wave packets' in diffraction of a *week electron beam*. This key phenomenon was discovered experimentally and suggests the *Born probabilistic interpretation* of the wave function.

5.1 Electron Beam and Uncertainty Principle

De Broglie identified free electron beam with the plane wave (3.2)

$$\psi(t,\mathbf{x}) = Ae^{i(\mathbf{kx}-\omega t)}. \tag{5.1}$$

Let us consider this identification in the framework of the Schrödinger theory.

5.1.1 Plane Wave as Electron Beam

The plane wave (3.2) is a solution of the *free* Schrödinger equation (without an external Maxwell field). Hence, the parameters satisfy the relations

$$\hbar\omega = \frac{\hbar^2}{2\mathrm{m}}\mathbf{k}^2 > 0. \tag{5.2}$$

By (4.19), the corresponding charge and electric current densities are given by

$$\rho = e\big|\psi(t,\mathbf{x})\big|^2 = e|A|^2, \qquad \mathbf{j} = \frac{e}{\mathrm{m}}\big[-i\hbar\nabla\psi(t,\mathbf{x})\big]\cdot\psi(t,\mathbf{x}) = \frac{e\hbar\mathbf{k}}{\mathrm{m}}|A|^2. \tag{5.3}$$

A. Komech, *Quantum Mechanics: Genesis and Achievements*,
DOI 10.1007/978-94-007-5542-0_5,

Also, the electron density, velocity and momentum should be as follows:

$$n := \rho/e = \left|\psi(t,\mathbf{x})\right|^2, \qquad \mathbf{v} = \mathbf{j}/\rho = \frac{\hbar\mathbf{k}}{\mathrm{m}}, \qquad \mathbf{p} = \mathrm{m}\mathbf{v} = \hbar\mathbf{k}. \tag{5.4}$$

Consequently, the energy density (i.e., the integrand in (3.61)) reads as

$$\mathcal{E} := \frac{1}{2\mathrm{m}}\left|[-i\hbar\nabla]\psi(t,\mathbf{x})\right|^2 = \frac{\hbar^2}{2\mathrm{m}}\mathbf{k}^2|A|^2 = \frac{\mathrm{m}\mathbf{v}^2}{2}n. \tag{5.5}$$

Correspondingly, the energy per *one electron* is given by

$$E = \frac{\mathcal{E}}{n} = \frac{\mathrm{m}\mathbf{v}^2}{2} = \frac{\hbar^2}{2\mathrm{m}}\mathbf{k}^2 = \hbar\omega, \tag{5.6}$$

in accordance with (5.2).

5.1.2 The Heisenberg Uncertainty Principle

The interpretation of a plane wave as a beam of free electrons cannot be completed by an assignment of the coordinates to electrons, since the charge density (5.3) is constant. For the wave function $\psi(x)$ with good decay at infinity, the mean value of the coordinate is defined by

$$\bar{\mathbf{x}} := \langle\psi, \hat{\mathbf{x}}\psi\rangle := \langle\psi, \mathbf{x}\psi\rangle \tag{5.7}$$

in accordance with Definition 3.13(ii). Hence the *uncertainty* of the mean value is the *mean square error*,

$$\Delta\mathbf{x} := \left|\left\langle\psi, (\hat{\mathbf{x}} - \bar{\mathbf{x}})^2\psi\right\rangle\right|^{1/2}/\|\psi\|, \tag{5.8}$$

where $\|\psi\|$ is the L^2-norm of ψ. Similarly, for the momentum

$$\bar{\mathbf{p}} := \langle\psi, \hat{\mathbf{p}}\psi\rangle := \langle\psi, -i\hbar\nabla\psi\rangle, \qquad \Delta\mathbf{p} := \left|\left\langle\psi, (\hat{\mathbf{p}} - \bar{\mathbf{p}})^2\psi\right\rangle\right|^{1/2}/\|\psi\|. \tag{5.9}$$

It is easy to construct a wave functions ψ so that either $\Delta\mathbf{x}(\psi)$ or $\Delta\mathbf{p}(\psi)$ is arbitrarily small. However, the errors cannot be small simultaneously, giving the exact **Uncertainty Principle**, as discovered by Heisenberg [81, 145]:

$$\Delta\mathbf{x}_j\,\Delta\mathbf{p}_{j'} \ge \frac{\hbar}{2}\delta_{jj'}, \quad j, j' = 1, 2, 3. \tag{5.10}$$

Example 5.1 Consider the 'plane wave' $\psi_R(\mathbf{x}) = \zeta(\mathbf{x}/R)e^{i\mathbf{kx}}$, where $\zeta \in C_0^\infty(\mathbb{R}^3)$, $\zeta(0) \neq 0$. Then $\Delta\mathbf{x}_j(\psi_R) \to \infty$ as $R \to \infty$, while $\Delta\mathbf{p}_j(\psi_R) \to 0$.

Proof of the Uncertainty Principle First, note that the operators $\alpha := \hat{\mathbf{x}}_j - \bar{\mathbf{x}}_j$ and $\beta := \hat{\mathbf{p}}_{j'} - \bar{\mathbf{p}}_{j'}$ are symmetric. Hence

$$\Delta \mathbf{x} = \frac{\|\alpha\psi\|}{\|\psi\|}, \qquad \Delta \mathbf{p} = \frac{\|\beta\psi\|}{\|\psi\|}. \tag{5.11}$$

Now the Cauchy–Schwarz inequality implies that

$$\Delta \mathbf{x} \Delta \mathbf{p} \geq \frac{|\langle \alpha\psi, \beta\psi\rangle|}{\|\psi\|^2} = \frac{|\langle(\beta\alpha+\alpha\beta)\psi, \psi\rangle/2 + \langle(\beta\alpha-\alpha\beta)\psi, \psi\rangle/2|}{\|\psi\|^2}. \tag{5.12}$$

It remains to observe that the operator $\beta\alpha + \alpha\beta$ is symmetric, while $\beta\alpha - \alpha\beta$ is antisymmetric. Then the first inner product in the last numerator is real, while the second is purely imaginary. Since $\beta\alpha - \alpha\beta = -i\hbar\delta_{jj'}$, (5.10) herewith follows. □

Remark 5.2 For an eigenfunction ψ of the operator $\hat{\mathbf{x}}$ ($\hat{\mathbf{p}}$, respectively), the uncertainty of the mean vanishes: $\Delta\hat{\mathbf{x}} = 0$ ($\Delta\hat{\mathbf{p}} = 0$, respectively). Hence, the Uncertainty Principle is related to noncommutativity of the operators $\hat{\mathbf{p}}$ and $\hat{\mathbf{x}}$, because it forbids the existence of their common eigenfunctions.

5.2 Diffraction of Electron Beams

The wave-particle duality was suggested theoretically by de Broglie in 1924. Experimentally, the duality was observed for the first time by Davisson and Germer in 1924–1927. The diffraction pattern is perfectly explained by the Schrödinger theory. However, the *reduction of wave packets* is an open problem apparently explained by the soliton type asymptotics for solutions of the Maxwell–Schrödinger equations.

5.2.1 Experimental Observations

- Davisson and Germer examined the scattering of electron beams (cathode rays) by the nickel crystal as the target to measure the intensity of the scattered beam at different directions. Their intention was to extend the Thomson formula (8.28) for the scattering cross section from the scattering by atom to the scattering by crystal, treating an electron beam as a flux of classical particles. Accordingly, Davisson and Germer expected to obtain a smooth slowly varying cross section (like the Thomson formula (8.28)).

  Very surprisingly, the observation showed an extraordinary peak at some directions in the intensity of the reflected electron beam. This peak resembled diffraction patterns in scattering of X-rays by crystals. This kind of diffraction was observed first by Laue in 1911–1914, who explained it as being due to the interference between X-rays reemitted by the crystal lattice; its merit was fitly recognized by the award to von Laue in 1914 of the Nobel Prize for physics.

Accordingly, Davisson and Germer treated the peak as a result of the interference between electron waves reemitted by the crystal lattice (Davisson shared the Nobel Prize in physics in 1937 with George Thomson).

- Similar diffraction pattern was observed in the diffraction of an electron beam by diffraction gratings [96].
- The 'gedanken' double-slit experiment is considered in all textbooks starting from 1927. However, this experiment was first performed in 1989 by Tonomura et al. using the set-up shown in Fig. 5.3 (see [184]).
- In 1949, Ehrenberg and Siday predicted the deflection of the diffraction pattern in experiments of Davisson and Germer under influence of a magnetic field (this was independently rediscovered by Aharonov and Bohm in 1959). The idea behind this is the Lorentz force (1.6), which should deflect electrons downwards, since the magnetic field is directed 'downwords', while the electron charge is negative (see Fig. 5.4). Chambers was first to verify this effect experimentally (1960, see [32]).

5.2.2 *The Davisson–Germer Experiment*

The incident electron beam excites each atom of the crystal and generates the reemitted wave of type (8.148). The diffraction pattern arises from the interference of the reemitted waves from all atoms of the crystal. Similarly, in the diffraction of light by gratings, the diffraction pattern arises from the interference of the secondary waves radiated by each 'scratch'. For the case of gratings, the Bragg rule reads

$$\frac{d\cos\alpha}{\lambda} - \frac{d\cos\beta}{\lambda} = n, \qquad n = 0, \pm 1, \pm 2, \dots, \tag{5.13}$$

where $d > 0$ is the grating spacing, and $\lambda = 2\pi/|\mathbf{k}|$ is the de Broglie's wavelength of the incident wave (5.1), α is the angle between the crystal surface and the wave vector **k**, which is the direction of the incident wave propagation (see Fig. 5.1). In the early Davisson–Germer experiments, the de Broglie wavelength $\lambda = 1.67\,\text{Å}$; this corresponds to the electrons of energy 54 eV (cf. (3.23)).

Davisson and Germer examined the diffraction of the reflected electron beam by crystals in the set-up shown in Fig. 5.1. The solutions β of the Bragg equation correspond to the directions along which the amplitude of the scattered wave attains its local maximum; this is because here the phases of the secondary waves coincide modulo 2π (see Fig. 5.1). The directions of local minimum of the amplitude corresponds to solutions β to the Bragg equation with $(n + 1/2)\lambda$ on the right. The corresponding Bragg rule for the crystal can be obtained similarly, taking into account the crystal structure.

In 1927, Thomson investigated the diffraction of a transmitted electron beam by a thin polycrystalline film in the set-up of Fig. 5.2; he observed the diffraction pattern corresponding to this configuration, [181].

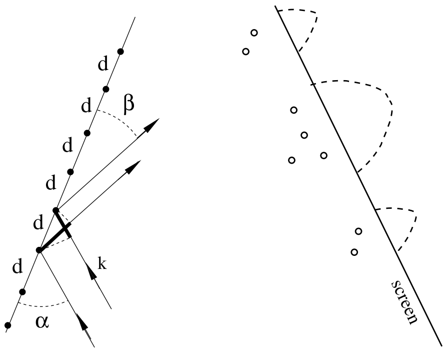


Fig. 5.1 Diffraction of reflected beam

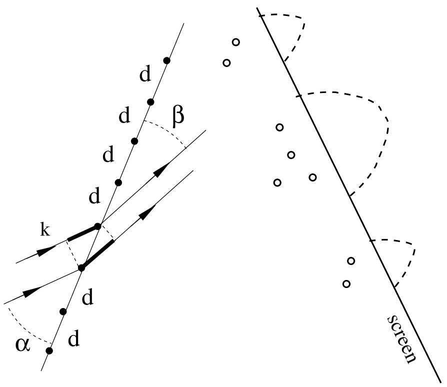


Fig. 5.2 Diffraction of transmitted beam

In all these observations, the diffraction pattern agrees with the corresponding Bragg rules and de Broglie's formulas for wavelength (3.15), (3.17), where the energy of electrons is determined from the voltage of the electron gun.

5.2.3 The Double-Slit Experiment

In 1989, Tonomura et al. observed the diffraction of an electron beam in a 'two-slit' arrangement. An electron beam emitted by the electron gun **S** is separated in two coherent beams by the electrostatic version of Fresnel's biprism in the arrangement of Fig. 5.3. The diffraction pattern in this case agrees with the Bragg rule of type (5.13) with the corresponding de Broglie's wavelength, see [184].

5.2.4 The Aharonov–Bohm Effect

The two-slit arrangement is superimposed by a magnetic field **M**. The shift in the set-up of Fig. 5.4 was first observed experimentally by Chambers [32, 153].

Fig. 5.3 Diffraction by double-slit

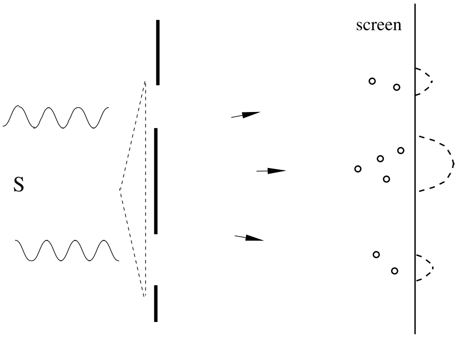
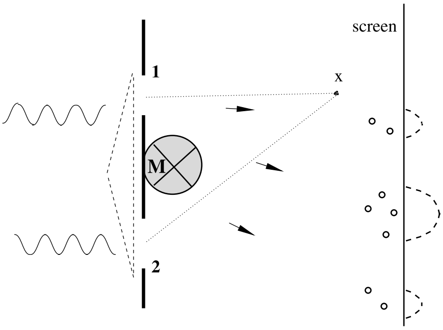


Fig. 5.4 Shift in magnetic field



5.2.5 *Diffraction of Electrons via Schrödinger Theory*

The diffraction of electron beams is perfectly described by the linear 'magnetic' Schrödinger equation (3.37),

$$i\hbar\partial_t\psi(t,\mathbf{x}) = \left(\frac{1}{2\mathrm{m}}\left[-i\hbar\nabla - \frac{e}{c}\mathbf{A}(\mathbf{x})\right]^2 + e\phi(\mathbf{x})\right)\psi(t,\mathbf{x}). \tag{5.14}$$

Namely, the observed wave field $\psi(t,\mathbf{x})$ is a solution of the scattering problem corresponding to the incident electron beam of type (5.1). In the arrangement of Fig. 5.4, the corresponding external magnetic field $\mathbf{B}(\mathbf{x}) = \operatorname{curl}\mathbf{A}(\mathbf{x})$ is localized in a small region screened from the incident electron beam.

Davisson–Germer Experiment

Let $\psi^0(t,\mathbf{x})$ and $\psi(t,\mathbf{x})$ be, respectively, the solutions of the scattering problem describing the arrangement of Figs. 5.3 and 5.4 with two slits 1 and 2. It is crucially important that the de Broglie's electron wavelength is negligible with respect to the distance d between the slits. For example,

(a) in the Tonomura et al. experiment (1989): $d = 10^{-4}$ cm $= 10^4$ Å, while $\lambda = 0.054$ Å for 'very cold' electrons with the energy 10 eV; see [184];
(b) in the Frabboni et al. experiment (2007): $d = 0.22 \times 10^4$ Å, while $\lambda = 0.0257$ Å; see [67].

In this short wavelength limit, we expect that the following 'splitting asymptotic' formulas hold,

$$\begin{aligned} \psi^0(t,\mathbf{x}) &\sim \psi_1^0(t,\mathbf{x}) + \psi_2^0(t,\mathbf{x}), \\ \psi(t,\mathbf{x}) &\sim \psi_1(t,\mathbf{x}) + \psi_2(t,\mathbf{x}), \quad \lambda/d \ll 1, \ \mathbf{x} \in Z; \end{aligned} \tag{5.15}$$

here Z is the zone between the scatterer (electrostatic biprism) and the screen, $\psi_1^0(t,\mathbf{x})$ and $\psi_1(t,\mathbf{x})$ are the solutions to the corresponding scattering problems (in the arrangement of Figs. 5.3 and 5.4, respectively) with the unique slit 1, and $\psi_2^0(t,\mathbf{x})$ and $\psi_2(t,\mathbf{x})$ are the solutions to the corresponding scattering problems with the unique slit 2. The splitting asymptotic formulas (5.15) hold in the Huygens–Kirchhoff diffraction approximation [5, 21, 173], which is valid, since $\lambda/d \ll 1$. The first splitting gives the interference, which perfectly explains the diffraction pattern in the Davisson–Germer experiments.

The Aharonov–Bohm Effect

A further step would be to find a relation between the solutions $\psi_j^0(t,\mathbf{x})$ and $\psi_j(t,\mathbf{x})$. For every $j = 1, 2$, they are related by the 'gauge transform' (4.25)

$$\psi_j(t,\mathbf{x}) = \psi_j^0(t,\mathbf{x}) e^{-i\frac{e}{c\hbar}\chi_j(\mathbf{x})}, \qquad \mathbf{A}(t,\mathbf{x}) = -\nabla\chi_j(\mathbf{x}), \quad \mathbf{x} \in Z \setminus \mathbf{M}, \tag{5.16}$$

leaving the electrostatic potential $\phi(\mathbf{x})$ unchanged. These relations hold, since the second equation admits the solution $\chi_j(\mathbf{x})$ outside the region $\mathbf{M}$, in which $\operatorname{curl}\mathbf{A}(\mathbf{x}) \equiv \mathbf{B}(\mathbf{x}) \neq 0$:

$$\chi_j(\mathbf{x}) = \int_{\Gamma_j(\mathbf{x})} \mathbf{A}(\mathbf{y})\,d\mathbf{y}, \quad \mathbf{x} \in Z \setminus \mathbf{M}. \tag{5.17}$$

The integral is taken over any path $\Gamma_j(\mathbf{x})$ connecting the slit j with the point $\mathbf{x}$, and does not depend on the path, inasmuch as $\operatorname{curl}\mathbf{A}(\mathbf{x}) = 0$, see Fig. 5.4.

The solutions with $j = 1, 2$ do not coincide, since the 'period' does not vanish (by the Stokes theorem):

$$\int_{\partial\mathbf{M}} \mathbf{A}(\mathbf{x})\,d\mathbf{x} = \Phi, \quad \Phi := \int_{\mathbf{M}} \mathbf{B}(\mathbf{x})\,d\mathbf{x} \neq 0 \tag{5.18}$$

where $\partial\mathbf{M}$ is oriented clockwise. Hence, $\chi_1(\mathbf{x}) = \chi_2(\mathbf{x}) + \Phi$, and the second splitting (5.15) together with (5.16) give

$$\begin{aligned}\psi(t,\mathbf{x}) &= \psi_1^0(t,\mathbf{x})e^{-i\frac{e}{c\hbar}\chi_1(\mathbf{x})} + \psi_2^0(t,\mathbf{x})e^{-i\frac{e}{c\hbar}\chi_2(\mathbf{x})} \\ &= \left[\psi_1^0(t,\mathbf{x}) + \psi_2^0(t,\mathbf{x})e^{-i\frac{e}{c\hbar}\Phi}\right]e^{-i\frac{e}{c\hbar}\chi_2(\mathbf{x})}. \end{aligned} \tag{5.19}$$

The presence of the phase factor $e^{-i\frac{e}{c\hbar}\Phi}$ exactly means a 'shift' of the diffraction pattern. Indeed, we should add the phase shift $\delta = \dfrac{e}{c\hbar}\Phi$ to the second term of the Bragg equation (5.13), where $\alpha = \pi/2$ and d is the distance between the slits 1 and 2,

$$-\frac{d\cos\beta}{\lambda} - \delta = n, \quad n = 0, \pm 1, \pm 2, \ldots. \tag{5.20}$$

Hence, the angles $\beta_n = \beta_n(\delta)$ change continuously with δ, resulting in the observed shift of the diffraction pattern.

5.3 Probabilistic Interpretation

The diffraction experiments for beams with small intensities suggest the probabilistic interpretation of the wave function. We discuss possible mechanism of the particle generation.

5.3.1 Reduction of Wave Packets and Probabilistic Interpretation

As we have seen, the linear Schrödinger equation (5.14) satisfactory explains the diffraction pattern in the electron diffraction and the shift in a magnetic field. However, the reduction of the wave packet remains a fundamental open question, which cannot be treated with the linear Schrödinger equation.

Namely, for an incident beam of small intensity, the *discrete registration of electrons* was observed with the *counting rate* corresponding to the diffraction pattern: see the discussion in [184]. This discrete registration, known as 'reduction of wave packets', cannot be explained by the linear equation (3.32). In 1927, Born suggested to interpret $|\psi(t,\mathbf{x})|^2$ as the *density of probability* in order to explain the electron diffraction in the Davisson–Germer experiment. The interpretation has been confirmed by validity of the Rutherford formula in the scattering of electrons (see Remark 8.8).

In particular, formulas (5.3) and (5.4) give the (mathematical) expectation of the charge density, current, density of particles, velocities and momentum.

5.3.2 Soliton-Type Asymptotics

To explain dynamically the reduction of wave packets, we should suggest an interpretation of elementary particles in terms of the coupled nonlinear Maxwell–Schrödinger equations (4.22).

For example, the elementary particles might correspond to traveling ('solitary') waves, which are the solutions of the type $(\psi(\mathbf{x}-\mathbf{v}t)e^{i\Phi(\mathbf{v},t,\mathbf{x})}, A(\mathbf{x}-\mathbf{v}t))$ to the coupled nonlinear Maxwell–Schrödinger equations (4.22). Respectively, the de Broglie's wave-particle duality should be treated as the soliton asymptotics

$$\big(\psi(t,\mathbf{x}), A(t,\mathbf{x})\big) \sim \sum_k \big(\psi_\pm^k\big(\mathbf{x}-\mathbf{v}_\pm^k t\big)e^{i\Phi(\mathbf{v}_\pm^k,t,\mathbf{x})}, A_\pm^k\big(\mathbf{x}-\mathbf{v}_\pm^k t\big)\big), \quad t\to\pm\infty. \tag{5.21}$$

The asymptotics should hold in the zone near the screen lying sufficiently far from the scatterer (crystal, grating or the electrostatic biprism), where the solitons probably cannot nucleate due to interaction with the scatterer.

Asymptotics (5.21) near the screen might be treated as the electron field decay to the collection of the electrons represented by bullets in Figs. 5.1–5.4. This treatment suggests that the density of the counted electrons should decrease with increasing distance from the screen.

5.3.3 Solitons and Reduction of Wave Packets

The diffraction pattern in the diffraction experiments agrees with the corresponding Bragg rules, which determine the location of maximum values of the charge density on the screen. This fact has led to the Born probabilistic interpretation for the charge density.

The soliton-like asymptotics (5.21) also agree with the Bragg rule for small wavelength. Namely, we suggest that the zone Z between the scatterer and the screen is divided into two zones:

- The 'wave zone' Z_w near the scatterer, where the wave field $\psi(t,\mathbf{x})$ admits the quasiclassical asymptotics (3.52), and
- The 'particle zone' Z_p near the screen, where the wave field $\psi(t,\mathbf{x})$ admits the soliton-type asymptotics (5.21).

The counting rate of solitons in the zone Z_p should be proportional to the absolute value of the electric current in the adjacent cell of the wave zone by the charge continuity if the phases of the solitons are random and the random process is ergodic.

It remains to explain why the maxima of the current density in the wave zone correspond to the maxima of the charge density for small wavelength. This follows from the quasiclassical asymptotics (3.52), since ∇S should vary slowly in the wave zone.

This correspondence explains why the counting rate is greatest when the wave function assumes its maximum. This seems to explain the correspondence between the counting rate of solitons and the diffraction pattern. Finally, the random nature of the process is provided by the microscopic fluctuations and uncontrolled imperfections of the scatterer and of the slits.

Remark 5.3 Note, however, that for large wavelength, the quasiclassical asymptotics (3.52) might break down. Respectively, the diffraction pattern in general should correspond to the maxima of the current density (instead of the charge density).

5.3.4 The Aharonov–Bohm Paradox

The Aharonov–Bohm effect looks paradoxical from the classical point of view, because the 'deviated particles' do not go through the region in which the magnetic field is present. This paradox can be removed by the above treatment, since the magnetic field is located in the wave zone Z_w and interacts with the nonlocal Schrödinger wave field, while the particles (= solitons) arise in the 'particle zone' near the screen which is far from the magnet.

5.3.5 The Known Results on Soliton Asymptotics

At the present time, soliton asymptotics (5.21) are proved for some model nonlinear hyperbolic partial differential equations. First results in this direction were established for the KdV equation and other *completely integrable equations*. For the KdV equation, any solution with sufficiently smooth and rapidly decaying initial data converges to a finite sum of soliton solutions moving to the right, and a dispersive wave moving to the left. A complete survey and proofs can be found in [54].

For nonintegrable equations, an analogue of soliton asymptotics (5.21) was obtained first by Soffer and Weinstein in the context of $U(1)$-invariant Schrödinger equation [166–170]. The asymptotics for translation invariant equations were obtained by Buslaev and Perelman [28–30] for the 1D Schrödinger equation, and by Miller, Pego and Weinstein, for the 1D modified KdV and RLW equations (see [140, 150, 151]). Recently the results were extended to the coupled Maxwell–Lorentz equations and relativistic nonlinear Ginzburg–Landau equations [8, 117, 118]. In all these papers, the asymptotics are proved for solutions sufficiently close to one soliton, and the sum in (5.21) contains only one term.

The soliton asymptotics (5.21) with one soliton for *all finite energy solutions* of nonintegrable equations were proved first in [110] for the *Abraham type model* of the *extended particle* coupled to wave equation. The result has been extended in [91] to Maxwell–Lorentz equations with extended charge.

Numerical experiments [115] demonstrate that soliton asymptotics (5.21) hold for general 1D relativistic-invariant equations, however the proof is still an open problem. The existence of solitons and the relativistic Einstein mass-energy identity for them are proved, respectively, in [9] and [53], for general relativistic nonlinear Klein–Gordon equations.

On the other hand, for the coupled Maxwell–Schrödinger and Maxwell–Dirac equations, the justification of the asymptotics is still an open question. For the coupled equations it is proved only the existence of solutions (see [22, 80]), and for the Maxwell–Dirac equations, the existence of solitons (see [61]).

5.3.6 Particle-Like Behavior of Solitons

In [120], *adiabatic effective dynamics* is established for solitons of a 3D wave equation coupled to a classical particle subject to a *slowly varying external potential*. The effective dynamics justifies the identification of solitons with particles, and explains the increment of the mass of the particle caused by its interaction with the field. The effective dynamics is extended in [68] to solitons of the nonlinear Schrödinger and Hartree equations. The first extension to relativistic-invariant equations is obtained in [135] for the nonlinear Klein–Gordon–Maxwell system with an external potential.

Chapter 6
The Eigenvalue Problem

Abstract The eigenvalue problem for the Hydrogen atom was solved by Schrödinger in 1926.

The solution relies on the separation of variables which is possible due to the spherical symmetry of the Schrödinger equation. The angular functions are chosen to be the spherical functions which are the eigenfunctions of the spherical Laplacian. The spherical functions are constructed by an analysis of the Lie algebra of the rotation group SO(3). The radial functions are obtained solving the radial differential equation applying the Sommerfeld method of factorization.

6.1 The Hydrogen Spectrum

Quantum stationary states for the hydrogen atom were obtained by separation of variables in the Schrödinger equation. The key role in the calculation is played by decomposing the space $L^2(S^2)$ into a sum of orthogonal eigenspaces of the spherical Laplacian (which we will construct in next section) and by solving the corresponding radial equation.

6.1.1 The Eigenvalue Problem

Ions of the hydrogen atom have a positive charge $-e$, being of valency one. Respectively, the atom has precisely one electron of negative charge e due to neutrality of the substance. Hence the normalization condition (3.35) should hold, and so $\psi \in \mathcal{E} := L^2(\mathbb{R}^3)$.

The Rutherford experiments demonstrate that the positive charge $|e|$ of the ions is concentrated in a relatively small region called 'nucleus', so its Maxwell potential is Coulombic $\phi = -e/|\mathbf{x}|$. The magnetic potential of the nucleus is assumed to be zero: $\mathbf{A}(t, \mathbf{x}) = 0$. Then the Schrödinger equation (3.37) for the hydrogen atom reads

$$i\hbar\partial_t\psi(t,\mathbf{x}) = H\psi(t,\mathbf{x}) := -\frac{\hbar^2}{2\mathrm{m}}\Delta\psi(t,\mathbf{x}) - \frac{e^2}{|\mathbf{x}|}\psi(t,\mathbf{x}). \tag{6.1}$$

A. Komech, *Quantum Mechanics: Genesis and Achievements*,
DOI 10.1007/978-94-007-5542-0_6,

In our model, the hydrogen nucleus is considered as fixed. This corresponds to the fact that the nucleus is heavy with respect to the electron with the mass ratio about 1836. Respectively, the nucleus potential is assumed to be static. The corresponding stationary states $\psi(t,\mathbf{x}) = \psi(\mathbf{x})e^{-iE/\hbar t}$ satisfy the stationary Schrödinger equation

$$E\psi(\mathbf{x}) = H\psi(\mathbf{x}). \tag{6.2}$$

Theorem 6.1 *The energies of quantum stationary states of the hydrogen atom are given by*

$$E_n = -2\pi\hbar c R/n^2, \quad n = 1, 2, \ldots, \tag{6.3}$$

where $R := \frac{\mathrm{m}e^4}{4\pi\hbar^3 c}$ *is the* ***Rydberg constant****.*

We will prove Theorem 6.1 in this lecture relying on spectral resolution of *spherical Laplacian* which we prove in next lecture.

Substituting values (3.36), we obtain $R \approx 108 \times 10^5\,\mathrm{cm}^{-1}$ (with $e = -4.803 \times 10^{-10}$ esu). Modern recommended value of the Rydberg constant:[1]

$$R = 109\,737.315\,685\,27\,(73)\ \mathrm{cm}^{-1}. \tag{6.4}$$

Formula (6.3) agrees with the Balmer empirical terms (1.2) with $B = 2\pi c R \approx 104.058 \times 10^{30}\ \mathrm{s}^{-1}$. Moreover, the lowest energy level E_1 should be the (negative) ionization energy of the Hydrogen atom which is known experimentally to be about -13.6 eV. Using the data of (3.36) for $\hbar$, c and (6.4) for R, it is found that

$$E_1 = -2\pi\hbar c R \approx -21.79 \times 10^{-12}\ \mathrm{erg} = -21.79 \times 10^{-19}\ \mathrm{J}. \tag{6.5}$$

Dividing by the electron charge $e = -1.602 \times 10^{-19}$ C (SI), we obtain $E_1/e = 13.60$ V.

6.1.2 Spherical Symmetry and Separation of Variables

The Rotation Invariance

The basic issue in solving the eigenvalue problem (6.2) is its spherical symmetry, which implies the angular momentum conservation. Namely, the Schrödinger operator H is invariant with respect to all rotations of the space $\mathbb{R}^3$. This means the commutation

$$H\hat{R}_k(\varphi) = \hat{R}_k(\varphi)H, \quad \varphi \in \mathbb{R},\ k = 1, 2, 3, \tag{6.6}$$

---

[1] CODATA: http://www.physicstoday.org/guide/fundconst.pdf.

where $(\hat{R}_k(\varphi)\psi)(\mathbf{x}) := \psi(R_k(\varphi)\mathbf{x})$ and $R_k(\varphi)$ is the space rotation around the unit vector $\mathbf{e}_k$ by angle φ (in radians) in the positive direction, $\mathbf{e}_1 = (1, 0, 0)$, etc. The commutations hold by (6.2), since the Laplacian Δ and the Coulomb potential are invariant under all rotations. Differentiating (6.6) in φ, we obtain

$$[H, \nabla_{\varphi_k}] = 0, \quad k = 1,\ 2,\ 3 \tag{6.7}$$

where

$$\nabla_{\varphi_k} := \frac{d}{d\varphi}\bigg|_{\varphi=0} \hat{R}_k(\varphi). \tag{6.8}$$

It follows that H also commutes with $\mathbf{H}_k := -i\nabla_{\varphi_k}$ and angular momentum operators $\hat{\mathbf{L}}_k = \hbar\mathbf{H}_k$

$$[H, \mathbf{H}_k] = 0, \qquad [H, \hat{\mathbf{L}}_k] = 0, \quad k = 1, 2, 3, \tag{6.9}$$

by (3.86). Hence the angular momenta $\mathbf{L}_k(t) := \langle\psi(t), \hat{\mathbf{L}}_k\psi(t)\rangle$ are conserved. The conservation of the corresponding classical angular momentum played a crucial role in determination of the hydrogen spectrum in the Bohr–Sommerfeld 'old quantum theory' (Sect. 14.2).

Spherical Harmonics

Now we proceed to explain our general strategy of the proof of Theorem 6.1. Commutation (6.7) suggests solving the spectral problem (6.2) by separation of variables. More precisely, the strategy depends on the following three general arguments:

I. Commutation (6.9) obviously implies that the operator $\mathbf{H}^2 := \mathbf{H}_1^2 + \mathbf{H}_2^2 + \mathbf{H}_3^2$ commutes with H:

$$\left[H, \mathbf{H}^2\right] = 0. \tag{6.10}$$

Hence, any eigenspace of the Schrödinger operator H is invariant with respect to $\mathbf{H}^2$ and any of the operators $\mathbf{H}_k$. Moreover, $\mathbf{H}^2$ also commutes with each operator $\mathbf{H}_k$:

$$\left[\mathbf{H}^2, \mathbf{H}_k\right] = 0, \quad k = 1,\ 2,\ 3. \tag{6.11}$$

Exercise 6.2 Verify (6.11). **Hint:** First, prove the commutation relations $[\mathbf{H}_k, \mathbf{H}_j] = -i\hbar\epsilon_{kjl}\mathbf{H}_l$, where ϵ_{kjl} is a totally antisymmetric tensor.

Since the operators $\mathbf{H}_k$ with $k = 1, 2, 3$ do not commute, they cannot be diagonalized simultaneously. On the other hand, operators H, $\mathbf{H}^2$, and $\mathbf{H}_k$ with any fixed k mutually commute. Hence, we could expect that there is a basis of common eigenfunctions.

II. We will diagonalize simultaneously $\mathbf{H}_3$ and $\mathbf{H}^2$. By (6.8), both these operators act only on the angular variables in spherical coordinates. Hence, these operators act in the Hilbert space $\mathcal{E}_1 := L^2(S)$, where S denotes the two-dimensional sphere $|\mathbf{x}| = 1$.

Theorem 6.3

(i) *For* $\mathcal{E}_1$, *there exist an orthonormal basis of spherical harmonics* $Y_l^m(\theta, \varphi)$, *which are common eigenfunctions of* $\mathbf{H}_3$ *and* $\mathbf{H}^2$:

$$\mathbf{H}_3 Y_l^m = m Y_l^m, \qquad \mathbf{H}^2 Y_l^m = l(l+1) Y_l^m, \quad m = -l, -l+1, \ldots, l; \tag{6.12}$$

here $l = 0, 1, 2, \ldots$.

(ii) $Y_l^m(\theta, \varphi) = F_l^m(\theta) e^{im\varphi}$, *where* $F_l^m(\theta)$ *are real functions.*

We will prove this theorem in next section. It is important that each space of the common eigenfunctions is one-dimensional, since eigenvalues depend on l and m. This suggests constructing eigenfunctions of the Schrödinger operator H in the form

$$\psi(\mathbf{x}) = R(r) Y_l^m(\theta, \varphi). \tag{6.13}$$

Each solution of the spectral problem (6.2) is a sum (or a series) of the solutions of the particular form (6.13) since the spherical functions Y_l^m form the basis in $\mathcal{E}_1$. We will see that solution of the spectral problem (6.12) relies on an investigation of commutation relations for the operators $\mathbf{H}_k$, $k = 1, 2, 3$; i.e., on the Lie algebra generated by them.

6.1.3 Spherical Coordinates

To determine the radial functions in (6.13), let us express the Laplace operator Δ in spherical coordinates r, θ, φ; by definition,

$$x^3 = r\cos\theta, \qquad x^1 = r\sin\theta\cos\varphi, \qquad x^2 = r\sin\theta\sin\varphi. \tag{6.14}$$

The operator Δ is symmetric in the Hilbert space $L^2(\mathbb{R}^3)$. Hence, it is defined uniquely by its quadratic form $(\Delta\psi, \psi)$, defined for $\psi \in D := \{\psi \in L^2(\mathbb{R}^3) : \psi^{(\alpha)} \in L^2(\mathbb{R}^3) \cap C(\mathbb{R}^3), |\alpha| \le 2\}$. In spherical coordinates

$$(\Delta\psi, \psi) = -(\nabla\psi, \nabla\psi) = -\int_0^\infty dr \int_0^\pi d\theta \int_0^{2\pi} d\varphi \left|\nabla\psi(r, \theta, \varphi)\right|^2 r^2 \sin\theta. \tag{6.15}$$

Geometrically, it is evident that

$$\nabla\psi(r, \theta, \varphi) = \mathbf{e}_r \nabla_r \psi + \mathbf{e}_\theta \frac{\nabla_\theta \psi}{r} + \mathbf{e}_\varphi \frac{\nabla_\varphi \psi}{r\sin\theta}, \tag{6.16}$$

where $\mathbf{e}_r$, $\mathbf{e}_\theta$, $\mathbf{e}_\varphi$ are the orthogonal unit vectors proportional to ∇_r, ∇_θ, ∇_φ, respectively. Therefore, (6.15) becomes

$$(\Delta\psi, \psi) = -\int_0^\infty dr \int_0^\pi d\theta \int_0^{2\pi} d\varphi \left(|\nabla_r \psi|^2 + \left|\frac{\nabla_\theta \psi}{r}\right|^2 + \left|\frac{\nabla_\varphi \psi}{r \sin\theta}\right|^2 \right) r^2 \sin\theta \tag{6.17}$$

Integrating by parts,

$$\begin{aligned}
&(\Delta\psi, \psi) \\
&= \int_0^\infty dr \int_0^\pi d\theta \int_0^{2\pi} d\varphi \left(r^{-2}\nabla_r r^2 \nabla_r \psi + \frac{\nabla_\theta \sin\theta \nabla_\theta \psi}{r^2 \sin\theta} + \frac{\nabla_\varphi^2 \psi}{r^2 \sin^2\theta} \right) \bar{\psi} r^2 \sin\theta \\
&= \left(r^{-2}\nabla_r r^2 \nabla_r \psi + \frac{\nabla_\theta \sin\theta \nabla_\theta \psi}{r^2 \sin\theta} + \frac{\nabla_\varphi^2 \psi}{r^2 \sin^2\theta}, \psi \right).
\end{aligned} \tag{6.18}$$

Therefore, the Laplace operator in spherical coordinates reads as follows:

$$\Delta\psi = r^{-2}\nabla_r r^2 \nabla_r \psi + \frac{\nabla_\theta \sin\theta \nabla_\theta \psi}{r^2 \sin\theta} + \frac{\nabla_\varphi^2 \psi}{r^2 \sin^2\theta} = r^{-2}\nabla_r r^2 \nabla_r \psi + r^{-2}\Lambda\psi; \tag{6.19}$$

here Λ is the differential operator on the sphere S in coordinates θ, φ:

$$\Lambda = \frac{\nabla_\theta \sin\theta \nabla_\theta}{\sin\theta} + \frac{\nabla_\varphi^2}{\sin^2\theta}. \tag{6.20}$$

Exercise 6.4 Verify the integration by parts in (6.18).

The operator Λ is called the **spherical Laplace operator**.

Exercise 6.5 Verify the identity

$$\Lambda = -\mathbf{H}^2. \tag{6.21}$$

Hint: Both sides are second order spherically symmetric elliptic operators.

6.1.4 The Radial Equation

Here we deduce Theorem 6.1 from Theorem 6.3 by substituting (6.13) into (6.2), taking into account (6.19), (6.21), (6.12) and applying the Sommerfeld method of factorization. Further, omitting $Y_l^m(\theta, \varphi)$, and multiplying by $-2m/\hbar^2$, we arrive at the *radial equation*

$$-\frac{2\mathrm{m}E}{\hbar^2} R(r) = r^{-2}\nabla_r r^2 \nabla_r R(r) - \frac{l(l+1)}{r^2} R(r) + \frac{2\mathrm{m}e^2}{\hbar^2 r} R(r), \quad r > 0. \tag{6.22}$$

Making $r \to \infty$, this equation becomes

$$-\frac{2\mathrm{m}E}{\hbar^2} R(r) \sim R_l''(r). \tag{6.23}$$

This suggests that $E < 0$ and that the asymptotics $R(r) \sim e^{-\gamma r}$ should hold as $r \to \infty$, where

$$\gamma = \sqrt{-2\mathrm{m}E}/\hbar > 0. \tag{6.24}$$

Respectively, we factorize $R(r) = e^{-\gamma r} F(r)$. Substituting into (6.22), this gives

$$F'' + \left[\frac{2}{r} - 2\gamma\right] F' + \left[\frac{d}{r} - \frac{l(l+1)}{r^2}\right] F = 0, \quad r > 0, \tag{6.25}$$

where $d = b - 2\gamma$ with $b = 2\mathrm{m}e^2/\hbar^2$. Finally, we introduce the new variable $\rho = 2\gamma r$. Hence (6.25) becomes

$$f'' + \left[\frac{2}{\rho} - 1\right] f' + \left[\frac{\lambda - 1}{\rho} - \frac{l(l+1)}{\rho^2}\right] f = 0, \quad \rho > 0, \tag{6.26}$$

where $f(\rho) = F(r)$ and $\lambda = b/(2\gamma)$. Now let us seek for f in the form

$$f(\rho) = \rho^s \big(a_0 + a_1 \rho + a_2 \rho^2 + \cdots\big) \equiv \rho^s L(\rho), \tag{6.27}$$

where $a_0 \neq 0$. Substituting (6.27) into (6.26), we get

$$\begin{aligned} &\rho^2 L'' + \left[2s\rho + \left[\frac{2}{\rho} - 1\right]\rho^2\right] L' \\ &\quad + \left[s(s-1) + \left[\frac{2}{\rho} - 1\right] s\rho + \left[\frac{\lambda - 1}{r} - \frac{l(l+1)}{\rho^2}\right]\rho^2\right] L = 0 \end{aligned} \tag{6.28}$$

for $\rho > 0$. Evaluating, this gives

$$\begin{aligned} &\rho^2 L'' + \rho\big[2(s+1) - \rho\big] L' + \big[\rho(\lambda - 1 - s) + s(s+1) - l(l+1)\big] L = 0, \\ &\rho > 0. \end{aligned} \tag{6.29}$$

Setting $\rho = 0$, it is found that $s(s+1) - l(l+1) = 0$. Hence, either $s = l$ or $s = -l - 1$. For $s = l$, Eq. (6.29) becomes

$$\rho L'' + \big[2(l+1) - \rho\big] L' + [\lambda - 1 - l] L = 0, \quad \rho > 0. \tag{6.30}$$

The case $s = -l - 1$ is forbidden since the corresponding eigenfunction $\psi(x)$ is not a function of finite energy because then $\nabla\psi(x) \notin L^2(\mathbb{R}^3)$ and $\int \phi(\mathbf{x})|\psi(\mathbf{x})|^2\, d\mathbf{x} = \infty$.

Putting $L(\rho) = a_0 + a_1\rho + a_2\rho^2 + \cdots$ and equating the coefficients with identical powers of ρ,

$$\begin{cases} \rho^0: & 2(l+1)a_1 + (\lambda - 1 - l)a_0 = 0, \\ \rho^1: & 2a_2 + 2(l+1)2a_2 - a_1 + (\lambda - 1 - l)a_1 = 0, \\ \rho^2: & 3 \cdot 2a_3 + 2(l+1)3a_3 - 2a_2 + (\lambda - 1 - l)a_2 = 0, \\ \dots \\ \rho^k: & (k+1)ka_{k+1} + 2(l+1)(k+1)a_{k+1} - ka_k + (\lambda - 1 - l)a_k = 0 \\ \dots \end{cases} \tag{6.31}$$

we obtain the recursive relation

$$a_{k+1} = \frac{k - (\lambda - 1 - l)}{(k+1)(k+2l+2)} a_k. \tag{6.32}$$

This shows that $a_{k+1}/a_k \sim 1/k$, provided all $a_k \neq 0$. Consequently, $|L(\rho)| \geq Ce^{\rho} = Ce^{2\gamma r}$, where $C > 0$. Hence, $R(r) = F(r)e^{-\gamma r} \to \infty$ as $r \to \infty$. This, however, contradicts the finite energy of a quantum stationary state. Therefore, $a_{\overline{k}+1} = 0$ and $a_{\overline{k}} \neq 0$ for some $\overline{k} = 0, 1, 2, \dots$. Hence, $\overline{k} - (\lambda - 1 - l) = 0$, and so

$$\lambda = \frac{b}{2\gamma} = \overline{k} + l + 1 = n = 1, 2, \dots. \tag{6.33}$$

Putting $b = 2\mathrm{m}e^2/\hbar^2$ and $\gamma = \sqrt{-2\mathrm{m}E}/\hbar$, we finally get

$$E = E_n := -\frac{2\pi\hbar R_1}{n^2}, \quad n = 1, 2, \dots, \tag{6.34}$$

where $R_1 := \mathrm{m}e^4/(4\pi\hbar^3)$. This proves (6.3).

6.1.5 *Eigenfunctions and Quantum Numbers*

From (6.13) and Theorem 6.3 we obtain eigenfunctions ψ with eigenvalues E_n. Indeed, in spherical coordinates

$$\begin{aligned} \psi_{lmn} &:= Ce^{-r/r_n} P_{nl}(r) F_l^m(\theta) e^{im\varphi}, \\ & n = 1, 2, \dots,\ 0 \leq l \leq n-1,\ m = -l, \dots, l. \end{aligned} \tag{6.35}$$

Here

$$r_n := 1/\gamma = \hbar/\sqrt{-2\mathrm{m}E_n} = \hbar^2 n/\left(\mathrm{m}e^2\right) \tag{6.36}$$

and $P_{nl}(r) = \rho^l L(\rho)$ is a polynomial function of degree $\overline{k} + l = n - 1 \geq l$, according to (6.33).

The numbers $l,\ m,\ n$ are called the *azimuthal quantum number*, the *magnetic quantum number*, and the *principal quantum number* of the stationary state, respectively.

Exercise 6.6 Calculate the multiplicity of eigenvalues E_n. **Hint:** Calculate $\sum_{0\le l\le n-1}(2l+1)$.

Note that, by (6.32), the signs of the coefficients a_k alternate (if $a_{\overline{k}} \in \mathbb{R}$), since $k \le \overline{k} = \lambda - 1 - l$. Hence, $L(\rho) = \rho^l(a_0 + \cdots + a_k\rho^k)$ does not vanish for $\rho < 0$, and $n-1 = k+l$ is the number of zeros of the radial function $P_{nl}(r)$ for $\rho \ge 0$ (one may prove that all the roots are real).

Groundstate ψ_1 is defined as the eigenfunction with the lowest energy $E_1 = -2\pi\hbar cR = -\mathrm{me}^4/(2\hbar^2)$:

$$\psi_1(\mathbf{x}) = \psi_{001} = C_1 e^{-|\mathbf{x}|/r_1}, \quad r_1 = \frac{\hbar^2}{me^2}. \tag{6.37}$$

It is spherically symmetric, and concentrated in a very small region of 'Bohr radius' $r_1 \sim 1\ \text{Å} = 10^{-8}$ cm.

6.2 The Spherical Spectral Problem

Proof of Theorem 6.3 depends on the rotational invariance of the Schrödinger operator and classification of irreducible representations of the Lie algebra of SO(3).

6.2.1 Hilbert–Schmidt Argument

We start with diagonalization of the operator $\mathbf{H}^2$.

Lemma 6.7 *The operator* $\mathbf{H}^2 := \mathbf{H}_1^2 + \mathbf{H}_2^2 + \mathbf{H}_3^2$ *in* $\mathcal{E}_1 := L^2(S)$ *is selfadjoint and admits the* ***spectral resolution***

$$E = \oplus_{l=0}^{\infty} L(l), \tag{6.38}$$

where $L(l) \subset C^\infty(S)$ *are finite-dimensional orthogonal subspaces of* E, *and* $\mathbf{H}^2|_{L(l)} = \lambda_l$ *as* $\lambda_l \to \infty$ *with* $l \to \infty$.

Proof Step (i) Each operator $\mathbf{H}_k := -i\nabla_{\varphi_k}$ is selfadjoint on $\mathcal{E}_1$. Indeed, the rotations $\hat{R}_k(\varphi)$ form a unitary group in $\mathcal{E}_1$, and hence its generators ∇_{φ_k} are skew-symmetric (see (6.8)). Therefore, $\mathbf{H}^2$ is a selfadjoint nonnegative operator on $\mathcal{E}_1$.

Step (ii) The operator $\mathbf{H}^2$ is a nonnegative elliptic second-order operator on S. This follows from (6.21) and (6.20). Hence, the operator $\mathbf{H}^2 + 1 : H^2(S) \to H^0(S)$ is invertible, where $H^k(S)$ denotes the Sobolev space on S (see [164]). Respectively, the operator $(\mathbf{H}^2+1)^{-1} : H^0(S) \to H^2(S)$ is selfadjoint and compact on $H^0(S) = L^2(S)$, by the Sobolev embedding theorem. Hence, resolution (6.38) holds by the Hilbert–Schmidt theorem for the operator $(\mathbf{H}^2+1)^{-1}$. □

We can assume that $\lambda_l \neq \lambda_{l'}$ for $l \neq l'$. Then all the spaces $L(l)$ are invariant with respect to rotations of the sphere, since $\mathbf{H}^2$ commutes with rotations. Similarly, we obtain

Corollary 6.8 *All spaces* $L(l)$ *are invariant with respect to* $\mathbf{H}_k$, $k = 1, 2, 3$.

6.2.2 Lie Algebra of Angular Momenta

The linear span of $\mathbf{H}_1$, $\mathbf{H}_2$ and $\mathbf{H}_3$ is a Lie algebra, because

$$[\mathbf{H}_1, \mathbf{H}_2] = i\mathbf{H}_3, \qquad [\mathbf{H}_2, \mathbf{H}_3] = i\mathbf{H}_1, \qquad [\mathbf{H}_3, \mathbf{H}_1] = i\mathbf{H}_2. \tag{6.39}$$

Exercise 6.9 Verify the commutation relations. **Hint:** $\mathbf{H}_k = -i(\mathbf{x} \times \nabla)_k$.

Lemma 6.10 *For each* $k = 1, 2, 3$,

(i) *All spaces* $L(l)$ *are invariant with respect to* $\mathbf{H}_k$.
(ii) $[\mathbf{H}_k, \mathbf{H}^2] = 0$.

Proof (i) The invariance holds by Corollary 6.8(ii). The commutations hold by (6.11). □

Let us denote $\mathbf{H}_\pm = \mathbf{H}_1 \pm i\mathbf{H}_2$.

Lemma 6.11 *All spaces* $L(l)$ *are invariant with respect to* $\mathbf{H}_\pm$. *Also,*

$$[\mathbf{H}_3, \mathbf{H}_\pm] = \pm\mathbf{H}_\pm, \tag{6.40}$$

$$\mathbf{H}^2 = \mathbf{H}_+\mathbf{H}_- + \mathbf{H}_3(\mathbf{H}_3 - 1) = \mathbf{H}_-\mathbf{H}_+ + \mathbf{H}_3(\mathbf{H}_3 + 1), \tag{6.41}$$

$$\mathbf{H}^2 = \frac{1}{2}(\mathbf{H}_+\mathbf{H}_- + \mathbf{H}_-\mathbf{H}_+) + \mathbf{H}_3^2. \tag{6.42}$$

Proof The invariance follows from Lemma 6.10(i); further, (6.40) and (6.41) follow from (6.39), and (6.42) is a consequence of (6.41). □

6.2.3 Irreducible Representations

Here we give complete classification of all possible triples of Hermitian operators satisfying commutation relations (6.39).

Proposition 6.12 (see [145]) *Let* E *be a nonzero finite-dimensional complex linear space with a Hermitian inner product. Suppose that*

(i) *Linear Hermitian operators* $\mathbf{H}_k$, *acting on* E, $k = 1, 2, 3$, *satisfy commutation relations* (6.39),
(ii) $\mathbf{H}^2 := \mathbf{H}_1^2 + \mathbf{H}_2^2 + \mathbf{H}_3^2$ *is a scalar*: $\mathbf{H}^2 = \alpha \geq 0$,
(iii) *The space* E *is* ***irreducible***; *i.e., it does not contain any nontrivial subspace that is invariant under all* $\mathbf{H}_k$.

Then there exists a ***spin number*** $J = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots$ *and an orthonormal basis for* E, $\{e_m : m = -J, -J+1, \ldots, J-1, \ J\}$, *such that* $\alpha = J(J+1)$, *and*

$$\mathbf{H}_1 e_m = \frac{s^+_{Jm}}{2} e_{m+1} + \frac{s^-_{Jm}}{2} e_{m-1}, \qquad \mathbf{H}_2 e_m = \frac{s^+_{Jm}}{2i} e_{m+1} - \frac{s^-_{Jm}}{2i} e_{m-1}, \qquad (6.43)$$
$$\mathbf{H}_3 e_m = m e_m,$$

where

$$s^{\pm}_{Jm} = \sqrt{(J \mp m)(J \pm m + 1)}. \qquad (6.44)$$

Proof We set $\mathbf{H}_\pm := \mathbf{H}_1 \pm i\mathbf{H}_2$ as before. Then all relations (6.40)–(6.42) hold by (6.39), as in Lemma 6.10. Since $\mathbf{H}_3$ is a Hermitian operator, there exists at least one eigenvectors e_M, and so we have $\mathbf{H}_3 e_M = M e_M$ with a real eigenvalue $M \in \mathbb{R}$. □

Lemma 6.13 *Let* e_m *be an eigenvector of* $\mathbf{H}_3$ *with eigenvalue* m. *Then either* $e_\pm := \mathbf{H}_\pm e_m = 0$ *or* $e_\pm$ *is an eigenvector of* $\mathbf{H}_3$ *with eigenvalue* $m \pm 1$.

Proof (6.40) implies that $\mathbf{H}_3 e_\pm = \mathbf{H}_3 \mathbf{H}_\pm e_m = \mathbf{H}_\pm \mathbf{H}_3 e_m \pm \mathbf{H}_\pm e_m = m e_\pm \pm e_\pm$. □

By this lemma, each vector $e_{M-k} := \mathbf{H}_-^k e_M$ with $k = 0, 1, \ldots$, is either zero or an eigenvector of $\mathbf{H}_3$ with eigenvalue $M - k$. Similarly, each vector $e_{M+k} := \mathbf{H}_+^k e_M$ with $k = 0, 1, \ldots$, is either an eigenvector of $\mathbf{H}_3$ with eigenvalue $M + k$ or zero.

Since E is finite dimensional, both sequences of the eigenvectors should terminate by zero: $e_{M_-} \neq 0$, but $e_M = 0$ for $M < M_-$, and similarly, $e_{M_+} \neq 0$, but $e_M = 0$ for $M > M_+$.

Let us show that M_- coincides with the unique nonpositive solution of the equation $\alpha = M(M-1)$, and similarly, M_+ coincides with the unique nonnegative solution of the equation $\alpha = M(M+1)$.

Namely, identities (6.41) imply that $\mathbf{H}_-^* \mathbf{H}_- = \alpha - \mathbf{H}_3(\mathbf{H}_3 - 1)$ and $\mathbf{H}_+^* \mathbf{H}_+ = \alpha - \mathbf{H}_3(\mathbf{H}_3 + 1)$. Hence, for any $m = M_-, M_- + 1, \ldots, M$,

$$\begin{aligned} 0 \leq \|\mathbf{H}_- e_m\|^2 &= \langle e_m, \mathbf{H}_-^* \mathbf{H}_- e_m \rangle = \langle e_m, \big(\alpha - \mathbf{H}_3(\mathbf{H}_3 - 1)\big) e_m \rangle \\ &= \big[\alpha - m(m-1)\big] \|e_m\|^2 \end{aligned} \qquad (6.45)$$

and similarly, for any $m = M, M+1, \ldots, M_+$

$$
\begin{aligned}
0 \le \|\mathbf{H}_+ e_m\|^2 &= \big\langle e_m, \mathbf{H}_+^* \mathbf{H}_+ e_m \big\rangle = \big\langle e_m, \big(\alpha - \mathbf{H}_3(\mathbf{H}_3 + 1)\big) e_m \big\rangle \\
&= \big[\alpha - m(m+1)\big] \|e_m\|^2. \qquad (6.46)
\end{aligned}
$$

As a result, the function $\alpha - m(m-1)$ is nonnegative for $m = M_-, \dots, M$, and vanishes at M_-, since $e_{M_- -1} := H_- e_{M_-} = 0$. Similarly, the function $\alpha - m(m+1)$ is nonnegative for $m = M, \dots, M_+$, and vanishes at M_+.

Therefore, $M_+ = -M_- =: J$, where $2J = 1, 2, \dots$, since m runs in the unit step from $-J$ to J. The vectors e_m, $m = -J, -J+1, \dots, J-1, J$, constitute the basis for the space E, since E is irreducible.

Further, $\alpha - J(J+1) = 0$, and so (6.45) and (6.46) imply that

$$
\begin{aligned}
\|\mathbf{H}_- e_m\|^2 &= \big[J(J+1) - m(m-1)\big] \|e_m\|^2, \\
\|\mathbf{H}_+ e_m\|^2 &= \big[J(J+1) - m(m+1)\big] \|e_m\|^2.
\end{aligned} \qquad (6.47)
$$

Therefore,

$$
\mathbf{H}_+ e_m = s_{Jm}^+ e_{m+1}, \qquad \mathbf{H}_- e_m = s_{Jm}^- e_{m-1}. \qquad (6.48)
$$

for normalized vectors e_m. Now formulas (6.43) follow.

Corollary 6.14 *We set* $Z_{Jm}^+ = s_{J,-J}^+ \cdots s_{J,\,m-1}^+$. *Then the vectors* $Y_J^m := \mathbf{H}_+^{J+m} e_{-J} / Z_{Jm}^+$, $m = -J, \dots, J$, *constitute an orthonormal basis for* E.

Definition 6.15 For $J = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \dots$, let $D(J)$ be the irreducible space E of Proposition 6.12 (with the operators $\mathbf{H}_k$ defined by (6.43)).

Example 6.16 For $J = \frac{1}{2}$, the operators $\mathbf{H}_k$ are represented by the matrices $\hat{\mathbf{s}}_k := \frac{1}{2}\sigma_k$ in the orthonormal basis $(Y_{1/2}^{-1/2}, Y_{1/2}^{1/2})$, where σ_k are the Pauli matrices:

$$
\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \qquad (6.49)
$$

Exercise 6.17 Deduce (6.49) from (6.43).

6.2.4 *Spherical Harmonics*

Now we can prove Theorem 6.3. We may assume that λ_l are distinct for distinct l. Then Proposition 6.12 implies that each eigenspace $L(l)$ is isomorphic to a direct sum of $M(l, J)$ spaces $D(J)$ with the same values of $J = J(l) = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \dots$. Indeed, distinct values of J are impossible, since the eigenvalue of $\mathbf{H}^2$ in $D(J)$ (which is equal to $J(J+1)$) is a strictly increasing function of $J \ge 0$. Hence, Theorem 6.3(i) will follow from the next lemma.

Lemma 6.18

(i) *$M(n, J) = 0$ for each n if J is half-integer.*
(ii) *For every integer $J = 1, 2, 3, \ldots$, there exists a unique $l = l(J)$ such that $M(l, J) = 1$.*

Proof (i) Each eigenspace $L(l)$ is a direct sum of subspaces $D(J)$. Let us consider one of these subspaces. By (6.8) and (6.43),

$$\mathbf{H}_3 e_m(\theta, \varphi) = -i\nabla_\varphi e_m(\theta, \varphi) = m e_m(\theta, \varphi) \tag{6.50}$$

for each eigenvector $e_m \in D(J)$ of $\mathbf{H}_3$. Therefore, $e_m(\theta, \varphi) = F_J^m(\theta)e^{im\varphi}$, an so m is an integer number, because the function $e_m(\theta, \varphi)$ should be single-valued. Hence J is also integer.

(ii) Let us consider the lowest eigenvalue $m = -J$ in the subspace $D(J)$. We have $\mathbf{H}_- e_{-J}(\theta, \varphi) = 0$ by Lemma 6.13. We will show below that, in spherical coordinates,

$$\mathbf{H}_- = -e^{-i\varphi}[\nabla_\theta - i\cot\theta\nabla_\varphi], \qquad \mathbf{H}_+ = e^{i\varphi}[\nabla_\theta + i\cot\theta\ \nabla_\varphi]. \tag{6.51}$$

Taking into account that $e_{-J}(\theta, \varphi) = F_J^{-J}(\theta)e^{-iJ\varphi}$ and since $\mathbf{H}_- e_{-J}(\theta, \varphi) = 0$, we arrive at the differential equation $(\nabla_\theta - J\cot\theta)F_J^{-J} = 0$. Hence, $F_J^{-J} = C\sin^J\theta$. This means that $-J$ is a simple eigenvalue. Hence, $M(l, J) \le 1$ for all l.

It remains to verify that $M(l, J) \ne 0$ for some l. This is equivalent to the existence of an eigenvector of $\mathbf{H}^2$ with eigenvalue $J(J+1)$. However, this is the case of the function $e_{-J}(\theta, \varphi) = C\sin^J\theta e^{-iJ\varphi}$ constructed above. Indeed, (6.41) implies that

$$\mathbf{H}^2 e_{-J} = \mathbf{H}_+\mathbf{H}_- e_{-J} + \mathbf{H}_3(\mathbf{H}_3 - 1)e_{-J} = J(J+1)e_{-J},$$

since $\mathbf{H}_- e_{-J} = 0$, and $\mathbf{H}_3 e_{-J} = -Je_{-J}$, because $\mathbf{H}_3 = -i\nabla_\varphi$. □

Now Theorem 6.3(i) is proved, inasmuch as the functions

$$Y_l^m(\theta, \varphi) := \mathbf{H}_+^{l+m} e_{-l}/Z_{lm}^+, \quad m = -l, \ldots, l \tag{6.52}$$

form an orthonormal basis for the space $D(l)$ by Corollary 6.14. Theorem 6.3(ii) also follows from (6.52), because

$$\mathbf{H}_+^{l+m} e_{-l} = \left(e^{i\varphi}[\nabla_\theta + i\cot\theta\ \nabla_\varphi]\right)^{l+m}\left(\sin^l\theta e^{-il\varphi}\right) = G_l^m(\theta)e^{im\varphi}, \tag{6.53}$$

where $G_l^m(\theta)$ is a real function. The latter follows from (6.53) by complex conjugation and the substitution $\varphi \mapsto -\varphi$.

6.2.5 *Angular Momenta in Spherical Coordinates*

Let us prove (6.51). First, we rewrite (6.16) as follows:

$$\boldsymbol{\nabla}\psi(r,\theta,\varphi)=\mathbf{e}_r\nabla_r\psi+\mathbf{e}_\theta\frac{\nabla_\theta\psi}{r}+\mathbf{e}_\varphi\frac{\nabla_\varphi\psi}{r\sin\theta}=\mathbf{e}_1\boldsymbol{\nabla}_1\psi+\mathbf{e}_3\boldsymbol{\nabla}_3\psi+\mathbf{e}_3\boldsymbol{\nabla}_3\psi; \tag{6.54}$$

here $\mathbf{e}_1 := (1,0,0)$, etc. It is geometrically evident that

$$\begin{cases}\mathbf{e}_r=(\mathbf{e}_1\cos\varphi+\mathbf{e}_2\sin\varphi)\sin\theta+\mathbf{e}_3\cos\theta,\\ \mathbf{e}_\theta=(\mathbf{e}_1\cos\varphi+\mathbf{e}_2\sin\varphi)\cos\theta-\mathbf{e}_3\sin\theta,\\ \mathbf{e}_\varphi=\ \mathbf{e}_2\cos\varphi-\mathbf{e}_1\sin\varphi.\end{cases} \tag{6.55}$$

Substituting into (6.54), this establishes

$$\begin{cases}\boldsymbol{\nabla}_1=\sin\theta\cos\varphi\nabla_r+\cos\theta\cos\varphi\dfrac{\nabla_\theta}{r}-\sin\varphi\dfrac{\nabla_\varphi}{r\sin\theta},\\ \boldsymbol{\nabla}_2=\sin\theta\sin\varphi\nabla_r+\cos\theta\sin\varphi\dfrac{\nabla_\theta}{r}+\cos\varphi\dfrac{\nabla_\varphi}{r\sin\theta},\\ \boldsymbol{\nabla}_3=\cos\theta\nabla_r-\sin\varphi\dfrac{\nabla_\theta}{r}.\end{cases} \tag{6.56}$$

Finally, substituting (6.56) and (6.14) into $\mathbf{H}_k=-i(\mathbf{x}\times\boldsymbol{\nabla})_k$, we get

$$\begin{cases}\mathbf{H}_1=i(\sin\varphi\nabla_\theta+\cot\theta\cos\varphi\nabla_\varphi),\\ \mathbf{H}_2=i(-\cos\varphi\nabla_\theta+\cot\theta\sin\varphi\nabla_\varphi),\\ \mathbf{H}_3=-i\nabla_\varphi.\end{cases} \tag{6.57}$$

Now, (6.51) follows from the first two formulas.

Chapter 7
Atom Radiation

Abstract A stationary atom radiation should be maintained by an external source, which provides the corresponding statistical equilibrium. The spectrum of the corresponding atom radiation can be calculated from the Maxwell equations. The calculation justifies the Rydberg–Ritz Combination Principle (1.1), and gives the intensity of spectral lines, and the *selection rules* in the case of a static uniform external magnetic field.

The formula for the intensities implies spectral distribution of the limit stationary Maxwell field ('black-body radiation').

7.1 Atom in Thermodynamical Equilibrium

Radiation obviously reduces the atom energy. Hence a stationary radiation should be maintained by an external source like electric discharge or heat bath, X-rays, etc. The corresponding equilibrium correlations for the wave function should agree with the Maxwell–Boltzmann–Gibbs distribution.

7.1.1 Relaxation to Equilibrium Distribution

Heat bath or X-rays source can be modeled by a *random incident wave* ($\phi_{\rm in}(t,\mathbf{x})$, $\mathbf{A}_{\rm in}(t,\mathbf{x})$) in the coupled Maxwell–Schrödinger equations (4.22). The number of atoms is of order 10^{23}, the atoms are identical and independent if we neglect their interactions. Respectively, the atoms of gas should be described by a *random solution* of the coupled equations with a static external Maxwell field and the random incident wave:

$$\begin{cases} \big[i\hbar\partial_t - e\phi_{\rm tot}(t,\mathbf{x})\psi(t,\mathbf{x})\big] = \dfrac{1}{2\mathrm{m}}\Big[-i\hbar\nabla - \dfrac{e}{c}\mathbf{A}_{\rm tot}(t,\mathbf{x})\Big]^2\psi(t,\mathbf{x}), \\ \dfrac{1}{4\pi}\Box\phi = e\big|\psi(t,\mathbf{x})\big|^2, \\ \dfrac{1}{4\pi}\Box\mathbf{A} = \dfrac{e}{\mathrm{m}c}\psi(t,\mathbf{x})\cdot\Big[-i\hbar\nabla - \dfrac{e}{c}\mathbf{A}_{\rm tot}(t,\mathbf{x})\Big]\psi(t,\mathbf{x}). \end{cases} \tag{7.1}$$

A. Komech, *Quantum Mechanics: Genesis and Achievements*,
DOI 10.1007/978-94-007-5542-0_7, © Springer Science+Business Media Dordrecht 2013

Here the total field is defined as

$$\phi_{\text{tot}} := \phi_{\text{ext}}(\mathbf{x}) + \phi_{\text{in}}(t,\mathbf{x}) + \phi(t,\mathbf{x}), \qquad \mathbf{A}_{\text{tot}} := \mathbf{A}_{\text{ext}}(\mathbf{x}) + \mathbf{A}_{\text{in}}(t,\mathbf{x}) + \mathbf{A}(t,\mathbf{x}). \tag{7.2}$$

The static external potentials $(\phi_{\text{ext}}(\mathbf{x}), \mathbf{A}_{\text{ext}}(\mathbf{x}))$ include the Coulombic potential of the hydrogen nucleus. The incident random wave $(\phi_{\text{in}}(t,\mathbf{x}), \mathbf{A}_{\text{in}}(t,\mathbf{x}))$ should be a *stationary random process*. Respectively, we suggest that the solution of equations (7.1) converges as $t \to \infty$ to the corresponding *stationary random process*, which does not depend on the initial state. This convergence is provisionally due to the radiation of energy to infinity.

Remark 7.1

(i) The convergence to a stationary random process for equations with random stationary coefficients is still an open problem.

(ii) The convergence was established for a number of linear hyperbolic equations with non random coefficients, and with initial measure satisfying the mixing Rosenblatt or Ibragimov–Linnik condition. Then the limit stationary measures are Gaussian (see e.g. [47–52]).

(iii) For nonlinear hyperbolic equations, the first (and unique) result on the convergence is obtained in [94]. In this case, the limit stationary measure is the canonical equilibrium Boltzmann–Gibbs distribution.

7.1.2 Perturbation Theory

Let us write the Schrödinger equation from (7.1) as

$$i\partial_t \psi = H\psi + H'(t)\psi, \tag{7.3}$$

where H is the 'unperturbed' Schrödinger operator

$$H = \frac{1}{2\mathrm{m}}\left[-i\hbar\nabla - \frac{e}{c}\mathbf{A}_{\text{ext}}(\mathbf{x})\right]^2 + e\phi_{\text{ext}}(\mathbf{x}) \tag{7.4}$$

and $H'(t)$ is the perturbation:

$$\begin{aligned} H'(t) = {} & \frac{e}{\mathrm{m}c}\mathbf{A}_*(t,\mathbf{x}) \cdot \left[i\hbar\nabla + \frac{e}{c}\mathbf{A}_{\text{ext}}(\mathbf{x})\right] + \frac{ie\hbar}{2\mathrm{m}c}\left[\nabla \cdot \mathbf{A}_*(t,\mathbf{x})\right] \\ & + \frac{1}{2\mathrm{m}}\left[\frac{e}{c}\mathbf{A}_*(t,\mathbf{x})\right]^2 + e\phi_*(t,\mathbf{x}), \end{aligned} \tag{7.5}$$

where

$$\mathbf{A}_*(t,\mathbf{x}) := \mathbf{A}(t,\mathbf{x}) + \mathbf{A}_{\text{in}}(t,\mathbf{x}), \qquad \phi_*(t,\mathbf{x}) := \phi(t,\mathbf{x}) + \phi_{\text{in}}(t,\mathbf{x}). \tag{7.6}$$

Equation (7.3) satisfies the charge conservation law

$$\|\psi(t)\| = 1 \tag{7.7}$$

in spite of the fact that a Maxwell field may depend on time. The normalization (7.7) means that the total charge of the hydrogen atom vanishes.

7.1.3 The Dirac 'Interaction Picture'

In many practical cases, the potentials $\mathbf{A}_*(t,\mathbf{x})$ and $\phi_*(t,\mathbf{x})$ are small with respect to $\mathbf{A}_{\text{ext}}(t,\mathbf{x})$ and $\phi_{\text{ext}}(t,\mathbf{x})$. Hence, it is natural to expand the random solution of the Schrödinger equation (7.3) in terms of eigenfunctions of the unperturbed Schrödinger operator H:

$$\psi(t,\mathbf{x}) = \sum c_n(t)\psi_n(\mathbf{x}) + \int_0^\infty c(\omega,t)\psi(\omega,\mathbf{x})\,d\omega. \tag{7.8}$$

Here $\psi_n(x)$ ($\psi(\omega,x)$, respectively) are the normalized eigenfunctions of the discrete (continuous) spectrum of H,

$$\begin{cases} H\psi_n = \hbar\omega_n\psi_n, & H\psi(\omega,\cdot) = \hbar\omega\psi(\omega,\cdot), \\ \langle\psi_n,\psi_{n'}\rangle = \delta_{nn'}, & \big\langle\psi(\omega,\cdot),\psi\big(\omega',\cdot\big)\big\rangle = \delta\big(\omega-\omega'\big). \end{cases} \tag{7.9}$$

The coefficients $c_n(t)$ and $c(\omega,t)$ also should be stationary random processes, since we suggest that for large time the solution is a stationary random processes.

If we neglect the potentials $\phi_*(t,\mathbf{x})$ and $\mathbf{A}_*(t,\mathbf{x})$, we obtain

$$c_n(t) = a_n e^{-i\omega_n t}, \qquad c(\omega,t) = a(\omega)e^{-i\omega t}. \tag{7.10}$$

Hence it is natural to write, in general case,

$$c_n(t) = a_n(t)e^{-i\omega_n t}, \qquad c_\omega(t) = a(\omega,t)e^{-i\omega t}. \tag{7.11}$$

Respectively, (7.12) reads

$$\psi(t,\mathbf{x}) = \sum a_n(t)e^{-i\omega_n t}\psi_n(\mathbf{x}) + \int_0^\infty a(\omega,t)e^{-i\omega t}\psi(\omega,\mathbf{x})\,d\omega. \tag{7.12}$$

Here the amplitudes $a_k(t)$ and $a(\omega,t)$ satisfy the following equations

$$i\hbar\dot{a}_n(t)e^{-i\omega_n t} = \big\langle H'(t)\psi(t),\psi_n\big\rangle, \qquad i\hbar\dot{a}(\omega,t)e^{-i\omega t} = \big\langle H'(t)\psi(t),\psi(\omega,\cdot)\big\rangle, \tag{7.13}$$

which are referred to Dirac's 'interaction picture' of the Schrödinger equation (7.3). Therefore, the amplitudes $a_n(t)$ and $a(\omega,t)$ vary slowly, since $H'(t)$ is a small perturbation.

7.1.4 Thermodynamical Equilibrium in Schrödinger Theory

For calculation of the atom radiation in thermal equilibrium we need a suitable postulate for the corresponding equilibrium distribution. The Boltzmann distribution (1.65) from the 'old quantum mechanics' means that the electron occupies exactly one stationary orbit at every moment of time with probability

$$p_n = \frac{1}{Z} e^{-\frac{\hbar\omega_n}{kT}}, \quad Z = \sum e^{-\frac{\hbar\omega_n}{kT}}, \tag{7.14}$$

since, by (1.54), $E_n = \hbar\omega_n$. An implementation of this distribution into the Schrödinger theory is not straightforward. For the orbit $|E_m\rangle$ corresponding to the wave function $\psi_m(x)e^{-\omega_m t}$, we have

$$\left|a_m(t)\right|^2 = 1, \quad \text{and} \quad a_n(t) = 0 \quad \text{for } n \neq m. \tag{7.15}$$

S1. Equilibrium Correlations One possible option in the Schrödinger theory is to identify the correlation functions

$$M\left|a_n(t)\right|^2 = \frac{1}{Z} e^{-\frac{\hbar\omega_n}{kT}}, \tag{7.16}$$

where M denotes the mathematical expectation. This postulate agrees with the Boltzmann distribution (7.14), since, by (7.15),

$$M\left|a_n(t)\right|^2 = p_n \cdot 1 + (1 - p_n) \cdot 0 = p_n \tag{7.17}$$

which coincides with (7.16).

S2. Phase Incoherence For calculation of intensity of spectral lines below we need the following orthogonality relations

$$\begin{aligned} &M\big[a_n(t) \otimes a_{n'}(t)\big] = 0, \\ &M\big[a_n(t) \otimes a_{n'}(t) \otimes a_{n''}(t) \otimes a_{n'''}(t)\big] = 0 \quad \text{if } n \neq n', n'', n''', \end{aligned} \tag{7.18}$$

where $\otimes$ is the tensor product of real vectors of $\mathbb{R}^2$ corresponding to the complex numbers. The relations express the 'phase incoherence'. These identities also correspond to the spirit of the old quantum mechanics, since the atoms are independent, while their amplitudes should be oscillatory with independent phases and zero mean, as in the following example.

Example 7.2 Let us consider the potentials

$$\phi_{\text{in}}(t, \mathbf{x}) = \hbar p \sin \nu t, \qquad \mathbf{A}_{\text{in}}(t, \mathbf{x}) \equiv 0, \qquad \mathbf{A}_{\text{ext}}(t, \mathbf{x}) \equiv 0, \tag{7.19}$$

where p and $\nu \neq 0$ are real. Then $H'(t) = e\phi_{\text{in}}(t, \mathbf{x})$, in the Born approximation, neglecting the contribution of the potential $\mathbf{A}(t, \mathbf{x})$. Hence, for the discrete component of (7.8),

$$i\dot{c}_n(t) = [\omega_n + ep \sin \nu t] c_n(t). \tag{7.20}$$

Integrating, this establishes, for $\nu \neq 0$,

$$c_n(t) = C_n e^{-i\omega_n t + ir \cos \nu t}, \quad r = \frac{ep}{\nu}. \tag{7.21}$$

Further, let us choose random initial state $c_n(0) = b_n e^{i\alpha_n}$ with some fixed amplitudes $b_n \in \mathbb{R}$ and uniformly distributed phases $\alpha_n \in [0, 2\pi]$, which are independent for different n. Then $C_n = b_n e^{i(\alpha_n - r)}$. Finally, $c_n(t)$ are independent stationary random processes, which can be written in the form

$$c_n(t) = a_n(t) e^{-i\omega_n t}, \qquad a_n(t) = b_n e^{i\theta_n(t)}, \tag{7.22}$$

where the random phases $\theta_n(t) = \alpha_n - r + r \cos \nu t$ are independent for different n, and each $\theta_n(t) \bmod 2\pi$ is uniformly distributed over $[0, 2\pi]$. Hence relation (7.18) obviously holds.

Remark 7.3 Suppose that the source potentials (7.19) is small: $|ep| \sim \varepsilon \ll 1$. Then the amplitudes $a_n(t)$ are *slowly varying*, since

$$\frac{|\dot{a}_n(t)|}{|a_n(t)|} = \left|\dot{\theta}_n(t)\right| \leq |r\nu| = |ep| \sim \varepsilon. \tag{7.23}$$

Hence,

$$a_n(t) \approx a_n(0), \quad |t| \leq \varepsilon^{-1}. \tag{7.24}$$

7.2 Atom Radiation

The Maxwell equations determine the radiation of a gas in a statistical equilibrium maintained by an external source. For large times, the Maxwell field coincides with the retarded potentials, which imply the Rydberg–Ritz combination rule and the intensity of spectral lines, thereby justifying the correspondence principle and the selection rules in an external Maxwell field with cylindrical symmetry.

7.2.1 Radiated Maxwell Field

The total Maxwell field (7.2) splits into three terms. The incident wave is absorbed by the heated cavity, which models the 'black-body' and does not contribute to the radiation. The external potentials are static, and their contribution to the radiation vanishes (see Remark 7.6(ii) below).

Finally, we will calculate the last term of (7.2) from the Maxwell equations in (7.1). We consider the Born approximation (i.e., we neglect the interaction terms containing $\mathbf{A}_*(t,\mathbf{x})$ in the right-hand side of last equation (7.1)),

$$\begin{aligned} &\frac{1}{4\pi}\Box\phi(t,\mathbf{x}) = e\big|\psi(t,\mathbf{x})\big|^2, \\ &\frac{1}{4\pi}\Box\mathbf{A}(t,\mathbf{x}) = \frac{e}{\mathrm{m}c}\psi(t,\mathbf{x})\cdot\Big[-i\hbar\nabla\psi(t,\mathbf{x}) - \frac{e}{c}\mathbf{A}_{\text{ext}}(\mathbf{x})\Big]\psi(t,\mathbf{x}). \end{aligned} \tag{7.25}$$

Long time asymptotics of the Maxwell potentials for bounded $|\mathbf{x}|$ are given by the retarded potentials (12.160):

$$\left\{\begin{aligned} \phi(t,\mathbf{x}) &\sim \phi_{\text{ret}}(t,\mathbf{x}) := \int\frac{\rho(t-|\mathbf{x}-\mathbf{y}|/c,\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}\,d\mathbf{y}, \\ \mathbf{A}(t,\mathbf{x}) &\sim \mathbf{A}_{\text{ret}}(t,\mathbf{x}) := \frac{1}{c}\int\frac{\mathbf{j}(t-|\mathbf{x}-\mathbf{y}|/c,\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}\,d\mathbf{y}, \end{aligned}\right| \quad t\to\infty,\ |\mathbf{x}|\le R. \tag{7.26}$$

We may assume that an atom is in the groundstate (6.37), which is located at the point $\mathbf{y}=0$. Then the corresponding charge and current densities decay rapidly:

$$\rho(t,\mathbf{y}),\mathbf{j}(t,\mathbf{y}) = \mathcal{O}\big(e^{-2|\mathbf{y}|/r_1}\big); \tag{7.27}$$

here $r_1 = \frac{\hbar^2}{\mathrm{m}e^2} > 0$ is the *atom radius*, $r_1 \sim 1\ \text{Å} = 10^{-8}\,\text{cm}$.

The densities are localized in the ball $|\mathbf{y}| \le r_1 \ll 1$, hence any macroscopic observation at a distance $|\mathbf{x}| \sim 1$ agrees with high precision with the approximations

$$\left\{\begin{aligned} \phi(t,\mathbf{x}) &\sim \frac{1}{|\mathbf{x}|}\int_{|\mathbf{y}|\le r_1}\rho\big(t-|\mathbf{x}-\mathbf{y}|/c,\mathbf{y}\big)\,d\mathbf{y}, \\ \mathbf{A}(t,\mathbf{x}) &\sim \frac{1}{c|\mathbf{x}|}\int_{|\mathbf{y}|\le r_1}\mathbf{j}\big(t-|\mathbf{x}-\mathbf{y}|/c,\mathbf{y}\big)\,d\mathbf{y}, \end{aligned}\right| \quad |\mathbf{x}|\sim 1. \tag{7.28}$$

7.2.2 The Rydberg–Ritz Combination Principle

Let us calculate the charge-current densities from (7.25),

$$\left\{\begin{aligned} \rho(t,\mathbf{y}) &= e\overline{\psi(t,\mathbf{y})}\psi(t,\mathbf{y}), \\ \mathbf{j}(t,\mathbf{y}) &= \frac{e}{\mathrm{m}}\mathrm{Re}\Big[\overline{\psi(t,\mathbf{y})}\Big[-i\hbar\nabla - \frac{e}{c}\mathbf{A}_{\text{ext}}(\mathbf{y})\Big]\psi(t,\mathbf{y})\Big], \end{aligned}\right. \tag{7.29}$$

by substituting expansion (7.12). The integral over the continuous spectrum in the RHS of (7.12) contributes only to the continuous spectrum of radiation. Hence, the discrete spectrum is completely determined by the sum over discrete eigenvalues. Substituting the sum into (7.29), we get the components of the charge and current densities, which are responsible for the discrete spectrum. Writing $c_k(t) = a_k(t)e^{-i\omega_k t}$, we obtain the components in the form,

$$\begin{cases} \rho_d\big(t - |\mathbf{x}-\mathbf{y}|/c, \mathbf{y}\big) := e\sum\limits_{nn'} \overline{a_n(t)}a_{n'}(t)e^{i(\omega_n-\omega_{n'})(t-|\mathbf{x}-\mathbf{y}|/c)}\overline{\psi_n(\mathbf{y})}\psi_{n'}(\mathbf{y}), \\ \mathbf{j}_d\big(t - |\mathbf{x}-\mathbf{y}|/c, \mathbf{y}\big) \\ \quad := \dfrac{e}{\mathrm{m}}\mathrm{Re}\sum\limits_{nn'} \overline{a_n(t)}a_{n'}(t)e^{i(\omega_n-\omega_{n'})t}\overline{\psi_n(\mathbf{y})}e^{-i\omega_n|\mathbf{x}-\mathbf{y}|/c}\left[-i\hbar\nabla_{\mathbf{y}} - \dfrac{e}{c}\mathbf{A}_{\text{ext}}(\mathbf{y})\right] \\ \quad\quad \times \left[\psi_{n'}(\mathbf{y})e^{i\omega_{n'}|\mathbf{x}-\mathbf{y}|/c}\right]. \end{cases} \tag{7.30}$$

Substituting into (7.28), we get the following approximation for the discrete component of radiation:

$$\left.\begin{cases} \phi_d(t,\mathbf{x}) \sim \dfrac{1}{|\mathbf{x}|}\sum \phi_{nn'}(t,\mathbf{x})e^{i\omega_{nn'}t}, \\ \mathbf{A}_d(t,\mathbf{x}) \sim \dfrac{1}{c|\mathbf{x}|}\mathrm{Re}\sum \mathbf{A}_{nn'}(t,\mathbf{x})e^{i\omega_{nn'}t}, \end{cases}\right| \quad |\mathbf{x}| \sim 1 \tag{7.31}$$

where $\omega_{nn'} := \omega_n - \omega_{n'}$, and

$$\begin{cases} \phi_{nn'}(t,\mathbf{x}) = e\overline{a_n(t)}a_{n'}(t)\displaystyle\int_{|\mathbf{y}|\le r_1} e^{-i\omega_{nn'}|\mathbf{x}-\mathbf{y}|/c}\overline{\psi_n(\mathbf{y})}\psi_{n'}(\mathbf{y})\,d\mathbf{y}, \\ \mathbf{A}_{nn'}(t,\mathbf{x}) \approx \dfrac{e}{\mathrm{m}}\overline{a_n(t)}a_{n'}(t)\displaystyle\int_{|\mathbf{y}|\le r_1} e^{-i\omega_{nn'}|\mathbf{x}-\mathbf{y}|/c}\overline{\psi_n(\mathbf{y})} \\ \qquad\qquad \times\left[-i\hbar\nabla - \dfrac{e}{c}\mathbf{A}_{\text{ext}}(\mathbf{y})\right]\psi_{n'}(\mathbf{y})\,d\mathbf{y}. \end{cases} \tag{7.32}$$

The last formula holds with an error proportional to $\hbar\omega_{n'}/c$, arising from the commutation of $e^{i\omega_{n'}|\mathbf{x}-\mathbf{y}|/c}$ with differentiations. The error is negligible if

$$\hbar|\omega_{n'}| \ll e\big|\mathbf{A}_{\text{ext}}(\mathbf{y})\big|, \quad |\mathbf{y}| \le r_1. \tag{7.33}$$

Corollary 7.4 *Let $a_n(t)$ be slowly varying functions as in* (7.24). *Then the Rydberg–Ritz Combination Principle holds by* (7.31): *the* ***discrete spectrum*** *of radiation is contained in the set* $\{\omega_{nn'} = \omega_n - \omega_{n'}\}$.

7.2.3 The Dipole Approximation

We will assume (7.24) and (7.33), and also that

$$|\omega_{nn'}|r_1/c \ll 2\pi \tag{7.34}$$

for all terms in (7.12); this means that $\lambda_{nn'} := 2\pi c/|\omega_{nn'}| \gg r_1$ for the corresponding wavelengths. Then the exponential function in integrands (7.32) can be replaced by $\exp(-i\omega_{nn'}|\mathbf{x}|/c)$, so (7.32) becomes

$$\begin{cases} \phi_{nn'}(t,\mathbf{x}) \approx e\overline{a_n(t)}a_{n'}(t)e^{-i\omega_{nn'}|\mathbf{x}|/c}\displaystyle\int \overline{\psi_n(\mathbf{y})}\psi_{n'}(\mathbf{y})\,d\mathbf{y}, \\ \mathbf{A}_{nn'}(t,\mathbf{x}) \approx \dfrac{e}{\mathrm{m}}\overline{a_n(t)}a_{n'}(t)e^{-i\omega_{nn'}|\mathbf{x}|/c}\displaystyle\int \overline{\psi_n(\mathbf{y})}\Big[-i\hbar\nabla - \frac{e}{c}\mathbf{A}_{\mathrm{ext}}(\mathbf{y})\Big]\psi_{n'}(\mathbf{y})\,d\mathbf{y}. \end{cases} \tag{7.35}$$

The first formula implies that $\phi_{nn'}(t,\mathbf{x}) = 0$ for $\omega_{nn'} \neq 0$, by the orthogonality of different eigenfunctions. On the other hand, for $\omega_{nn'} = 0$, the first formula shows that $\phi_{nn'}(t,\mathbf{x})$ does not depend on $\mathbf{x}$, hence the electrostatic potential $\phi_d(t,\mathbf{x}) \sim C/|\mathbf{x}|$ as $|\mathbf{x}| \to \infty$, by (7.31). Therefore,

$$\big|\nabla\phi_d(t,\mathbf{x})\big| = \mathcal{O}\big(|\mathbf{x}|^{-2}\big), \quad |\mathbf{x}| \to \infty. \tag{7.36}$$

Hence, the potential does not contribute to the energy radiation of the discrete component through the large sphere. Similarly, the external potential ϕ_{ext} of the atomic nucleus does not contribute to the energy radiation through the large sphere. Consequently, the radiation is totally due to the magnetic potential

$$\mathbf{A}_d(t,\mathbf{x}) = \frac{1}{c|\mathbf{x}|}\sum \mathrm{Re}\big[\overline{a_n(t)}a_{n'}(t)e^{i\omega_{nn'}(t-|\mathbf{x}|/c)}\mathbf{J}_{nn'}\big], \quad t \to \infty, \tag{7.37}$$

where $\mathbf{J}_{nn'}$ stands for the last integral in (7.35) up to a factor,

$$\mathbf{J}_{nn'} = \frac{e}{\mathrm{m}}\int \overline{\psi_n(\mathbf{y})}\Big[-i\hbar\nabla - \frac{e}{c}\mathbf{A}_{\mathrm{ext}}(\mathbf{y})\Big]\psi_{n'}(\mathbf{y})\,d\mathbf{y}. \tag{7.38}$$

We set

$$\mathbf{p}_{nn'}(t) = \mathrm{Im}\Bigg[\frac{\overline{a_n(t)}a_{n'}(t)}{\omega_{nn'}}e^{i\omega_{nn'}t}\mathbf{J}_{nn'}\Bigg]. \tag{7.39}$$

With this notation, the magnetic potential (7.37) reads

$$\mathbf{A}_d(t,\mathbf{x}) = \frac{1}{c|\mathbf{x}|}\sum_{nn'}\dot{\mathbf{p}}_{nn'}\big(t - |\mathbf{x}|/c\big), \tag{7.40}$$

since the coefficients $a_n(t)$, $a_{n'}(t)$ are slowly varying in time. Comparing with the Hertzian dipole radiation (12.120), it is found that the term nn' of the radiated field is identical to radiation of the dipole with momentum (7.39). We will show that the radiated energy is the sum of the dipoles radiation under condition (7.18).

Remark 7.5 Definition (7.39) implies that $\mathbf{p}_{n'n}(t) = \mathbf{p}_{nn'}(t)$, since $\mathbf{J}_{n'n}(t) = \overline{\mathbf{J}_{nn'}(t)}$ by (7.38).

7.2.4 Energy Flow at Infinity

Magnetic field can be calculated similarly to the case of Hertzian dipole, as in Sect. 12.10: if $r := |\mathbf{x}| \to \infty$, we obtain

$$
\begin{aligned}
\mathbf{B}(t,\mathbf{x}) = \nabla \times \mathbf{A}_d(t,\mathbf{x}) &\approx \frac{1}{c}\nabla \times \sum \frac{\dot{\mathbf{p}}_{nn'}(t-r/c)}{r} \\
&= \frac{1}{c}\sum \dot{\mathbf{p}}_{nn'}(t-r/c) \times \frac{\mathbf{x}}{r^3} + \frac{1}{cr}\nabla(t-r/c) \times \sum \ddot{\mathbf{p}}_{nn'}(t-r/c) \\
&\sim \frac{1}{c^2}\sum \frac{\ddot{\mathbf{p}}_{nn'}(t-r/c) \times \mathbf{x}}{r^2} + \mathcal{O}\big(r^{-2}\big).
\end{aligned}
$$

In a similar fashion, for the electric field, we have

$$
\mathbf{E}(t,\mathbf{x}) = -\frac{1}{c}\dot{\mathbf{A}}_d(t,\mathbf{x}) - \nabla\phi(t,\mathbf{x}) \sim -\frac{1}{c^2}\sum \frac{\ddot{\mathbf{p}}_{nn'}(t-r/c)}{r} + \mathcal{O}\big(|\mathbf{x}|^{-2}\big), \quad (7.41)
$$

by (7.36). Now for energy flow (Poynting vector) we find

$$
\begin{aligned}
&\mathbf{S}(t,\mathbf{x}) = \frac{c}{4\pi}\mathbf{E} \times \mathbf{B} \sim \frac{1}{4\pi c^3 r^3}\sum \ddot{\mathbf{p}}_{n_1 n_1'}(t-r/c) \times \big[\ddot{\mathbf{p}}_{n_2 n_2'}(t-r/c) \times \mathbf{x}\big], \\
&|\mathbf{x}| \gg r_1.
\end{aligned}
\quad (7.42)
$$

7.2.5 Intensity of Spectral Lines

The power of radiation is the integral of energy flow (7.42) over the sphere of a large radius $R > 0$,

$$
\begin{aligned}
P_R(t) &:= \int_{|\mathbf{x}|=R} \mathbf{S}(t,\mathbf{x}) \cdot \mathbf{n}\, d\mathbf{x} \\
&\sim \frac{1}{4\pi c^3 R^4}\sum \int_{|\mathbf{x}|=R} \big[\ddot{\mathbf{p}}_{n_1 n_1'}(t-R/c) \times \mathbf{x}\big] \cdot \big[\ddot{\mathbf{p}}_{n_2 n_2'}(t-R/c) \times \mathbf{x}\big]\, d\mathbf{x},
\end{aligned}
\quad (7.43)
$$

where $\mathbf{n} := \mathbf{x}/|\mathbf{x}|$. By (7.39) and (7.24),

$$
\ddot{\mathbf{p}}_{nn'}(t) \approx \mathrm{Re}\big[a_n(t)\overline{a_{n'}(t)}\omega_{nn'}e^{-i\omega_{nn'}t}\mathbf{J}_{nn'}\big] = -\omega_{nn'}^2 \mathbf{p}_{nn'}(t). \quad (7.44)
$$

Hence, (7.18) implies, in this approximation,

$$
M\big[\ddot{\mathbf{p}}_{n_1 n_1'}(t-R/c) \times \mathbf{x}\big] \cdot \big[\ddot{\mathbf{p}}_{n_2 n_2'}(t-R/c) \times \mathbf{x}\big] = 0, \quad (7.45)
$$

provided $(n_1, n_1') \neq (n_2, n_2')$ and $(n_1, n_1') \neq (n_2', n_2)$, where M stands for the mathematical expectation. Moreover, contributions to (7.43) of the terms with $(n_1, n_1') =$

(n_2, n_2') and $(n_1, n_1') = (n_2', n_2)$ coincide when averaged over time. Therefore, applying M and averaging in time in (7.43), we obtain, for large R, the mean power of radiation

$$\begin{aligned} MP_R(t) &\sim \frac{1}{4\pi c^3} 2 \sum M \int_{|\mathbf{n}|=1} \overline{\left[\ddot{\mathbf{p}}_{nn'}(t-R/c) \times \mathbf{n}\right]^2}\, d\mathbf{n} \\ &= \frac{1}{2\pi c^3} \sum M \overline{\ddot{\mathbf{p}}^2_{nn'}(t-R/c)} \int_{|\mathbf{n}|=1} \sin^2 \chi \, d\mathbf{n} \\ &= \frac{1}{2\pi c^3} \sum M \overline{\ddot{\mathbf{p}}^2_{nn'}(t-R/c)} \frac{8\pi}{3} \\ &= \frac{4}{3c^3} \sum M \overline{\ddot{\mathbf{p}}^2_{nn'}(t-R/c)} \approx \frac{4}{3c^3} \sum \omega^4_{nn'} M \overline{\mathbf{p}^2_{nn'}(t-R/c)}, \end{aligned} \tag{7.46}$$

by (7.44). Here χ denotes the angle between $\ddot{\mathbf{p}}_{nn'}(t-R/c)$ and $\mathbf{n}$, and the bar stands for the averaging in time. The averaging is justified, since we consider a stationary process. Finally, substituting (7.39), we obtain

$$MP_R \approx \frac{4}{3c^3} \sum \overline{\mathbf{p}^2_{nn'}} \omega^4_{nn'} = \frac{2}{3c^3} \sum \omega^2_{nn'} |\mathbf{J}_{nn'}|^2 M\left[|a_n|^2 |a_{n'}|^2\right], \qquad R \to \infty, \tag{7.47}$$

where $\overline{\mathbf{p}^2_{nn'}}$ is the time average of $\overline{\mathbf{p}^2_{nn'}(t)}$.

Remark 7.6

(i) The external *static* magnetic potential $\mathbf{A}_{\text{ext}}(\mathbf{x})$ does not contribute to electric field (7.41), and hence its contribution to the radiation vanishes when averaged over time.
(ii) The last formula means that the energy flow per unit time is positive, constant, and (asymptotically) is independent of the radius of the sphere, meaning the energy flow to infinity.
(iii) We suggest that this energy radiation to infinity is responsible for the global attraction (3.41) to quantum stationary states and for the convergence of the solution of (7.1) to the stationary random process.

Formula (7.47) gives the *total* mean power of radiation to infinity though its angular distribution is not necessarily isotropic. The energy radiation is isotropic if so is the random source field. Then formula (7.47) is the (constant) angular density of the mean power of radiation multiplied by 4π.

Furthermore, it is possible to experimentally separate the radiation components with each fixed wavelength $\lambda = 2\pi c/\omega$ and measure the corresponding angular density of the mean energy flow per unit time in each component. The density is given by the sum (7.47) over nn' with $|\omega_n - \omega_{n'}| = \omega$.

Exercise 7.7 Verify the last identity in (7.47).

Corollary 7.8 *The* **intensity of a spectral line** $\omega = \omega_n - \omega_{n'}$ *is given by*

(i) *By formula* (*cf.* [160, (36.22)]),

$$I_r(\omega_{nn'}) = \frac{8}{3c^3}\overline{\mathbf{p}^2_{nn'}}\omega^4_{nn'} = \frac{4}{3c^3}\omega^2_{nn'}|\mathbf{J}_{nn'}|^2 M\big[|a_n|^2|a_{n'}|^2\big] \tag{7.48}$$

for simple eigenvalues ω_n *and* $\omega_{n'}$.

(ii) *By the formula*

$$I_r(\omega) = \frac{2}{3c^3}\sum_{|\omega_n-\omega_{n'}|=\omega} \omega^2_{nn'}|\mathbf{J}_{nn'}|^2 M\big[|a_n|^2|a_{n'}|^2\big] \tag{7.49}$$

for multiple eigenvalues.

Remark 7.9

(i) Formula (7.48) contains the factor $M[|a_n|^2|a_{n'}|^2]$; this is in contrast to [160, (36.13)], where the factor is set to one.
(ii) Each term in (7.49) and multiplicities of the corresponding eigenvalues can be defined experimentally from splitting of the spectral lines in a weak magnetic field (see [171]).

7.2.6 The Correspondence Principle

As we pointed out above, the radiation field (7.40) is identical to the sum of the Hertzian dipole radiation fields of type (12.120) corresponding to *dipole moments* (7.39). Let us show that the double expression (7.39) coincides with the dipole moment of the corresponding atomic charge distribution.

Proposition 7.10 *For each* n, n',

$$2\mathbf{p}_{nn'}(t) = e\int \mathbf{x}\big|\psi_{nn'}(t,\mathbf{x})\big|^2\,d\mathbf{x}, \tag{7.50}$$

$$\psi_{nn'}(t,\mathbf{x}) := a_n(t)e^{-i\omega_n t}\psi_n(\mathbf{x}) + a_{n'}(t)e^{-i\omega_{n'}t}\psi_{n'}(\mathbf{x}). \tag{7.51}$$

Proof (i) First let us prove that

$$\mathbf{J}_{nn'} = ie\omega_{nn'}\langle\mathbf{x}\rangle_{nn'}, \quad \langle\mathbf{x}\rangle_{nn'} := \int \mathbf{x}\overline{\psi_n(\mathbf{x})}\psi_{n'}(\mathbf{x})\,d\mathbf{x}. \tag{7.52}$$

for each n, n'. Indeed, the commutation relations (3.92) imply that $[(\hat{\mathbf{p}} - \frac{e}{c}\mathbf{A}(t,\hat{\mathbf{x}}))^2, \hat{\mathbf{x}}] = -2i\hbar(\hat{\mathbf{p}} - \frac{e}{c}\mathbf{A}(t,\hat{\mathbf{x}}))$. Hence, $[H,\hat{\mathbf{x}}] = -\frac{i\hbar}{\mathrm{m}}(\hat{\mathbf{p}} - \frac{e}{c}\mathbf{A}(t,\hat{\mathbf{x}}))$, where H is the Schrödinger operator (7.4). Now definition (7.38) implies that

$$\begin{aligned}\mathbf{J}_{nn'} &= \frac{e}{\mathrm{m}}\Big\langle\Big(\hat{\mathbf{p}} - \frac{e}{c}\mathbf{A}(t, \hat{\mathbf{x}})\Big)\psi_{n'}, \psi_n\Big\rangle = \frac{e}{\mathrm{m}}\Big\langle\frac{i\mathrm{m}}{\hbar}[H, \hat{\mathbf{x}}]\psi_{n'}, \psi_n\Big\rangle \\ &= \frac{ie}{\hbar}\langle\hat{\mathbf{x}}\psi_{n'}, H\psi_n\rangle - \langle\hat{\mathbf{x}}H\psi_{n'}, \psi_n\rangle. \end{aligned} \tag{7.53}$$

Hence, (7.52) follows, because $H\psi_n = \hbar\omega_n\psi_n$ and $H\psi_{n'} = \hbar\omega_{n'}\psi_{n'}$, by (7.9).

(ii) Formula (7.52) allows one to rewrite (7.39) as

$$\mathbf{p}_{nn'}(t) = e\mathrm{Re}\big[\overline{a_n(t)}a_{n'}(t)e^{i\omega_{nn'}t}\langle\mathbf{x}_{nn'}\rangle\big]. \tag{7.54}$$

It remains to obtain the double expression for the integral in (7.50). Substituting $\psi_{nn'}(t, \mathbf{x})$ from (7.51), we obtain

$$e\int \mathbf{x}\big|\psi_{nn'}(t, \mathbf{x})\big|^2 d\mathbf{x} = 2e\mathrm{Re}\Big[\overline{a_n(t)}a_{n'}(t)e^{i\omega_{nn'}t}\int \mathbf{x}\overline{\psi_n(\mathbf{x})}\psi_{n'}(\mathbf{x})\, d\mathbf{x}\Big], \tag{7.55}$$

since

$$\int \mathbf{x}\big|\psi_n(t, \mathbf{x})\big|^2 d\mathbf{x} = \int \mathbf{x}\big|\psi_{n'}(t, \mathbf{x})\big|^2 d\mathbf{x} = 0. \tag{7.56}$$

The last identities hold for eigenfunctions (6.35). Moreover, these identities hold for any radial potential with an appropriate choice of eigenfunctions. Indeed, the corresponding Schrödinger operator commutes with the reflection $\mathbf{x} \mapsto -\mathbf{x}$, and so eigenfunctions can be chosen to be either even or odd. This proves (7.50). ☐

Exercise 7.11 Verify (7.56) for eigenfunctions (6.35). **Hint:** Use formulas (6.52) and (6.53).

Corollary 7.12

(i) *Using* (7.52), *formula* (7.48) *can be written as follows*:

$$I_r(\omega_{nn'}) = \frac{4}{3c^3}\omega_{nn'}^4|\mathbf{x}_{nn'}|^2 M\big[|a_n|^2|a_{n'}|^2\big] \tag{7.57}$$

(ii) *Using* (7.50), *formula* (7.57) *can be written similarly to* (2.40):

$$I_r(\omega_{nn'}) = \frac{2}{3c^3}\overline{(2\mathbf{p}_{nn'})^2}\omega_{nn'}^4; \tag{7.58}$$

this is the intensity of the classical Hertzian dipole radiation (12.127) *with the dipole moment* (7.50). *This identity confirms the Bohr correspondence principle used in identification* (2.40). *The factor* 2 *in* (7.50) *is due to the fact that sum* (7.40) *contains two identical terms* $\mathbf{p}_{n'n}(t) = \mathbf{p}_{nn'}(t)$, *by Remark* 7.5. *The correspondence principle applies*, *since condition* (7.34) *holds for large quantum numbers* n, n', *by* (6.3).

7.2.7 Selection Rules

A number of summands in (7.47) may vanish due to some symmetry conditions. Then, generally, the spectrum of atom radiation is a proper subset of the set of all differences $\{\omega_{nn'} := \omega_n - \omega_{n'}\}$ predicted by the Rydberg–Ritz Combination Principle.

For example, let us consider the case of the *radial* electrostatic potential $\phi_{\rm ext}(\mathbf{x}) = \phi_{\rm ext}(|\mathbf{x}|)$ and of the static *uniform* magnetic field $\mathbf{B} = (0, 0, B)$. Then $\mathbf{A}_{\rm ext}(\mathbf{x}) = \mathbf{B} \times \mathbf{x}/2$, and the eigenfunctions are given by (6.35) (see Sect. 9.1.1).

Let us consider a fixed spectral line $\omega_{kk'} = \omega_k - \omega_{k'}$, corresponding to some frequencies ω_k and $\omega_{k'}$ and eigenfunctions

$$\psi_k = R_{nl}(r) F_l^m(\theta) e^{im\varphi}, \qquad \psi_{k'} = R_{n'l'}(r) F_{l'}^{m'}(\theta) e^{im'\varphi}, \tag{7.59}$$

where $k := (l, m, n)$ and $k' := (l', m', n')$.

Lemma 7.13 *Assume that condition* (7.34) *hold. Let* $\phi_{\rm ext}(\mathbf{x}) = \phi_{\rm ext}(|\mathbf{x}|)$, *and* $\mathbf{A}_{\rm ext}(\mathbf{x}) = B \times \mathbf{x}/2$. *Then*

$$\mathbf{J}_{kk'} = 0 \quad \textit{if either } l' \neq l \pm 1 \textit{ or } m' \neq m,\ m \pm 1. \tag{7.60}$$

Proof It suffices to prove that

$$\int \mathbf{x}\overline{\psi_k(\mathbf{x})}\psi_{k'}(\mathbf{x})\, d\mathbf{x} = 0, \qquad t \in \mathbb{R} \tag{7.61}$$

if either $l' \neq l \pm 1$ or $m' \neq m, m \pm 1$. For $l' \neq l \pm 1$ the identity is proved [20, 191]. Here we check the case $m' \neq m, m \pm 1$. Let us rewrite the last integral in spherical coordinates. Then, by (7.59),

$$\int \mathbf{x}\overline{\psi_k(\mathbf{x})}\psi_{k'}(\mathbf{x})\, d\mathbf{x} = \int_0^\infty \overline{R_{nl}(r)} R_{n'l'}(r) r^3\, dr \\ \times \int_0^\pi \left[\int_0^{2\pi} \frac{\mathbf{x}}{r} \overline{F_l^m(\theta) e^{im\varphi}} F_{l'}^{m'}(\theta) e^{im'\varphi} d\varphi\right] \sin\theta\, d\theta. \tag{7.62}$$

Finally, the inner integral vanishes if $m' \neq m, m \pm 1$, because

$$\frac{\mathbf{x}}{r} = (\cos\theta\cos\varphi, \cos\theta\sin\varphi, \sin\theta).$$

□

Formula (7.48) implies that the intensity of the spectral line $\omega_{kk'}$ vanishes under conditions (7.60), so this line is not presented in the spectrum in the Born approximation.

7.3 Black Body Radiation

The Kirchhoff–Planck radiation law,

$$I(\omega)=\frac{\hbar\omega^3}{\pi^2c^3}\frac{e^{-\frac{\hbar\omega}{kT}}}{1-e^{-\frac{\hbar\omega}{kT}}}, \tag{7.63}$$

was a starting point in the genesis of quantum mechanics, leading to the discovery of the Heisenberg and Schrödinger dynamical equations and their coupling to the Maxwell field. However, afterwords one should justify the radiation law as an inherent property of the dynamical equations.

In 1916, Einstein [60] derived the Planck formula (1.24) without *explicit usage* of the discretization hypothesis (1.42), see Sect. 1.4.5.

Dirac implemented the Einstein ideas into the context of coupled Maxwell–Schrödinger equations [43]. The calculation involved relied on the Dirac nonstationary perturbation theory and the Boltzmann–Gibbs distribution (7.16). Namely, *slow evolution* of coefficients $a_n(t)$ from (7.10) is governed by Eqs. (7.13), which describe the corresponding energy exchange between different eigenmodes in (7.12). The exchange is provided by perturbation (7.5), which we will approximate by the first term,

$$H'(t)\approx\frac{e}{\mathrm{m}c}\mathbf{A}_*(t,\mathbf{x})\cdot\left[i\hbar\nabla+\frac{e}{c}\mathbf{A}_{\mathrm{ext}}(\mathbf{x})\right]; \tag{7.64}$$

this results in the relaxation to equilibrium distribution (7.63). The distribution is provisionally an approximation to the correlation function of a random stationary solution of the dynamical equations (7.13).

7.3.1 Correlations of the Radiation Field

According to (7.6), the 'radiation field' is defined by $\mathbf{A}_*(t,\mathbf{x})=\mathbf{A}(t,\mathbf{x})+\mathbf{A}_{\mathrm{in}}(t,\mathbf{x})$, where $\mathbf{A}_{\mathrm{in}}(t,\mathbf{x})$ is the 'incident wave' while $\mathbf{A}(t,\mathbf{x})$ is the 'outgoing wave'. The radiation field is a *stationary random process* which is a solution of homogeneous Maxwell equation with zero charge and current densities for $|\mathbf{x}|\gg r_1$, where r_1 is the atom size which is about $1\,\mathring{\mathrm{A}}=10^{-8}$ cm. Hence, the field admits the Fourier expansion

$$\mathbf{A}_*(t,\mathbf{x})\sim\sum_{\pm}\int e^{i(\mathbf{kx}/c\pm|\mathbf{k}|t)}\hat{\mathbf{A}}_*^{\pm}(\mathbf{k})\,d\mathbf{k},\qquad |\mathbf{x}|\gg r_1. \tag{7.65}$$

For the sake of concreteness, let us consider fields of the type

$$\mathbf{A}_*(t,\mathbf{x})\sim 2\int\cos\bigl(\mathbf{kx}/c-|\mathbf{k}|t\bigr)\hat{\mathbf{A}}(\mathbf{k})\,d\mathbf{k},\qquad |\mathbf{x}|\gg r_1, \tag{7.66}$$

with *translation invariant and isotropic correlations*

$$M\hat{\mathbf{A}}_i(\mathbf{k}) \otimes \overline{\hat{\mathbf{A}}_j\big(\mathbf{k}'\big)} = \delta_{ij}\, Q_*(\mathbf{k})\delta\big(\mathbf{k}-\mathbf{k}'\big), \quad i, j = 1, 2, 3. \tag{7.67}$$

Here the matrix $Q_*(\mathbf{k}) \geq 0$ is independent of time, since the radiation field is stationary.

7.3.2 *The Dipole Approximation*

We consider 'long wave' fields $\mathbf{A}_*(t, \mathbf{x})$ with the wave vectors $\mathbf{k}$ satisfying the condition

$$\frac{|\mathbf{k}| r_1}{c} \ll 2\pi, \tag{7.68}$$

which is similar to (7.34). The condition means that the wavelength $\lambda = 2\pi c/|\mathbf{k}|$ is much greater than r_1. For example, in all measurements of the spectral density up to the year 1900, the wavelength was $\lambda > 0.4\ \mu\text{m} = 0.4 \times 10^{-4}$ cm, while $r_1 \approx 1\ \text{Å} = 10^{-8}$ cm.

Under condition (7.68), equations (7.13) read as

$$i\hbar\dot{a}_n(t)e^{-i\omega_n t} = \frac{e}{\mathrm{m}c}\mathbf{A}_*(t, 0) \cdot \Big\langle \Big[i\hbar\nabla + \frac{e}{c}\mathbf{A}_{\text{ext}}(\mathbf{x}) \Big] \psi(t, \mathbf{x}), \psi_n(\mathbf{x}) \Big\rangle, \tag{7.69}$$

since the wave function $\psi(t, \mathbf{x})$ is concentrated near $\mathbf{x} = 0$. Substituting expansion (7.12) with $c_n(t) = a_n(t)e^{-i\omega_n t}$, we obtain the equations

$$\begin{aligned} i\hbar\dot{a}_n(t) &= \frac{e}{\mathrm{m}c}\mathbf{A}_*(t, 0) \cdot \sum_{n'} e^{-i\omega_{nn'}t} a_{n'}(t) \Big\langle \Big[i\hbar\nabla + \frac{e}{c}\mathbf{A}_{\text{ext}}(\mathbf{x}) \Big] \psi_{n'}(\mathbf{x}, t), \psi_n(\mathbf{x}) \Big\rangle \\ &= -\frac{ie}{c}\mathbf{A}_*(t, 0) \cdot \sum_{n'} \omega_{nn'} \langle \mathbf{x} \rangle_{nn'} e^{i\omega_{nn'}t} a_{n'}(t) \end{aligned} \tag{7.70}$$

in accordance with (7.38) and (7.52), where the sum should include the corresponding integral over the continuous spectrum. The system can be written in vector form as follows:

$$\dot{\mathbf{a}}(t) = K(t)\mathbf{a}(t), \quad \mathbf{a}(t) = \big\{a_n(t) : n = 1, 2, \ldots\big\}. \tag{7.71}$$

The operator $K(t)$ is skew-symmetric in l^2, since

$$K_{nn'}(t) = -\frac{e}{\hbar c}\omega_{nn'}\mathbf{A}_*(t, 0) \cdot \langle \mathbf{x} \rangle_{nn'} e^{i\omega_{nn'}t}, \tag{7.72}$$

where $\omega_{nn'} = -\omega_{n'n}$ and $\overline{\langle \mathbf{x} \rangle}_{nn'} = \langle \mathbf{x} \rangle_{n'n}$ by (7.52). This skewsymmetry reflects the charge conservation $\sum |a_n(t)|^2 = \text{const}$, where the sum includes the corresponding integral over the continuous spectrum.

7.3.3 Equilibrium Correlations and Spectral Density

To find spectral function (7.63), we calculate below the mean energy of the electron field

$$ME(t) := M\langle\psi(t), H\psi(t)\rangle = \sum_n \hbar\omega M\big|a_n(t)\big|^2, \tag{7.73}$$

where M stands for the mathematical expectation. Expression (7.73) involves the equilibrium correlation function (7.67) which we will express in the spectral density (1.27). First, the Fourier integral (7.66) can be written as follows:

$$\mathbf{A}_*(t,0) = 2\int_0^\infty \cos\omega t\tilde{\mathbf{A}}(\omega)\,d\omega, \quad \tilde{\mathbf{A}}(\omega) = \int_{|\mathbf{k}|=\omega} \hat{\mathbf{A}}(\mathbf{k})\,d\mathbf{k}, \tag{7.74}$$

where

$$M\tilde{\mathbf{A}}_i(\omega)\overline{\tilde{\mathbf{A}}_j(\omega')} = \delta_{ij}q_*(\omega)\delta(\omega-\omega'), \quad q_*(\omega) = \int_{|\mathbf{k}|=\omega} Q_*(\mathbf{k})\,d\mathbf{k}. \tag{7.75}$$

On the other hand, we can relate the correlations with the equilibrium spectral density $I(\omega)$ defined by

$$\frac{1}{8\pi}M\big[\tilde{\mathbf{E}}_\omega(t,\mathbf{x})\cdot\tilde{\mathbf{E}}_{\omega'}(t,\mathbf{x}) + \tilde{\mathbf{B}}_\omega(t,\mathbf{x})\cdot\tilde{\mathbf{B}}_{\omega'}(t,\mathbf{x})\big] = I(\omega)\delta(\omega-\omega') \tag{7.76}$$

(cf. (1.27)), where $\tilde{\mathbf{E}}_\omega(t,\mathbf{x})$ and $\tilde{\mathbf{B}}_\omega(t,\mathbf{x})$ are the Fourier components (1.28) of the electric and magnetic field, corresponding to the potential $\mathbf{A}_*$. The Maxwell fields are given by

$$\mathbf{B}(t,\mathbf{x}) = \operatorname{curl}\mathbf{A}_*(t,\mathbf{x}), \qquad \mathbf{E}(t,\mathbf{x}) = -\frac{1}{c}\dot{\mathbf{A}}_*(t,\mathbf{x}). \tag{7.77}$$

Hence, for the components of (1.28), we have

$$\begin{aligned}\tilde{\mathbf{E}}_\omega(t,\mathbf{x}) &= -\frac{2}{c}\int_{|\mathbf{k}|=\omega} |\mathbf{k}|\sin\big(\mathbf{kx}/c - |\mathbf{k}|t\big)\hat{\mathbf{A}}(\mathbf{k})\,d\mathbf{k}\\ \tilde{\mathbf{B}}_\omega(t,\mathbf{x}) = \overline{\tilde{\mathbf{B}}_\omega(t,\mathbf{x})} &= -\frac{2}{c}\int_{|\mathbf{k}|=\omega} \sin\big(\mathbf{kx}/c - |\mathbf{k}|t\big)\mathbf{k}\times\overline{\hat{\mathbf{A}}(\mathbf{k})}\,d\mathbf{k},\end{aligned}$$

where the bar now denotes the complex conjugation. Hence taking into account (7.76), (7.67) and (7.75), we obtain

$$I(\omega) = \frac{3\omega^2}{2\pi c^2}\int_{|\mathbf{k}|=\omega} Q_*(\mathbf{k})\,d\mathbf{k} = \frac{3\omega^2}{2\pi c^2}q_*(\omega), \tag{7.78}$$

because $\mathbf{k}\perp\mathbf{A}_c(\mathbf{k})$ and since the time average $\overline{\sin^2|\mathbf{k}|t} = 1/2$. The averaging is justified, since we are dealing with a stationary process. Finally, the equilibrium correlation functions are given by

$$q_*(\omega) = \frac{2\pi c^2}{3\omega^2} I(\omega). \tag{7.79}$$

7.3.4 Nonstationary Perturbation Theory

It remains to calculate $q_*(\omega)$. Then $I(\omega)$ is given by (7.78). We will obtain $q_*(\omega)$ calculating the correlation functions

$$q_{nn}(t) := M\big|a_n(t)\big|^2, \tag{7.80}$$

where $\{a_n(t) : n = 1, 2, \ldots\}$ is the solution to the system (7.71). The system can be solved by successive approximations

$$\dot{a}_n^{(N)}(t) = \sum_{n'} K_{nn'}(t) a_{n'}^{(N-1)}(t), \quad a_n^{(N)}(0) = a_n(0),\ N = 1, 2, \ldots \tag{7.81}$$

with $a_{n'}^{(0)}(t) \equiv a_{n'}^{(0)}(0)$. The first approximation reads

$$a_n^{(1)}(t) = a_n(0) + \int_0^t \sum_{n_1} K_{nn_1}(t_1) a_{n_1}(0)\, ds. \tag{7.82}$$

Similarly, the second approximation is as follows

$$\begin{aligned} a_n^{(2)}(t) &= a_n(0) + \int_0^t \sum_{n_1} K_{nn_1}(t_1) a_{n_1}^{(1)}(t_1)\, dt_1 \\ &= a_n(0) + \int_0^t \sum_{n_1} K_{nn_1}(t_1) \left[a_{n_1}(0) + \int_0^{t_1} \sum_{n_2} K_{n_1 n_2}(t_2) a_{n_2}(0)\, dt_2 \right] dt_1. \end{aligned} \tag{7.83}$$

In all these formulas the sum should include integral over continuous spectrum.

7.3.5 Correlations in First Approximation

Using (7.82) and the first orthogonality condition (7.18) at $t = 0$, we obtain the first approximation to the correlation functions (7.80):

$$
\begin{aligned}
q^{(1)}_{nn}(t) = M\Big\langle a_n(0) + \int_0^t \sum_{n_1'} K_{nn_1'}\big(t_1'\big)a_{n_1'}(0)\,dt_1', \\
a_n(0) + \int_0^t \sum_{n_1''} K_{nn_1''}\big(t_1''\big)a_{n_1''}(0)\,dt_1''\Big\rangle \\
= q_{nn}(0) + 2q_{nn}(0)M\mathrm{Re}\left\{\int_0^t K_{nn}\big(t'\big)\,dt'\right\} \\
+ \sum_{n_1} q_{n_1n_1}(0)M\left|\int_0^t K_{nn_1}(t_1)\,dt_1\right|^2, \qquad (7.84)
\end{aligned}
$$

where we assume that $a_n(0)$ are independent of the potential $\mathbf{A}_*(t,0)$ from expressions (7.72) for $K_{nn}(t)$. Further, the integral in curly brackets vanishes, since $K_{nn}(t') \equiv 0$. Hence, substituting formula (7.72) for matrix elements $K_{nn_1}(t)$, it is found that

$$
q^{(1)}_{nn}(t) = q_{nn}(0) + \frac{e^2}{(\hbar c)^2}\sum_{n_1}\omega^2_{nn_1}q_{n_1n_1}(0)M\left|\int_0^t \mathbf{A}_*(t_1,0)\cdot\langle\mathbf{x}\rangle_{nn_1}e^{i\omega_{nn_1}t_1}\,dt_1\right|^2. \qquad (7.85)
$$

Finally, substituting the Fourier integral representations (7.74) for Maxwell potentials, this gives

$$
\begin{aligned}
q^{(1)}_{nn}(t) &= q_{nn}(0) + \frac{4e^2}{(\hbar c)^2}\sum_{n_1}\omega^2_{nn_1}q_{n_1n_1}(0)\big|\langle\mathbf{x}\rangle_{nn_1}\big|^2 \\
&\quad\times\int_0^\infty q_*(\omega)\left|\int_0^t \cos(\omega t_1)e^{i\omega_{nn_1}t_1}\,dt_1\right|^2 d\omega \\
&= q_{nn}(0) + \frac{e^2}{(\hbar c)^2}\sum_{n_1}\omega^2_{nn_1}q_{n_1n_1}(0)\big|\langle\mathbf{x}\rangle_{nn_1}\big|^2 \\
&\quad\times\int_0^\infty q_*(\omega)t^2\big|D_{nn_1}(\omega,t)\big|^2\,d\omega, \qquad (7.86)
\end{aligned}
$$

by (7.75), where

$$
D_{nn_1}(\omega,t) := D\big((\omega_{nn_1}+\omega)t\big) + D\big((\omega_{nn_1}-\omega)t\big), \quad D(z) = \frac{e^{iz}-1}{iz}. \qquad (7.87)
$$

Now the key observation is that

$$
t\big|D_{nn_1}(\omega,t)\big|^2 \to 2\pi\big[\delta(\omega_{nn_1}+\omega)+\delta(\omega_{nn_1}-\omega)\big], \quad t\to\infty, \qquad (7.88)
$$

since

$$\int \left|D(z)\right|^2 dz = 4\int \frac{\sin^2\frac{z}{2}}{z^2}\,dz = 2\pi. \tag{7.89}$$

Hence, (7.86) can be written as follows:

$$q^{(1)}_{nn}(t) \approx q_{nn}(0) + \frac{2\pi e^2 t}{(\hbar c)^2}\sum_{n_1}\omega^2_{nn_1} q_{n_1n_1}(0)\left|\langle\mathbf{x}\rangle_{nn_1}\right|^2 q_*\big(|\omega_{nn_1}|\big), \qquad t \gg 1. \tag{7.90}$$

7.3.6 *The Dirac–Einstein Theory of Transitions*

In 1926, Dirac employed the nonstationary perturbation theory [43] to fitted the Einstein theory of radiation into the context of the Schrödinger theory.

We further consider the case of an atom with two quantum stationary states ψ_1 and ψ_2 and corresponding eigenvalues $\omega_1 < \omega_2$. Then the sum in (7.90) contains only one term, since $\mathbf{J}_{nn} = 0$ by (7.52). Hence, asymptotics (7.90) reads

$$\left.\begin{aligned} q^{(1)}_{11}(t) &\approx q_{11}(0) + t\frac{2\pi e^2}{(\hbar c)^2}\omega^2_{12}q_{22}(0)q_*\big(|\omega_{12}|\big)\left|\langle\mathbf{x}\rangle_{12}\right|^2, \\ q^{(1)}_{22}(t) &\approx q_{22}(0) + t\frac{2\pi e^2}{(\hbar c)^2}\omega^2_{12}q_{11}(0)q_*\big(|\omega_{12}|\big)\left|\langle\mathbf{x}\rangle_{12}\right|^2, \end{aligned}\right| \quad t \gg 1. \tag{7.91}$$

Spontaneous and Induced Transitions

The Einstein theory is based on Bohr's postulates on transitions between quantum stationary states ψ_1 and ψ_2 with energies $\hbar\omega_1 < \hbar\omega_2$ and the Einstein hypothesis on *photons*.

I. Spontaneous Emission Formula (7.57) for the intensity of spectral line ω_{21} is *treated* as follows: the coefficient

$$A_{2\to 1} := \frac{4e^2}{3c^3}\omega^4_{21}\left|\langle\mathbf{x}\rangle_{12}\right|^2/\hbar\omega_{21} = \frac{4e^2}{3c^3\hbar}\omega^3_{21}\left|\langle\mathbf{x}\rangle_{12}\right|^2 \tag{7.92}$$

is the *rate of spontaneous emission of photons* in transitions $\psi_2 \mapsto \psi_1$. Here $\hbar\omega_{21}$ is the photon energy according to the Einstein theory of photoelectric effect. 'Spontaneous' means that this emission process is independent of the radiation field $\mathbf{A}_*$.

II. Induced Emission The coefficient of t in the first formula of (7.91) is *treated* as the rate of transitions $\psi_2 \mapsto \psi_1$ *induced* by harmonic waves with frequency ω_{21}.

Then the atomic energy decreases by $\hbar\omega_2 - \hbar\omega_1 = \hbar\omega_{21}$ with every such a transition; *hence* the transition is *induced by the emission of the photon*. Thus, rate of the *induced emission* of photons is given by

$$B_{2\to 1} = \frac{2\pi e^2}{(\hbar c)^2}\omega_{12}^2 q_{22}(0) q_*\big(|\omega_{12}|\big)\big|\langle \mathbf{x}\rangle_{12}\big|^2. \tag{7.93}$$

III. Induced Absorption Similarly, the coefficient in the second formula (7.91) is treated as the rate of the *induced absorption* of photons

$$C_{1\to 2} = \frac{2\pi e^2}{(\hbar c)^2}\omega_{12}^2 q_{11}(0) q_*\big(|\omega_{12}|\big)\big|\langle \mathbf{x}\rangle_{12}\big|^2. \tag{7.94}$$

Statistical Equilibrium

The balance of the emitted and absorbed photons should hold,

$$p_2 A_{2\to 1} + B_{2\to 1} = C_{1\to 2}, \tag{7.95}$$

where p_2 is the probability of the state ψ_2, since $A_{2\to 1}$ is the 'conditional' rate of transitions from ψ_2. Substituting here the expressions for the coefficients, and setting $p_2 = q_{22}(0)$, we have, by (7.14) and (7.16),

$$\frac{4e^2}{3c^3\hbar}\omega_{21}^3\big|\langle \mathbf{x}\rangle_{12}\big|^2 q_{22}(0) = \frac{2\pi e^2}{(\hbar c)^2}\omega_{12}^2 q_*(\omega_{21})\big|\langle \mathbf{x}\rangle_{12}\big|^2\big[q_{11}(0) - q_{22}(0)\big]. \tag{7.96}$$

Substituting further $q_{nn}(0) = \frac{1}{Z}e^{-\frac{\hbar\omega_n}{kT}}$ from (7.16), we obtain

$$\frac{1}{Z}e^{-\frac{\hbar\omega_2}{kT}}\frac{4}{3c^3\hbar}\omega_{21} = \frac{2\pi}{(\hbar c)^2}q_*(\omega_{21})\left[\frac{1}{Z}e^{-\frac{\hbar\omega_1}{kT}} - \frac{1}{Z}e^{-\frac{\hbar\omega_2}{kT}}\right]. \tag{7.97}$$

Hence,

$$q_*(\omega_{21}) = \frac{2\hbar}{3\pi c}\omega_{21}\frac{e^{-\frac{\hbar\omega_2}{kT}}}{e^{-\frac{\hbar\omega_1}{kT}} - e^{-\frac{\hbar\omega_2}{kT}}}. \tag{7.98}$$

Finally, substituting expression (7.79), we arrive at the Planck formula (1.24)

$$I(\omega_{21}) = \frac{\hbar\omega_{21}^3}{\pi^2 c^3}\frac{1}{e^{\frac{\hbar\omega_{21}}{kT}} - 1}. \tag{7.99}$$

Remark 7.14 Formula (7.99) implies that the equilibrium spectral density does not depend on the substance which confirms the Kirchhoff radiation law.

Second Order Approximation

The stationarity should imply the identity

$$q_{11}^{(1)}(t) + q_{22}^{(1)}(t) = \text{const}, \tag{7.100}$$

which expresses the charge conservation. Unfortunately, the left hand side grows linearly like t by (7.91). So we should compensate this error, which is due to perturbation theory. The Einstein interpretation of the expressions (7.93) and (7.94) suggests introducing the ‘counterterms’, which take into account the ‘inverse transitions’: namely, the transitions 1 → 2 in the first equation, and 2 → 1, in the second one. Then we obtain the corrected asymptotic formulas

$$\begin{cases} q_{11}^{(c)}(t) \approx q_{11}(0) + t\dfrac{2\pi e^2}{(\hbar c)^2}\omega_{12}^2 q_*(\omega_{21})\big|\langle \mathbf{x}\rangle_{12}\big|^2\big[q_{22}(0) - q_{11}(0)\big], \\ q_{22}^{(c)}(t) \approx q_{22}(0) + t\dfrac{2\pi e^2}{(\hbar c)^2}\omega_{12}^2 q_*(\omega_{21})\big|\langle \mathbf{x}\rangle_{12}\big|^2\big[q_{11}(0) - q_{22}(0)\big], \end{cases} \tag{7.101}$$

which satisfy the charge conservation law (7.100).

Let us justify the corrected asymptotics (7.101) by the second approximation (7.83). Using (7.83) and the first orthogonality condition (7.18) at $t = 0$, we have, similarly to (7.84),

$$\begin{aligned} q_{nn}^{(2)}(t) = M\Big\langle & a_n(0) + \int_0^t \sum_{n_1'} K_{nn_1'}\big(t_1'\big)\Big[a_{n_1'}(0) + \int_0^{t_1'} \sum_{n_2'} K_{n_1'n_2'}\big(t_2'\big)a_{n_2'}(0)\,dt_2'\Big]\,dt_1', \\ & a_n(0) + \int_0^t \sum_{n_1''} K_{nn_1''}\big(t_1''\big)\Big[a_{n_1''}(0) + \int_0^{t_1''} \sum_{n_2''} K_{n_1''n_2''}\big(t_2''\big)a_{n_2''}(0)dt_2''\Big]\,dt_1''\Big\rangle \\ = q_{nn}(0) & + 2q_{nn}(0)\mathrm{Re}\Big\{M\int_0^t \sum_n K_{nn}(t_1)\,dt_1 \\ & + M\int_0^t \sum_{n_1} K_{nn_1}(t_1)\Big[\int_0^{t_1} K_{n_1 n}(t_2)\,dt_2\Big]\,dt_1\Big\} \\ & + \sum_{n_1} q_{n_1 n_1}(0)M\Big|\int_0^t K_{nn_1}\big(t_1'\big)\,dt_1'\Big|^2, \end{aligned} \tag{7.102}$$

up to terms of second order, where we assume that $a_n(0)$ are independent of the potential $\mathbf{A}(t,0)$ from expressions (7.72) for $K_{nn'}(t)$. Let us calculate the expression $B(t)$ in curly brackets: by substituting formula (7.72) for the matrix elements $K_{nn'}(t)$, we obtain, similarly to (7.85),

$$\begin{aligned}B(t) = -\frac{e^2}{(\hbar c)^2} M \sum_{n_1} \omega_{nn_1}^2 \int_0^t \mathbf{A}_*(t_1, 0) \cdot \langle \mathbf{x} \rangle_{nn_1} e^{i\omega_{nn_1} t_1} \\ \times \left[\int_0^{t_1} \mathbf{A}_*(t_2, 0) \cdot \langle \mathbf{x} \rangle_{n_1 n} e^{i\omega_{n_1 n} t_2}\, dt_2 \right] dt_1,\end{aligned} \tag{7.103}$$

because $K_{nn}(t_1) = 0$. Substituting the Fourier integral representations (7.74) for Maxwell potentials, we obtain, similarly to (7.86),

$$\begin{aligned}B(t) = -\frac{4e^2}{(\hbar c)^2} \sum_{n_1} \omega_{nn_1}^2 \left| \langle \mathbf{x} \rangle_{nn_1} \right|^2 \int_0^\infty q_*(\omega) \int_0^t \cos \omega t_1 e^{i\omega_{nn_1} t_1} \\ \times \left[\int_0^{t_1} \cos \omega t_2 e^{i\omega_{n_1 n} t_2}\, dt_2 \right] dt_1\, d\omega,\end{aligned} \tag{7.104}$$

by (7.75). Let us calculate the integral in t_1, t_2. Changing variables $r = t_1 - t_2$ and $s = t_1 + t_2$, this gives

$$\begin{aligned}I(t) &:= \int_0^t \cos \omega t_1 e^{i\omega_{nn_1} t_1} \left[\int_0^{t_1} \cos \omega t_2 e^{i\omega_{n_1 n} t_2}\, dt_2 \right] dt_1 \\ &= \frac{1}{4} \int_0^{2t} \left\{ \int_0^t [\cos \omega r + \cos \omega s] e^{i\omega_{nn_1} r}\, dr \right\} ds \\ &= \frac{1}{4} \left\{ 2t \left[\frac{e^{i(\omega + \omega_{nn_1})t} - 1}{2i(\omega + \omega_{nn_1})} + \frac{e^{i(\omega - \omega_{nn_1})t} - 1}{2i(\omega - \omega_{nn_1})} \right] + \frac{\sin 2\omega t}{\omega} \frac{e^{i\omega_{nn_1} t}}{i\omega_{nn_1}} \right\}.\end{aligned} \tag{7.105}$$

Now the key observation is that

$$\frac{e^{i\omega t} - 1}{i\omega} \to \pi \delta(\omega), \quad t \to \infty \tag{7.106}$$

since

$$\int \frac{\sin z}{z} = \pi. \tag{7.107}$$

Hence, the following asymptotics hold

$$I(t) \sim \frac{\pi t}{4} \left[\delta(\omega + \omega_{nn_1}) + \delta(\omega - \omega_{nn_1}) \right], \quad t \to \infty. \tag{7.108}$$

Therefore, (7.104) implies the asymptotics

$$B(t) \sim -\frac{\pi e^2 t}{(\hbar c)^2} \sum_{n_1} \omega_{nn_1}^2 \left| \langle \mathbf{x} \rangle_{nn_1} \right|^2 q_*\left(|\omega_{nn_1}| \right). \tag{7.109}$$

Hence, taking into account (7.90), it is found from (7.102) that

$$q_{nn}^{(2)}(t) \approx q_{nn}(0) + t\,\frac{2\pi e^2}{(\hbar c)^2}\sum_{n_1}\omega_{nn_1}^2\big|\langle\mathbf{x}\rangle_{nn_1}\big|^2 q_*\big(|\omega_{nn_1}|\big)\big[q_{n_1n_1}(0)| - q_{nn}(0)\big],$$

$$t \gg 1, \tag{7.110}$$

instead of (7.91). Asymptotics (7.110) satisfies the charge conservation (7.100), and coincide with the corrected system (7.101) in the case of a two-level atom.

Chapter 8
Scattering of Light and Particles

Abstract Scattering of light and electron beam by Hydrogen atom should be described by the coupled Maxwell–Schrödinger equations. However, the coupled equations are nonlinear, and so the calculations can be done only by perturbation procedure neglecting the self action; i.e., in the Born approximation. This approximation leads to some inconsistency breaking the charge conservation law, which should be fixed in a true nonlinear approach.

The corresponding scattering cross sections are similar to the classical ones given by the Thomson and Rutherford formulas respectively.

The calculation of the scattering of light relies on the limiting amplitude principle and the limiting absorption principle, which allow to explain the Einstein's rules for the photoelectric effect.

8.1 The Classical Scattering of Light

Scattering of light by matter is well known from everyday observations. It is well explained by interaction of the electromagnetic wave with atomic electrons. The scattering by nucleus is negligible, since it is relatively heavy. Incident electromagnetic wave is modeled as a plane wave satisfying the free Maxwell equations, and a matter, as classical, point-like particle obeying the Lorentz equation. Inconsistency inherent to the concept of point-like charged particles is circumvented by neglecting the self-interaction. Application of the Hertzian dipole radiation formula gives the Thomson differential cross-section.

8.1.1 The Incident Plane Wave

In 1861, Maxwell identified light with electromagnetic waves. For the sake of concreteness, we choose the incident *plane wave* in the form

$$\phi_0(t,\mathbf{x}) = 0, \qquad \mathbf{A}_0(t,\mathbf{x}) = A\Theta\big(ct - x^1\big)\sin k\big(x^1 - ct\big)\mathbf{e}_3, \tag{8.1}$$

where Θ is the Heaviside function, k is the wave number, and $\mathbf{e}_3 = (0,0,1)$ is the polarization. The incident wave is a solution of the Maxwell equations in the *free*

A. Komech, *Quantum Mechanics: Genesis and Achievements*,
DOI 10.1007/978-94-007-5542-0_8,

space with $\rho(t,\mathbf{x})=0$ and $\mathbf{j}(t,\mathbf{x})=0$:

$$\Box\phi_0(t,\mathbf{x})=0, \qquad \Box\mathbf{A}_0(t,\mathbf{x})=0, \quad (t,\mathbf{x})\in\mathbb{R}^4. \tag{8.2}$$

According to (12.44), the corresponding Maxwell fields are expressed through the potentials:

$$\mathbf{E}_0(t,\mathbf{x})=-\nabla\phi_0(t,\mathbf{x})-\frac{1}{c}\dot{\mathbf{A}}_0(t,\mathbf{x}), \qquad \mathbf{B}_0(t,\mathbf{x})=\operatorname{curl}\mathbf{A}_0(t,\mathbf{x}). \tag{8.3}$$

For potentials (8.1),

$$\begin{aligned}\mathbf{E}_0(t,\mathbf{x})&=kA\Theta\big(ct-x^1\big)\cos k\big(x^1-ct\big)\mathbf{e}_3,\\ \mathbf{B}_0(t,\mathbf{x})&=-kA\Theta\big(ct-x^1\big)\cos k\big(x^1-ct\big)\mathbf{e}_2.\end{aligned} \tag{8.4}$$

We will consider the scattering of wave (8.4) by a classical electron. The corresponding energy flux (i.e., the Poynting vector) is as follows:

$$\mathbf{S}_0(t,\mathbf{x})=\frac{c}{4\pi}\mathbf{E}_0(t,\mathbf{x})\times\mathbf{B}_0(t,\mathbf{x})=\frac{c\mathbf{E}_0^2}{4\pi}\Theta\big(ct-x^1\big)\cos^2 k\big(x^1-ct\big)\mathbf{e}_1; \tag{8.5}$$

here $\mathbf{E}_0:=kA\mathbf{e}_3$. The energy flux is directed along $\mathbf{e}_1$ and its *intensity* is given by

$$I_0:=\lim_{T\to\infty}\frac{1}{T}\left|\int_0^T\mathbf{S}_0(t,\mathbf{x})\,dt\right|=\frac{c\mathbf{E}_0^2}{8\pi}, \quad \mathbf{x}\in\mathbb{R}^3. \tag{8.6}$$

8.1.2 The Scattering Problem

The scattering is described by the Maxwell equations for the fields

$$\begin{cases}\operatorname{div}\mathbf{E}(t,\mathbf{x})=4\pi e\delta\big(\mathbf{x}-\mathbf{x}(t)\big), & \operatorname{curl}\mathbf{E}(t,\mathbf{x})=-\frac{1}{c}\dot{\mathbf{B}}(t,\mathbf{x}),\\ \operatorname{div}\mathbf{B}(t,\mathbf{x})=0, & \operatorname{curl}\mathbf{B}(t,\mathbf{x})=\frac{1}{c}\dot{\mathbf{E}}(t,\mathbf{x})+\frac{4\pi}{c}e\dot{\mathbf{x}}\delta\big(\mathbf{x}-\mathbf{x}(t)\big)\end{cases} \tag{8.7}$$

coupled to the Lorentz equation for the electron trajectory

$$\mathrm{m}\ddot{\mathbf{x}}(t)=e\left[\mathbf{E}\big(t,\mathbf{x}(t)\big)+\frac{1}{c}\dot{\mathbf{x}}(t)\times\mathbf{B}\big(t,\mathbf{x}(t)\big)\right], \quad t\in\mathbb{R}. \tag{8.8}$$

Initial condition for the electron does not matter. For example,

$$\mathbf{x}(t)=0, \qquad \dot{\mathbf{x}}(t)=0, \quad t<0. \tag{8.9}$$

Initial condition for the fields depends on the incident wave (8.1), since for $t < 0$

$$\begin{aligned} \mathbf{E}(t,\mathbf{x}) &= \Theta\big(ct - x^1\big)\mathbf{E}_0(t,\mathbf{x}) - e\frac{\mathbf{x}}{|\mathbf{x}|^2}, \\ \mathbf{B}(t,\mathbf{x}) &= \Theta\big(ct - x^1\big)\mathbf{B}_0(t,\mathbf{x}), \end{aligned} \tag{8.10}$$

where $-e\mathbf{x}/|\mathbf{x}|^2$ is the static Coulomb field generated by the electron with position (8.9). Functions (8.10), together with the electron trajectory (8.9), give solution of the nonlinear system of equations (8.7), (8.8) for $t < 0$, since the RHS of (8.8) is identically zero for $t < 0$.

8.1.3 Neglecting Self-Action

Let us split the solution of (8.7) like

$$\left\{\begin{aligned} \mathbf{E}(t,\mathbf{x}) &= \Theta\big(ct - x^1\big)\mathbf{E}_0(t,\mathbf{x}) + \mathbf{E}_r(t,\mathbf{x}) \\ \mathbf{B}(t,\mathbf{x}) &= \Theta\big(ct - x^1\big)\mathbf{B}_0(t,\mathbf{x}) + \mathbf{B}_r(t,\mathbf{x}) \end{aligned}\right| \qquad t \in \mathbb{R}, \tag{8.11}$$

where $\mathbf{E}_r(t,\mathbf{x}), \mathbf{B}_r(t,\mathbf{x})$ are the *radiated fields*. Then the Maxwell equations (8.7) read

$$\left\{\begin{aligned} &\operatorname{div}\mathbf{E}_r(t,\mathbf{x}) = 4\pi e\delta\big(\mathbf{x} - \mathbf{x}(t)\big), \qquad \operatorname{curl}\mathbf{E}_r(t,\mathbf{x}) = -\frac{1}{c}\dot{\mathbf{B}}_r(t,\mathbf{x}), \\ &\operatorname{div}\mathbf{B}_r(t,\mathbf{x}) = 0, \qquad \operatorname{curl}\mathbf{B}_r(t,\mathbf{x}) = \frac{1}{c}\dot{\mathbf{E}}_r(t,\mathbf{x}) + \frac{4\pi}{c}e\dot{\mathbf{x}}\delta\big(\mathbf{x} - x(t)\big) \end{aligned}\right. \tag{8.12}$$

because the incident wave (8.4) is a solution of the homogeneous Maxwell equations. From the initial conditions (8.10), it follows that

$$\mathbf{E}_r(t,\mathbf{x}) = -e\frac{\mathbf{x}}{|\mathbf{x}|^2}, \qquad \mathbf{B}_r(t,\mathbf{x}) = 0, \quad t < 0. \tag{8.13}$$

The Lorentz equation (8.8) now reads

$$\begin{aligned} \mathrm{m}\ddot{\mathbf{x}}(t) = e\Big[&\Theta\big(ct - \mathbf{x}^1(t)\big)\mathbf{E}_0\big(t,\mathbf{x}(t)\big) + \mathbf{E}_r\big(t,\mathbf{x}(t)\big) \\ &+ \frac{1}{c}\dot{\mathbf{x}}(t) \times \Big(\Theta\big(ct - \mathbf{x}^1(t)\big)\mathbf{B}_0\big(t,\mathbf{x}(t)\big) + \mathbf{B}_r\big(t,\mathbf{x}(t)\big)\Big)\Big], \quad t \in \mathbb{R}. \end{aligned} \tag{8.14}$$

Unfortunately, the problem (8.12), (8.14) is not well posed. Namely, it is clear from (8.12) and (8.13) that the solutions $\mathbf{E}_r(t,\mathbf{x}), \mathbf{B}_r(t,\mathbf{x})$ of (8.12) are infinite at $(t,\mathbf{x}(t))$. Therefore, the RHS of Eq. (8.14) is not well defined.

To make the problem well posed it is necessary to replace the point-like electron by the **extended electron**, as suggested by M. Abraham [2]. For this model, the well-posedness is proved in [111].

Here we employ another traditional approach to make the problem well posed; this approach is similar to the Born approximation. Namely, omitting the radiation fields at the RHS of (8.14), we write

$$\mathrm{m}\ddot{\mathbf{x}}(t)=e\Big[\Theta\big(ct-\mathbf{x}^1(t)\big)\mathbf{E}_0\big(t,\mathbf{x}(t)\big)+\frac{1}{c}\,\dot{\mathbf{x}}(t)\times\Theta\big(ct-\mathbf{x}^1(t)\big)\mathbf{B}_0\big(t,\mathbf{x}(t)\big)\Big],\quad t\in\mathbb{R}. \tag{8.15}$$

Then we substitute the solution $\mathbf{x}(t)$ obtained into the RHS of the Maxwell equations (8.12) to calculate the radiated fields $\mathbf{E}_r(t,\mathbf{x}), \mathbf{B}_r(t,\mathbf{x})$.

Let us assume that the electron velocities are small compared to the speed of light:

$$\beta:=\max_{t\in\mathbb{R}}\big|\dot{\mathbf{x}}(t)\big|/c\ll 1. \tag{8.16}$$

Then we can neglect the contribution of the magnetic field to the RHS of (8.15). Hence we obtain the equation

$$\mathrm{m}\ddot{\mathbf{x}}(t)=e\mathbf{E}_0\big(t,\mathbf{x}(t)\big)=kA\Theta\big(ct-x^1(t)\big)\cos k\big(x^1(t)-ct\big)\mathbf{e}_3,\quad t>0. \tag{8.17}$$

Consequently, $x^1(t)\equiv x^2(t)\equiv 0$, by the initial conditions (8.9), and so the equation becomes

$$\mathrm{m}\ddot{\mathbf{x}}(t)=kA\cos kct\;\mathbf{e}_3,\quad t>0. \tag{8.18}$$

Now the initial conditions (8.9) define the trajectory $\mathbf{x}(t)$ uniquely:

$$\mathbf{x}(t)=\frac{eA}{\mathrm{m}kc^2}(1-\cos kct)\mathbf{e}_3. \tag{8.19}$$

Note that condition (8.16) is equivalent to

$$\frac{|e|A}{\mathrm{m}c^2}\ll 1. \tag{8.20}$$

This relation means that amplitude of oscillations $\frac{|e|A}{\mathrm{m}kc^2}$ is small compared to wavelength $\lambda=2\pi/k$ of incident wave.

8.1.4 The Dipole Approximation: The Thomson Formula

We still should solve the Maxwell equations (8.12) to determine the radiation fields $\mathbf{E}_r(t,\mathbf{x}), \mathbf{B}_r(t,\mathbf{x})$. Our goal is to calculate the energy flux at infinity, i.e., to find the Poynting vector $\mathbf{S}_r(t,\mathbf{x}):=(c/4\pi)\mathbf{E}_r(t,\mathbf{x})\times\mathbf{B}_r(t,\mathbf{x})$ as $|x|\to\infty$. We will use the traditional *dipole approximation* for calculation of the radiation fields (Sect. 12.10); this leads to the well-known *Thomson formula*.

For this purpose, let us expand the charge density in the Maxwell equations (8.12) in formal Taylor series:

$$e\delta\big(\mathbf{x}-\mathbf{x}(t)\big)=e\delta(\mathbf{x})+e\mathbf{x}(t)\cdot\nabla\delta(\mathbf{x})+\frac{1}{2}e\big(\mathbf{x}(t)\cdot\nabla\big)^2\delta(\mathbf{x})+\cdots,\qquad t>0.\tag{8.21}$$

Here, the first term is static, and the corresponding radial Maxwell field does not contribute to the energy flux. The second term corresponds to the Hertzian dipole with the dipole moment $\mathbf{p}(t):=e\mathbf{x}(t)$. The subsequent terms give small contributions to energy radiation to infinity, because $|\mathbf{x}(t)|$ is small by (8.20) ($k=2\pi/\lambda\gg 1$ for visible light with $\lambda\sim 10^{-4}$ cm).

Hence we can use the Hertz formula (12.124) for the dipole radiation:

$$\mathbf{S}_r(t,\mathbf{x})\sim\mathbf{n}\frac{\sin^2\chi}{4\pi c^3|\mathbf{x}|^2}\ddot{\mathbf{p}}^2\big(t-|\mathbf{x}|/c\big),\qquad |\mathbf{x}|\to\infty,\tag{8.22}$$

where χ is the angle between $\ddot{\mathbf{p}}(t-|\mathbf{x}|/c)\sim\mathbf{e}^3$ and $\mathbf{n}:=\mathbf{x}/|\mathbf{x}|$. By (8.18),

$$\ddot{\mathbf{p}}(t)=e\ddot{\mathbf{x}}(t)=\frac{e^2}{\mathrm{m}}\mathbf{E}_0\cos kct,\qquad \mathbf{E}_0=kA\mathbf{e}_3.\tag{8.23}$$

Hence,

$$\ddot{\mathbf{p}}^2(t)=\left(\frac{e^2}{\mathrm{m}}\right)^2\mathbf{E}_0^2\cos^2 kct.\tag{8.24}$$

Let θ be the angle between $\mathbf{n}$ and $\mathbf{e}_1$ and let φ e the *azimuthal* angle between $\mathbf{e}_3$ and the plane $(\mathbf{n},\mathbf{e}_1)$. Then $\cos\chi=\cos\varphi\sin\theta$ and $\sin^2\chi=1-\cos^2\varphi\sin^2\theta$. Therefore,

$$\mathbf{S}_r(t,\mathbf{x})\sim\mathbf{n}\frac{1-\cos^2\varphi\sin^2\theta}{4\pi c^3|\mathbf{x}|^2}\left(\frac{e^2}{\mathrm{m}}\right)^2\mathbf{E}_0^2\cos^2 kc\big(t-|\mathbf{x}|/c\big),\qquad |\mathbf{x}|\to\infty.\tag{8.25}$$

Hence, the corresponding *mean in time intensity* is obtained replacing $\cos^2 kc(t-|\mathbf{x}|/c)$ by $1/2$:

$$\begin{aligned}I_r(\mathbf{x}):=\lim_{T\to\infty}\frac{1}{T}\left|\int_0^T\mathbf{S}_r(t,\mathbf{x})\,dt\right|&\approx\frac{1-\cos^2\varphi\sin^2\theta}{8\pi|\mathbf{x}|^2}\left(\frac{e^2}{\mathrm{m}c^2}\right)^2\mathbf{E}_0^2\\&=\left(\frac{e^2}{\mathrm{m}c^2}\right)^2\frac{\sin^2\chi}{|\mathbf{x}|^2}I_0,\end{aligned}\tag{8.26}$$

where I_0 is the intensity (8.6) of the incident wave. Therefore, the *mean intensity per unit angle* $\mathcal{I}_r=\lim_{|\mathbf{x}|\to\infty}I_r(\mathbf{x})|\mathbf{x}|^2$ is given by

$$\mathcal{I}_r(\varphi,\theta)\approx\left(\frac{e^2}{\mathrm{m}c^2}\right)^2\big(1-\cos^2\varphi\sin^2\theta\big)I_0.\tag{8.27}$$

Finally, the *differential cross-section* is as follows:

$$D(\varphi,\theta) := \frac{\mathcal{I}_r(\varphi,\theta)}{I_0} \approx \left(\frac{e^2}{\mathrm{m}c^2}\right)^2 \left(1-\cos^2\varphi\sin^2\theta\right); \tag{8.28}$$

this is the *Thomson formula*.

Depending on φ, the differential cross-section (8.28) is not invariant with respect to rotations around $\mathbf{e}_1$. This reflects the fact that an incident wave is linearly polarized. If we consider light with random polarization, then the differential cross-section is given by (8.28) with $1/2$ instead of $\cos^2\varphi$.

Finally, for the *total cross-section*, we have the following expression:

$$T := \int \mathcal{I}_r(\varphi,\theta)\,d\Omega/I_0 = \int D(\varphi,\theta)\,d\Omega \approx \left(\frac{e^2}{\mathrm{m}c^2}\right)^2 \int \sin^2\chi\,d\Omega = \left(\frac{e^2}{\mathrm{m}c^2}\right)^2 \frac{8\pi}{3}. \tag{8.29}$$

Remark 8.1 Now it is clear why we can neglect scattering of the light by the nucleus: mass of the nucleus is about 1836 electron mass (see (1.10)), so its classical scattering cross section (8.28), (8.29) should be about 1836^{-2} of the corresponding electron cross section.

8.2 Quantum Scattering of Light

Quantum scattering of light by Hydrogen atom is well described by interaction of the electromagnetic wave with the Schrödinger wave field, because the scattering by the nucleus is negligible, as explained above.

We will calculate the energy flux for the radiated Maxwell field, and find the corresponding differential cross-section in the *first order approximation* for small amplitude of the incident wave.

8.2.1 The Scattering Problem

We want to describe the scattering of the plane wave (8.4) by the Hydrogen atom in its ground state. By (6.3) and (6.37), the hydrogen ground state energy is $E_1 = -2\pi\hbar cR = -\mathrm{m}e^4/(2\hbar^2)$, and the corresponding wave function is $\psi_1(\mathbf{x}) = C_1 e^{-|\mathbf{x}|/r_1}$ (we assume that the atom is situated at the origin). Then the corresponding solution of the Schrödinger equation is as follows:

$$\psi_1(t,\mathbf{x}) = C_1 e^{-|\mathbf{x}|/r_1} e^{-i\omega_1 t}, \quad \omega_1 = \frac{E_1}{\hbar} = -\frac{\mathrm{m}e^4}{2\hbar^3}, \; r_1 = \frac{\hbar^2}{\mathrm{m}e^2}. \tag{8.30}$$

The scattering is described by the coupled Maxwell–Schrödinger equations in the Born approximation (4.27):

$$\big[i\hbar\partial_t - e\phi_n(\mathbf{x})\big]\psi(t,\mathbf{x}) = \frac{1}{2\mathrm{m}}\Big[-i\hbar\nabla - \frac{e}{c}\mathbf{A}_0(t,\mathbf{x})\Big]^2\psi(t,\mathbf{x}) \tag{8.31}$$

$$\begin{cases} \dfrac{1}{4\pi}\Box\phi(t,\mathbf{x}) = \rho(t,\mathbf{x}) = e\big|\psi(t,\mathbf{x})\big|^2, \\ \dfrac{1}{4\pi}\Box\mathbf{A}(t,\mathbf{x}) = \dfrac{\mathbf{j}(t,\mathbf{x})}{c} = \dfrac{e}{\mathrm{m}c}\Big[-i\hbar\nabla - \dfrac{e}{c}\mathbf{A}_0(t,\mathbf{x})\Big]\psi(t,\mathbf{x})\cdot\psi(t,\mathbf{x}), \end{cases} \tag{8.32}$$

where $\phi_n(\mathbf{x}) = -e/|\mathbf{x}|$ is the Coulomb potential of the nucleus, and $\mathbf{A}_0$ is the incident wave (8.1):

$$\mathbf{A}_0(t,\mathbf{x}) = A\Theta\big(ct - x^1\big)\sin k\big(x^1 - ct\big)\mathbf{e}_3. \tag{8.33}$$

In our model (8.31), (8.32), the hydrogen nucleus is considered as fixed. This corresponds to the fact that nucleus is heavy with respect to the electron. Respectively, the nucleus potential $\phi_n(\mathbf{x})$ is static, radial, and does not contribute to the scattered energy flux (cf. Remark 8.1).

8.2.2 The Atomic Form Factor

We consider the incident Maxwell wave with small amplitudes $|A| \ll 1$, and suppose that the atom was in its groundstate (cf. (8.9)):

$$\psi(t,\mathbf{x}) = \psi_1(t,\mathbf{x}), \quad t < 0. \tag{8.34}$$

In the *zero order approximation* (in A), the wave function is unperturbed,

$$\psi(t,\mathbf{x}) = \psi_1(t,\mathbf{x}), \quad t \in \mathbb{R}. \tag{8.35}$$

Hence, the corresponding approximation to the radiation field is given by solutions of the Maxwell equations (8.32) with $\psi(t,\mathbf{x}) = \psi_1(t,\mathbf{x})$. Then the charge density $\rho(t,\mathbf{x})$ is static and radial. Hence, the corresponding Maxwell field also is static and radial, and also does not contribute to the scattered energy flux.

Therefore, it suffices to solve the equation for $\mathbf{A}$ with the current

$$\frac{\mathbf{j}(t,\mathbf{x})}{c} = \frac{e}{\mathrm{m}c}\Big[-i\hbar\nabla\psi_1(t,\mathbf{x})\cdot\psi_1(t,\mathbf{x}) - \frac{e}{c}\mathbf{A}_0(t,\mathbf{x})\psi_1(t,\mathbf{x})\cdot\psi_1(t,\mathbf{x})\Big]. \tag{8.36}$$

Here, the first term on the RHS is zero, since the corresponding eigenfunction $e^{-|x|/r_1}$ is real:

$$i\hbar\nabla\psi_1(t,\mathbf{x})\cdot\psi_1(t,\mathbf{x})$$
$$:=\mathrm{Re}\big(i\hbar\nabla\psi_1(t,\mathbf{x})\overline{\psi}_1(t,\mathbf{x})\big)=|C_1|^2\mathrm{Re}\big(i\hbar\big(\nabla e^{-|x|/r_1}\big)e^{-|x|/r_1}\big)=0. \quad (8.37)$$

Therefore, the current reduces to

$$\frac{\mathbf{j}(t,\mathbf{x})}{c}=-\frac{e^2}{\mathrm{m}c^2}\mathbf{A}_0(t,\mathbf{x})\big|\psi_1(\mathbf{x})\big|^2. \quad (8.38)$$

Let us split the solution as

$$\mathbf{A}(t,\mathbf{x})=\mathbf{A}_0(t,\mathbf{x})+\mathbf{A}_r(t,\mathbf{x}), \quad (8.39)$$

where $\mathbf{A}_r(t,\mathbf{x})$ is the *radiated field* in the *first order approximation in A*. Then (8.32) becomes

$$\Box\mathbf{A}_r(t,\mathbf{x})=-4\pi\frac{e^2}{\mathrm{m}c^2}A\Theta\big(ct-x^1\big)\sin k\big(x^1-ct\big)\big|\psi_1(\mathbf{x})\big|^2\mathbf{e}_3$$
$$=-4\pi\frac{e^2}{\mathrm{m}c^2}\,\mathrm{Im}\,A\Theta\big(ct-x^1\big)e^{ik(x^1-ct)}\big|\psi_1(\mathbf{x})\big|^2\mathbf{e}_3=:f(t,\mathbf{x}), \quad (8.40)$$

because $\mathbf{A}_0$ is a solution of the homogeneous equation.

The radiation field can be characterized uniquely by an initial condition. In our setting, the radiation field is of finite energy; i.e.,

$$E_r(t):=\int_{\mathbb{R}^3}\big(\big|\dot{\mathbf{A}}_r(t,\mathbf{x})\big|^2+\big|\nabla\mathbf{A}_r(t,\mathbf{x})\big|^2\big)\,dx<\infty, \quad t\in\mathbb{R}. \quad (8.41)$$

Then the long time asymptotics of the radiation field does not depend on initial condition. Namely, in view of Theorem 12.23, condition (8.41) at $t=0$ implies that the radiated field is asymptotically given by the retarded potential,

$$\mathbf{A}_r(t,\mathbf{x})\sim\int\frac{f(t-|\mathbf{x}-\mathbf{y}|/c,\mathbf{y})\,d\mathbf{y}}{4\pi|\mathbf{x}-\mathbf{y}|}, \quad t\to\infty. \quad (8.42)$$

Hence, the *limiting amplitude principle* of type (13.113) holds

$$\mathbf{A}_r(t,\mathbf{x})\sim\mathrm{Im}\big[a_r(\mathbf{x})e^{-ikct}\big], \quad t\to\infty. \quad (8.43)$$

This follows from (8.42), because

$$\int\frac{f(t-|\mathbf{x}-\mathbf{y}|/c,\mathbf{y})\,d\mathbf{y}}{4\pi|\mathbf{x}-\mathbf{y}|}=-\frac{e^2}{\mathrm{m}c^2}A\,\mathrm{Im}\Big[\int\frac{e^{ik(\mathbf{y}^1-c(t-|\mathbf{x}-\mathbf{y}|/c))}|\psi_1(\mathbf{y})|^2\,d\mathbf{y}}{|\mathbf{x}-\mathbf{y}|}\Big]\mathbf{e}_3$$
$$=-\frac{e^2}{\mathrm{m}c^2}A\,\mathrm{Im}\Big[e^{-ikct}\int\frac{e^{ik(\mathbf{y}^1+|\mathbf{x}-\mathbf{y}|)}|\psi_1(\mathbf{y})|^2\,d\mathbf{y}}{|\mathbf{x}-\mathbf{y}|}\Big]\mathbf{e}_3 \quad (8.44)$$

for large t. Let us find asymptotics of the integral as $|\mathbf{x}| \to \infty$. For any fixed $\mathbf{y} \in \mathbb{R}^3$,

$$|\mathbf{x} - \mathbf{y}| = |\mathbf{x}| - \mathbf{y} \cdot \mathbf{n} + o(1), \quad |\mathbf{x}| \to \infty \tag{8.45}$$

where $\mathbf{n} = \mathbf{n}(x) = \mathbf{x}/|\mathbf{x}|$. Writing $\mathbf{y}^1 = \mathbf{y} \cdot \mathbf{e}_1$, we get

$$\mathbf{A}_r(t,\mathbf{x}) \sim -\frac{e^2}{\mathrm{m}c^2} A \,\mathrm{Im}\left[e^{-ikct} \frac{e^{ik|\mathbf{x}|}}{|\mathbf{x}|} \int e^{ik\mathbf{y}\cdot(\mathbf{e}^1 - \mathbf{n})} \big|\psi_1(\mathbf{y})\big|^2 \, d\mathbf{y} \right] \mathbf{e}_3, \quad |\mathbf{x}| \to \infty. \tag{8.46}$$

Next, we evaluate the last integral. We set $K := k|\mathbf{e}^1 - \mathbf{n}|$ and denote by θ the angle between $\mathbf{n}$ and $\mathbf{e}_1$. Then

$$\begin{aligned} K = K(k,\theta) &= k\sqrt{(1-n_1)^2 + n_2^2 + n_3^2} = k\sqrt{2\big(1-n_1\big)} \\ &= k\sqrt{2(1-\cos\theta)} = 2k \sin\frac{\theta}{2}. \end{aligned} \tag{8.47}$$

Let α be the angle between $\mathbf{y}$ and $\mathbf{e}_1 - \mathbf{n}$, and let φ be the azimuthal angle round $\mathbf{e}_1 - \mathbf{n}$. Finally, let us take into account that the ground state $\psi_1(\mathbf{y}) = \psi_1^r(|\mathbf{y}|)$ is spherically symmetric. Then integral (8.46) becomes

$$\begin{aligned} &\int_0^\infty |\mathbf{y}|^2 d|\mathbf{y}| \int_0^\pi \sin\alpha \, d\alpha \int_0^{2\pi} d\varphi \, e^{iK\cos\alpha|\mathbf{y}|} \big|\psi_1(\mathbf{y})\big|^2 \\ &= 4\pi \int_0^\infty \frac{\sin K|\mathbf{y}|}{K|\mathbf{y}|} \big|\psi_1^r\big(|\mathbf{y}|\big)\big|^2 |\mathbf{y}|^2 \, d|\mathbf{y}| =: F_a(k,\theta), \end{aligned} \tag{8.48}$$

which is called the *atomic form factor* corresponding to the ground state $\psi_1(\mathbf{x}) = C_1 e^{-|\mathbf{x}|/r_1}$. Since $F_a(k,\theta)$ is real, asymptotics (8.46) becomes

$$\mathbf{A}_r(t,\mathbf{x}) \sim -\frac{e^2}{\mathrm{m}c^2} A \frac{\sin k(|\mathbf{x}| - ct)}{|\mathbf{x}|} F_a(k,\theta) \mathbf{e}_3, \quad |\mathbf{x}| \to \infty. \tag{8.49}$$

Exercise 8.2 Calculate integral (8.48).

8.2.3 *The Energy Flux*

We still have to calculate the Maxwell field and the Poynting vector corresponding to last vector potential. It suffices to compare expression (8.49) with the vector potential of the Hertzian dipole (formula (12.120) of Sect. 12.10):

$$\mathbf{A}(t,\mathbf{x}) = \frac{1}{c} \frac{\dot{\mathbf{p}}(t - r/c)}{r}. \tag{8.50}$$

It is identical to (8.49) with $F_a(k,\theta) = 1$ if

$$\mathbf{p}(t) := -\frac{e^2}{\mathrm{m}c^2 k} A \cos kct \, \mathbf{e}_3. \tag{8.51}$$

Therefore, the energy flux $\mathbf{S}(t, \mathbf{x})$ corresponding to (8.49) is given, up to $\mathcal{O}(|x|^{-3})$, by the Hertz formula (8.22), with additional factor $|F_a(\theta)|^2$. This follows from the fact that angle θ is zero order homogeneous function of $\mathbf{x}$, hence any differentiation of the form factor $F_a(\theta)$ in $\mathbf{x}$ gives an additional factor with decay $\mathcal{O}(|\mathbf{x}|^{-1})$. Finally, for function (8.51), $\ddot{\mathbf{p}}(t)$ coincides with (8.23). In our case, (8.25) gives

$$\mathbf{S}(t, \mathbf{x}) \sim \mathbf{n}\big|F_a(\theta)\big|^2 \frac{1 - \cos^2\varphi \sin^2\theta}{4\pi c^3 |\mathbf{x}|^2} \left(\frac{e^2}{\mathrm{m}}\right)^2 E_0^2 \cos^2 kc\big(t - |\mathbf{x}|/c\big), \quad |\mathbf{x}| \to \infty. \tag{8.52}$$

Hence, the intensity per unit angle and the differential cross-section also contain the additional factor $|F_a(\theta)|^2$. Finally, the differential cross-section coincides with the Thomson formula (8.28) up to the atomic form factor:

$$D(k, \varphi, \theta) = \big|F_a(k, \theta)\big|^2 \left(\frac{e^2}{\mathrm{m}c^2}\right)^2 \big(1 - \cos^2\varphi \sin^2\theta\big). \tag{8.53}$$

8.3 Polarization and Dispersion

In the previous section, we were concerned with the scattering of light of small amplitude A by the Hydrogen atom, solving the Maxwell equations with the unperturbed ground state.

Now we analyze the change of the ground state in the first order in A, solving the Schrödinger equation. For small light frequencies, this change results in the corresponding modification of atomic charges and currents; this means *polarization* and *magnetization* of the atom. Thus we obtain the corresponding *permittivity* and *electric susceptibility*, and the *refraction coefficient*.

8.3.1 The First Order Approximation

To calculate the first order correction to the ground state, we expand the solution of (8.31) for small amplitudes:

$$\psi(t, \mathbf{x}) = \psi_1(t, \mathbf{x}) + Aw(t, \mathbf{x}) + \mathcal{O}\big(A^2\big), \quad |A| \ll 1; \tag{8.54}$$

here $\psi_1(t, \mathbf{x})$ is the ground state (8.30). Now condition (8.34) implies that

$$w(t, \mathbf{x}) = 0, \quad t < 0. \tag{8.55}$$

Substituting (8.54) into (8.31), we obtain, in the first order in A,

$$A\big[i\hbar\partial_t - e\phi_n(\mathbf{x})\big]w(t, \mathbf{x}) = A\frac{1}{2\mathrm{m}}[-i\hbar\nabla]^2 w(t, \mathbf{x}) + \frac{i\hbar e}{\mathrm{m}c}\mathbf{A}_0(t, \mathbf{x}) \cdot \nabla\psi_1(t, \mathbf{x}), \tag{8.56}$$

since $\psi_1(t,\mathbf{x})$ is a solution of Eq. (8.31) with $\mathbf{A}_0 = 0$. By (8.1) and (8.30), we have the following splitting for the source term in the RHS of (8.56),

$$\begin{aligned}\frac{i\hbar e}{\mathrm{m}c}\mathbf{A}_0(t,\mathbf{x})\cdot\nabla\psi_1(t,\mathbf{x}) &= \frac{i\hbar e}{\mathrm{m}c}A\sin k\big(x^1-ct\big)\mathbf{e}_3\cdot\nabla\psi_1(\mathbf{x})e^{-i\omega_1 t}\\ &= \frac{A\hbar e}{2\mathrm{m}c}\big[e^{ik(x^1-ct)}-e^{-ik(x^1-ct)}\big]e^{-i\omega_1 t}\nabla_3\psi_1(\mathbf{x})\\ &= \psi_+(\mathbf{x})e^{-i(\omega_1+\omega)t}-\psi_-(\mathbf{x})e^{-i(\omega_1-\omega)t}, \quad t > x^1-c,\end{aligned} \tag{8.57}$$

where $\omega := kc$. Now we apply the *limiting amplitude principle* (13.113):

$$w(t,\mathbf{x}) = w_+(\mathbf{x})e^{-i(\omega_1+\omega)t}-w_-(\mathbf{x})e^{-i(\omega_1-\omega)t}+\sum_l C_l\psi_l(\mathbf{x})e^{-i\omega_l t}+r(t,\mathbf{x}), \tag{8.58}$$

where $w_\pm(\mathbf{x})$ are the *limiting amplitudes*, and $r(t,\cdot)\to 0$ as $t\to\infty$, in an appropriate norm. Here $\psi_l(\mathbf{x})$ denote eigenfunctions of the discrete spectrum of the Schrödinger operator

$$H = \frac{1}{2\mathrm{m}}[-i\hbar\nabla]^2 + e\phi_n(\mathbf{x}). \tag{8.59}$$

The asymptotics (8.58) hold, provided $\omega_1 \pm \omega \neq \omega_l$ for all l. The sum over the discrete spectrum on the RHS of (8.58) vanishes by (8.55) and Lemma 13.30(iii) (see also Remark 13.31(i)). Hence, in the first order approximation,

$$w(t,\mathbf{x}) \sim w_+(\mathbf{x})e^{-i(\omega_1+\omega)t}-w_-(\mathbf{x})e^{-i(\omega_1-\omega)t}, \quad t\to+\infty. \tag{8.60}$$

For $w_\pm$, we get equations of type (13.108)

$$\left[\omega_1\pm\omega+\frac{e^2}{\hbar|\mathbf{x}|}\right]w_\pm(\mathbf{x})+\frac{\hbar}{2\mathrm{m}}\Delta w_\pm(\mathbf{x}) = \frac{\psi_\pm(\mathbf{x})}{\hbar A} = \frac{e}{2\mathrm{m}c}e^{\pm ikx^1}\nabla_3\psi_1(\mathbf{x}). \tag{8.61}$$

Here we consider the incident light of small frequencies:

$$|\omega| < |\omega_1|. \tag{8.62}$$

For the Hydrogen atom, we have $|\omega_1| = \frac{\mathrm{m}e^4}{2\hbar^3} \approx 20.5\times 10^{15}\ \mathrm{s}^{-1}$, by (8.30) with the data of (3.36). Hence, bound (8.62) holds for wave numbers $|k| < k_1 := |\omega_1|/c \approx 68\times 10^7\ \mathrm{m}^{-1}$ or wavelengths

$$\lambda > 2\pi/k_1 = 0.91176\times 10^{-5}\ \mathrm{cm} = 911.76\ \text{Å}. \tag{8.63}$$

The frequency bound implies that $\omega_1\pm\omega < 0$; hence the values $\hbar(\omega_1\pm\omega)$ do not belong to the continuous spectrum of the stationary Schrödinger equation (8.59). Therefore, Eq. (8.61) implies that

$$w_\pm \in L^2\big(\mathbb{R}^3\big). \tag{8.64}$$

This means a decay of the limiting amplitudes $w_\pm(\mathbf{x})$ at infinity. Actually, the amplitudes decay exponentially,

$$\left|w_\pm(\mathbf{x})\right| \le Ce^{-\varepsilon_\pm|\mathbf{x}|}, \tag{8.65}$$

where $\varepsilon_\pm > 0$. In fact, we can neglect the term with $\frac{e^2}{\hbar|\mathbf{x}|}$ in Eq. (8.61), since it is relatively small and decays at infinity. Then we obtain

$$[\Delta + z_\pm]w_\pm(\mathbf{x}) = f_\pm(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^3, \tag{8.66}$$

where $z_\pm = 2m(\omega_1 \pm \omega)/\hbar < 0$ and $|f_\pm(\mathbf{x})| \le Ce^{-\varepsilon|\mathbf{x}|}$ with $\varepsilon = -1/r_1 > 0$. Hence, $w_\pm = E_\pm * f_\pm$, where $E_\pm(\mathbf{x})$ is the fundamental solution $E_\pm(\mathbf{x}) = -e^{-\kappa_\pm|\mathbf{x}|}/(4\pi|\mathbf{x}|)$ with $\kappa_\pm := \sqrt{-z_\pm} > 0$:

$$w_\pm(\mathbf{x}) = -\int \frac{e^{-\kappa_\pm|\mathbf{x}-\mathbf{y}|}}{4\pi|\mathbf{x}-\mathbf{y}|} f_\pm(\mathbf{y})\,d\mathbf{y}. \tag{8.67}$$

As a result, decay (8.65) holds with $\varepsilon_\pm = \min(\varepsilon, \kappa_\pm) > 0$.

8.3.2 The Modified Ground State

Further, let us calculate the modified ground state using the spectral resolution of the Schrödinger operator in Eq. (8.61). First, expanding the RHS, we have

$$\frac{e}{2\mathrm{m}c} e^{\pm ikx^1} \nabla_3 \psi_1(\mathbf{x}) = \sum\nolimits_l a_l^\pm \psi_l(\mathbf{x}), \tag{8.68}$$

where $\sum_l$ includes the sum over the discrete spectrum and the integral over the continuous spectrum of the Schrödinger operator (8.59). Hence, the solutions $w_\pm$ are of the form

$$w_\pm(\mathbf{x}) = \sum\nolimits_l \frac{a_l^\pm \psi_l(\mathbf{x})}{\omega_1 \pm \omega - \omega_l}. \tag{8.69}$$

Therefore, (8.60) becomes

$$w(t,\mathbf{x}) \sim \sum\nolimits_l \frac{a_l^+ \psi_l(\mathbf{x})}{\omega_1 + \omega - \omega_l} e^{-i(\omega_1+\omega)t} - \sum\nolimits_l \frac{a_l^- \psi_l(\mathbf{x})}{\omega_1 - \omega - \omega_l} e^{-i(\omega_1-\omega)t}, \quad t \to \infty. \tag{8.70}$$

Let us calculate the coefficients $a_l^\pm$. Formally,

$$a_l^\pm = \frac{e}{2\mathrm{m}c} \int \overline{\psi}_l(\mathbf{x}) e^{\pm ikx^1} \nabla_3 \psi_1(\mathbf{x})\,d\mathbf{x} \tag{8.71}$$

if the eigenfunctions of the discrete spectrum are orthogonal and normalized, and the eigenfunctions of the continuous spectrum obey ‘delta function normalization’.

We assume that $kr_1 \ll 1$. Then we can substitute $e^{\pm ikx^1} = 1$. Hence, $a_l^\pm$ are approximately identical, and so

$$a_l^\pm \approx i\frac{1}{2c\hbar}\mathbf{J}_{l1}^3 \tag{8.72}$$

according to definition (7.38). Using formula (7.52), we get

$$a_l^\pm \approx a_l := -\frac{e\omega_{l1}}{2c\hbar}\int \mathbf{x}^3\overline{\psi}_l(\mathbf{x})\psi_1(\mathbf{x})\,d\mathbf{x} = \frac{e\omega_{1l}}{2c\hbar}\mathbf{x}_{l1}^3, \tag{8.73}$$

where $\omega_{1l} := \omega_1 - \omega_l$, and $\mathbf{x}_{l1}^3 := \int \mathbf{x}^3\psi_1(\mathbf{x})\overline{\psi}_l(\mathbf{x})\,d\mathbf{x}$. Finally, (8.70) becomes,

$$w(t,\mathbf{x}) \sim \sum\nolimits_l a_l\psi_l(\mathbf{x})\left(\frac{e^{-i\omega t}}{\omega_{1l}+\omega} - \frac{e^{i\omega t}}{\omega_{1l}-\omega}\right)e^{-i\omega_1 t}, \quad t \to \infty, \tag{8.74}$$

and so the expansion (8.54) for large times reads

$$\begin{aligned}\psi(t,\mathbf{x}) &= \left[\psi_1(\mathbf{x}) + A\sum\nolimits_l a_l\psi_l(\mathbf{x})\left(\frac{e^{-i\omega t}}{\omega_{1l}+\omega} - \frac{e^{i\omega t}}{\omega_{1l}-\omega}\right)\right]e^{-i\omega_1 t} + \mathcal{O}\big(A^2\big)\\ &= \big[\psi_1(\mathbf{x}) + A\Sigma(t,\mathbf{x})\big]e^{-i\omega_1 t} + \mathcal{O}\big(A^2\big).\end{aligned} \tag{8.75}$$

8.3.3 The Kramers–Kronig Formula

Let us calculate the corresponding electric dipole moment. First, the charge density is given by

$$\begin{aligned}\rho(t,\mathbf{x}) &= e\psi(t,\mathbf{x})\overline{\psi}(t,\mathbf{x}) = e\big(\psi_1(\mathbf{x}) + A\Sigma(t,\mathbf{x})\big)\big(\overline{\psi}_1(\mathbf{x}) + A\overline{\Sigma}(t,\mathbf{x})\big) + \mathcal{O}\big(A^2\big)\\ &= e\big|\psi_1(t,\mathbf{x})\big|^2 + eA\big[\Sigma^+e^{i\omega t} + \Sigma^-e^{-i\omega t}\big] + \mathcal{O}\big(A^2\big),\end{aligned} \tag{8.76}$$

where

$$\Sigma^+ = \sum\nolimits_l\left(\frac{\psi_1\overline{a}_l\overline{\psi}_l}{\omega_{1l}+\omega} - \frac{a_l\psi_l\overline{\psi}_1}{\omega_{1l}-\omega}\right), \qquad \Sigma^- = \overline{\Sigma^+}. \tag{8.77}$$

Therefore, the electric dipole moment is as follows:

$$\mathbf{p}(t) := \int \mathbf{x}\rho(t,\mathbf{x})\,d\mathbf{x} = \mathbf{p}_{11} + \mathbf{P}(t) + \mathcal{O}\big(A^2\big); \tag{8.78}$$

here $\mathbf{p}_{11} := \int \mathbf{x}e|\psi_1(t,\mathbf{x})|^2\,d\mathbf{x} = 0$ by spherical symmetry, and

$$\mathbf{P}(t) = eA\left[\sum\nolimits_l\left(\frac{\overline{a}_l\mathbf{x}_{l1}}{\omega_{1l}-\omega} - \frac{a_l\bar{\mathbf{x}}_{l1}}{\omega_{1l}+\omega}\right)e^{i\omega t} + \sum\nolimits_l\left(\frac{a_l\bar{\mathbf{x}}_{l1}}{\omega_{1l}-\omega} - \frac{\overline{a}_l\mathbf{x}_{l1}}{\omega_{1l}+\omega}\right)e^{-i\omega t}\right], \tag{8.79}$$

where $\mathbf{x}_{l1} := \int \mathbf{x}\psi_1\overline{\psi}_l\,d\mathbf{x}$. By symmetry arguments, the vector $\mathbf{P}(t)$ is directed along $\mathbf{e}_3$. Indeed, the invariance of $\mathbf{P}(t)$ with respect to the reflection $\mathbf{x}^2 \mapsto -\mathbf{x}^2$ is obvious. The invariance with respect to the reflection $\mathbf{x}^1 \mapsto -\mathbf{x}^1$ follows from (8.68)–(8.70), since we finally set $k = 0$. Therefore, substituting a_l from (8.73), and projecting $\mathbf{x}_{l1}$ onto $\mathbf{e}_3$, we get

$$\mathbf{P}(t) \approx A\mathbf{e}_3 \frac{2\omega e^2}{c\hbar} \sum\nolimits_l \frac{\omega_{1l}|\mathbf{x}_{l1}^3|^2}{\omega_{1l}^2 - \omega^2} \cos\omega t. \tag{8.80}$$

At last, averaging (8.78) with respect to all orientations of the atom, we obtain, in the first order approximation,

$$\bar{\mathbf{p}}(t) := A\mathbf{e}_3 \frac{2ke^2}{\hbar} \sum_l \frac{\omega_{1l}\overline{|\mathbf{x}_{l1}^3|^2}}{\omega_{1l}^2 - \omega^2} \cos\omega t, \tag{8.81}$$

since $\mathbf{p}_{11} = 0$ and $\omega = kc$.

Now we can express permittivity (12.203) of the hydrogen in its ground state ψ_1. Let $\mathbf{E}(t)$ be the electric field at position $\mathbf{x} = 0$ of the atom. By (8.4), we have $\mathbf{E}(t) = kA\cos\omega t\,\mathbf{e}_3$. Hence, (8.81) can be written as

$$\bar{\mathbf{p}}(t) := \frac{2e^2}{3\hbar} \sum\nolimits_l \frac{\omega_{1l}\overline{|\mathbf{x}_{l1}^3|^2}}{\omega_{1l}^2 - \omega^2} \mathbf{E}(t). \tag{8.82}$$

Therefore, the permittivity of atomic hydrogen in its ground state is given by the Kramers–Kronig formula,

$$\chi_e(\omega) = N\big|\bar{\mathbf{p}}(t)\big|/\big|\mathbf{E}(t)\big| = N\frac{2e^2}{3\hbar} \sum\nolimits_l \frac{\omega_{1l}\overline{|\mathbf{x}_{l1}^3|^2}}{\omega_{1l}^2 - \omega^2}, \tag{8.83}$$

where N is number of atoms per unit volume. It is worth noting that formula (8.83) has the same analytic structure as its analogue (14.37) established first in the framework of ‘Old Quantum Mechanics’ by Kramers and Heisenberg [121–123].

More precisely, the formula should be rewritten as the sum over the discrete spectrum and the integral over the continuous spectrum

$$\chi_e(\omega) = N\frac{2e^2}{3\hbar}\left[\sum_{l=1}^{L} \frac{\omega_{1l}\overline{|\mathbf{x}_{l1}^3|^2}}{\omega_{1l}^2 - (\omega + i0)^2} + \int_0^\infty \frac{\omega_{1\nu}\overline{|\mathbf{x}_{\nu 1}^3|^2}}{(\omega_1 - \nu)^2 - (\omega + i0)^2}\,d\nu\right], \tag{8.84}$$

where the supplement $i0$ is necessary by the limiting absorption principle (13.117).

This famous formula implies many important conclusions which agree with experimental observations (see details in [7]):

I. It allows to express the electric susceptibility $\varepsilon = 1 + 4\pi\chi_e$ (see (12.204)) and hence the refraction coefficient $n = \sqrt{\varepsilon\mu} \sim \sqrt{\varepsilon}$ by (12.206) in case of the magnetic susceptibility $\mu \sim 1$.

II. It implies the singularity of $\varepsilon(\omega)$ at the resonances ω_k, which is in a good agreement with experimental observations.
III. It explains the anomalous dispersion near the eigenfrequencies ω_k, i.e. weaker polarizability coefficient $\varepsilon(\omega)$ for higher frequencies $|\omega|$ near ω_k.
IV. The function (8.84) is holomorphic in the upper complex half plane $\operatorname{Im}\omega > 0$. This fact implies the integral *dispersion relations* between real and imaginary parts of $\chi_e(\omega)$ and $\varepsilon(\omega)$ discovered by Kramers and Kronig [124–126]. This theory was developed later in the framework of Quantum Field Theory, see the survey in [156, v. III].

8.4 Photoelectric Effect

Now we consider the incident light of high frequency in case when the condition (8.62) fails. In this case, the electron cloud of an atom is not modified, but rather it is completely destroyed, and the atom is *ionized*. The emitted electrons produce the *photocurrent* with specific angular distribution.

This ionization was observed first by Hertz in 1887 and studied experimentally by Lenard in 1902–1905. First theoretical explanation has been provided by Einstein in 1905, who suggested the corpuscular theory of light, by introducing the 'photons', which are particles of light. In the framework of the Schrödinger theory, this effect has been described first by Wentzel, who calculated the angular distribution of photocurrent. The calculation relies on the perturbation procedure applied to the coupled Maxwell–Schrödinger equations.

Hertz discovered the discharge of negatively charged *electroscope* under the electromagnetic radiation of very short wavelength (like visible or ultraviolet light). This discharge was treated as an emission of electrons from metals due to their absorption of energy from electromagnetic radiation.

Lenard systematically studied the behavior of 'photoelectrons', (i.e., emitted electrons) in external electric and magnetic fields. His conclusions were the following:

L1. The saturation photocurrent is proportional to the intensity of incident light.

L2. The photocurrent is not zero only for sufficiently small wavelength; i.e., for high frequencies:

$$|\omega| > \omega_{\text{red}}, \tag{8.85}$$

where ω_{red} is called the *red bound* of the photoelectric effect, which depends on the substance, but *is independent of the intensity of light*.

L3. The photocurrent vanishes if the stopping voltage U_{stop} is applied; the minimal U_{stop} also depends on the substance, but *is independent of the intensity of light*. Moreover, the minimal U_{stop} increases for decreasing wavelength of the incident light.

This independence of ω_{red} and minimal U_{stop} of the intensity of light was the main difficulty in theoretical explanation of the Lenard observations. This independence seemed to constitute a new mysterious phenomenon, which newer occurred in classical physics.

In 1905, Einstein proposed a revolutionary interpretation, by developing the discretization (1.42) for the energy of the Maxwell field oscillators with steps $\hbar\omega$. Namely, he suggested that the matter absorbs the light energy also by the discrete portions $\hbar\omega$. This corresponds to the treatment of light with frequency ω as a beam of particles, called 'photons', with energy $\hbar\omega$. *Einstein's rules* for the photoelectric effect are the following:

E1. The flux of photons is proportional to the intensity of the incident light.

E2. The maximum energy of photoelectrons is given by

$$\frac{\mathrm{m}v_{\max}^2}{2} = \hbar\omega - W, \tag{8.86}$$

where W is the *work function* of the substance. Hence, the emission of electron is possible only if $\hbar\omega - W > 0$; consequently, the redbound $\omega_{\text{red}} = W/\hbar$ is independent of the intensity of light—this agrees with Lenard's observations!

E3. Respectively, the stopping voltage should satisfy the inequality

$$-eU_{\text{stop}} > \hbar\omega - W, \tag{8.87}$$

where $e < 0$. Thus the minimal U_{stop} is also independent of the intensity of light.

Formula (8.86) formally represents the energy conservation in the absorption of photon by electron. However, let us stress that (8.86) is a theoretical interpretation of formula (8.87), which is verified experimentally and gives the minimum stopping voltage $-(\hbar\omega - W)/e$. Moreover, formula (8.87) allows to measure the Planck constant $\hbar$ with high precision.

Thus the 'Einstein rules' **E1**–**E3** give the complete explanation of Lenard's observations. In 1922, Einstein was awarded the 1921 Nobel Prize in Physics for his theory of the photoelectric effect, relying on the revolutionary *corpuscular theory of light*.

In 1927 G. Wentzel calculated angular distribution of the photocurrent applying the first order perturbation approach to the coupled Maxwell–Schrödinger equations. Wentzel's calculations [171, Vol. II] and Einstein's rules for the photoelectric effect are justified by the limiting amplitude principle.

Namely, the photoelectric effect is caused by slow decay of the limiting amplitude at infinity for $|\omega| > |\omega_1|$. The slow decay results in a nonvanishing current to infinity; this means the photoelectric effect. Thus, $\omega_{\text{red}} = |\omega_1|$. Moreover, the photoelectron energy is given by (8.86), and the stopping voltage satisfies (8.87).

Unfortunately, the perturbation approach is not selfconsistent and should be considered, rather, as a hint for explaining the atomic ionization. The corresponding rigorous theory of ionization was developed recently [35–38, 157]. However, this theory implies the atomic ionization for any light frequency $\omega \neq 0$.

For second quantized models a perturbation treatment of the atomic ionization and of relation (8.86) were given in [6, 76, 197].

This being so, a dynamical nonperturbation explanation of Einstein's rules for photoelectric effect remains an open challenging problem.

8.4.1 Radiation in Continuous Spectrum

Now we apply the perturbation approach (8.54) to the problem of scattering of light with large frequencies:

$$|\omega| > |\omega_1|, \tag{8.88}$$

in contrast to (8.62). Let us recall that for the Hydrogen atom this condition means that the light wavelenght $\lambda < 911.76$ Å, see (8.63).

For simplicity of notation, we assume that $\omega > 0$. Hence $\omega_1 - \omega < 0$, but

$$\omega_1 + \omega = \omega - |\omega_1| > 0 \tag{8.89}$$

by (8.88). Consequently, $\hbar(\omega_1 + \omega)$ belongs to the continuous spectrum of the stationary Schrödinger equation (8.59). Therefore, the solution $w_+(\mathbf{x}) \notin L^2$, as distinct from (8.64). This means a slow decay of the limiting amplitude:

$$\big|w_+(\mathbf{x})\big| \sim \frac{a(\mathbf{n}(\mathbf{x}))}{|\mathbf{x}|}, \quad |\mathbf{x}| \to \infty,$$

where $\mathbf{n}(\mathbf{x}) := \mathbf{x}/|\mathbf{x}|$. We will calculate the amplitude $a(\mathbf{n})$ and obtain the main term of the radiation in the form

$$Aw_+(\mathbf{x})e^{-i(\omega_1+\omega)t} \sim A\frac{a(\varphi,\theta)}{|\mathbf{x}|}e^{i[k_r|\mathbf{x}|-(\omega+\omega_1)t]}, \quad |\mathbf{x}| \to \infty. \tag{8.90}$$

On the other hand, $\hbar(\omega_1 - \omega) < 0$ does not belong to the continuous spectrum of the Schrödinger equation. Hence, $w_-(x)$ decays exponentially, similarly to (8.65):

$$\big|w_-(\mathbf{x})\big| \le Ce^{-\varepsilon_-|\mathbf{x}|}, \quad \mathbf{x} \in \mathbb{R}, \tag{8.91}$$

where $\varepsilon_- > 0$.

We will deduce from (8.90) and (8.91) the following asymptotics for the limiting stationary electric current at infinity,

$$\mathbf{j}(t,\mathbf{x}) \sim A^2\frac{e\hbar k_r}{\mathrm{m}}\frac{a^2(\varphi,\theta)}{|\mathbf{x}|^2}\mathbf{n}(\mathbf{x}), \quad |\mathbf{x}| \to \infty. \tag{8.92}$$

The formula was obtained by Wentzel in 1927 (see [190]) with amplitude

$$a(\varphi,\theta) = C\sin\theta\cos\varphi, \tag{8.93}$$

where $C \neq 0$. Hence, formula (8.92) describes a non-zero electric current from the atom to infinity. Indeed, asymptotics (8.92) imply that *total electric current to infinity* does not vanish, i.e.,

$$J_\infty := \lim_{R\to\infty} \int_{|\mathbf{x}|=R} \mathbf{j}(t,\mathbf{x})\, dS(\mathbf{x}) \neq 0. \tag{8.94}$$

8.4.2 The Limiting Amplitude

Let us calculate the limiting amplitude $w_+(\mathbf{x})$. First, we rewrite Eq. (8.61) as follows

$$\left[\nabla^2 + k_r^2(\omega)\right] w_+(\mathbf{x}) = \frac{e}{\hbar c} e^{ikx^1} \nabla_3 \psi_1(\mathbf{x}) - \frac{2e^2 \mathrm{m}}{\hbar^2 |\mathbf{x}|} w_+(\mathbf{x}), \tag{8.95}$$

where $k_r(\omega) := \sqrt{2\mathrm{m}(\omega_1 + \omega)/\hbar} > 0$. In the first approximation, we can neglect the last term on the RHS, because it is small and decays at infinity. Then we get the Helmholtz equation of type (8.66),

$$\left[\Delta + k_r^2(\omega)\right] w_+(\mathbf{x}) = f_+(\mathbf{x}) := \frac{e}{\hbar c} e^{ikx^1} \nabla_3 \psi_1(\mathbf{x}). \tag{8.96}$$

Hence the exponential decay (8.91) does not hold for $w_+(\mathbf{x})$. This is obvious in the Fourier space, where (8.96) becomes

$$\hat{w}_+(\mathbf{k}) = \frac{\hat{f}_+(\mathbf{k})}{-\mathbf{k}^2 + k_r^2(\omega)}. \tag{8.97}$$

The denominator vanishes on the sphere $|\mathbf{k}| = k_r(\omega)$, while $\hat{f}_+(\mathbf{k}) \sim \mathbf{k}_3 \hat{\psi}_1(\mathbf{k}_1 + k, \mathbf{k}_2, \mathbf{k}_3)$ is zero only for $\mathbf{k}_3 = 0$.

Exercise 8.3 Verify that $\hat{f}_+(\mathbf{k}) \neq 0$ for $\mathbf{k}_3 \neq 0$. **Hint:** Calculate $\hat{\psi}_1(\mathbf{k})$ in the spherical coordinates.

Hence $w_+(\mathbf{x})$ cannot decay exponentially. Now the solution is given by convolution with fundamental solution (13.111)

$$w_+(\mathbf{x}) = -\int \frac{e^{ik_r(\omega)|\mathbf{x}-\mathbf{y}|}}{4\pi|\mathbf{x}-\mathbf{y}|} f_+(\mathbf{y})\, d\mathbf{y}. \tag{8.98}$$

This follows from the limiting absorption principle (13.117), because the fundamental solution $\frac{e^{ik_r(\omega+i\varepsilon)|\mathbf{x}|}}{4\pi|\mathbf{x}|}$ is a tempered distribution for small $\varepsilon > 0$, since $\operatorname{Im} k_r(\omega + i\varepsilon) > 0$ for the fixed branch $k_r(\omega) > 0$.

Now we can calculate asymptotics (8.90). To do so we substitute expression (8.96) for f_+ into (8.98). By partial integration, we obtain

$$
\begin{aligned}
w_+(\mathbf{x}) &= -\frac{e}{\hbar c}\int \nabla_{\mathbf{y}_3}\frac{e^{ik_r|\mathbf{x}-\mathbf{y}|}}{4\pi|\mathbf{x}-\mathbf{y}|}e^{ik\mathbf{y}^1}\psi_1(\mathbf{y})\,d\mathbf{y}\\
&= \frac{ik_r e}{\hbar c}\int \frac{e^{ik_r|\mathbf{x}-\mathbf{y}|}(\mathbf{x}^3-\mathbf{y}^3)}{4\pi|\mathbf{x}-\mathbf{y}|^2}e^{ik\mathbf{y}^1}\psi_1(\mathbf{y})\,d\mathbf{y}+\mathcal{O}\big(|\mathbf{x}-\mathbf{y}|^{-2}\big).
\end{aligned}
\tag{8.99}
$$

Recall that θ denotes the angle between $\mathbf{n} := \mathbf{x}/|\mathbf{x}|$ and $\mathbf{e}_1$, and φ stands for the *azimuthal* angle between $\mathbf{e}_3$ and the plane $(\mathbf{n}, \mathbf{e}_1)$. Then $\mathbf{x}^3 = \sin\theta\cos\varphi|\mathbf{x}|$. Hence, (8.99) implies (8.90) with angular distribution (8.93), because the ground state $\psi_1(\mathbf{y})$ decays rapidly at infinity. The constant C in (8.93) is given by

$$
C = C(k) = \frac{ik_r e}{4\pi\hbar c}\int e^{ik\mathbf{y}^1}\psi_1(\mathbf{y})\,d\mathbf{y}. \tag{8.100}
$$

It does not vanish for the ground state (8.30).

8.4.3 Angular Distribution: The Wentzel Formula

Our aim here is to derive (8.92). Equations (8.32) imply that, in the first approximation, the current is given by

$$
\mathbf{j} := -\frac{e}{\mathrm{m}}\big[i\hbar\nabla\psi(t,\mathbf{x})\big]\cdot\psi(t,\mathbf{x}). \tag{8.101}
$$

Further, $w_-(\mathbf{x})$ decays exponentially at infinity by (8.91), as well as the eigenfunction $\psi_1(\mathbf{x})$. Therefore, (8.54) and (8.60) imply the asymptotics

$$
\psi(t,\mathbf{x}) \sim A w_+(\mathbf{x})e^{-i(\omega_1+\omega)t}, \quad |\mathbf{x}| \to \infty. \tag{8.102}
$$

Substituting into (8.101), and using asymptotics (8.90), we obtain the Wentzel formula (8.92) with amplitude (8.93).

8.4.4 Derivation of Einstein's Rules

Now we can explain Lenard's observations and Einstein's rules for the photoelectric effect:

E1. By (8.92), the saturation photocurrent is proportional to A^2, which in turn is proportional to the intensity of incident light, by (14.31) and (14.26), (14.27).

E2. Asymptotics (8.90) imply that the energy per one photoelectron is given by the *Einstein formula* (8.86). Indeed, for large $|\mathbf{x}|$, the radiated wave (8.90) is locally close to the plane wave (5.1) with $\omega - |\omega_1|$ instead of ω. Hence, formula (5.6) implies that the energy per one photoelectron is given by $E = \hbar(\omega - |\omega_1|)$, which is equivalent to (8.86) with

$$W = \hbar\omega_1. \tag{8.103}$$

E3. Application of stopping voltage is equivalent to the corresponding modification of the scalar potential in the Schrödinger equation (8.31): $\phi_n(\mathbf{x}) \mapsto \tilde{\phi}(\mathbf{x}) = \phi_n(\mathbf{x}) + \phi_{\text{stop}}(\mathbf{x})$, where $\phi_{\text{stop}}(\mathbf{x})$ is a slowly varying potential, and $\phi_{\text{stop}}(\mathbf{x}) = U_{\text{stop}} > 0$ in a macroscopic region containing the atom. Therefore, the ground state energy $\hbar\omega_1$ changes to $\hbar\tilde{\omega}_1$, and $\hbar\tilde{\omega}_1 \approx \hbar\omega_1 + eU_{\text{stop}}$ with high precision. Indeed, by the Courant minimax principle,

$$\hbar\omega_1 = \min_{\|\psi\|=1} (\psi, H\psi). \tag{8.104}$$

We can assume that $0 \le \phi_{\text{stop}}(\mathbf{x}) \le U_{\text{stop}}$ for $\mathbf{x} \in \mathbb{R}^3$. Then

$$\hbar\tilde{\omega}_1 = \min_{\|\psi\|=1} \big(\psi, \big[H + e\phi_{\text{stop}}(\mathbf{x})\big]\psi\big) \ge \hbar\omega_1 + eU_{\text{stop}} \tag{8.105}$$

since $e < 0$. On the other hand, the unperturbed ground state $\psi_1(\mathbf{x})$ is localized in a very small region of the size about $1\,\text{Å} = 10^{-8}$ cm, where $e\phi_{\text{stop}}(\mathbf{x}) = U_{\text{stop}}$. Hence,

$$\big(\psi_1, \big[H + e\phi_{\text{stop}}(\mathbf{x})\big]\psi_1\big) \approx \hbar\omega_1 + eU_{\text{stop}}. \tag{8.106}$$

Therefore,

$$\hbar\tilde{\omega}_1 \approx \hbar\omega_1 + eU_{\text{stop}}. \tag{8.107}$$

For the eigenstates with highest numbers, the localization gets progressively worse, and the eigenfunctions of the continuous spectrum are not localized at all. Respectively, the shift of the highest eigenvalues is smaller and smaller, and the continuous spectrum of the modified Schrödinger operator remains unchanged.

Finally, the potential prevents the photoelectric effect if the spectral condition (8.88) fails for the modified ground state; i.e., $0 < \omega < |\tilde{\omega}_1|$ or

$$\hbar\omega < |\hbar\omega_1 + eU_{\text{stop}}| = \hbar|\omega_1| - eU_{\text{stop}}, \tag{8.108}$$

since $\omega_1 < 0$, while $e < 0$ and we define $U_{\text{stop}} > 0$. In other words,

$$-eU_{\text{stop}} > \hbar\omega - \hbar|\omega_1|, \tag{8.109}$$

which is equivalent to (8.87) by (8.103).

8.4.5 Further Improvements

The Wentzel calculation takes into account the interaction of the Maxwell and Schrödinger fields in the first order of approximation. The second order correction was obtained by Sommerfeld and Shur [174]. The corresponding corrected formula reads as follows (see [171, Vol. II]):

$$\mathbf{j}(t,\mathbf{x}) \sim \frac{\sin^2\theta\cos^2\varphi(1+4\beta\cos\theta)}{|\mathbf{x}|^2}\mathbf{n}(\mathbf{x}), \quad |\mathbf{x}| \to \infty. \tag{8.110}$$

Here $\beta = \frac{v}{c}$, where v is the velocity of photoelectrons. The formula means an *increment* of the scattering amplitude for angles $0 < \theta < \frac{\pi}{2}$ and a *decrement* of the scattering amplitude for angles $\frac{\pi}{2} < \theta < \pi$. This means a *forwardshift of scattering* due to *pressure of the incident light upon the outgoing photocurrent*, as predicted by Wentzel [190].

Fisher and Sauter obtained the formula which holds in each order (see [171, Vol. II]):

$$\mathbf{j}(t,\mathbf{x}) \sim \frac{\sin^2\theta\cos^2\varphi}{(1-\beta\cos\theta)^4|\mathbf{x}|^2}\mathbf{n}(\mathbf{x}), \quad |\mathbf{x}| \to \infty. \tag{8.111}$$

8.4.6 Atomic Ionization and Photoelectric Effect

Unfortunately, the perturbation theory of the photoelectric effect is not selfconsistent in contrast to the case (8.62). The difference is caused by the fact that for small frequencies ω the limit stationary regime is close to the ground state, while for (8.88) the limit regime is quite different.

For instance, the stationary nonvanishing photocurrent (8.94) contradicts the charge conservation law, because the atomic charge is finite. The contradiction is provided by the perturbation strategy, which leaves unchanged ψ_1 on the right of (8.56), while it should be substituted by the solution of (8.31), (8.32). This 'self-action' should result in a decay of the photocurrent until the negative atomic charge will be exhausted, i.e.,

$$\int_{|\mathbf{x}|<R} \left|\psi(t,\mathbf{x})\right|^2 d\mathbf{x} \to 0, \quad t \to \infty \tag{8.112}$$

for all $R > 0$. Thus, formula (8.94) does not justify the photoelectric effect, but it rather suggests the *atomic ionization* (8.112), as established in [35–38].

On the other hand, a self-consistent justification of the photoelectric effect should rely on the stationary photocurrent, since the stopping voltage is concerned exactly the stationary picture. To maintain the stationary photocurrent, one needs either an external source (galvanic element, etc.) to reimburse the charge decay, or a different model with infinite charge (e.g., crystal).

Finally, the ionization occurs at any nonzero light frequency ω, according to the results of [35–38]. Thus, a satisfactory nonperturbation explanation of the Einstein rules of the photoelectric effect still remains an open problem.

8.5 The Classical Scattering of Charged Particles

In 1911, Rutherford experimentally studied the scattering of beams of *alpha particles* using a very thin gold foil. The results proved to be very surprising, testifying a concentration of positive charge in a point nucleus. Rutherford calculated the corresponding *differential cross section*, which was found to be in agreement with the experimental data.

8.5.1 The Rutherford Scattering

Homogeneous beam of classical particles with charge Q and mass M falls onto a heavy nucleus with charge $|e|Z > 0$. The Rutherford formula holds both in the repulsive case, when $Q > 0$, and in the attractive case, when $Q < 0$. The repulsive case corresponds, for example, to the scattering of α-particles with $Q = 2|e|$, and the attractive case corresponds, for example, to the scattering of electrons with $Q = e < 0$.

The mass of the nucleus being much larger than M, we may assume that the nucleus is fixed at the origin. For an incident particle moving along the trajectory $\mathbf{x} = \mathbf{x}(t)$, $-\infty < t < \infty$, we assume that particle comes from infinity with nonzero velocity; i.e.,

$$\lim_{t\to-\infty} \left|\mathbf{x}(t)\right| = \infty, \qquad \lim_{t\to-\infty} \left|\dot{\mathbf{x}}(t)\right| = v > 0. \tag{8.113}$$

This trajectory is known to be a hyperbola (see Sect. 14.1). Let us choose coordinates in space in such a way that $\mathbf{x}^3(t) \equiv 0$, and

$$\lim_{t\to-\infty} \dot{\mathbf{x}}(t) = (v, 0, 0), \qquad \lim_{t\to-\infty} \mathbf{x}^1(t) = -\infty, \qquad \lim_{t\to-\infty} \mathbf{x}^2(t) = b, \tag{8.114}$$

where b is the *impact* parameter. Let us further use the conventional polar coordinates in the plane $\mathbf{x}^3 = 0$,

$$\mathbf{x}^1 = r\cos\theta, \qquad \mathbf{x}^2 = r\sin\theta \tag{8.115}$$

and denote by $r(t)$, $\theta(t)$ the trajectory of a particle in these coordinates. From the initial scattering conditions (8.114) we have

$$\lim_{t\to-\infty} \theta(t) = \pi, \qquad \lim_{t\to-\infty} r(t)\sin\theta(t) = b. \tag{8.116}$$

8.5.2 The Angle of Scattering

Let us calculate the final scattering angle

$$\overline{\theta} := \lim_{t\to\infty} \theta(t). \tag{8.117}$$

Lemma 8.4 *The final scattering angle satisfies the equation*

$$\cot\frac{\overline{\theta}}{2} = \frac{Mbv^2}{Q|e|Z}. \tag{8.118}$$

Proof First, according to the law of conservation of the angular momentum and energy,

$$r^2(t)\dot{\theta}(t) = bv, \qquad \frac{M}{2}\big(\dot{r}^2(t) + r^2(t)\dot{\theta}^2(t)\big) + \frac{Q|e|Z}{r(t)} = \frac{M}{2}v^2. \tag{8.119}$$

If we substitute here

$$\dot{r}(t) := \frac{dr}{dt} = \frac{dr}{d\theta}\frac{d\theta}{dt} = r'\dot{\theta}, \tag{8.120}$$

then the energy conservation becomes

$$\frac{M}{2}\dot{\theta}^2(t)\big(\big|r'\big(\theta(t)\big)\big|^2 + r^2(t)\big) + \frac{Q|e|Z}{r(t)} = \frac{M}{2}v^2. \tag{8.121}$$

Now let us introduce the **Clairaut substitution** $r = 1/\rho$. Then $r' = -\rho'/\rho^2$ and momentum conservation gives $\dot{\theta}(t) = vb\rho^2$. Therefore, (8.121) reads

$$\frac{M}{2}b^2v^2\big(\big|\rho'\big|^2 + \rho^2\big) + Q|e|Z\rho = \frac{M}{2}v^2. \tag{8.122}$$

Let us differentiate this expression in θ. Dividing by ρ', we arrive at the **Clairaut equation**

$$\rho'' + \rho = C := -\frac{Q|e|Z}{Mbv^2}. \tag{8.123}$$

The general solution of this equation is as follows:

$$\rho(\theta) = A\cos\theta + B\sin\theta + C. \tag{8.124}$$

Now the initial scattering conditions (8.116) give,

$$\lim_{\theta\to\pi} \rho(\theta) = 0, \qquad \lim_{\theta\to\pi} \frac{\rho(\theta)}{\sin\theta} = \frac{1}{b}. \tag{8.125}$$

If we substitute here (8.124), we obtain $-A + C = 0$ and $B = 1/b$, and hence

$$\rho(\theta) = C(1 + \cos\theta) + \frac{1}{b}\sin\theta. \tag{8.126}$$

At last, for the final scattering angle, it follows from (8.117) that $\rho(\overline{\theta}) = 0$. Hence,

$$C(1 + \cos\overline{\theta}) + \frac{1}{b}\sin\overline{\theta} = 0. \tag{8.127}$$

This implies (8.118). □

Remark 8.5 The solution $\overline{\theta} \in (-\pi, \pi)$ of Eq. (8.118) exists and is unique. For $b > 0$, the repulsive case corresponds to $Q > 0$ and $\overline{\theta} \in (0, \pi)$, while the attractive case corresponds to $Q < 0$ and $\overline{\theta} \in (-\pi, 0)$.

8.5.3 Differential Cross Section: Rutherford Formula

Now let us assume that the incident particles constitute a beam with flux density of n particles per cm^2s in the direction $\mathbf{e}_1$. Let $N = N(b, b + db)$ be the number of incident particles per second with the impact parameter within the interval $[b, b + db]$. By the axial symmetry, we have, for infinitesimal db,

$$N(b, b + db) = n2\pi b\, db. \tag{8.128}$$

The particles are scattered into the spatial angle $d\Omega = 2\pi \sin\overline{\theta}\, d\overline{\theta}$.

Definition 8.6 The differential cross section of the scattering is defined by

$$D(\overline{\theta}) := \frac{N/d\Omega}{n} = \frac{b db}{\sin\overline{\theta}\, d\overline{\theta}}. \tag{8.129}$$

Let us calculate the cross section. Rewriting (8.118) as

$$b^2 = \left(\frac{Q|e|Z}{Mv^2}\right)^2 \cot^2\frac{\overline{\theta}}{2} \tag{8.130}$$

and differentiating, it is found that

$$2bdb = \left(\frac{Q|e|Z}{Mv^2}\right)^2 2\cot\frac{\overline{\theta}}{2}\frac{1}{\sin^2\frac{\overline{\theta}}{2}}\frac{d\overline{\theta}}{2}. \tag{8.131}$$

Substituting this into (8.129), we arrive at the *Rutherford formula*

$$D(\overline{\theta}) = \frac{(\frac{Q|e|Z}{Mv^2})^2}{4\sin^4\frac{\overline{\theta}}{2}}. \tag{8.132}$$

8.6 The Quantum Scattering of Electrons

A quantum analogue of the Rutherford scattering is the scattering of an electron beam by the hydrogen atom in its ground state $\psi_1(t,\mathbf{x})$, as given by (8.30). The incident electron beam is described by the plane wave of type (5.1):

$$\psi_{\text{in}}(t,\mathbf{x}) = Ae^{i(\mathbf{kx}-\omega t)}, \quad \mathbf{k} = (k,0,0) \neq 0, \tag{8.133}$$

moving in the direction $\mathbf{e}_1$ (if $k>0$). The wave is a solution of the *free* Schrödinger equation. Hence (see (5.2)),

$$\hbar\omega = \frac{\hbar^2}{2\text{m}}\mathbf{k}^2 > 0. \tag{8.134}$$

The corresponding stationary electric current density is given by (5.3):

$$\mathbf{j}_{\text{in}} := \frac{e}{\text{m}}\big[-i\hbar\nabla\psi_{\text{in}}(t,\mathbf{x})\big]\cdot\psi_{\text{in}}(t,\mathbf{x}) = \frac{e\hbar\mathbf{k}}{\text{m}}|A|^2. \tag{8.135}$$

We will assume that the incident plane wave is a small perturbation of the ground state; i.e.,

$$|A| \ll C_1 \tag{8.136}$$

(see (8.30)). Hence, the total wave field, is approximately a solution of the Schrödinger equation (8.31) with $\mathbf{A}_0 = 0$,

$$\big[i\hbar\partial_t - e\phi_n(\mathbf{x})\big]\psi(t,\mathbf{x}) = \frac{1}{2\text{m}}[-i\hbar\nabla]^2\psi(t,\mathbf{x}). \tag{8.137}$$

Here $\phi_n(\mathbf{x}) = -e/|\mathbf{x}|$ is the nucleus potential; its decay at infinity is rather slow, resulting in the divergence in intermediate calculations. Let us modify $\phi_n(\mathbf{x})$ as follow:

$$\phi_\varepsilon(\mathbf{x}) = -\frac{e}{|\mathbf{x}|}e^{-\varepsilon|\mathbf{x}|}; \tag{8.138}$$

here $\varepsilon > 0$. Later we will make $\varepsilon \to 0$.

8.6.1 *The Radiated Wave*

Let us decompose the total wave field into three terms,

$$\psi(t,\mathbf{x}) = \psi_{\text{in}}(t,\mathbf{x}) + \psi_1(t,\mathbf{x}) + \psi_{\text{r}}(t,\mathbf{x}), \tag{8.139}$$

where $\psi_{\text{r}}(t,\mathbf{x})$ is a small radiated wave. Substituting (8.139) into the Schrödinger equation (8.137), this gives

$$\big[i\hbar\partial_t - e\phi_\varepsilon(\mathbf{x})\big]\big[\psi_{\text{in}}(t,\mathbf{x})+\psi_{\text{r}}(t,\mathbf{x})\big] = \frac{1}{2\text{m}}[-i\hbar\nabla]^2\big[\psi_{\text{in}}(t,\mathbf{x})+\psi_{\text{r}}(t,\mathbf{x})\big], \tag{8.140}$$

since $\psi_1(t, \mathbf{x})$ is an exact solution. We rewrite this equation as follows:

$$\left(i\hbar\partial_t - e\phi_\varepsilon(\mathbf{x}) - \frac{1}{2\mathrm{m}}[-i\hbar\nabla]^2\right)\psi_\mathrm{r}(t, \mathbf{x}) = e\phi_\varepsilon(\mathbf{x})\psi_\mathrm{in}(t, \mathbf{x}) = e\phi_\varepsilon(\mathbf{x})Ae^{i(\mathbf{kx}-\omega t)}. \tag{8.141}$$

Since $\omega > 0$, it follows by (8.134) that the frequency ω lies in the continuous spectrum of the Schrödinger operator. Now we apply the limiting amplitude principle (13.113) to obtain the long-time asymptotics for $\psi_\mathrm{r}(t, \mathbf{x})$:

$$\psi_\mathrm{r}(t, \mathbf{x}) \sim \psi_\mathrm{r}(\mathbf{x})e^{-i\omega t}, \quad t \to \infty; \tag{8.142}$$

here the contribution of the discrete spectrum is neglected, since the eigenfunctions decay rapidly at infinity, as distinct from $\psi_\mathrm{r}(\mathbf{x})$.

Substituting these asymptotics into equation (8.141), we obtain the stationary equation for the limiting amplitude

$$\left(\hbar\omega - e\phi_\varepsilon(\mathbf{x}) - \frac{1}{2\mathrm{m}}[-i\hbar\nabla]^2\right)\psi_\mathrm{r}(\mathbf{x}) = eA\phi_\varepsilon(\mathbf{x})e^{i\mathbf{kx}}. \tag{8.143}$$

Neglecting $e\phi_\varepsilon(\mathbf{x})$ in the LHS, we finally get the Helmholtz equation

$$\left(\Delta + \mathbf{k}^2\right)\psi_\mathrm{r}(\mathbf{x}) = \frac{2\mathrm{m}eA}{\hbar^2}\phi_\varepsilon(\mathbf{x})e^{i\mathbf{kx}}, \tag{8.144}$$

because $\mathbf{k}^2 = \frac{2\mathrm{m}\omega}{\hbar}$ by (8.134). This is an equation of type (8.96). Therefore, the solution is given by a convolution similar to (8.98),

$$\psi_\mathrm{r}(\mathbf{x}) = -\frac{2\mathrm{m}eA}{\hbar^2}\int \frac{e^{ik|\mathbf{x}-\mathbf{y}|}}{4\pi|\mathbf{x}-\mathbf{y}|}\phi_\varepsilon(\mathbf{y})e^{i\mathbf{ky}}\,d\mathbf{y}. \tag{8.145}$$

Convolution (8.145) is similar to the last integral (8.44). Evaluating it by method (8.45)–(8.48), we obtain

$$\psi_\mathrm{r}(\mathbf{x}) \sim -A\frac{e^{ik|\mathbf{x}|}}{|\mathbf{x}|}f(k, \theta), \quad |\mathbf{x}| \to \infty, \tag{8.146}$$

where θ is the polar angle between $\mathbf{x}$ and $\mathbf{e_1}$, and

$$f(k, \theta) = \frac{2\mathrm{m}e}{\hbar^2}\int_0^\infty \frac{\sin K|\mathbf{y}|}{K|\mathbf{y}|}\phi(\mathbf{y})|\mathbf{y}|^2\,d|\mathbf{y}|, \quad K := 2k\sin\frac{\theta}{2}. \tag{8.147}$$

Now (8.142) becomes, for large $|\mathbf{x}|$,

$$\psi_\mathrm{r}(t, \mathbf{x}) \sim -A\frac{e^{ik|\mathbf{x}|}}{|\mathbf{x}|}f(k, \theta)e^{-i\omega t}, \quad t \to \infty. \tag{8.148}$$

The limiting amplitude ψ_r decays slowly at infinity, its L^2 norm being infinite. This corresponds to the fact that the frequency $\omega > 0$ lies in the continuous spectrum. Physically, this means that the electron flux goes off to infinity, as we shall see below.

8.6.2 The Differential Cross Section

The stationary electric current, corresponding to radiated wave $\psi_{\mathrm{r}}(t,\mathbf{x})$ for large t and $|\mathbf{x}|$, is given by (8.148),

$$\begin{aligned}\mathbf{j}_{\mathrm{r}}(\mathbf{x}) &:= \frac{e}{\mathrm{m}}\big[-i\hbar\nabla\psi_{\mathrm{r}}(t,\mathbf{x})\big]\cdot\psi_{\mathrm{r}}(t,\mathbf{x}) \sim \frac{e\hbar k\mathbf{n}(\mathbf{x})}{\mathrm{m}}\frac{|f(k,\theta)|^2}{|\mathbf{x}|^2}|A|^2,\\ \mathbf{n}(\mathbf{x}) &:= \frac{\mathbf{x}}{|\mathbf{x}|},\end{aligned}\tag{8.149}$$

similarly to (8.135). The angular density of the stationary current of radiation is defined by

$$j_{\mathrm{r}}^{a}(\mathbf{n}) := \lim_{R\to\infty}\mathbf{j}_{\mathrm{r}}(R\mathbf{n})\cdot\mathbf{n}R^2,\quad \mathbf{n}\in\mathbb{R}^3,\ |\mathbf{n}|=1.\tag{8.150}$$

The current (8.149) is radial at infinity, hence the corresponding angular density is given by

$$j_{\mathrm{r}}^{a}(\mathbf{n}) = \frac{e\hbar k}{\mathrm{m}}\big|f(k,\theta)\big|^2|A|^2.\tag{8.151}$$

In the quantum case, Definition 8.6 should be modified as follows.

Definition 8.7 The differential cross section of the quantum scattering is defined by

$$D(\mathbf{n}) := \frac{j_{\mathrm{r}}^{a}(\mathbf{n})}{|\mathbf{j}_{\mathrm{in}}|},\quad \mathbf{n}\in\mathbb{R}^3,\ |\mathbf{n}|=1.\tag{8.152}$$

Formulas (8.151) and (8.135) imply that

$$D(\mathbf{n}) = \big|f(k,\theta)\big|^2.\tag{8.153}$$

Finally, substituting (8.138) into (8.147), we get

$$f(k,\theta) = \frac{2\mathrm{m}e^2}{K\hbar^2}\int_0^\infty \sin K|\mathbf{y}|e^{-\varepsilon|\mathbf{y}|}\,d|\mathbf{y}| = \frac{2\mathrm{m}e^2}{\hbar^2(K^2+\varepsilon^2)}.\tag{8.154}$$

Here $K = 2k\sin\frac{\theta}{2}\neq 0$ if $\theta\neq 0$ and $k\neq 0$. Then we can drop ε^2 as $\varepsilon\to 0$, to obtain

$$f(k,\theta) = \frac{2\mathrm{m}e^2}{\hbar^2K^2} = \frac{\mathrm{m}e^2}{2\hbar^2k^2\sin^2\frac{\theta}{2}}.\tag{8.155}$$

We rewrite this expression using (5.4):

$$f(k,\theta) = \frac{e^2}{2\mathrm{m}\mathbf{v}^2\sin^2\frac{\theta}{2}}.\tag{8.156}$$

Now (8.153) reads

$$D(\mathbf{n}) = \frac{(\frac{e^2}{\mathrm{m}\mathbf{v}^2})^2}{4\sin^4\frac{\theta}{2}}, \tag{8.157}$$

which coincides with the classical Rutherford formula (8.132) with $Q = e$ and $Z = 1$.

Remark 8.8 The agreement of (8.157) with the classical formula (8.132) was considered by Born as a crucial confirmation of *probabilistic interpretation* of quantum mechanics [20]: $|\psi(t, \mathbf{x})|^2$ is the **density of probability for the particle registration**, and the current (8.149) is the **flux of the probability** multiplied by e.

Chapter 9
Atom in Magnetic Field

Abstract Energy levels of the Hydrogen atom in a uniform magnetic field depend on the magnetic quantum number, while the eigenfunctions keep their form in a special basis. This dependence perfectly explains the (normal) Zeeman splitting of spectral lines.

The magnetic moment of the Hydrogen atom in a stationary state expresses as an integral of the corresponding electric current. The integral splits in two summands, of which one is responsible for diamagnetism resulting in the Langevin formula for diamagnetic susceptibility, and the second one is responsible for paramagnetism.

9.1 The Normal Zeeman Effect

In 1895, Zeeman discovered the splitting of spectral lines of atoms and molecules in a magnetic field. In the Schrödinger theory, the splitting and polarization have the same manifestation as in the classical Zeeman model (see Sect. 14.5). Namely, the *unperturbed spectral lines*, corresponding to the zero magnetic field $B = 0$, split into the *normal triplet* for $B \neq 0$.

The Hydrogen atom in a static external magnetic potential $\mathbf{A}(\mathbf{x})$ is described by the Schrödinger equation of type (3.37)

$$i\hbar\partial_t\psi(t,\mathbf{x}) = \frac{1}{2\mathrm{m}}\left(-i\hbar\nabla - \frac{e}{c}\mathbf{A}(\mathbf{x})\right)^2\psi(t,\mathbf{x}) - \frac{e^2}{|\mathbf{x}|}\psi(t,\mathbf{x}), \quad (t,\mathbf{x}) \in \mathbb{R}^4. \quad (9.1)$$

For the static potential, the Lorentz gauge condition (12.45) is equivalent to the Coulomb gauge condition

$$\nabla\mathbf{A}(\mathbf{x}) = 0. \quad (9.2)$$

Hence, on evaluating, we get

$$\begin{aligned} i\hbar\partial_t\psi(t,\mathbf{x}) = &-\frac{1}{2\mathrm{m}}\hbar^2\Delta\psi(t,\mathbf{x}) + i\frac{\hbar e}{\mathrm{m}c}\mathbf{A}(\mathbf{x})\cdot\nabla\psi(t,\mathbf{x}) \\ &- \frac{1}{2\mathrm{m}}\hbar^2\mathbf{A}^2(\mathbf{x})\psi(t,\mathbf{x}) - \frac{e^2}{|\mathbf{x}|}\psi(t,\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^3. \end{aligned} \quad (9.3)$$

A. Komech, *Quantum Mechanics: Genesis and Achievements*,
DOI 10.1007/978-94-007-5542-0_9,

Let us assume that the potential $\mathbf{A}(\mathbf{x})$ is small:

$$\big|\mathbf{A}(\mathbf{x})\big| \ll 1. \tag{9.4}$$

Then we can neglect term with $\mathbf{A}^2(\mathbf{x})$ in (9.8), obtaining

$$i\hbar\partial_t\psi(t,\mathbf{x}) = -\frac{1}{2\mathrm{m}}\hbar^2\Delta\psi(t,\mathbf{x}) + i\,\frac{\hbar e}{\mathrm{m}c}\mathbf{A}(\mathbf{x})\cdot\nabla\psi(t,\mathbf{x}) - \frac{e^2}{|\mathbf{x}|}\,\psi(t,\mathbf{x}), \quad \mathbf{x}\in\mathbb{R}^3. \tag{9.5}$$

9.1.1 The Hydrogen Spectrum in a Magnetic Field

Now we consider the uniform static magnetic field $\mathbf{B} = (0,0,B)$. In this case we have the external *static* Maxwell potentials (cf. (3.68))

$$\phi(\mathbf{x}) = -\frac{e}{|\mathbf{x}|}, \qquad \mathbf{A}(\mathbf{x}) = \frac{1}{2}\mathbf{B}\times\mathbf{x} = \frac{1}{2}B\big(-\mathbf{x}^2,\mathbf{x}^1,0\big), \tag{9.6}$$

which are axially symmetric with respect to the rotations round $\mathbf{e}_3$. Therefore, $i\hbar\mathbf{A}(\mathbf{x})\cdot\nabla = -B\hat{\mathbf{L}}_3/2$, by (9.6). Hence (9.5) reads

$$i\hbar\partial_t\psi(t,\mathbf{x}) = H\psi(t,\mathbf{x}) - \frac{e}{2\mathrm{m}c}B\hat{\mathbf{L}}_3\psi(t,\mathbf{x}), \quad \mathbf{x}\in\mathbb{R}^3, \tag{9.7}$$

where H is the Schrödinger operator (6.1). Hence, the corresponding stationary equation is as follows

$$E\psi(\mathbf{x}) = H\psi(\mathbf{x}) + \omega_{\mathcal{L}}\hat{\mathbf{L}}_3\psi(\mathbf{x}), \quad \mathbf{x}\in\mathbb{R}^3; \tag{9.8}$$

here $\omega_{\mathcal{L}} := -eB/(2\mathrm{m}c)$ is the *Larmor frequency*. The key observation is that the operators $\hat{\mathbf{L}}_3$ and H commute, inasmuch as the potentials are axially symmetric. Hence, there is a complete set of common eigenfunctions. Namely, (6.35) are common eigenfunctions, because $\hat{\mathbf{L}}_3 = -i\hbar\partial_\varphi$ by (3.86), and

$$H\psi_{lmn} = (E_n + m\hbar\omega_{\mathcal{L}})\psi_{lmn}, \quad n=1,2,\ldots,\ m=-l,\ldots,l,\ l\le n-1, \tag{9.9}$$

where E_n are given by (6.3). Therefore, the energy levels in the uniform magnetic field are given by

$$E_{mn} := -2\pi\hbar c R/n^2 + m\hbar\omega_{\mathcal{L}}, \quad n=1,2,\ldots,\ m=-l,\ldots,l,\ l\le n-1 \tag{9.10}$$

9.1.2 The Normal Splitting of Spectral Lines

Now the energy level splitting, the intensities and polarization of the Hydrogen radiation can be computed by substituting (9.10) and (6.35) into formulas (7.48) and

(7.52), (7.37). Namely, formula (6.35) can be written as

$$\psi_j = C e^{-r/r_n} P_{nl}(r) F_l^m(\theta) e^{im\varphi}, \quad j = (lmn), \tag{9.11}$$

and the corresponding eigenvalues are given by (9.10),

$$\omega_j = -\frac{2\pi c R}{n^2} + m\omega_{\mathcal{L}}. \tag{9.12}$$

If $B = 0$, we also have $\omega_{\mathcal{L}} = 0$. Hence, the *unperturbed spectral lines* are

$$\omega^0_{jj'} = 2\pi R\left[\frac{1}{n'^2} - \frac{1}{n^2}\right], \quad j = (lmn),\ j' = \big(l'm'n'\big).$$

Next consider the case $B \neq 0$. From (7.48) it follows that the intensity of the spectral line $\omega_{jj'}$ is proportional to $|\mathbf{J}_{jj'}|^2$, where

$$\mathbf{J}_{jj'} \sim \int_0^\infty R_{nl}(r)\overline{R_{n'l'}}(r) r^3\, dr \int_0^\pi \left[\int_0^{2\pi} \frac{\mathbf{x}}{r} F_l^m(\theta)\overline{F_{l'}^{m'}}(\theta) e^{i(m-m')\varphi}\, d\varphi\right] \sin\theta\, d\theta. \tag{9.13}$$

by (7.52) and (7.62). According to (7.60), $\mathbf{J}_{jj'}$ is nonzero only if $m' = m, m \pm 1$ and $l' = l \pm 1$. Hence, the unperturbed spectral line $\omega^0_{jj'}$ generates the *triplet*

$$\begin{cases} m' = m\colon & \omega_{jj'} = \omega^0_{jj'}, \\ m' = m-1\colon & \omega_{jj'} = \omega^0_{jj'} + \omega_{\mathcal{L}}, \\ m' = m+1\colon & \omega_{jj'} = \omega^0_{jj'} - \omega_{\mathcal{L}}; \end{cases} \tag{9.14}$$

this is precisely like in the Lorentz model (14.43). To find the corresponding polarizations, we write the vector $\mathbf{x}$ in the form

$$\mathbf{x} = \sin\theta\big(e^{-i\varphi}\mathbf{e}_+ + e^{i\varphi}\mathbf{e}_-\big) + \cos\theta\mathbf{e}_3, \quad \mathbf{e}_\pm := \frac{1}{2}(\mathbf{e}_1 \pm i\mathbf{e}_2). \tag{9.15}$$

Inserting this back into (9.13), it is found that for $m' = m$, only the $\mathbf{e}_3$ component contributes, while for $m' = m \pm 1$, only the $\mathbf{e}_\mp$ component contributes. Hence, using (7.37), we obtain the radiation in the form

$$\begin{cases} m' = m\colon & \mathbf{A}_{jj'}(t,\mathbf{x}) \sim \dfrac{1}{|\mathbf{x}|}\operatorname{Re} c_j\overline{c}_{j'}\mathbf{e}_3 e^{-i\omega_0(t-|\mathbf{x}|/c)}, \\ m' = m-1\colon & \mathbf{A}_{jj'}(t,\mathbf{x}) \sim \dfrac{1}{|\mathbf{x}|}\operatorname{Re} c_j\overline{c}_{j'}\mathbf{e}_+ e^{-i(\omega_0+\omega_{\mathcal{L}})(t-|\mathbf{x}|/c)}, \\ m' = m+1\colon & \mathbf{A}_{jj'}(t,\mathbf{x}) \sim \dfrac{1}{|\mathbf{x}|}\operatorname{Re} c_j\overline{c}_{j'}\mathbf{e}_- e^{-i(\omega_0-\omega_{\mathcal{L}})(t-|\mathbf{x}|/c)}. \end{cases}$$

Note that the resulting polarizations of the radiation fields are exactly the same as in the classical Zeemann model (see Sect. 14.5).

The spectral lines (9.12) are independent of the *orbital quantum number* l. Respectively, the radiations induced by pairs $j=(lmn)$, $j'=(l'm'n')$ with fixed n,n', $m'=m,m\pm 1$ and all possible quantum numbers l,l' contribute to the same frequencies (9.14).

On the other hand, in many cases the atom spectra in a magnetic field demonstrate *multiplet structure* different in significant way from the triplet structure (9.14). This situation suggests that the eigenvalues may depend on l,l'. This **anomalous Zeeman effect** cannot be explained by the Schrödinger equation (9.1). The explanation is provided by the **Pauli** equation, which takes into account the **electron spin** (see Sect. 10.3).

9.2 Diamagnetism and Paramagnetism

The magnetic moment of the Hydrogen atom in a uniform magnetic field $\mathbf{B}$ is known to be the sum of two terms $\mathbf{m}=\mathbf{m}'+\chi_m\mathbf{B}$ where $\chi_m<0$; as a result, $\chi_m\mathbf{B}$ describes the diamagnetism, and $\mathbf{m}'$ describes the paramagnetism of an atom. The calculation leads to the *Langevin formula* for the *magnetic susceptibility* of the hydrogen atom (see Lecture 12.13).

9.2.1 The Magnetic Moment

The magnetic moment of an atom in a stationary state is defined by (12.192):

$$\mathbf{m}=\frac{1}{2c}\int \mathbf{y}\times\mathbf{j}(\mathbf{y})\,d\mathbf{y}, \tag{9.16}$$

where the electric current $\mathbf{j}$ is defined, in the Born approximation (8.32), as follows:

$$\mathbf{j}(t,\mathbf{x})=\frac{e}{\mathrm{m}}\left[-i\hbar\nabla-\frac{e}{c}\mathbf{A}(\mathbf{x})\right]\psi(t,\mathbf{x})\cdot\psi(t,\mathbf{x}). \tag{9.17}$$

We choose coordinates in such a way that $\mathbf{B}=(0,0,B)$ with $B\geq 0$. Then the corresponding vector potential $\mathbf{A}(\mathbf{x})$ is given by (9.6). The stationary solutions in the uniform magnetic field are given by (9.11), (9.12):

$$\psi(t,\mathbf{x})=a_{lmn}(r,\theta)e^{im\varphi}e^{-i\omega_{mn}t}, \tag{9.18}$$

where a_{lmn} is a real function by Theorem 6.3(ii).

Let us express the gradient operator and the vector potential $\mathbf{A}(\mathbf{x})$ in the spherical coordinates:

$$\begin{aligned}\nabla&=\mathbf{e}_r\nabla_r\psi+\mathbf{e}_\theta\frac{\nabla_\theta}{r}+\mathbf{e}_\varphi\frac{\nabla_\varphi}{r\sin\theta},\\ \mathbf{A}(\mathbf{x})&=\frac{1}{2}B\big(-x^2,x^1,0\big)=\frac{1}{2}B|\mathbf{x}|\sin\theta\,\mathbf{e}_\varphi,\end{aligned} \tag{9.19}$$

by (6.54) and (6.55). Using (9.17) and (9.18) this gives,

$$\mathbf{j}(t,\mathbf{x}) = \mathbf{e}_r \frac{e}{\mathrm{m}}\big[-i\hbar\nabla_r a_{lmn}(r,\theta)\big]\cdot a_{lmn}(r,\theta) + \mathbf{e}_\theta \frac{e}{\mathrm{m}}\left[-i\hbar\frac{\nabla_\theta}{r} a_{lmn}(r,\theta)\right]\cdot a_{lmn}(r,\theta)$$

$$+\mathbf{e}_\varphi \frac{e}{\mathrm{m}}\left[m\hbar\frac{1}{r\sin\theta}\right]\big|\psi(t,\mathbf{x})\big|^2 - \mathbf{e}_\varphi\frac{e^2}{2\mathrm{m}c}B|\mathbf{x}|\sin\theta\big|\psi(t,\mathbf{x})\big|^2. \tag{9.20}$$

The first and second term on the RHS are zero, because the function $a_{nml}(r,\theta)$ is real. Finally,

$$\begin{aligned}\mathbf{j}(t,\mathbf{x}) &= \mathbf{e}_\varphi \frac{e}{\mathrm{m}}\left[m\hbar\frac{1}{r\sin\theta}\right]\big|\psi(t,\mathbf{x})\big|^2 - \mathbf{e}_\varphi\frac{e^2}{2\mathrm{m}c}B|\mathbf{x}|\sin\theta\big|\psi(t,\mathbf{x})\big|^2\\ &= \mathbf{j}'(\mathbf{x}) + \mathbf{j}''(\mathbf{x}).\end{aligned} \tag{9.21}$$

Now the magnetic moment (9.16) becomes

$$\mathbf{m} = \frac{1}{2c}\int \mathbf{y}\times\mathbf{j}'(\mathbf{y})\,d\mathbf{y} + \frac{1}{2c}\int \mathbf{y}\times\mathbf{j}''(\mathbf{y})\,d\mathbf{y} = \mathbf{m}' + \mathbf{m}''. \tag{9.22}$$

9.2.2 The Langevin Formula

The currents $\mathbf{j}'(\mathbf{x})$, $\mathbf{j}''(\mathbf{x})$ are obviously axially symmetric with respect to axis $O\mathbf{x}^3$; also

$$(\mathbf{y}\times\mathbf{e}_\varphi)_3 = |\mathbf{y}|\mathbf{e}_3\sin\theta. \tag{9.23}$$

Hence, the direction of $\mathbf{m}''$ is opposite to $\mathbf{B}$, and so $\mathbf{m}''$ describes the *diamagnetism* of an atom. The corresponding diamagnetic susceptibility (per atom) is defined by the formula $\mathbf{m}'' = \chi_m\mathbf{B}$; hence

$$\chi_m := -\frac{\mathbf{m}_3''}{B} = -\frac{e^2}{4\mathrm{m}c^2}\int |\mathbf{y}|^2\sin^2\theta|\psi|^2\,d\mathbf{y} = -\frac{e^2}{4\mathrm{m}c^2}\Theta_3, \tag{9.24}$$

where Θ_3 denotes the moment of inertia of distribution $|\psi|^2$ with respect to the axis $O\mathbf{x}^3$. For the spherically symmetric stationary states, we have

$$\Theta_3 = \frac{2}{3}\overline{\Theta}, \quad \overline{\Theta} := \int |\mathbf{y}|^2|\psi|^2\,d\mathbf{y}. \tag{9.25}$$

Then (9.24) becomes the *Langevin formula* (cf. Sect. 14.6)

$$\chi_m = -\frac{e^2\overline{\Theta}}{6\mathrm{m}c^2}. \tag{9.26}$$

For spherically non-symmetric states, this formula also holds *in the mean*, because of the *random orientation* of atoms with respect to direction of the magnetic field (if the field is small).

9.2.3 Paramagnetism

The moment $\mathbf{m}'$ is different from zero even in the absence of a magnetic field,

$$\begin{aligned}\mathbf{m}' &= m\hbar \frac{e}{2c\mathrm{m}} \int \frac{(\mathbf{y} \times \mathbf{e}_\varphi)_3 \mathbf{e}_3}{|\mathbf{y}| \sin\theta} \left|\psi(t, \mathbf{y})\right|^2 d\mathbf{y} \\ &= m\hbar \mathbf{e}_3 \frac{e}{2c\mathrm{m}} \int \left|\psi(t, \mathbf{y})\right|^2 d\mathbf{y} = \frac{e}{2c\mathrm{m}} m\hbar \mathbf{e}_3, \end{aligned} \tag{9.27}$$

by (9.23). Hence, $\mathbf{m}'$ should describe the *paramagnetism* of an atom for $m > 0$. However, its mathematical expectation is zero, due to the random orientation of an atom, because the probabilities of the values m and $-m$ are identical for $B = 0$. On the other hand, the probabilities are not identical, if the magnetic field does not vanish. The corresponding *statistical* theory was developed by Langevin (see [7]).

Chapter 10
Electron Spin and Pauli Equation

Abstract The concept of the *electron spin* was developed empirically as a specific component of the angular momentum. It was suggested by many experimental facts and theoretical problems: the anomalous Zeeman splitting of spectra, the Einstein–de Haas and Stern–Gerlach experiments, as well as by some open questions in the Bohr theory of periodic table.

The electron spin was introduced in 1925 by Goudsmit and Uhlenbeck to explain the double splitting of spectral lines in the Stern–Gerlach experiments. This concept appeared to be a realization of the mysterious two-valued degree of freedom introduced by Pauli in 1924 to fix the Bohr theory of the periodic system.

Pauli introduced a new term into the Schrödinger equation, corresponding to the interaction of the spin with magnetic field. This Pauli equation was further completed by the Russell–Saunders spin-orbital coupling. The completed equation implies the Stern–Gerlach double splitting, and the Landé formula for the *gyromagnetic ratio*, as obtained in the 'old quantum theory'. This Landé formula perfectly explains the Einstein–de Haas experiments and the anomalous Zeeman effect.

In 1927, Sommerfeld applied this spin concept to clarify some open questions in the Drude theory of metals.

10.1 Concept of Electron Spin

The concept of spin was suggested by the anomalous Zeeman effect, the Einstein–de Haas experiment, Bohr's theory of periodic table, and the Stern–Gerlach experiment. This concept allowed to clarify some questions in the Drude theory of metals.

10.1.1 The Anomalous Zeeman Effect

In 1895, Zeeman observed that, in an external magnetic field, most of the spectral lines split into different number of lines: two, five, etc., see [20]. This *anomalous* Zeeman effect contradicts the Schrödinger theory, which is known to predict only the 'normal' Zeeman splitting into three lines (9.14). This contradiction suggests the existence of an additional magnetic moment responsible for the anomalous splitting.

A. Komech, *Quantum Mechanics: Genesis and Achievements*,
DOI 10.1007/978-94-007-5542-0_10,

10.1.2 The Einstein–de Haas Experiment

In 1915, Einstein considered an iron bar positioned vertically in the Earth's gravitational field; the bar was attached to a vertical string allowing rotation about its axis. The bar was magnetized by a vertical external weak magnetic field. The field orients the elementary Ampère molecular currents, so that their magnetic moments increase the total magnetic field. Therefore, the corresponding elementary angular momenta are also oriented in the same direction, and the sum of the microscopic angular momenta increases. On the other hand, the total angular momentum of the bar (macroscopic+microscopic) is conserved by the axial symmetry. Hence, the bar, as a whole, must also change its macroscopic angular momentum.

Einstein and de Haas have measured the changes in the magnetic moment and in the macroscopic angular momentum of the bar to study their ratio. The ratio is predicted to be $\frac{e}{2\mathrm{m}c}$ according to the classical theory and the Schrödinger theory as well. However, this value contradicts the experimental observations (see [20] and [171, Vol. I]), and suggests the existence of an additional magnetic moment of atoms, which is responsible for the anomalous ratio.

The Classical Theory

In detail, the total angular momentum and the total magnetic moment of all electrons of the bar are given by

$$\mathbf{L} := \sum \mathrm{m}\mathbf{x}_k \times \mathbf{v}_k, \qquad \mathbf{m} := \frac{1}{2c} \sum e\mathbf{x}_k \times \mathbf{v}_k, \tag{10.1}$$

where $\mathbf{x}_k$ and $\mathbf{v}_k$ are the electron positions and velocities. Therefore,

$$\mathbf{m} = \frac{e}{2\mathrm{m}c}\mathbf{L}. \tag{10.2}$$

Initially, the magnetization of the bar is negligible, since the magnetic moments of its atoms cancel out, due to their random orientations. The external magnetic field reorients orbits of the atomic electrons increasing magnitude of the bar magnetization according to the Ampère's conjecture.

By (10.2), the magnitude of the total orbital angular momentum of the electrons in the bar also increases. On the other hand, the total angular momentum of the bar is conserved by the axial symmetry. Therefore, the bar acquires a macroscopic rotation with opposite angular momentum $\mathbf{L}^b = -\mathbf{L}$.

The magnitudes of $\mathbf{L}_3^b$ and $\mathbf{m}_3$ can be measured experimentally: $\mathbf{L}_3^b$ by the torsion vibration of the string, and $\mathbf{m}_3$, by the residual magnetism. The experiment was performed by Einstein and de Haas in 1915. However, the result was very discouraging:

$$\mathbf{m}_3 = g\frac{e}{2\mathrm{m}c}\mathbf{L}_3 \tag{10.3}$$

with the *Landé factor* $g \approx 2$, which contradicts (10.2).

The Schrödinger Quantum Theory

Relation (10.2) also holds in the Schrödinger theory. Indeed, (9.16) and (9.17) imply that

$$\mathbf{m} = \frac{1}{2c}\int \mathbf{y} \times \frac{e}{\mathrm{m}}[-i\hbar\nabla]\psi(t,\mathbf{y}) \cdot \psi(t,\mathbf{y})\,d\mathbf{y}$$
$$= \frac{e}{2\mathrm{m}c}\int [-i\hbar\mathbf{y}\times\nabla]\psi(t,\mathbf{y}) \cdot \psi(t,\mathbf{y})\,d\mathbf{y} = \frac{e}{2\mathrm{m}c}\mathbf{L}. \qquad (10.4)$$

In particular, for the stationary state (9.18), we have the magnetization (cf. (9.27))

$$\mathbf{m}_3 = \frac{e}{2\mathrm{m}c} m\hbar, \qquad (10.5)$$

while the corresponding energy equals const $+ m\hbar\omega_{\mathcal{L}}$, according to (9.10). A week external magnetic field causes transitions to the stationary states with lower values of the energy; this corresponds to greater values of m, since $\omega_{\mathcal{L}} < 0$. Therefore, magnetization (9.27) increases with transitions. This mechanism substitutes the Ampère's conjecture on reorientation of atomic orbits in the quantum approach, and in particular, explains the magnetization of metallic needles in the Earth's magnetic field.

10.1.3 The Stern–Gerlach Experiment

In 1922, Stern and Gerlach sent a beam of silver atoms through a strongly inhomogeneous magnetic field. Later, similar experiments were performed with hydrogen atoms. The atoms are in the ground state (8.30), which is non-degenerate, according to the Schrödinger equation (9.7). The term $\hat{\mathbf{L}}_3\psi(t,\mathbf{x})$ in equation vanishes, by the spherical symmetry of the ground state. Hence, the ground state also satisfies equation (9.7) with $B \neq 0$. However, Stern and Gerlach observed splitting of a beam into two components. This contradicts the identity of all atoms in the ground state, and suggests the following conclusions:

(i) For $B = 0$, the eigenspace of lowest energy has dimension at least two.
(ii) For $B \neq 0$, the eigenspace splits in two distinct eigenspaces.

This suggests the existence of an additional magnetic moment of the electron, which does not vanish for the ground state.

10.1.4 Borh's Theory of Periodic Table

In 1920–1924, Bohr put forward the theory of the Mendeleev periodic table in the framework of the 'old quantum mechanics'. However, the corresponding periods were found to be two-times shorter than those required in chemistry.

In 1924, Pauli explained the needed doubling of the periods by presence of a new two-valued degree of freedom, which should be combined to known quantum numbers in accordance with the Exclusion Principle.

10.1.5 The Spin Conjecture

To explain the Stern–Gerlach experiments, Goudsmit and Uhlenbeck introduced in 1925 the new conjecture, resulting from analysis of the Einstein–de Haas and Stern–Gerlach experiments and the Pauli two-valued degree of freedom [95]:

> *The electron has the intrinsic spin momentum with values* $\pm\hbar/2$ *and the magnetic moment with magnitude* $|e|\hbar/2\mathrm{m}c$.

Here the term 'intrinsic' means that the spin angular momentum is not related to an orbital rotation of a particle (in contrast to the angular momentum), and the spin magnetic moment is not related to the convection current.

Let us comment on the motivations behind this conjecture.

I. One may expect a suitable modification $\tilde{H}$ for the Schrödinger operator H and $\hat{\mathbf{J}} = (\hat{\mathbf{J}}_1, \hat{\mathbf{J}}_2, \hat{\mathbf{J}}_3)$ for the angular momentum operator $\hat{\mathbf{L}}$. The operators $\tilde{H}$ and $\hat{\mathbf{J}}$ should commute, because of the spherical symmetry of the hydrogen atom; the commutation relations for $\hat{\mathbf{J}}_1, \hat{\mathbf{J}}_2, \hat{\mathbf{J}}_3$ should be similar to (6.39). Therefore, the two-dimensional eigenspace of lowest energy for $\tilde{H}$ with $B = 0$ should be also an eigenspace for $\hat{\mathbf{J}}^2$ with an eigenvalue $\hbar^2 J(J+1)$, where $J = 0,\ 1/2,\ 1, \ldots$ according to Proposition 6.12.

II. The dimension of the eigenspace is $2J + 1$. Hence, the Stern–Gerlach double splitting suggests $J = \frac{1}{2}$. Respectively, the extremal values for each component of the spin momentum should be $\pm\frac{\hbar}{2}$, by the last formula of (6.43).

III. The value for magnetic moment was chosen to provide the Landé factor $g = 2$ suggested by the Einstein–de Haas observations (10.3).

10.1.6 The Sommerfeld Theory of Metals

In 1927, Sommerfeld extended the Pauli Exclusion Principle to the electron gas in metals. This extension is highly nontrivial: Pauli suggested the principle for the *microscopic* system of $N \sim 10^2$ electrons in an atom, while Sommerfeld extended the principle to a *macroscopic* system of electrons in a metal with $N \sim 10^{23}$. This extension had clarified open questions on the specific heat and conductivity in the Drude theory of metals [175]. This great success confirmed the concept of spin and the Pauli Exclusion Principle.

10.2 The Pauli Equation

As we have seen in previous section, the Schrödinger theory is incapable of explaining the anomalous Zeeman effect and the Einstein–de Haas and Stern–Gerlach experiments. In 1927, Pauli introduced an additional term into the Schrödinger equation, which corresponds to energy of the spin angular momentum, as suggested by Goudsmit and Uhlenbeck. The Pauli equation explains the Stern–Gerlach experiment, but it cannot explain the anomalous Zeeman effect and the Einstein–de Haas experiment.

10.2.1 The Additional Magnetic Moment

From (10.4) it follows that the last term in (9.7) can be written as follows:

$$-\frac{e}{2\mathrm{m}c}\mathbf{B}\hat{\mathbf{L}}\psi(t,\mathbf{x}) = -\mathbf{B}\hat{\mathbf{m}}\psi(t,\mathbf{x}); \tag{10.6}$$

here $\hat{\mathbf{m}}$ is the quantum observable, corresponding to the magnetic moment. This corresponds to the fact that $-\mathbf{B}\hat{\mathbf{m}}$ should correspond to the energy of the magnetic moment $\hat{\mathbf{m}}$ in the magnetic field $\mathbf{B}$. The Goudsmit–Uhlenbeck conjecture means that the additional magnetic moment corresponds to the additional *spin angular momentum* $\hat{\mathbf{s}}$ of the electron involving the Landé factor 2; i.e.,

$$\hat{\mathbf{m}} = 2\frac{e}{2\mathrm{m}c}\hat{\mathbf{s}}. \tag{10.7}$$

Finally, equation (9.7) should be modified as

$$i\hbar\partial_t\psi(t,\mathbf{x}) = -\frac{1}{2\mathrm{m}}\hbar^2\Delta\psi(t,\mathbf{x}) + e\phi\big(|\mathbf{x}|\big)\psi(t,\mathbf{x}) - \frac{e}{2\mathrm{m}c}\mathbf{B}[\hat{\mathbf{L}} + 2\hat{\mathbf{s}}]\psi(t,\mathbf{x}). \tag{10.8}$$

10.2.2 The Spin Momentum

We still should identify the spin angular momentum $\hat{\mathbf{s}}$. First, let us analyze the introduction of the electron *orbital angular momentum* $\hat{\mathbf{L}}$ for the spherically symmetric scalar potential $\phi(|\mathbf{x}|)$:

$$\begin{aligned} i\hbar\partial_t\psi(t,\mathbf{x}) = H\psi(t,\mathbf{x}) := -\frac{1}{2\mathrm{m}}\hbar^2\Delta\psi(t,\mathbf{x}) + e\phi\big(|\mathbf{x}|\big)\psi(t,\mathbf{x}), \\ \mathbf{x} \in \mathbb{R}^3. \end{aligned} \tag{10.9}$$

The orbital angular momentum is a conserved quantity, which corresponds to the invariance of Lagrangian (4.7) with respect to the **regular representation** of the rotation group SO(3). This representation acts on the phase space $\mathcal{E} := L^2(\mathbb{R}^3)$ by $R_g\psi(\mathbf{x}) := \psi(g^{-1}\mathbf{x})$ for $\mathbf{x} \in \mathbb{R}^3$, where $g \in \mathrm{SO}(3)$ and $\psi \in \mathcal{E}$. The proof of this conservation relies on the commutation $[H, \hat{\mathbf{L}}_k] = 0$. An alternative proof follows from the Noether theorem.

This analysis suggests that the spin momentum should correspond to another action of the rotation group, which is different from the regular representation. To construct this **spinor representation**, we denote its generators by $\mathbf{h}_k$. Then, similarly to (6.39), we necessarily have

$$[\mathbf{h}_1, \mathbf{h}_2] = i\mathbf{h}_3, \qquad [\mathbf{h}_2, \mathbf{h}_3] = i\mathbf{h}_1, \qquad [\mathbf{h}_3, \mathbf{h}_1] = i\mathbf{h}_2. \tag{10.10}$$

The double splitting of the ground state in the Stern–Gerlach experiment suggests that the dimension of the unperturbed 'ground eigenspace' must be two. Therefore, the action of the group SO(3) in the unperturbed ground eigenspace is a two-dimensional representation S_g. Such an *irreducible* two-dimensional representation is given by Proposition 6.12: it corresponds to generators $\mathbf{h}_k := \dfrac{1}{2}\sigma_k$, where σ_k are the **Pauli matrices** (see (6.49))

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{10.11}$$

Exercise 10.1 Verify the commutation relations (10.10) for generators $\mathbf{h}_k = \frac{1}{2}\sigma_k$.

Exercise 10.2 Verify the identities

$$\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = 1. \tag{10.12}$$

10.2.3 Uniform Magnetic Field

Further, we set $\hat{\mathbf{s}}_k = \hbar\mathbf{h}_k$, similarly to $\hat{\mathbf{L}}_k = \hbar\mathbf{H}_k$. Then the spin angular momenta

$$\mathbf{s}_k = \langle\psi, \hat{\mathbf{s}}_k\psi\rangle, \quad k = 1,\ 2,\ 3 \tag{10.13}$$

are conserved quantities, provided that the *modified Schrödinger operator* commutes with $\mathbf{h}_k$. The simplest way to make the ground eigenspace two-dimensional is to define the modified phase space as the tensor product $\mathcal{E}^{\otimes} := \mathcal{E} \otimes \mathbb{C}^2$, and consider the tensor product of the regular representation R_g in $\mathcal{E}$ and the spinor representation S_g in $\mathbb{C}^2$. By definition of the tensor product, all generators $\mathbf{H}_k$ commute with all generators $\mathbf{h}_j$, hence all $\hat{\mathbf{L}}_k$ commute with all $\hat{\mathbf{s}}_j$.

With these new treatment of the wave function and the spin operators $\hat{\mathbf{s}}$, (10.8) becomes the **Pauli equation**,

$$\begin{aligned} i\hbar\partial_t\Psi(t,\mathbf{x}) &= \mathcal{P}\Psi(t,\mathbf{x}) \\ &:= -\frac{1}{2\mathrm{m}}\hbar^2\Delta\Psi(t,\mathbf{x}) + e\phi\big(|\mathbf{x}|\big)\Psi(t,\mathbf{x}) \\ &\quad - \frac{e}{2\mathrm{m}c}\mathbf{B}[\hat{\mathbf{L}} + 2\hat{\mathbf{s}}]\Psi(t,\mathbf{x}), \end{aligned} \tag{10.14}$$

where $\Psi(t,\mathbf{x}) = (\psi_1(t,\mathbf{x}), \psi_2(t,\mathbf{x})) \in \mathcal{E}^{\otimes} = \mathcal{E} \otimes \mathbb{C}^2$, $\hat{\mathbf{s}} = -\frac{\hbar}{2}\sigma$.

Definition 10.3

(i) For the Pauli Equation, the total angular momentum is defined as

$$\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{s}}. \tag{10.15}$$

(ii) The total magnetic moment is

$$\hat{\mathbf{m}} = \frac{e}{2\mathrm{m}c}[\hat{\mathbf{L}} + 2\hat{\mathbf{s}}]. \tag{10.16}$$

For $\mathbf{B} = (0, 0, B)$, equation (10.14) reads as follows:

$$\begin{aligned} i\hbar\partial_t\Psi(t,\mathbf{x}) &= \mathcal{P}\Psi(t,\mathbf{x}) \\ &= -\frac{1}{2\mathrm{m}}\hbar^2\Delta\Psi(t,\mathbf{x}) + e\phi\big(|\mathbf{x}|\big)\Psi(t,\mathbf{x}) \\ &\quad - \frac{eB}{2\mathrm{m}c}[\hat{\mathbf{L}}_3 + 2\hat{\mathbf{s}}_3]\Psi(t,\mathbf{x}). \end{aligned} \tag{10.17}$$

Remark 10.4 This so-defined operator $\mathcal{P}$ with $\mathbf{B} = (0, 0, B)$ commutes with $\hat{\mathbf{L}}^2 := \hat{\mathbf{L}}_1^2 + \hat{\mathbf{L}}_2^2 + \hat{\mathbf{L}}_3^2$, $\hat{\mathbf{L}}_3$, $\hat{\mathbf{s}}^2 := \hat{\mathbf{s}}_1^2 + \hat{\mathbf{s}}_2^2 + \hat{\mathbf{s}}_3^2$, $\hat{\mathbf{s}}_3$ and $\hat{\mathbf{J}}^2 := \hat{\mathbf{J}}_1^2 + \hat{\mathbf{J}}_2^2 + \hat{\mathbf{J}}_3^2$, $\hat{\mathbf{J}}_3$. Therefore, the corresponding mean values $\mathbf{L}^2, \mathbf{L}_3, \mathbf{s}^2 = 3/4, \mathbf{s}_3$ and $\mathbf{J}^2, \mathbf{J}_3 = \mathbf{L}_3 + \mathbf{s}_3$, are conserved.

10.2.4 General Maxwell Field

A natural extension of the Pauli equation for general external Maxwell field reads

$$\begin{aligned} &\big[i\hbar\partial_t - e\phi(t,\mathbf{x})\big]\Psi(t,\mathbf{x}) \\ &= \frac{1}{2\mathrm{m}}\Big[-i\hbar\nabla - \frac{e}{c}\mathbf{A}(t,\mathbf{x})\Big]^2\Psi(t,\mathbf{x}) - \frac{e}{\mathrm{m}c}\mathbf{B}(t,\mathbf{x})\hat{\mathbf{s}}(t,\mathbf{x})\Psi(t,\mathbf{x}). \end{aligned} \tag{10.18}$$

It corresponds to the Lagrangian density (cf. (4.7))

$$\begin{aligned} \mathcal{L}_P(x,\Psi,\nabla\Psi) &= \big[i\hbar\nabla_0 - e\phi(x)\big]\Psi\cdot\Psi - \frac{1}{2\mathrm{m}}\Big|\Big[-i\hbar\nabla - \frac{e}{c}\mathbf{A}(x)\Big]\Psi\Big|^2 \\ &\quad - \frac{e}{\mathrm{m}c}\big[\mathbf{B}(x)\hat{\mathbf{s}}\Psi\big]\cdot\Psi, \end{aligned} \tag{10.19}$$

where $\mathbf{B}(x) = \operatorname{curl} \mathbf{A}(x)$ and '$\cdot$' stands for the inner product on $\mathbb{C}^2 \equiv \mathbb{R}^4$. This suggests the following Lagrangian density for the coupled Maxwell–Pauli equations (cf. (4.18))

$$\mathcal{L}_{MP} = \big[i\hbar\nabla_0 - e\phi(x)\big]\Psi \cdot \Psi - \frac{1}{2\mathrm{m}}\left|\left[-i\hbar\boldsymbol{\nabla} - \frac{e}{c}\mathbf{A}(x)\right]\Psi\right|^2 - \frac{e}{\mathrm{m}c}\big[\mathbf{B}(x)\hat{\mathbf{s}}\Psi\big] \cdot \Psi - \frac{1}{16\pi}\mathcal{F}^{\alpha\beta}\mathcal{F}_{\alpha\beta}, \tag{10.20}$$

where $\mathcal{F}^{\mu\nu} := \partial^\mu \mathcal{A}^\nu - \partial^\nu \mathcal{A}^\mu$ and $\mathcal{F}_{\mu\nu} := \partial_\mu \mathcal{A}_\nu - \partial_\nu \mathcal{A}_\mu$. The corresponding Euler–Lagrange equations are the coupled Maxwell–Pauli equations (cf. (4.19))

$$\begin{cases} \big[i\hbar\partial_t - e\phi(t,\mathbf{x})\big]\Psi(t,\mathbf{x}) = \dfrac{1}{2\mathrm{m}}[-i\hbar\boldsymbol{\nabla} - \dfrac{e}{c}\mathbf{A}(t,\mathbf{x})]^2\Psi(t,\mathbf{x}) \\ \qquad\qquad\qquad\qquad - \dfrac{e}{\mathrm{m}c}\mathbf{B}(t,\mathbf{x})\hat{\mathbf{s}}\Psi(t,\mathbf{x}), \\ \dfrac{1}{4\pi}\nabla_\alpha\mathcal{F}^{\alpha\beta}(t,\mathbf{x}) = \left(\rho, \dfrac{\mathbf{j}}{c}\right), \end{cases} \tag{10.21}$$

where

$$\begin{aligned} \rho &:= e\big|\Psi(t,\mathbf{x})\big|^2, \\ \frac{\mathbf{j}}{c} &:= \frac{e}{\mathrm{m}c}\left[-i\hbar\boldsymbol{\nabla} - \frac{e}{c}\mathbf{A}(t,\mathbf{x})\right]\Psi(t,\mathbf{x}) \cdot \Psi(t,\mathbf{x}) + \frac{e}{\mathrm{m}c}\operatorname{curl}\big(\big[\hat{\mathbf{s}}\Psi(t,\mathbf{x})\big] \cdot \Psi(t,\mathbf{x})\big). \end{aligned} \tag{10.22}$$

10.2.5 The Stern–Gerlach Double Splitting

It is easy to construct quantum stationary states corresponding to the Pauli equation (10.17) for the hydrogen atom. Namely, we introduce the vector-functions

$$\begin{aligned} \Psi^+_{nml} &= \begin{pmatrix} \psi_{nml} \\ 0 \end{pmatrix}, \\ \Psi^-_{nml} &= \begin{pmatrix} 0 \\ \psi_{nml} \end{pmatrix}, \quad n = 1,2,3,\ldots, l \le n-1,\ m = -l,\ldots,l, \end{aligned} \tag{10.23}$$

where ψ_{lmn} are given by (6.35). They are eigenfunctions of the operator $\mathcal{P}$ corresponding to the energies

$$E^{\pm}_{mn} := -2\pi\hbar R/n^2 + m\hbar\omega_{\mathcal{L}} \pm \hbar\omega_{\mathcal{L}}. \tag{10.24}$$

In particular, the Schrödinger ground state ψ_{001} splits into two stationary states $\Psi^{\pm}_{001}$ with distinct energies $E^{\pm}_{01} := -2\pi\hbar R \pm \hbar\omega_{\mathcal{L}}$ and distinct spin magnetic moments $\pm\frac{e\hbar}{2\mathrm{m}c}$. This explains the double splitting of the beam in the Stern–Gerlach experiment.

Remark 10.5 In the classical interpretation, the splitting term $\pm\hbar\omega_{\mathcal{L}}$ of (10.24) means that the 'projection of the spin magnetic moment' to the axis Ox^3 is $\pm\frac{e}{2\mathrm{m}c}$. The magnitude of the projection is attained 'instantly' with the magnetic field, which contradicts the classical concept (assuming nonzero atomic moment of inertia). In the quantum context, the instant reaction is not surprising, since it should be interpreted just as the (instant) bifurcation of quantum stationary states.

10.3 The Anomalous Zeeman Effect

Formula (10.24) shows that the Pauli equation is still incapable of explaining the anomalous Zeeman effect. Indeed, the selection rules $m \mapsto m, m \pm 1$, give the splitting $\omega_0 \mapsto \omega_0, \omega_0 \pm \omega_{\mathcal{L}}$ of an unperturbed spectral line ω_0, as in the Schrödinger theory. Hence, the splitting just gives the normal Zeeman triplet (9.14), as before.

This situation is related to the fact that the Pauli equation (10.14) does not take into account the interaction between the orbital and spin angular momenta. The interaction was suggested by the great success of the phenomenological *Landé vector model* in the 'old quantum mechanics' (see Sect. 14.7). Namely, the (classical) orbital angular momentum is related to the orbital motion of the electron. This motion results in the circular current, which generates the magnetic field acting on the spin magnetic moment.

10.3.1 The Spin-Orbital Coupling

The interaction between orbital and spin angular momenta was introduced in 1926 by Thomas [180] and Frenkel [66]. This interaction appears as a correction to the energy of a particle, which is due to the interaction of the spin with magnetic field $\mathbf{B}_*(\mathbf{x}_*)$, arising in moving frame of an electron. This magnetic field is expressed in the electrostatic *radial potential* $\phi(|\mathbf{x}|)$, by the Lorentz formulas (12.74): we have, in the first order approximation in $\beta = v/c$,

$$\begin{aligned}\mathbf{B}_*(\mathbf{x}_*) &= \frac{1}{c}\mathbf{E}(\mathbf{x}) \times \mathbf{v} = -\frac{1}{c}\nabla\phi\big(|\mathbf{x}|\big) \times \mathbf{v} = -\frac{1}{\mathrm{m}c}\phi'\big(|\mathbf{x}|\big)\frac{\mathbf{x}}{|\mathbf{x}|} \times \mathbf{p}\\ &= -\frac{1}{\mathrm{m}c}\frac{\phi'(|\mathbf{x}|)}{|\mathbf{x}|}\mathbf{L}. \qquad (10.25)\end{aligned}$$

The interaction of this magnetic field with the spin magnetic moment produces the energy correction $-\dfrac{e}{\mathrm{m}c}\mathbf{s}\mathbf{B}_*(\mathbf{x}_*)$. Hence, this gives the correction

$$\frac{e}{\mathrm{m}c}\hat{\mathbf{s}}\mathbf{B}_*(\mathbf{x}_*)\Psi(t,\mathbf{x}) = \frac{e}{\mathrm{m}^2c^2}\frac{\phi'(|\mathbf{x}|)}{|\mathbf{x}|}\hat{\mathbf{s}}\mathbf{L}\Psi(t,\mathbf{x})$$

to the RHS of (10.14). However, this correction does not agree with experimental observations, which suggest an additional factor $1/2$. Thomas and Frenkel argued for the additional factor using a sophisticated analysis of the Larmor precession of the spin (see the details in [93]). The corresponding rigorous argument, providing the Thomas factor $1/2$ (and the Landé factor $g = 2$ as well), is automatically provided by the relativistic Dirac theory (see Sect. 11.7). Finally, the modified Eq. (10.14) reads

$$\begin{aligned} i\hbar\partial_t\Psi(t,\mathbf{x}) &= \tilde{\mathcal{P}}\Psi(t,\mathbf{x}) \\ &:= -\frac{1}{2\mathrm{m}}\hbar^2\Delta\Psi(t,\mathbf{x}) + e\phi\big(|\mathbf{x}|\big)\Psi(t,\mathbf{x}) \\ &\quad + \frac{e}{2\mathrm{m}^2c^2}\frac{\phi'(|\mathbf{x}|)}{|\mathbf{x}|}\hat{\mathbf{s}}\hat{\mathbf{L}}\Psi(t,\mathbf{x}) - \frac{e}{2\mathrm{m}c}\mathbf{B}[\hat{\mathbf{L}} + 2\hat{\mathbf{s}}]\Psi(t,\mathbf{x}). \end{aligned} \tag{10.26}$$

The correction term with $\hat{\mathbf{s}}\hat{\mathbf{L}}$ received the name the **Russell–Saunders spin-orbital coupling** after their papers (1925), which developed the Landé vector model for many-electron atoms.

We will see below that stationary energies for equation (10.26) depend on the orbital quantum number l, in contrast to (10.24). This dependence perfectly describes the anomalous Zeeman effect, as well as the Einstein–de Haas experiment.

For $\mathbf{B} = (0, 0, B)$, we can write (10.26) in the form

$$i\hbar\partial_t\Psi(t,\mathbf{x}) = \tilde{\mathcal{P}}_0\Psi(t,\mathbf{x}) - B\hat{\mathbf{m}}_3\Psi(t,\mathbf{x}), \tag{10.27}$$

where $\tilde{\mathcal{P}}_0$ corresponds to $B = 0$. We will apply the perturbation theory to calculate the spectrum of equation (10.27) for small $|B|$.

We shall also calculate the Landé factor for quantum stationary states of equation (10.27). The factor explains the Einstein–de Haas experiment and the anomalous Zeeman effect. The magnetic moment for a state Ψ is defined as

$$\mathbf{m} = \langle\Psi, \hat{\mathbf{m}}\Psi\rangle. \tag{10.28}$$

Finally, the Landé factor g for a state Ψ is defined by

$$\frac{\mathbf{m}_3}{\mathbf{J}_3} = g\frac{e}{2\mathrm{m}c}; \tag{10.29}$$

this factor (or rather its maximum value) was measured in the Einstein–de Haas experiment.

10.3.2 Quantum Numbers

The momenta $\hat{\mathbf{L}}_3$ and $\hat{\mathbf{s}}_3$ do not commute with $\hat{\mathbf{s}}\hat{\mathbf{L}}$ (and hence with $\tilde{\mathcal{P}}$). Thus, we cannot use the classification of type (10.23) for the stationary states by ‘quantum numbers' $n, l, m, \pm 1$, which correspond to eigenvalues of operators $\tilde{\mathcal{P}}, \hat{\mathbf{L}}^2, \hat{\mathbf{L}}_3, \hat{\mathbf{s}}_3$. We should choose another quantum numbers.

Exercise 10.6 Verify that the operators $\mathcal{P}_m, \hat{\mathbf{L}}^2, \hat{\mathbf{J}}^2, \hat{\mathbf{J}}_3$ commute with each other.
Hints:

(i) (10.26) implies that

$$\tilde{\mathcal{P}} = -\frac{1}{2\mathrm{m}}\hbar^2\Delta + e\phi\big(|\mathbf{x}|\big) + \frac{e}{2\mathrm{m}^2c^2}\frac{\phi'(|\mathbf{x}|)}{|\mathbf{x}|}\hat{\mathbf{s}}\hat{\mathbf{L}} - \frac{e}{2\mathrm{m}c}B[\hat{\mathbf{L}}_3 + 2\hat{\mathbf{s}}_3]. \tag{10.30}$$

(ii) $\hat{\mathbf{L}}\hat{\mathbf{s}}$ commutes with $\hat{\mathbf{L}}^2$, since each $\hat{\mathbf{L}}_n$ and $\hat{\mathbf{s}}_n$ commutes with $\hat{\mathbf{L}}^2$.
(iii) $\hat{\mathbf{J}}^2$ commutes with $\hat{\mathbf{L}}^2$, since $\hat{\mathbf{J}}^2 = \hat{\mathbf{L}}^2 + 2\hat{\mathbf{s}}\hat{\mathbf{L}} + \hat{\mathbf{s}}^2$.
(iv) $\hat{\mathbf{s}}\hat{\mathbf{L}}$ commutes with $\hat{\mathbf{J}}^2$ and $\hat{\mathbf{J}}_3$, since $2\hat{\mathbf{s}}\hat{\mathbf{L}} = \hat{\mathbf{J}}^2 - \hat{\mathbf{L}}^2 - \hat{\mathbf{s}}^2$.

Therefore, the operators $\tilde{\mathcal{P}}, \hat{\mathbf{L}}^2, \hat{\mathbf{J}}^2, \hat{\mathbf{J}}_3$ in space $\mathcal{E}^{\otimes} = \mathcal{E} \otimes \mathbb{C}^2$ can be simultaneously diagonalized. Let us recall that possible eigenvalues of these operators are, respectively, $E, \hbar^2 L(L+1), \hbar^2 J(J+1)$ and $\hbar M$, where $J = 0, \frac{1}{2}, 1, \frac{3}{2}, \dots$ and $M = -J, \dots, J$ (see Sect. 6.2).

10.3.3 The Landé Formula

Let $|E, L, J, M\rangle$ be the common eigenvector, corresponding to the eigenvalues E, $\hbar^2 L(L+1)$, $\hbar^2 J(J+1)$, $\hbar M$ of the operators $\mathcal{P}, \hat{\mathbf{L}}^2, \hat{\mathbf{J}}^2, \hat{\mathbf{J}}_3$, respectively.

Theorem 10.7 *The Landé factor $g = g(\Psi)$, corresponding to the stationary state $\Psi = |E, L, J, M\rangle$, is given by*

$$g = \frac{3}{2} + \frac{3/4 - L(L+1)}{2J(J+1)}. \tag{10.31}$$

Proof We have $\hat{\mathbf{J}}^2\Psi = \hbar^2 J(J+1)\Psi$ and $\hat{\mathbf{J}}_3\Psi = \hbar M\Psi$. Therefore Ψ is an element e_M of the corresponding canonical basis $e_{-J}, \dots, e_J$, as constructed in Proposition 6.12, and so all basis vectors $e_{M'}$ are obtained from e_M by application of the operators $\hat{\mathbf{J}}_\pm := \hat{\mathbf{J}}_1 \pm i\hat{\mathbf{J}}_2$.

For a linear operator A in $\mathcal{E}^{\otimes}$, let $A^{M,M'} = \langle A e_M, e_{M'}\rangle$ be its matrix element. Then (10.15) and (10.16) imply that

$$\mathbf{m}_3 = \hat{\mathbf{m}}_3^{M,M} = \frac{e}{2\mathrm{m}c}\big(\hat{\mathbf{J}}_3^{M,M} + \hat{\mathbf{s}}_3^{M,M}\big) = \frac{e}{2\mathrm{m}c}\big(\hbar M + \hat{\mathbf{s}}_3^{M,M}\big). \tag{10.32}$$

It remains to find the matrix element $\hat{\mathbf{s}}_3^{M,M}$ and calculate g from definition (10.29). Let us collect the commutators

$$\left\{\begin{array}{lll} [\hat{\mathbf{J}}_1,\hat{\mathbf{s}}_1]=0, & [\hat{\mathbf{J}}_1,\hat{\mathbf{s}}_2]=i\hbar\hat{\mathbf{s}}_3, & [\hat{\mathbf{J}}_1,\hat{\mathbf{s}}_3]=-i\hbar\hat{\mathbf{s}}_2, \\ [\hat{\mathbf{J}}_2,\hat{\mathbf{s}}_2]=0, & [\hat{\mathbf{J}}_2,\hat{\mathbf{s}}_3]=i\hbar\hat{\mathbf{s}}_1, & [\hat{\mathbf{J}}_2,\hat{\mathbf{s}}_1]=-i\hbar\hat{\mathbf{s}}_3, \\ [\hat{\mathbf{J}}_3,\hat{\mathbf{s}}_3]=0, & [\hat{\mathbf{J}}_3,\hat{\mathbf{s}}_1]=i\hbar\hat{\mathbf{s}}_2, & [\hat{\mathbf{J}}_3,\hat{\mathbf{s}}_2]=-i\hbar\hat{\mathbf{s}}_1, \end{array}\right. \tag{10.33}$$

where the second and the third line follow from the first one by cyclic permutations. Let us use (10.33) to calculate the commutators of $\hat{\mathbf{J}}_\pm$ with $\hat{\mathbf{s}}_\pm := \hat{\mathbf{s}}_1 \pm i\hat{\mathbf{s}}_2$. We have

$$[\hat{\mathbf{J}}_-,\hat{\mathbf{s}}_+]=-2\hbar\hat{\mathbf{s}}_3, \qquad [\hat{\mathbf{J}}_+,\hat{\mathbf{s}}_+]=0. \tag{10.34}$$

The first formula implies the identity

$$\hat{\mathbf{J}}_-^{M,M+1}\hat{\mathbf{s}}_+^{M+1,M}-\hat{\mathbf{s}}_+^{M,M-1}\hat{\mathbf{J}}_-^{M-1,M}=-2\hbar\hat{\mathbf{s}}_3^{M,M}, \tag{10.35}$$

because all matrix elements $\hat{\mathbf{J}}_-^{M',M''}$ with $M''\neq M'+1$ vanish by (6.48). The nonzero matrix elements of $\hat{\mathbf{J}}_\pm$ are known from (6.48) and (6.44):

$$\hat{\mathbf{J}}_+^{M+1,M}=\hbar\sqrt{(J-M)(J+M+1)}, \qquad \hat{\mathbf{J}}_-^{M-1,M}=\hbar\sqrt{(J+M)(J-M+1)}. \tag{10.36}$$

On the other hand, $\hat{\mathbf{s}}_+$ can be obtained from the second identity (10.34): taking matrix element $(\cdot)^{M+1,M-1}$ of both sides, we get

$$\hat{\mathbf{J}}_+^{M+1,M}\hat{\mathbf{s}}_+^{M,M-1}-\hat{\mathbf{s}}_+^{M+1,M}\hat{\mathbf{J}}_+^{M,M-1}=0. \tag{10.37}$$

Then (10.36) implies that

$$\frac{\hat{\mathbf{s}}_+^{M+1,M}}{\sqrt{(J-M)(J+M+1)}}=\frac{\hat{\mathbf{s}}_+^{M,M-1}}{\sqrt{(J-M+1)(J+M)}}=:a. \tag{10.38}$$

Substituting this into (10.35) and using (10.36) for the matrix elements of $\hat{\mathbf{J}}_-$, it follows that

$$\hat{\mathbf{s}}_3^{M,M}=aM. \tag{10.39}$$

It remains to calculate a. We start with the identity

$$\hat{\mathbf{J}}^2=(\hat{\mathbf{L}}+\hat{\mathbf{s}})^2=\hat{\mathbf{L}}^2+2\hat{\mathbf{s}}\hat{\mathbf{L}}+\hat{\mathbf{s}}^2=\hat{\mathbf{L}}^2+2\hat{\mathbf{s}}\hat{\mathbf{J}}-\hat{\mathbf{s}}^2. \tag{10.40}$$

Hence

$$(\hat{\mathbf{s}}\hat{\mathbf{J}})^{M,M}=\hbar^2\frac{J(J+1)-L(L+1)+3/4}{2}, \tag{10.41}$$

because $\hat{\mathbf{s}}^2=3/4$. On the other hand, the same matrix element can be expressed from a different identity

$$2\hat{\mathbf{s}}\hat{\mathbf{J}}=\hat{\mathbf{s}}_+\hat{\mathbf{J}}_-+\hat{\mathbf{s}}_-\hat{\mathbf{J}}_++2\hat{\mathbf{s}}_3\hat{\mathbf{J}}_3, \tag{10.42}$$

which directly follows from the definitions of $\hat{\mathbf{s}}_\pm$ and $\hat{\mathbf{J}}_\pm$ and from the commutations $[\hat{s}_k, \hat{\mathbf{J}}_k] = 0$. Now

$$(\hat{\mathbf{s}}\hat{\mathbf{J}})^{M,M} = \frac{1}{2}\hat{\mathbf{s}}_+^{M,M-1}\hat{\mathbf{J}}_-^{M-1,M} + \frac{1}{2}\hat{\mathbf{s}}_-^{M,M+1}\hat{\mathbf{J}}_+^{M+1,M} + \hbar M \hat{\mathbf{s}}_3^{M,M}. \tag{10.43}$$

By (10.41), we have

$$\hbar^2 \frac{J(J+1) - L(L+1) + 3/4}{2} = \frac{1}{2}\hat{\mathbf{s}}_+^{M,M-1}\hat{\mathbf{J}}_-^{M-1,M} + \frac{1}{2}\hat{\mathbf{s}}_-^{M,M+1}\hat{\mathbf{J}}_+^{M+1,M} + \hbar M \hat{\mathbf{s}}_3^{M,M}. \tag{10.44}$$

At last, the matrix elements $\hat{\mathbf{s}}_+^{M+1,M}$ and $\hat{\mathbf{s}}_-^{M,M+1}$ are complex conjugates. Therefore, using (10.38), this establishes

$$\hat{\mathbf{s}}_+^{M+1,M} = a\sqrt{(J-M)(J+M+1)} = \hat{\mathbf{s}}_-^{M,M+1}, \tag{10.45}$$

because the constant a is real by (10.39).

Finally, substituting (10.45), (10.39) and (10.36) into (10.44), gives an equation for a, which implies that

$$a = \hbar \frac{J(J+1) - L(L+1) + 3/4}{2J(J+1)}. \tag{10.46}$$

Now, from (10.39) and (10.32) we obtain

$$\mathbf{m}_3 = \frac{e}{2\mathrm{m}c}(\hbar M + aM) = \frac{e}{2\mathrm{m}c}(1 + a/\hbar)\hbar M = \frac{e}{2\mathrm{m}c}(1 + a/\hbar)\mathbf{J}_3. \tag{10.47}$$

Substituting (10.46) for a, we get the ratio (10.29), where g coincides with (10.31). ☐

Remark 10.8

(i) Our proof follows the calculations of [133].
(ii) Landé [128] was the first to obtain formula (10.31) using the phenomenological *vector model* and the Bohr correspondence principle (see Sect. 14.7).

10.3.4 Applications of the Landé Formula

The Landé formula (10.31) implies the following fundamental results.

The Einstein–de Haas Experiment

Formula (10.31) for the Landé factor in (10.29) is in excellent agreement with experimental observations of the ratio (10.29) in the Einstein–de Haas experiment (see [12, 20, 34] and [171, Vol. I]). This agreement was one of the greatest successes of quantum theory.

The Anomalous Zeeman Effect

Formula (10.31) allows one, *inter alia*, to explain the multiplet structure in the anomalous Zeeman effect by splitting of the eigenvalues for (10.27) in the magnetic field $(0, 0, B)$ with small $|B|$. Namely, applying the perturbation theory (3.43), we obtain the resulting correction to the frequency $\omega = E/\hbar$ of the eigenstate $|E, L, J, M\rangle$,

$$\Delta\omega_{L,J,M} = -\hat{\mathbf{m}}_3^{M,M} B/\hbar = -\mathbf{m}_3 B/\hbar = -g\frac{e}{2\mathrm{m}c} MB = -g\omega_{\mathcal{L}} M \tag{10.48}$$

accurate to order $\mathcal{O}(B^2)$. Hence, the splitted spectral lines now become

$$\begin{aligned}\omega_{jj'} &= \omega^0_{jj'} - \omega_{\mathcal{L}}\big[g(j)M - g\big(j'\big)M'\big] \\ &= \omega^0_{jj'} - \omega_{\mathcal{L}} g(j)\big[M - M'\big] - \omega_{\mathcal{L}}\big[g(j) - g\big(j'\big)\big]M',\end{aligned} \tag{10.49}$$

where $j = (E, L, J, M)$, $j' = (E', L', J', M')$ and $\omega^0_{jj'}$ is the 'unperturbed' spectral line corresponding to $B = 0$.

Finally, we should take into account the selection rules

$$J \to J' = J \pm 1, \qquad M \to M' = M, M \pm 1$$

which follow similarity to (7.60) by following two arguments:

(i) The contribution of the last term (10.22) into the second integral of (7.26) decays like $|\mathrm{x}|^{-2}$.

(ii) The contribution of the first term depends only on the angular factors of the eigenfunctions according to (7.62), while the eigenfunctions have the same structure (7.59) with new radial functions $R_{nl}(r), R_{n'l'}(r) \in \mathbb{C}^2$.

Now, in contrast to (9.14), the magnitude of splitting depends on all quantum numbers L, J, M, L', J', M', by (10.31). As a result, the splitted spectral lines do not reduce to the normal triplet (9.14).

Chapter 11
Relativistic Quantum Mechanics

Abstract In 1928, Dirac introduced the new wave equation, which was a relativistic covariant generalization of the Schrödinger equation. In this case, the wave function has 4 complex components and the coefficients of the equation are 4×4 *Dirac matrices*.

Dirac calculated the first two approximations of this equation in the limit as $c \to \infty$. The approximation up to order $1/c$ coincides with the Pauli equation, while the second approximation displays, up to order $1/c^2$, the Russell–Sounders spin-orbital coupling, as well as some other effects.

The resulting equation admits the Lagrangian and Hamiltonian formulations, which provide the corresponding conserved observables and the continuity equation for charge and current. The angular momentum automatically includes the spin component with the factor $1/2$, as predicted by Goudsmit and Uhlenbeck.

A few months later, Gordon and Darwin independently solved the Dirac spectral problem for Hydrogen. Now the energies depend on the angular momentum, in contrast to the nonrelativistic case. This dependence perfectly explains the 'fine structure' of the spectrum.

The energy for the Dirac equation is not bounded from above and from below, suggesting instability of solutions. This problem was solved in quantum electrodynamics by imposing anticommutation relations for the electron field.

11.1 The Free Dirac Equation

It is also worth extending the Special Relativity of Einstein to the quantum mechanical theory. Containing the first order derivatives in time and the second order derivatives in $\mathbf{x}$, the Schrödinger equation is obviously non-invariant with respect to the Lorentz group. One possible approach is to employ the Klein–Gordon equation

$$\frac{\hbar^2}{c^2}\partial_t^2\psi(x) = \hbar^2\Delta\psi(x) - \mathrm{m}^2c^2\psi(x), \quad x \in \mathbb{R}^4, \tag{11.1}$$

which is Lorentz-invariant, like the wave equation (12.26). However, this approach results in negative energies, a thing which is physically unsatisfactory. Namely, the

A. Komech, *Quantum Mechanics: Genesis and Achievements*,
DOI 10.1007/978-94-007-5542-0_11,

Klein–Gordon equation follows from the relativistic energy-momentum relation

$$\frac{E^2}{c^2} = \mathbf{p}^2 + \mathrm{m}^2 c^2 \tag{11.2}$$

by the canonical quantization (3.26). In the Fourier transform,

$$\hat{\psi}(t, \mathbf{p}) := \int e^{i\frac{\mathbf{px}}{\hbar}} \psi(t, \mathbf{x})\, d\mathbf{x},$$

the Klein–Gordon equation becomes the ordinary differential equation with the parameter $\mathbf{p} \in \mathbb{R}^3$,

$$\hbar^2 \partial_0^2 \hat{\psi}\big(x^0, \mathbf{p}\big) = -\mathbf{p}^2 \hat{\psi}\big(x^0, \mathbf{p}\big) - \mathrm{m}^2 c^2 \hat{\psi}\big(x^0, \mathbf{p}\big), \quad x^0 \in \mathbb{R}, \tag{11.3}$$

where $x^0 := ct$ and $\partial_0 := \partial/\partial_{x^0}$. Hence, the solutions are linear combinations of $e^{i\frac{E}{\hbar}t}$, where $E/c = \pm\sqrt{\mathbf{p}^2 + \mathrm{m}^2 c^2}$ as in (11.2). The solutions with $E/c = -\sqrt{\mathbf{p}^2 + \mathrm{m}^2 c^2}$ seem to correspond to negative energies (unbounded from below), and hence the physical interpretation of the Klein–Gordon equation requires an additional argument.

This is why Dirac introduced in [45] the relativistic invariant equation of the first order in time, like the Schrödinger equation, to avoid having negative roots. The relativistic invariance also requires the first order in space; i.e.,

$$\gamma^\alpha P_\alpha \psi(x) = \mathrm{m}c\psi(x), \quad x \in \mathbb{R}^4 \tag{11.4}$$

with $P := (i\hbar\partial_0, i\hbar\nabla)$. The corresponding energy-momentum relation can be written as follows

$$\gamma(p) = \mathrm{m}c, \quad p \in \mathbb{R}^4; \tag{11.5}$$

here $\gamma(p)$ is the linear form $\gamma^\alpha p_\alpha$ and $p := (E/c, -\mathbf{p})$, according to (3.26).

The last requirement is the 'correspondence principle': equation (11.4) should imply the Klein–Gordon equation (11.1). Namely, applying operator $\gamma^\alpha P_\alpha$ to both sides of (11.4), we get

$$\big[\gamma^\alpha P_\alpha\big]^2 \psi(x) = \mathrm{m}^2 c^2 \psi(x), \quad x \in \mathbb{R}^4. \tag{11.6}$$

Hence, the correspondence principle is equivalent to algebraic identity

$$\big[\gamma(p)\big]^2 = p_0^2 - \mathbf{p}^2, \quad p = (p_0, \mathbf{p}) \in \mathbb{R}^4 \tag{11.7}$$

since (11.2) can be written as $p_0^2 - \mathbf{p}^2 = \mathrm{m}^2 c^2$. Dirac's extra idea was the choice of the coefficients γ^α in a matrix algebra since scalar coefficients do not exist. The existence of scalar coefficients would mean that the polynomial $p_0^2 - \mathbf{p}^2$ is reducible which is not true.

Exercise 11.1 Verify that (11.7) is impossible with scalar coefficients γ^α.

Lemma 11.2 *In* 2×2 *block form, the matrices*

$$\gamma(p) = \begin{pmatrix} p_0 & \sigma \cdot \mathbf{p} \\ -\sigma \cdot \mathbf{p} & -p_0 \end{pmatrix} \tag{11.8}$$

satisfy identity (11.7), *where* $\sigma := (\sigma_1, \sigma_2, \sigma_3)$ *are the Pauli spin matrices.*

Proof Direct multiplication of 2×2 block matrices shows that

$$\gamma^2(p) = \begin{pmatrix} p_0^2 - (\sigma \cdot \mathbf{p})^2 & 0 \\ 0 & p_0^2 - (\sigma \cdot \mathbf{p})^2 \end{pmatrix} \tag{11.9}$$

It remains to invoke the identity $(\sigma \cdot \mathbf{p})^2 = \mathbf{p}^2$. □

Now we can calculate matrices $\gamma^\alpha = \gamma(e_\alpha)$, where $e_0 = (1, 0, 0, 0)$, etc. From (11.8), we get

$$\gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad \gamma^k = \begin{pmatrix} 0 & \sigma_k \\ -\sigma_k & 0 \end{pmatrix}, \quad k = 1, 2, 3. \tag{11.10}$$

Lemma 11.3 *Matrices* γ^α *satisfy the relations*

$$\left(\gamma^0\right)^2 = 1, \qquad \left(\gamma^j\right)^2 = -1, \quad j = 1, 2, 3; \qquad \gamma^\alpha\gamma^\beta + \gamma^\beta\gamma^\alpha = 0, \quad \alpha \neq \beta. \tag{11.11}$$

Proof Rewriting (11.7) in the form

$$\gamma^2(p) = g(p) := p_0^2 - \mathbf{p}^2, \quad p \in \mathbb{R}^4, \tag{11.12}$$

we get

$$\gamma(p)\gamma(q) + \gamma(q)\gamma(p) = 2g(p, q), \tag{11.13}$$

where $g(p, q) = p_0 q_0 - \mathbf{pq}$ is the corresponding symmetric bilinear form. In particular, for $p = e_\alpha$ and $q = e_\beta$, we obtain

$$\gamma^\alpha\gamma^\beta + \gamma^\beta\gamma^\alpha = 2g(e_\alpha, e_\beta),$$

which implies (11.11). □

Matrices (11.10) are known as the ordinary, or standard representation (solution) of Eqs. (11.11). It is easily checked that the matrices are not unique solutions: for example, we can replace γ^α by $-\gamma^\alpha$ for certain indexes α. Below we shall prove the Pauli Theorem: the matrices γ^α are unique up to a change of the orthonormal basis e_α.

Let us rewrite the Dirac equation (11.4) in the 'Schrödinger form'. First, it is equivalent to

$$i\hbar\gamma^0\partial_t\psi(x) = c\big(\mathrm{m}c - i\hbar\gamma^k\partial_k\big)\psi(x). \tag{11.14}$$

Second, multiplying by γ^0, we obtain

$$i\hbar\partial_t\psi(x) = \gamma^0\big(\mathrm{m}c^2 - ic\hbar\gamma^k\partial_k\big)\psi(x). \tag{11.15}$$

We set

$$\alpha_k := \gamma^0\gamma^k = \begin{pmatrix} 0 & \sigma_k \\ \sigma_k & 0 \end{pmatrix}, \quad k = 1, 2, 3; \qquad \beta := \gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{11.16}$$

Then (11.15) reads

$$i\hbar\partial_t\psi(x) = H_D^0\psi(x) := \big(\mathrm{m}c^2\beta - ic\hbar\alpha^k\partial_k\big)\psi(x), \tag{11.17}$$

where the operator H_D^0 is called the *free Dirac Hamiltonian*. The matrices β and α^k are Hermitian, and hence operator H_D^0 is symmetric in the complex Hilbert space $L^2(\mathbb{R}^3, \mathbb{C}^4)$.

11.2 Pauli's Theorem

The relativistic covariance of the Dirac equation should rely on a complete description of all possible solutions to the commutation relations (11.11). The following *Pauli's theorem* states that all solutions of (11.11) are equivalent to standard Dirac matrices (11.10).

Theorem 11.4 *Let γ^α, $\alpha = 0, \ldots, 3$ be operators on a finite-dimensional complex vector space V of dimension* $\dim V \le 4$. *Suppose that they satisfy relations* (11.11). *Then* $\dim V = 4$, *and the operators γ^α, for a suitable choice of a basis, can be exhibited in the matrix form* (11.10).

Proof Step (i) The key idea of the proof is the following simple characterization of basis vectors. Namely, for the standard Dirac matrices (11.10), the matrix $\gamma^1\gamma^2$ is diagonal, and hence it commutes with γ^0, which is also diagonal: γ^0 and $\gamma^1\gamma^2$ are block diagonal matrices:

$$\gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad \gamma^1\gamma^2 = \begin{pmatrix} -i\sigma_3 & 0 \\ 0 & -i\sigma_3 \end{pmatrix}. \tag{11.18}$$

Therefore, the basis vectors $e_0, \ldots, e_3$ are common eigenvectors of the matrices γ^0 and $\gamma^1\gamma^2$ with eigenvalues 1 and $-i$, 1 and i, -1 and $-i$, -1 and i, respectively.

Step (ii) Now we apply this observation to general matrices γ^α of Theorem 11.4. Using (11.11), it follows from the anticommutation relations that the matrices γ^0 and $\gamma^1\gamma^2$ commute with each other:

$$\gamma^0\gamma^1\gamma^2 = \gamma^1\gamma^2\gamma^0. \tag{11.19}$$

Hence there exists at least one common eigenvector v_1 for both of those (since V is a *complex vector space*!):

$$\gamma^0 v_1 = \alpha v_1 \quad \text{and} \quad \gamma^1\gamma^2 v_1 = \beta v_1, \tag{11.20}$$

where α and β are suitable complex numbers.

Step (iii) We have $\alpha^2 = 1$, since $(\gamma^0)^2 = 1$, and similarly, $\beta^2 = -1$, since $(\gamma^1\gamma^2)^2 = -1$. Hence, $\alpha = \pm 1$ and $\beta = \pm i$. Let us check that all four combinations of the signs are possible for suitable eigenvectors v_1. Namely,

$$\gamma^0\gamma^3 v_1 = -\gamma^3\gamma^0 v_1 = -\alpha\gamma^3 v_1$$

and

$$\gamma^1\gamma^2\gamma^3 v_1 = \gamma^3\gamma^1\gamma^2 v_1 = \beta\gamma^3 v_1.$$

Hence, the vector $v_3 := \gamma^3 v_1$ is also a common eigenvector with eigenvalues $-\alpha$ and β. Similarly, $v_2 := -\gamma^3\gamma^1 v_1$ and $v_4 := -\gamma^1 v_1$ are, respectively, common eigenvectors of γ^0 and $\gamma^1\gamma^2$ with eigenvalues α, $-\beta$ and $-\alpha$, $-\beta$, respectively. Since all four possible signs are disposed of, we may permute the four vectors to ensure that $\alpha = 1$ and $\beta = -i$.

Step (iv) The vectors v_1, v_2, v_3, v_4 are linearly independent, and hence generate V. In the basis v_1, v_2, v_3, v_4, the operators γ^0 and $\gamma^1\gamma^2$ are block diagonal matrices of form (11.18). Moreover, in this basis, the operators γ^1 and γ^3 have the form

$$\gamma^1 = \begin{pmatrix} 0 & \sigma_1 \\ -\sigma_1 & 0 \end{pmatrix}, \qquad \gamma^3 = \begin{pmatrix} 0 & \sigma_3 \\ -\sigma_3 & 0 \end{pmatrix}, \tag{11.21}$$

which coincide with (11.10). Hence, $\gamma^2 = -\gamma^1(\gamma^1\gamma^2)$ also has the desired form. □

Exercise 11.5 Verify that v_1, v_2, v_3, v_4 are linearly independent. **Hint:** $\gamma^0 = 1$ in the span (v_1, v_2), and $\gamma^0 = -1$, in the span (v_3, v_4), while $\gamma^1\gamma^2 = -i$ in the span (v_1, v_3), and $\gamma^1\gamma^2 = i$, in the span (v_2, v_4).

Exercise 11.6 Check (11.21).

Corollary 11.7 *For any Lorentz transformation* Λ (*see Sect.* 12.3.1), *there exists a nondegenerate matrix* $\Gamma(\Lambda) \in GL(4, \mathbb{C})$ *such that*

$$\gamma(\Lambda p) = \Gamma(\Lambda)\gamma(p)\Gamma^{-1}(\Lambda), \quad p \in \mathbb{R}^4. \tag{11.22}$$

Proof From (11.12) we have

$$\gamma^2(\Lambda p) = g(\Lambda p) = g(p), \quad p \in \mathbb{R}^4, \tag{11.23}$$

since Λ is a Lorentz transformation. Hence, the matrices $\gamma(\Lambda e_\alpha)$ satisfy relations (11.11), and also $\gamma^\alpha := \gamma(e_\alpha)$. Therefore, by Pauli's theorem,

$$\gamma(\Lambda e_\alpha) = \Gamma(\Lambda)\gamma(e_\alpha)\Gamma^{-1}(\Lambda), \quad \alpha = 0, \dots, 3, \tag{11.24}$$

where $\Gamma(\Lambda)$ is an invertible operator in $\mathbb{R}^4$ (which maps the vector e_α into v_α, $\alpha = 0, \dots, 3$). Then (11.22) follows by linearity. □

11.3 Lorentz Covariance

The Einstein postulate (12.24) of Special Relativity extends to the quantum mechanics as follows:

The laws of quantum mechanics take identical form

in all Inertial Frames. (11.25)

The next theorem ensures the existence of the corresponding transformation for wave functions, leaving the Dirac equation invariant. Thus the Dirac equation provides relativistic invariant quantum theory.

We consider two frames of reference related by the Lorentz transformation: $x' = \Lambda x$.

Theorem 11.8 *Let $\psi(x)$ be a solution of the Dirac equation* (11.4). *Suppose that*

$$\psi'(x') := \Gamma(\Lambda^\#)\psi(x), \quad x \in \mathbb{R}^4, \tag{11.26}$$

where $x' = \Lambda x$, $\Lambda^\# := (\Lambda^t)^{-1}$, and Λ^t is the transposed matrix Λ. Then the function $\psi'(x')$ is also a solution of the Dirac equation.

Proof In the Fourier transform, the Dirac equation (11.4) reads

$$\hat\psi(p) := \int e^{\frac{ipx}{\hbar}} \psi(x)\, dx, \quad p \in \mathbb{R}^4 \tag{11.27}$$

where $px := p_\alpha x^\alpha$. Then (11.4) implies

$$\gamma(p)\hat\psi(p) = \mathrm{m}c\hat\psi(p), \quad p \in \mathbb{R}^4. \tag{11.28}$$

The Fourier transform translates (11.26) into

$$\hat\psi'(p') = \Gamma(\Lambda^\#)\hat\psi(p), \quad p \in \mathbb{R}^4, \tag{11.29}$$

where $p = \Lambda^t p'$.

Exercise 11.9 Verify (11.29). **Hint:** We formally have

$$\hat{\psi}'(p') := \int e^{\frac{ip'x'}{\hbar}} \psi'(x')\, dx' = \int e^{\frac{ip'\Lambda x}{\hbar}} \Gamma(\Lambda^{\#})\psi(x) |\det \Lambda|\, dx$$

$$= \Gamma(\Lambda^{\#}) \int e^{\frac{i\Lambda^t p' x}{\hbar}} \psi(x)\, dx = \Gamma(\Lambda^{\#})\hat{\psi}(p), \tag{11.30}$$

because $|\det \Lambda| = 1$ for the Lorentz transformation Λ.

Now we express (11.28) in terms of the wave function $\hat{\psi}'(p')$:

$$\gamma(p)\Gamma^{-1}(\Lambda^{\#})\hat{\psi}'(p') = \mathrm{m}c\Gamma^{-1}(\Lambda^{\#})\hat{\psi}'(p'), \quad p' \in \mathbb{R}^4. \tag{11.31}$$

This identity is equivalent to the Dirac equation (11.28) if and only if

$$\Gamma(\Lambda^{\#})\gamma(p)\Gamma^{-1}(\Lambda^{\#}) = \gamma(p'), \quad p' \in \mathbb{R}^4. \tag{11.32}$$

Finally, this equation is equivalent to (11.22) with $\Lambda^{\#}$ instead of Λ, since $p' = \Lambda^{\#} p$. It remains to note that (11.22) with $\Lambda^{\#}$ holds, inasmuch as $\Lambda^{\#}$ also belongs to the Lorentz group. □

This theorem implies that transform (11.26) leaves the Dirac equation unchanged. In other words, the Dirac equation is *covariant* with respect to the Lorentz group.

Exercise 11.10 Verify that $\Lambda^{\#}$ is a Lorentz transformation for any $\Lambda \in L$.

Formal calculations in (11.30) are justified by properties of the Fourier transform of tempered distributions. This framework is necessary, because of the charge conservation law (see (11.56) below), since it implies that the integrals (11.30) never converge, for nonzero solutions.

11.4 The Angular Momentum

The conserved orbital momentum for the Schrödinger equation is defined by (3.82) and (3.83): $\mathbf{L} = \mathbf{L}(\psi) := \langle \psi, \hat{\mathbf{L}}\psi \rangle$; here $\hat{\mathbf{L}} = -i\hbar\mathbf{x} \times \nabla$. For solutions of the Dirac equation, the orbital momentum is generally not conserved, because the operator $\hat{\mathbf{L}}$ does not commute with the Dirac operator H_D^0. Hence, for the Dirac equation, the definition of the angular momentum requires a modification.

Definition 11.11 For the Dirac equation, the angular momentum is defined by $\mathbf{J} = \mathbf{J}(\psi) = \langle \psi, \hat{\mathbf{J}}\psi \rangle$, where

$$\hat{\mathbf{J}} := \hat{\mathbf{L}} + \frac{1}{2}\hbar\Sigma, \quad \Sigma := \begin{pmatrix} \sigma & 0 \\ 0 & \sigma \end{pmatrix}. \tag{11.33}$$

Theorem 11.12 *The angular momentum* **J** *is conserved for solutions of the Dirac equation,*

$$\mathbf{J}\big(\psi(t,\cdot)\big) = \text{const}, \qquad t \in \mathbb{R}. \tag{11.34}$$

Proof Differentiating, we obtain similarly to the Heisenberg equation (3.93),

$$\begin{aligned}\frac{d}{dt}\mathbf{J}\big(\psi(t)\big) &= \big\langle\dot\psi(t), \hat{\mathbf{J}}\psi(t)\big\rangle + \big\langle\psi(t), \hat{\mathbf{J}}\dot\psi(t)\big\rangle \\ &= -\left\langle\frac{i}{\hbar}H_D^0\psi(t), \hat{\mathbf{J}}\psi(t)\right\rangle - \left\langle\psi(t), \hat{\mathbf{J}}\frac{i}{\hbar}H_D^0\psi(t)\right\rangle \\ &= -\frac{i}{\hbar}\big[\big\langle H_D^0\psi(t), \hat{\mathbf{J}}\psi(t)\big\rangle - \big\langle\psi(t), \hat{\mathbf{J}}H_D^0\psi(t)\big\rangle\big] \\ &= -\frac{i}{\hbar}\big\langle\psi(t), \big[H_D^0, \hat{\mathbf{J}}\big]\psi(t)\big\rangle. \end{aligned} \tag{11.35}$$

It remains to verify the commutation

$$\big[H_D^0, \hat{\mathbf{J}}\big] = 0. \tag{11.36}$$

First note that we already have the commutators $[\hat{\mathbf{L}}_k, p_j] = i\hbar\epsilon_{kjl}p_l$ at our disposal, where $p_j := i\hbar\partial_j$ and ϵ_{kjl} is totally antisymmetric tensor. Therefore,

$$\big[\hat{\mathbf{L}}_k, H_D^0\big] = -c\hbar\gamma^0\gamma^j[\mathbf{L}_k, p_j] = -ic\hbar\gamma^0\gamma^j\epsilon_{kjl}p_l. \tag{11.37}$$

This shows that H_D^0 does not commute with the orbital angular momentum operators $\hat{\mathbf{L}}_k$, and hence the orbital momentum **L** is generally not conserved. It remains to calculate commutators $[\Sigma, H_D^0]$. Note that

$$\gamma^0\gamma^l = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}0 & \sigma_l\\ -\sigma_l & 0\end{pmatrix} = \begin{pmatrix}0 & \sigma_l\\ \sigma_l & 0\end{pmatrix}.$$

Hence,

$$\big[\Sigma_k, H_D^0\big] = \big[\Sigma_k, \mathrm{mc}^2\gamma^0 - c\gamma^0\gamma^l p_l\big] = -c\begin{pmatrix}0 & [\sigma_k, \sigma_l]\\ [\sigma_k, \sigma_l] & 0\end{pmatrix}p_l.$$

The commutation relations for the Pauli spin matrices allow us to reduce this to

$$-2ic\epsilon_{klj}\begin{pmatrix}0 & \sigma_j\\ \sigma_j & 0\end{pmatrix}p_l = -2ic\epsilon_{klj}\gamma^0\gamma^j p_l.$$

Multiplying by $\hbar/2$ and adding (11.37), we get commutation (11.36) by antisymmetry of ϵ_{klj}. □

This theorem justifies the Goudsmit–Uhlenbeck conjecture on electron spin as an intrinsic property of dynamical equations. The term $\frac{1}{2}\hbar\Sigma$ in (11.33) represents an intrinsic spinor angular momentum of the relativistic electron.

Moreover, we will see below that the coupling of the Dirac equation to the magnetic field provides the automatically correct Landé factor $g = 2$ for the spinor angular momentum in a nonrelativistic approximation.

These facts were the great triumph of the Dirac theory, suggesting that the electron spin has a genuine relativistic nature.

11.5 Negative Energies

The Dirac equation (11.17) is a Hamiltonian system with the Hamilton functional

$$\mathcal{H}^0(\psi) := \langle \psi(\mathbf{x}), H_D^0 \psi(\mathbf{x}) \rangle, \tag{11.38}$$

as analogous to the Schrödinger equation. It is conserved; i.e.,

$$\mathcal{H}^0\big(\psi(t,\cdot)\big) = \text{const}, \quad t \in \mathbb{R}, \tag{11.39}$$

for solutions of (11.15).

Exercise 11.13 Verify (11.39). **Hint:** Differentiate (11.38) and use the Dirac equation (11.17) and the symmetry of the Dirac operator H_D^0.

Let us check that energy for the Dirac equation is not bounded from above and from below. Namely, the quadratic form (11.38) is not positive definite. To see this it is useful to split each Dirac spinor into a pair of two-component vectors

$$\psi(\mathbf{x}) = \begin{pmatrix} \psi_+(\mathbf{x}) \\ \psi_-(\mathbf{x}) \end{pmatrix}, \tag{11.40}$$

where $\psi_+ := \binom{\psi_1}{\psi_2}$ and $\psi_- := \binom{\psi_3}{\psi_4}$. Let us define the Fourier transform

$$\hat{\psi}(\mathbf{p}) = \int e^{i\mathbf{px}/\hbar} \psi(\mathbf{x})\, d\mathbf{x}, \quad \mathbf{p} \in \mathbb{R}^3. \tag{11.41}$$

With application of the Plancherel identity the quadratic form (11.38) can be written

$$\begin{aligned}
\mathcal{H}^0(\psi) &= (2\pi)^{-3} \left\langle \hat{\psi}(\mathbf{p}), c\gamma^0 \begin{pmatrix} \mathrm{m}c & -\sigma\cdot\mathbf{p} \\ \sigma\cdot\mathbf{p} & \mathrm{m}c \end{pmatrix} \hat{\psi}(\mathbf{p}) \right\rangle \\
&= c(2\pi)^{-3} \left\langle \hat{\psi}(\mathbf{p}), \begin{pmatrix} \mathrm{m}c & -\sigma\cdot\mathbf{p} \\ -\sigma\cdot\mathbf{p} & -\mathrm{m}c \end{pmatrix} \hat{\psi}(\mathbf{p}) \right\rangle \\
&= c(2\pi)^{-3} \big[\mathrm{m}c \langle \hat{\psi}_+(\mathbf{p}), \hat{\psi}_+(\mathbf{p}) \rangle - 2 \langle \hat{\psi}_+(\mathbf{p}), \sigma\cdot\mathbf{p}\hat{\psi}_-(\mathbf{p}) \rangle \\
&\quad - \mathrm{m}c \langle \hat{\psi}_-(\mathbf{p}), \hat{\psi}_-(\mathbf{p}) \rangle \big].
\end{aligned}$$

In particular,

$$\begin{aligned}\mathcal{H}^0\begin{pmatrix}\psi_+(\mathbf{x})\\ 0\end{pmatrix} &= (2\pi)^{-3}\mathrm{m}c^2\big\langle\hat{\psi}_+(\mathbf{p}),\hat{\psi}_+(\mathbf{p})\big\rangle,\\ \mathcal{H}^0\begin{pmatrix}0\\ \psi_-(\mathbf{x})\end{pmatrix} &= -(2\pi)^{-3}\mathrm{m}c^2\big\langle\hat{\psi}_-(\mathbf{p}),\hat{\psi}_-(\mathbf{p})\big\rangle.\end{aligned}\tag{11.42}$$

The negative energy might lead to an instability of the Dirac dynamics due to a possible transition of the solution of the states $\binom{0}{\psi_-(\mathbf{x})}$, as a result of interaction via the Maxwell field (see the next section). On the other hand, this instability has never been proved.

Dirac suggested that the transition of all particles is forbidden by the Pauli Exclusion Principle, because almost all states with negative energy were occupied long time ago. On the other hand, by the Dirac theory, transitions for certain particles are possible, and the 'negative states' can be interpreted as states with positive energy for *antiparticles* which are *positrons* (i.e., positively charged electrons) [13]. Positrons were discovered experimentally in *cosmic rays* by Anderson in 1932 (for his discovery, Anderson received a half share of the 1936 Nobel Prize in physics).

The satisfactory solution of the problem of negative energies is provided by the *second quantization* in Quantum Electrodynamics.

11.6 The Maxwell–Dirac Equations

Interaction of the Dirac field with an external Maxwell field is defined by analogy with the Schrödinger equation, in order to keep the gauge invariance (4.25). We obtain the corresponding Hamiltonian and Lagrangian formulations and the coupled Dirac-Maxwell equations. Finally, we will obtain the charge conjugation, which transform the Dirac equation for electrons with negative charge $e<0$ into similar equation for *positrons* with positive charge $-e>0$.

11.6.1 The 'Minimal Coupling' to External Maxwell Field

The free Dirac equation (11.4) can be written as

$$\gamma(P)\psi(x)=\mathrm{m}c\psi(x),\tag{11.43}$$

where the differential operator $\gamma(P):=\gamma^\alpha P_\alpha$ is called the *Dirac operator* and $P_\alpha:=i\hbar\nabla_\alpha$.

We recall the Schrödinger equation (3.32) with the Maxwell field:

$$\big[i\hbar\partial_t-e\phi(t,\mathbf{x})\big]\psi(t,\mathbf{x})=\frac{1}{2\mathrm{m}}\Big[-i\hbar\boldsymbol{\nabla}-\frac{e}{c}\mathbf{A}(t,\mathbf{x})\Big]^2\psi(t,\mathbf{x}).\tag{11.44}$$

In the notation with P_α and $\mathcal{A}_\alpha = (\phi, -\mathbf{A})$ (see (12.49)), we have

$$c\left[P_0 - \frac{e}{c}\mathcal{A}_0(t,\mathbf{x})\right]\psi(t,\mathbf{x}) = \frac{1}{2\mathrm{m}}\sum_{k=1}^{3}\left[P_k - \frac{e}{c}\mathcal{A}_k(t,\mathbf{x})\right]^2\psi(t,\mathbf{x}). \tag{11.45}$$

This suggests the following generalization of the Dirac equation (11.43) for a relativistic electron in an external Maxwell field:

$$\gamma\left(P - \frac{e}{c}\mathcal{A}(x)\right)\psi(x) = \mathrm{m}c\psi(x), \quad x \in \mathbb{R}^4. \tag{11.46}$$

11.6.2 Gauge Transformation

We recall that the gauge transformation (see (12.16))

$$\phi(t,\mathbf{x}) \mapsto \phi(t,\mathbf{x}) + \frac{1}{c}\dot{\chi}(t,\mathbf{x}), \qquad \mathbf{A}(t,\mathbf{x}) \mapsto \mathbf{A}(t,\mathbf{x}) - \nabla\chi(t,\mathbf{x}), \tag{11.47}$$

where $\chi(t,\mathbf{x})$ is an arbitrary real function, does not change the Maxwell fields corresponding to potentials the $\phi(t,\mathbf{x})$ and $\mathbf{A}(t,\mathbf{x})$. In our notation,

$$\begin{aligned} \phi(x) &\mapsto \phi'(x) := \phi(x) + \tfrac{1}{i\hbar}P_0\chi(x), \\ \mathcal{A}_k(x) &\mapsto \mathcal{A}'_k(x) := \mathcal{A}_k(x) + \tfrac{1}{i\hbar}P_k\chi(x). \end{aligned} \tag{11.48}$$

Theorem 11.14 *Let $\psi(x)$ be a solution of the Dirac equation (11.46) with potentials $\mathcal{A}(x)$. Then $\psi'(x) := \exp(\frac{e\chi(x)}{i\hbar c})\psi(x)$ satisfies Eq. (11.46) with potentials $\mathcal{A}'(x)$.*

Proof This result follows by a direct calculation, because

$$P\exp\left(\frac{e\chi(x)}{i\hbar c}\right) = \exp\left(\frac{e\chi(x)}{i\hbar c}\right)\frac{eP\chi(x)}{i\hbar c}. \qquad \square$$

11.6.3 The Hamiltonian and Lagrangian Formalism

Similarly to (11.17), we can write (11.46) in the form

$$\begin{aligned} i\hbar\partial_t\psi(x) = \left[H_D(t)\psi\right](x) := \left[e\mathcal{A}_0(x) - c\alpha^k\left[i\hbar\partial_k - \frac{e}{c}\mathcal{A}_k(x)\right] + \mathrm{m}c^2\beta\right]\psi(x), \\ t = x_0/c, \end{aligned} \tag{11.49}$$

where the operator $H_D(t)$ is called the *Dirac Hamiltonian*. This operator is symmetric in the complex Hilbert space $L^2(\mathbb{R}^3, \mathbb{C}^4)$, as similar to the free Dirac Hamiltonian (11.17). Hence, (11.49) is the Hamilton equation with the Hamiltonian functional

$$\mathcal{H}\big(\psi(\cdot), t\big) = \big\langle \psi(\mathbf{x}), H_D(t)\psi(\mathbf{x})\big\rangle, \quad \psi \in L^2\big(\mathbb{R}^3, \mathbb{C}^4\big), \tag{11.50}$$

where $\langle \cdot, \cdot \rangle$ is the inner product on $L^2(\mathbb{R}^3, \mathbb{C}^4)$.

Let us find a Lagrangian density to rewrite (11.49) as the corresponding variational Euler–Lagrange equation of type (4.10). First, we write (11.49) as

$$L_D\psi(x) := i\hbar\partial_t\psi(x) - \big[H_D(t)\psi\big](x) = 0, \quad x \in \mathbb{R}^4, \tag{11.51}$$

similarly to Eq. (3.32). The linear operator L_D is symmetric in the complex Hilbert space $L^2(\mathbb{R}^4, \mathbb{C}^4)$, and hence the corresponding operator $\mathbb{L}_D$ in the real Hilbert space $L^2(\mathbb{R}^4, \mathbb{R}^8)$ is also symmetric. Hence, we can apply the methods of Sect. 4.1.3 to determine a Lagrangian density for the Dirac equation. Namely, (11.49) is equivalent to the variational equation (4.10), with action given by

$$S_\Omega(\psi) = \langle \Psi, \mathbb{L}\Psi \rangle = \int_\Omega \psi(x)^\dagger L_D\psi(x)\, dx, \tag{11.52}$$

where $\Psi(x) \in \mathbb{R}^8$ is the real form of the complex wave function $\psi(x) \in \mathbb{C}^4$ and $\psi(x)^\dagger$ denotes the conjugate transpose of $\psi(x)$. Hence, the corresponding Lagrangian density can be expressed in the complex field $\psi(x)$ as follows:

$$\begin{aligned}
&\mathcal{L}_D(x, \psi, \nabla\psi) \\
&\quad := \mathrm{Re}\big[\psi^\dagger L_D\psi\big] \\
&\quad = \mathrm{Re}\left[\psi^\dagger\left(i\hbar\partial_t\psi - e\mathcal{A}_0(x)\psi + c\alpha^k\left[i\hbar\partial_k\psi - \frac{e}{c}\mathcal{A}_k(x)\psi\right] - \mathrm{m}c^2\beta\psi\right)\right] \\
&\quad = c\,\mathrm{Re}\,\psi^\dagger\left[\alpha^\nu\left[i\hbar\partial_\nu\psi - \frac{e}{c}\mathcal{A}_\nu(x)\psi\right] - \mathrm{m}c\beta\psi\right] \\
&\quad = c\,\mathrm{Re}\,\overline{\psi}\left[\gamma^\nu\left[i\hbar\partial_\nu\psi - \frac{e}{c}\mathcal{A}_\nu(x)\psi\right] - \mathrm{m}c\psi\right],
\end{aligned} \tag{11.53}$$

where $\alpha_0 := I$, and $\overline{\psi} := \psi^\dagger\gamma^0$.

11.6.4 Charge and Current

The interaction term in the Lagrangian density (11.53) can be written in the form

$$-e\psi^\dagger\alpha^\nu\psi\mathcal{A}_\nu = -e\overline{\psi}\gamma^\nu\psi\mathcal{A}_\nu = -\mathcal{J}^\nu\mathcal{A}_\nu, \tag{11.54}$$

as analogous to (12.56). This suggests the identification $\mathcal{J}^\nu = (\rho, \mathbf{j}/c)$ where

$$\begin{aligned} \rho(x) &= e\psi^\dagger(x)\psi(x) = e\overline{\psi}(x)\gamma^0\psi(x), \\ \mathbf{j}^k(x) &= ec\psi^\dagger(x)\alpha^k\psi(x) = ec\overline{\psi}(x)\gamma^k\psi(x) \end{aligned} \tag{11.55}$$

which agrees with (12.59) and (13.99). Note that the charge density $\rho(x)$ is non-positive, because $e < 0$.

Lemma 11.15 *For any solution $\psi(x)$ of the Dirac equation (11.46), the four-current density satisfies the continuity equation*

$$\dot{\rho}(t, \mathbf{x}) + \operatorname{div} \mathbf{j}(t, \mathbf{x}) = 0. \tag{11.56}$$

Proof This identity follows from Corollary 13.26 and Lemma 13.28, because the Lagrangian density (11.53) is invariant under phase rotations $\psi \mapsto e^{is}\psi$ and since it is gauge invariant (see Sect. 13.4.4). □

Exercise 11.16 Prove (11.56) by direct computation. **Hints:**

(i) Differentiating, we obtain $\hbar\partial_t\rho = \hbar e[(\partial_t\psi)^\dagger\psi + \psi^\dagger(\partial_t\psi)] = ie[(H_D(t)\psi)^\dagger\psi - \psi^\dagger(H_D(t)\psi)]$.

(ii) Substitute (11.49) and use that potentials $\mathcal{A}_\nu$ are real, considering that the matrices β and α^k are Hermitian.

11.6.5 The Coupled Equations

The arguments of Sect. 4.2.1 suggest the following Lagrangian density $\mathcal{L}_{MD}$ for the coupled Maxwell–Dirac equations,

$$\begin{aligned} &\mathcal{L}_{MD}(x, \psi, \nabla\psi, \mathcal{A}_\mu, \nabla\mathcal{A}_\mu) \\ &\quad = \operatorname{Re}\left[c\psi^\dagger\alpha^\nu\left[i\hbar\partial_\nu\psi - \frac{e}{c}\mathcal{A}_\nu\psi\right] - \mathrm{m}c^2\psi^\dagger\beta\psi\right] - \frac{1}{16\pi}\mathcal{F}^{\mu\nu}\mathcal{F}_{\mu\nu}, \end{aligned} \tag{11.57}$$

where $\mathcal{F}^{\mu\nu} := \partial^\mu\mathcal{A}^\nu - \partial^\nu\mathcal{A}^\mu$ and $\mathcal{F}_{\mu\nu} := \partial_\mu\mathcal{A}_\nu - \partial_\nu\mathcal{A}_\mu$. The Lagrangian density differs from (11.53) only in the last term, which depends on the derivatives $\nabla\mathcal{A}_\mu$, but is independent of the potentials $\mathcal{A}_\mu$. Hence, formulas (11.55) for the charge and current densities remain true. As a result, the corresponding Euler–Lagrange equations read as follows (cf. (4.19)):

$$\begin{cases} i\hbar\partial_t\psi(x) = \left[e\mathcal{A}_0(x) - c\alpha^k\left[i\hbar\partial_k - \frac{e}{c}\mathcal{A}_k(x)\right] + \mathrm{m}c^2\beta\right]\psi(x), \\ \nabla_\mu\mathcal{F}^{\mu\nu}(x) = 4\pi e\psi^\dagger(x)\alpha^\nu\psi(x), \quad \nu = 0, \ldots, 3. \end{cases} \tag{11.58}$$

System (11.58) describes the dynamics of a Dirac wave field ψ in its 'own' induced Maxwell potentials $\mathcal{A}_\nu(x)$ generated by charges and currents of the wave field.

Now let us introduce the external potentials $\mathcal{A}_\nu^{\text{ext}}(x)$ of the Maxwell field generated by some external sources. The introduction is formalized by the Lagrangian density

$$\begin{aligned}\mathcal{L}_{MD} = \operatorname{Re}\Big[c\psi^\dagger\alpha^\nu\Big(i\hbar\partial_\nu\psi - \frac{e}{c}\big(\mathcal{A}_\nu(x) + \mathcal{A}_\nu^{\text{ext}}(x)\big)\psi\Big) - \mathrm{m}c^2\psi^\dagger\beta\psi\Big]\\ - \frac{1}{16\pi}\mathcal{F}^{\mu\nu}\mathcal{F}_{\mu\nu}.\end{aligned} \tag{11.59}$$

The corresponding equations are as follows:

$$\begin{cases} i\hbar\partial_t\psi(x) = \Big[e\big(\mathcal{A}_0(x) + \mathcal{A}_0^{\text{ext}}(x)\big) - c\alpha^k\Big[i\hbar\partial_k - \frac{e}{c}\big(\mathcal{A}_k(x) + \mathcal{A}_k^{\text{ext}}(x)\big)\Big] \\ \qquad\qquad + \mathrm{m}c^2\beta\Big]\psi(x) \\ \nabla_\mu\mathcal{F}^{\mu\nu}(x) = 4\pi e\psi^\dagger(x)\alpha^\nu\psi(x), \quad \alpha = 0, \ldots, 3. \end{cases} \tag{11.60}$$

11.6.6 Gauge Transformation for Coupled Equations

The *gauge transformation* (11.48) can be completed by the corresponding transformation of the wave function:

$$\left|\begin{array}{l} \phi(t,\mathbf{x}) \mapsto \phi(t,\mathbf{x}) + \frac{1}{c}\dot{\chi}(t,\mathbf{x}), \;\; \mathbf{A}(t,\mathbf{x}) \mapsto \mathbf{A}(t,\mathbf{x}) - \nabla\chi(t,\mathbf{x}), \\ \psi(t,\mathbf{x}) \mapsto e^{-i\frac{e}{c\hbar}\chi(t,\mathbf{x})}\psi(t,\mathbf{x}) \end{array}\right. \tag{11.61}$$

for any real function $\chi(t,\mathbf{x})$. It is easy to check that the new functions are also solution of Eqs. (11.58) or (11.60), provided that so are the old ones. Moreover, transformations (11.61) preserve the electric charge and current densities in (11.58) and (11.60).

11.6.7 Charged Antiparticles

There exists a one-to-one correspondence between the solutions of the Dirac equations with charges e and $-e$, respectively. Namely, consider the Dirac equation (11.46) with $-e$ instead of e:

$$\gamma\Big(P + \frac{e}{c}\mathcal{A}(x)\Big)\psi(x) = \mathrm{m}c\psi(x), \quad x \in \mathbb{R}^4. \tag{11.62}$$

This equation describes the positron wave field, corresponding to particles with positive charge $-e = |e|$. We will find an isomorphism between the solutions of (11.46) and (11.62).

Definition 11.17 The **charge conjugation** transforms each wave function $\psi(x)$ with values in $\mathbb{C}^4$ into $\psi_c(x) := \gamma^2\overline{\psi}(x)$.

Theorem 11.18 *Let* ψ *satisfy the Dirac equation* $\gamma(P - \frac{e}{c}\mathcal{A}(x))\psi(x) = \mathrm{m}c\psi(x)$ *with mass* m *and charge* e. *Then* ψ_c *satisfies the Dirac equation* (11.62) *with mass* m *and charge* $-e$.

Proof Step (i) We first check that, for any vector $p \in \mathbb{C}^4$,

$$\gamma^2\gamma(\overline{p})\gamma^2 = \overline{\gamma(p)}. \tag{11.63}$$

In other words, all the matrices γ^α are real, except γ^2, which is purely imaginary. Hence

$$\overline{\gamma^\alpha} = \begin{cases} \gamma^\alpha, & \alpha \neq 2, \\ -\gamma^\alpha, & \alpha = 2. \end{cases} \tag{11.64}$$

On the other hand, anticommutation relations for the Dirac matrices imply

$$\gamma^2\gamma^j\gamma^2 = \begin{cases} \gamma^\alpha, & \alpha \neq 2, \\ -\gamma^\alpha, & \alpha = 2. \end{cases} \tag{11.65}$$

This gives (11.63).

Step (ii) Conjugating the Dirac equation for ψ, we obtain

$$\gamma^2\gamma\left(\overline{P - \frac{e}{c}\mathcal{A}(x)}\right)\gamma^2\overline{\psi} = \mathrm{m}c\overline{\psi}. \tag{11.66}$$

Here $\frac{e}{c}\mathcal{A}(x)$ is real, while P includes the imaginary factor i. Hence, multiplying (11.66) by γ^2, this establishes

$$\gamma\left(P + \frac{e}{c}\mathcal{A}(x)\right)\psi_c = \mathrm{m}c\psi_c, \tag{11.67}$$

as claimed. □

11.7 Nonrelativistic Approximations

In the limit as $c \to \infty$, the Lorentz transformations (12.25) become the Galilean ones, (12.23), and the retarded potentials become the Coulomb ones. It would be advantageous to obtain in this limit the nonrelativistic Pauli equation as a suitable approximation to the relativistic Dirac theory.

Dirac introduced Eq. (11.49) for the Hydrogen atom in his seminal paper [45] and obtained the first two approximations in the limit as $c \to \infty$: up to terms $1/c$ and up to $1/c^2$. He discovered that, in the first approximation, stationary Dirac equation reduces to the corresponding stationary Pauli equation, justifying the correct (maximum) Landé factor $g = 2$ for the spin. Moreover, in the next approximation, up to terms of the order $1/c^2$, there is a manifestation of the Russell–Saunders spin-orbital coupling with the correct Thomas factor $1/2$, and also of some other effects. These discoveries are regarded as a great triumph of the Dirac relativistic theory.

First, we rewrite the Dirac equation (11.49) similarly to (11.44):

$$(i\hbar\partial_t - e\phi)\psi = \left(mc^2\beta + c\boldsymbol{\alpha}\cdot\left[\mathbf{P} - \frac{e}{c}\mathbf{A}\right]\right)\psi, \tag{11.68}$$

where $\boldsymbol{\alpha} := (\alpha^1, \alpha^2, \alpha^3)$ and $\mathbf{P} := -i\hbar\nabla$. Substituting for α matrices (11.16) and considering the splitting of the wave function (11.40), the Dirac equation reduces to coupled equations

$$\begin{cases} (i\hbar\partial_t - e\phi)\psi_+ = c\sigma\cdot\left(\mathbf{P} - \frac{e}{c}\mathbf{A}\right)\psi_- + \mathrm{m}c^2\psi_+, \\ (i\hbar\partial_t - e\phi)\psi_- = c\sigma\cdot\left(\mathbf{P} - \frac{e}{c}\mathbf{A}\right)\psi_+ - \mathrm{m}c^2\psi_-. \end{cases} \tag{11.69}$$

We consider the stationary states $\psi(x) = \psi(\mathbf{x})e^{-iEt/\hbar}$ in the case of static external potentials $\phi(\mathbf{x})$ and $\mathbf{A}(\mathbf{x})$. Our goal is to find the asymptotic expansion of the energy levels E in the nonrelativistic limit as $c \to \infty$. We consider these levels close to the free electron energy at rest. More precisely, we assume that

$$E \approx mc^2, \qquad \left|e\phi(\mathbf{x})\right| \ll mc^2. \tag{11.70}$$

Substituting $\psi(x) = \psi(\mathbf{x})e^{-iEt/\hbar}$ into (11.69), we obtain the corresponding stationary eigenvalue problem

$$\begin{cases} \left(E - e\phi - \mathrm{m}c^2\right)\psi_+(\mathbf{x}) = c\sigma\cdot\left(\mathbf{P} - \frac{e}{c}\mathbf{A}\right)\psi_-(\mathbf{x}), \\ \left(E - e\phi + \mathrm{m}c^2\right)\psi_-(\mathbf{x}) = c\sigma\cdot\left(\mathbf{P} - \frac{e}{c}\mathbf{A}\right)\psi_+(\mathbf{x}). \end{cases} \tag{11.71}$$

In the last equation, the factor $E - e\phi + \mathrm{m}c^2$ is large, by our assumption (11.70). Hence, the component ψ_+ is 'large' with respect to ψ_- (the opposite relation holds under the assumption $E \approx -mc^2$). Let us eliminate the small component ψ_- from the first equation by using the second one:

$$\left(E - e\phi - \mathrm{m}c^2\right)\psi_+(\mathbf{x}) = \sigma\cdot\left(\mathbf{P} - \frac{e}{c}\mathbf{A}\right)\left[\frac{c^2}{E - e\phi + \mathrm{m}c^2}\sigma\cdot\left(\mathbf{P} - \frac{e}{c}\mathbf{A}\right)\right]\psi_+(\mathbf{x}). \tag{11.72}$$

Setting $E_0 = E - mc^2$, we expand the difference E_0 in powers of $1/c$. First, we expand the factor

$$\begin{aligned}\frac{c^2}{E - e\phi + \mathrm{mc}^2} &= \frac{c^2}{E_0 - e\phi + 2\mathrm{mc}^2} = \frac{1}{2m}\frac{2mc^2}{2\mathrm{mc}^2 + E_0 - e\phi}\\ &= \frac{1}{2m}\left[1 - \frac{E_0 - e\phi}{2mc^2} + \cdots\right]. \end{aligned} \tag{11.73}$$

11.7.1 The Order $1/c$

Keeping only the leading term in (11.73), we obtain from (11.72) that

$$(E_0 - e\phi)\psi_+ = \frac{1}{2\mathrm{m}}\left(\sigma \cdot \left(\mathbf{P} - \frac{e}{c}\mathbf{A}\right)\right)^2 \psi_+, \tag{11.74}$$

neglecting terms of order $1/c^2$. Let us evaluate the RHS.

Lemma 11.19 *The identity*

$$\left(\sigma \cdot \left(\mathbf{P} - \frac{e}{c}\mathbf{A}\right)\right)^2 = \left(\mathbf{P} - \frac{e}{c}\mathbf{A}\right)^2 + \frac{e}{c}\hbar\sigma \cdot \mathbf{B} \tag{11.75}$$

holds, where $\mathbf{B} := \operatorname{curl}\mathbf{A}$ *is the magnetic field.*

Proof The standard identity for the Pauli matrices reads

$$\left(\sigma\left(\mathbf{P} - \frac{e}{c}\mathbf{A}\right)\right)^2 = \left(\mathbf{P} - \frac{e}{c}\mathbf{A}\right) \cdot \left(\mathbf{P} - \frac{e}{c}\mathbf{A}\right) + i\sigma \cdot \left(\left(\mathbf{P} - \frac{e}{c}\mathbf{A}\right) \times \left(\mathbf{P} - \frac{e}{c}\mathbf{A}\right)\right). \tag{11.76}$$

Note that its proof does not depend on commutation of components of the vector $\mathbf{P} - \frac{e}{c}\mathbf{A}$. On the other hand, the vector product of $\mathbf{P} - \frac{e}{c}\mathbf{A}$ by itself does not vanish, since the components do not commute. For example, let us calculate the first component:

$$\begin{aligned}&\left(P_2 - \frac{e}{c}A_2\right)\left(P_3 - \frac{e}{c}A_3\right) - \left(P_3 - \frac{e}{c}A_3\right)\left(P_2 - \frac{e}{c}A_2\right)\\ &= \left[P_2 - \frac{e}{c}A_2, P_3 - \frac{e}{c}A_3\right]. \end{aligned} \tag{11.77}$$

The commutator obviously reduces to

$$-\frac{e}{c}\big([P_2, A_3] + [A_2, P_3]\big) = i\frac{e\hbar}{c}(\partial_2 A_3 - \partial_3 A_2) = i\frac{e\hbar}{c}(\operatorname{curl} A)_1 = i\frac{e\hbar}{c}\mathbf{B}_1. \tag{11.78}$$

□

Now (11.74) becomes

$$E_0\psi_+ = \left[\frac{1}{2\mathrm{m}}\left(\mathbf{P}-\frac{e}{c}\mathbf{A}\right)^2 - \frac{e\hbar}{2mc}\sigma\cdot\mathbf{B} + e\phi\right]\psi_+, \tag{11.79}$$

which is the stationary eigenvalue problem corresponding to the Pauli equation (10.18).

11.7.2 The Order $1/c^2$

Equation (11.79) neglects the terms of order $1/c^2$ in (11.72). Let us calculate the next approximation of order $1/c^2$ for the case of zero magnetic potential, when $\mathbf{A}=0$.

In this case, (11.79) becomes the electrostatic problem with an external static potential $\phi(\mathbf{x})$ (for example, Coulombic potential of the hydrogen kernel). Note that all terms of order $1/c$ vanish for the zero magnetic potential. We will show that E_0, up to the order $1/c^2$, is the eigenvalue of the problem

$$\begin{aligned} E_0\varphi(\mathbf{x}) &= \left[\frac{\mathbf{P}^2}{2\mathrm{m}} + e\phi(\mathbf{x})\right]\varphi(\mathbf{x}) \\ &\quad + \left[-\frac{\mathbf{P}^4}{8m^3c^2} - \frac{e\hbar^2}{8m^2c^2}\nabla\cdot\mathbf{E}(x) - \frac{e\hbar\sigma\cdot(\mathbf{E}\times\mathbf{P})}{4m^2c^2}\right]\varphi(\mathbf{x}). \end{aligned} \tag{11.80}$$

Comparing this with (10.26) shows that the spin-orbital coupling appears exactly in the Russell–Saunders form, as conjectured by Frenkel and Thomas, because

$$-\frac{e\hbar\sigma\cdot(\mathbf{E}\times\mathbf{P})}{4m^2c^2} = \frac{e}{2\mathrm{m}^2c^2}\frac{\phi'(|\mathbf{x}|)}{|\mathbf{x}|}\hat{\mathbf{L}}\hat{\mathbf{s}}$$

for the radial potential $\phi(|\mathbf{x}|)$. This agreement with the Russell–Saunders coupling was also a great triumph of the Dirac theory.

The remaining terms are discussed in [12, Sect. 12], [92, Sect. 2.2.4], [158, Sect. 3-3], and [159, 160]; here we note that the term with $\nabla\cdot\mathbf{E}$ is known as the *Darwin term* responsible for the *zitterbewegung*, and term with $\mathbf{P}^4$ is known to give the relativistic correction to the electron mass.

Proof of (11.80) Let us keep the second term on the right hand side of (11.73). Then (11.72) (with $\mathbf{A}=0$) gives that

$$(E_0 - e\phi)\psi_+ = \frac{1}{2m}\sigma\cdot\mathbf{P}\left[1 - \frac{E_0 - e\phi}{2mc^2}\right]\sigma\cdot\mathbf{P}\psi_+. \tag{11.81}$$

Collecting the terms involving E_0, we obtain

$$\begin{aligned} E_0\left[1+\frac{\mathbf{P}^2}{4m^2c^2}\right]\psi_+ &= \frac{1}{2m}\sigma\cdot\mathbf{P}\left[1+\frac{e\phi}{2mc^2}\right]\sigma\cdot\mathbf{P}\psi_+ + e\phi\psi_+ \\ &= \frac{1}{2m}\left[1+\frac{e\phi}{2mc^2}\right]\mathbf{P}^2\psi_+ + \frac{e}{4m^2c^2}\sigma\cdot\mathbf{P}\phi\;\sigma\cdot\mathbf{P}\psi_+ + e\phi\psi_+ \\ &= \frac{1}{2m}\left[1+\frac{e\phi}{2mc^2}\right]\mathbf{P}^2\psi_+ + \frac{ie\hbar}{4m^2c^2}\sigma\cdot\mathbf{E}\sigma\cdot\mathbf{P}\psi_+ + e\phi\psi_+, \end{aligned} \tag{11.82}$$

because $\sigma\cdot\mathbf{P}\phi = -i\hbar\sigma\cdot\mathbf{E} = i\hbar\sigma\cdot\mathbf{E}$, which is a consequence of the identities $\mathbf{P} = -i\hbar\nabla$ and $\mathbf{E} = -\nabla\phi$. By (11.70), we may neglect the term with $\frac{e\phi}{2mc^2}$. Therefore, applying identity of type (11.76), it follows that

$$\begin{aligned} E_0\left[1+\frac{\mathbf{P}^2}{4m^2c^2}\right]\psi_+ = \frac{1}{2m}\mathbf{P}^2\psi_+ + \frac{ie\hbar}{4m^2c^2}\big[\mathbf{E}\cdot\mathbf{P}\psi_+ + i\sigma\cdot(\mathbf{E}\times\mathbf{P}\psi_+)\big] \\ + e\phi\psi_+. \end{aligned} \tag{11.83}$$

This equation has the form $E_0B\psi_+ = H\psi_+$, where $B := 1+\mathbf{P}^2/(4m^2c^2)$ is positive selfadjoint operator in the Hilbert space $L^2(\mathbb{R}^3)\otimes\mathbb{C}^2$. Hence, $E_0\varphi = \tilde{H}\varphi$, where $\varphi := B^{1/2}\psi_+$, and $\tilde{H} := B^{-1/2}HB^{-1/2}$. Writing $B^{-1/2} = 1-\mathbf{P}^2/(8m^2c^2)+\cdots$, we approximate (11.83) by

$$\begin{aligned} E_0\varphi = \big[1-\mathbf{P}^2/\big(8m^2c^2\big)\big]\bigg\{\frac{1}{2m}\left[1+\frac{e\phi}{2mc^2}\right]\mathbf{P}^2 \\ + \frac{ie\hbar}{4m^2c^2}\big[\mathbf{E}\cdot\mathbf{P}+i\sigma\cdot(\mathbf{E}\times\mathbf{P})\big]+e\phi\bigg\} \\ \times\big[1-\mathbf{P}^2/\big(8m^2c^2\big)\big]\varphi. \end{aligned}$$

Neglecting the terms of higher orders ($1/c^4, \ldots$), this gives

$$\begin{aligned} E_0\varphi &= \tilde{H}_1\varphi \\ &:= \left[\frac{\mathbf{P}^2}{2\mathrm{m}}+e\phi(\mathbf{x})\right]\varphi + \left\{-\frac{\mathbf{P}^4}{8m^3c^2}+\frac{e\hbar}{4m^2c^2}\big[\hbar\mathbf{E}\cdot\nabla-\sigma\cdot(\mathbf{E}\times\mathbf{P})\big]\right\}\varphi. \end{aligned} \tag{11.84}$$

Finally, we should symmetrize the operator on the right. Namely, operator (11.82) is selfadjoint, while the term $\mathbf{E}\cdot\mathbf{P}$ in (11.83) is not. All other terms in (11.83) are selfadjoint: in particular, the term $\sigma(\mathbf{E}\times\mathbf{P})$ is selfadjoint, because $\operatorname{curl}\mathbf{E}(\mathbf{x}) = 0$. The symmetry breaking arose because we neglected the term with $\frac{e\phi}{2mc^2}$. Restoring the term with $\frac{e\phi}{2mc^2}$ in (11.83), we can neglect its contribution to (11.84) again relying on (11.70). As a result, we obtain the symmetrized version involving the operator $[\tilde{H}_1 + \tilde{H}_1^*]/2$ instead of $\tilde{H}_1$. This version coincides with (11.80), because $[\mathbf{E} \cdot \nabla + (\mathbf{E} \cdot \nabla)^*]/2 = -\nabla \cdot \mathbf{E}/2$.

Remark 11.20 Both Hamiltonians (11.80) and (11.84) generate the identical quadratic forms. Hence, they provide the identical corrections $\sim 1/c^2$ to energy levels of unperturbed Pauli operator. This follows from the known formulas of perturbation theory (3.43).

11.8 Spinor Spherical Harmonics

Dirac introduced equation (11.68) for the Hydrogen atom in [45] and obtained two nonrelativistic approximations from previous section. However, he did not solve the corresponding spectral problem. A few months later, Gordon [74] and Darwin [39] solved independently the spectral problem for the hydrogen.

The solution relies on the separation of variables and the Sommerfeld factorization method in solving the corresponding radial equation; this method was applied by Schrödinger in the nonrelativistic case (see Sect. 6.1.4). For the Dirac equation, the reduction to the radial problem is more involved, requiring an application of the Clebsch–Gordan method to develop the Lie algebras arguments from Lecture 6.2.

11.8.1 Spherical Symmetry

For the hydrogen atom, the corresponding four-potential of the nucleus is $\mathcal{A} = (\phi, 0, 0, 0)$ with $\phi = -e/|\mathbf{x}|$. Then the corresponding Dirac equation (11.68) reads

$$i\hbar\partial_t \psi = H_D \psi := \big(mc^2\beta + c\boldsymbol{\alpha}\mathbf{P} + e\phi\big)\psi. \tag{11.85}$$

We are going to determine all quantum stationary states that are solutions of the form $\psi(\mathbf{x})e^{-iEt/\hbar}$ with finite charge (11.55):

$$\int |\psi(\mathbf{x})|^2 \, d\mathbf{x} < \infty. \tag{11.86}$$

Substituting $\psi = \psi(\mathbf{x})e^{-iEt/\hbar}$ into the Dirac equation (11.85), we arrive at the corresponding stationary eigenvalue problem

$$E\psi = \big(mc^2\beta + c\boldsymbol{\alpha}\mathbf{P} + e\phi\big)\psi, \tag{11.87}$$

which can be reduced to the coupled equations of type (11.71),

$$\begin{cases} \big(E - \mathrm{mc}^2 - e\phi\big)\psi_+ = c\sigma\mathbf{P}\psi_-, \\ \big(E + \mathrm{mc}^2 - e\phi\big)\psi_- = c\sigma\mathbf{P}\psi_+. \end{cases} \tag{11.88}$$

The key argument for the solution is the spherical symmetry of the problem. Namely, the nucleus potential $\phi(\mathbf{x})$ is spherically symmetric. Hence, the angular momentum (11.33) is conserved, as in the free case, because

$$[\hat{\mathbf{J}}, H_D] = 0. \tag{11.89}$$

Indeed, this follows from the commutation $[\hat{\mathbf{J}}, \phi] = 0$, since $H_D = H_D^0 + e\phi$; the commutation $[\hat{\mathbf{J}}, H_D^0] = 0$ is proved in (11.36).

11.8.2 Separation of Variables

It is worth pointing out that we have solved the spectral problem for the nonrelativistic Schrödinger equation by employing the general strategy of separation of variables (Sects. 6.1 and 6.2). Now we are going to develop this strategy for relativistic problem (11.87) in the similar way. In this case, the role of the orbital angular momentum **L** is played by the total angular momentum **J**, since it is conserved. Hence, the strategy now should be modified accordingly:

I. First, (11.89) implies that the operator $\hat{\mathbf{J}}^2 := \hat{\mathbf{J}}_1^2 + \hat{\mathbf{J}}_2^2 + \hat{\mathbf{J}}_3^2$ commutes with H_D:

$$\big[\hat{\mathbf{J}}^2, H_D\big] = 0. \tag{11.90}$$

Second, H_D also commutes with each $\hat{\mathbf{J}}_n$. Hence, each eigenspace of the Dirac operator H_D is invariant with respect to each operator $\hat{\mathbf{J}}_n$ and $\hat{\mathbf{J}}^2$. Moreover, the operator $\hat{\mathbf{J}}^2$ also commutes with each operator $\hat{\mathbf{J}}_n$; for example,

$$\big[\hat{\mathbf{J}}^2, \hat{\mathbf{J}}_3\big] = 0. \tag{11.91}$$

Exercise 11.21 Verify (11.91). **Hint:** First, prove the commutation relations $[\hat{\mathbf{J}}_k, \hat{\mathbf{J}}_j] = -i\hbar\epsilon_{kjl}\hat{\mathbf{J}}_l$, where ϵ_{kjl} is the totally antisymmetric tensor. The relations follow from the similar ones for the orbital and spinor angular momenta, and from the commutation of momenta.

Hence, one may expect that there is a basis of common eigenfunctions for the operators H_D, $\hat{\mathbf{J}}_3$ and $\hat{\mathbf{J}}^2$. First, it would be helpful to simultaneously diagonalize $\hat{\mathbf{J}}^2$ and $\hat{\mathbf{J}}_3$.

II. Condition (11.86) means that we consider the eigenvalue problem (11.87) in the Hilbert space $\mathcal{E} := L^2(\mathbb{R}^3) \otimes \mathbb{C}^4$. On the other hand, both operators $\hat{\mathbf{J}}_3$ and $\hat{\mathbf{J}}^2$ act only on spinor variables and angular variables in spherical coordinates. Hence, the operators also act in the Hilbert space $\mathcal{E}_1 := L^2(S) \otimes \mathbb{C}^4$, where S denotes the two-dimensional sphere $|\mathbf{x}| = 1$.

We split the space $\mathcal{E}_1$ into the sum $\mathcal{E}_1 = \mathcal{E}_1^+ \oplus \mathcal{E}_1^-$, corresponding to the two-component representation (11.40), where $\mathcal{E}_1^\pm = L^2(S) \otimes \mathbb{C}^2$. Both operators $\hat{\mathbf{J}}_3$ and $\hat{\mathbf{J}}^2$ are of the block form

$$\hat{\mathbf{J}}_3 = \begin{pmatrix} \hat{\mathbf{L}}_3 + \hat{\mathbf{s}}_3 & 0 \\ 0 & \hat{\mathbf{L}}_3 + \hat{\mathbf{s}}_3 \end{pmatrix}, \qquad \hat{\mathbf{J}}^2 = \begin{pmatrix} (\hat{\mathbf{L}} + \hat{\mathbf{s}})^2 & 0 \\ 0 & (\hat{\mathbf{L}} + \hat{\mathbf{s}})^2 \end{pmatrix} \tag{11.92}$$

where $\hat{\mathbf{s}} := \frac{1}{2}\hbar\sigma$. Hence, the action of the operators is identical in each summand $\mathcal{E}_1^\pm$. In next section, we will prove the following result.

Proposition 11.22

(i) *In the space $\mathcal{E}_1^\pm$, there exists an orthonormal basis of spinor spherical harmonics $\mathcal{S}_{jk}(\theta,\varphi)$, which are common eigenfunctions of $\hat{\mathbf{J}}_3$ and $\hat{\mathbf{J}}^2$:*

$$\begin{aligned} &\hat{\mathbf{J}}_3 \mathcal{S}_{jk}(\theta,\varphi) = \hbar k \mathcal{S}_{jk}(\theta,\varphi), \\ &\hat{\mathbf{J}}^2 \mathcal{S}_{jk}(\theta,\varphi) = \hbar^2 j(j+1) \mathcal{S}_{jk}(\theta,\varphi), \quad k = -j, -j+1, \ldots, j, \end{aligned} \tag{11.93}$$

where $j = 1/2, 3/2, \ldots$.

(ii) *The space of solutions of* (11.93) *is two-dimensional for any fixed j, k, except for the one-dimensional space corresponding to $j = 1/2$, $k = -1/2$.*

III. This proposition suggests that we could construct eigenfunctions of the Dirac operator H_D by separating variables; namely, in the form

$$\psi_E = \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix} = \begin{pmatrix} R_+^+(r)\mathcal{S}_{jk}^+(\theta,\varphi) + R_+^-(r)\mathcal{S}_{jk}^-(\theta,\varphi) \\ R_-^+(r)\mathcal{S}_{jk}^+(\theta,\varphi) + R_-^-(r)\mathcal{S}_{jk}^-(\theta,\varphi) \end{pmatrix}, \tag{11.94}$$

where $\mathcal{S}_{jk}^+$, $\mathcal{S}_{jk}^-$ are some solutions of (11.93).

IV. Each solution of the spectral problem (11.87) is a sum (or a series) of solutions of particular form (11.94) since $\mathcal{S}_{jk}^\pm$ form the basis in $\mathcal{E}_1$. The construction of solutions (11.94) relies on the investigation of the commutation relations for $\hat{\mathbf{J}}_k$, $\hat{\mathbf{L}}_k$ and $\hat{\mathbf{s}}_k$, $k = 1, 2, 3$; i.e., on the corresponding Lie algebra. The radial functions will be obtained by solving the corresponding radial equations via the factorization method as in Sect. 6.1.4.

11.8.3 Clebsch–Gordan's Theorem

We prove Proposition 11.22. In (11.93), the operators $\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{s}}$ act on the spaces $\mathcal{E}_1^\pm = L^2(S) \otimes \mathbb{C}^2$. The operator $\hat{\mathbf{L}}$ acts on $L^2(S)$, while $\hat{\mathbf{s}}$ acts on the second factor $\mathbb{C}^2$. As a result, $\hat{\mathbf{L}}$ commutes with $\hat{\mathbf{s}}$. Hence, $\hat{\mathbf{J}}^2 = \hat{\mathbf{L}}^2 + 2\hat{\mathbf{L}}\hat{\mathbf{s}} + 3/4$ is the selfadjoint second order elliptic operator on the sphere S. Therefore, there exists an orthonormal basis of eigenfunctions for $\hat{\mathbf{J}}^2$ in $L^2(S) \otimes \mathbb{C}^2$.

Moreover, $\hat{\mathbf{J}}$ is the generator of the tensor product of the regular and spinor representations of the rotation group SO(3). Hence, the eigenvalues of $\hat{\mathbf{J}}^2$ are given by Proposition 6.12. It remains to check that only halfinteger spins J are possible with the multiplicity two for $J \geq 3/2$ and one for $J = 1/2$.

First, we know the spectral decomposition (6.38) of the operator $\hat{\mathbf{L}}^2$ in the space $L^2(S)$:

$$L^2(S) = \bigoplus_{l=0}^{\infty} L(l). \tag{11.95}$$

Here $L(l)$ are finite-dimensional orthogonal eigenspaces of $\hat{\mathbf{L}}^2$ corresponding to the eigenvalues $\hbar^2 l(l+1)$, $l = 0, 1, \ldots$. Second, we will use the structure of the spaces $L(l)$, as established in the proof of Proposition 6.12 and Lemma 6.18. Namely, in $L(l)$, there is an orthonormal basis $e_{-l}, \ldots, e_l$, where $e_m = \hat{\mathbf{H}}_+^{m+l} e_{-l}$ (here $\hat{\mathbf{H}}_+ := \hat{\mathbf{H}}_1 + i\hat{\mathbf{H}}_2$ with $\hat{\mathbf{H}}_k := \hbar^{-1}\hat{\mathbf{L}}_k$) and (cf. (11.93)),

$$\hat{\mathbf{L}}_3 e_{lm} = \hbar m e_{lm}, \qquad \hat{\mathbf{L}}^2 e_{lm} = \hbar^2 l(l+1) e_{lm}, \quad m = -l, \ldots, l. \tag{11.96}$$

We recall that $e_{lm} = Y_l^m$, where Y_l^m are the spherical harmonics (6.12). Similarly, in $\mathbb{C}^2$, there is an orthonormal basis $f_{-1/2}$, $f_{1/2}$, where $f_{1/2} = \hat{\mathbf{s}}_+ f_{-1/2}$, and

$$\hat{\mathbf{s}}_3 f_s = \hbar s f_s, \qquad \hat{\mathbf{s}}^2 f_s = \hbar^2 s(s+1) f_s, \quad s = -1/2,\ 1/2. \tag{11.97}$$

Here $f_{-1/2} = \binom{0}{1}$ and $f_{1/2} = \binom{1}{0}$. Therefore, we have

$$L^2(S) \otimes \mathbb{C}^2 = \oplus_{l=0}^{\infty} \mathcal{F}(l), \quad \mathcal{F}(l) := L(l) \otimes \mathbb{C}^2, \tag{11.98}$$

and the tensor products $e_{lm} \otimes f_s$ with $m = -l, \ldots, l$ and $s = -1/2, 1/2$ form an orthonormal basis in $\mathcal{F}(l)$. Relations (11.96), (11.97) imply that

$$\hat{\mathbf{J}}_3 e_{lm} \otimes f_s = \hbar(m+s) e_{lm} \otimes f_s. \tag{11.99}$$

So, the multiplicity of the eigenvalue is two if $m + s \geq 3/2$, and only the halfinteger spins J are possible. However, the vectors $e_{lm} \otimes f_s$ generally are not the eigenvectors for $\hat{\mathbf{J}}^2$. It turns out that $\hat{\mathbf{J}}^2$ has two different eigenvalues $l \pm 1/2$ in $\mathcal{F}(l)$ if $l \geq 1$. This follows from Clebsch–Gordan's theorem [71] known as the law of 'addition of angular momenta'. In our particular case this theorem sounds as follows.

Lemma 11.23 *For any* $l = 0, 1, \ldots$

(i) *Each space* $\mathcal{F}(l)$ *is an orthogonal sum of two eigenspaces* $\mathcal{F}_\pm(l)$ *of the operator* $\hat{\mathbf{J}}^2$,

$$\mathcal{F}(l) = \mathcal{F}_+(l) \oplus \mathcal{F}_-(l), \qquad \hat{\mathbf{J}}^2|_{\mathcal{F}_\pm(l)} = \hbar^2 j_\pm(j_\pm + 1), \tag{11.100}$$

where $j_\pm = l \pm 1/2$ *and* $\dim \mathcal{F}_\pm(l) = 2j_\pm + 1$;

(ii) *For* $l \geq 1$, *in space* $\mathcal{F}_\pm(l)$ *there exists a basis of functions* $\mathcal{S}_k^\pm$, $k = -j_\pm, -j_\pm + 1, \ldots, j_\pm$, *satisfying* (11.93) *with* $j = j_\pm$;

(iii) *For* $l = 0$, *we have* $\mathcal{F}_-(0) = 0$, *and also, for* $\mathcal{F}_+(0)$ *there exists a basis* $\{\mathcal{S}_k^+ : k = -1/2, 1/2\}$, *satisfying* (11.93) *with* $j = 1/2$.

Proof Let $\hat{\mathbf{J}}_+ := \hat{\mathbf{J}}_1 + i\hat{\mathbf{J}}_2$. Then we can choose $\mathcal{F}_\pm(l)$ with the basis

$$\mathcal{S}_k^\pm = \mathcal{S}_{j_\pm,k}^\pm := \hat{\mathbf{J}}_+^{k+j_\pm} \mathcal{S}_{-j_\pm}^\pm, \quad k = -j_\pm, -j_\pm + 1, \ldots, j_\pm, \tag{11.101}$$

where $\mathcal{S}_{-j_\pm}^\pm$ is the eigenvector of $\hat{\mathbf{J}}_3$ with eigenvalue $-\hbar j_\pm$.

Obviously, $\mathcal{S}_{-j_+}^+ = e_{-l} \otimes f_{-1/2}$ is unique up to a factor, and so it remains to construct the vector $\mathcal{S}_{-j_-}^-$. This vector should be orthogonal to $\mathcal{F}_+(l)$ and should lie in the subspace $F \subset \mathcal{F}(l)$, which consists of all eigenvectors of $\hat{\mathbf{J}}_3$ with eigenvalue $-\hbar j_-$ in the space $\mathcal{F}(l)$. We should choose a nonzero vector $\mathcal{S}_{-j_-}^- \in F$, which is orthogonal to $\mathcal{F}_+(l)$. We separately consider two cases: $l \geq 1$ and $l = 0$.

The case $l \geq 1$. In this case, the eigenspace F is the two-dimensional linear span of $e_{-l} \otimes f_{1/2}$ and $e_{-l+1} \otimes f_{-1/2}$. The intersection $\mathcal{F}_+(l) \cap F$ is the one-dimensional linear span of $\hat{\mathbf{J}}_+ \mathcal{S}_{-j_+}^+$. Hence, $\mathcal{S}_{-j_-}^- \in F$ is determined uniquely (up to a factor) as the vector orthogonal to $\hat{\mathbf{J}}_+ \mathcal{S}_{-j_+}^+$.

The case $l = 0$. Now the space F is the one-dimensional span of $e_0 \otimes f_{1/2}$, since the vector e_1 does not exist in this case. Hence, $\mathcal{F}_-(0) = 0$. □

Proof of Proposition 11.22 (i) Functions (11.101) are solutions of (11.93) with $j = j_\pm$; they form an orthogonal basis for $\mathcal{E}_1(l)$. (ii) From (11.100) and (11.98) it follows that every solution of (11.93) is a linear combination of $\mathcal{S}_{j,k}^+$ and $\mathcal{S}_{j,k}^-$, provided $(j, k) \neq (1/2, -1/2)$. Otherwise, any solution is proportional to $\mathcal{S}_{1/2,-1/2}^+$, because the function $\mathcal{S}_{1/2,-1/2}^-$ does not exist. □

It is worth noting that $\mathcal{F}_\pm(l)$ are the eigenspaces also for $\sigma\mathbf{L}$; namely, we have the following result.

Lemma 11.24 $\sigma\hat{\mathbf{L}}$ *takes value* $\hbar l$ *resp.* $-\hbar(l+1)$ *on space* $\mathcal{F}_+(l)$ *resp.* $\mathcal{F}_-(l)$.

Proof This follows from the second equation of (11.93) with $j = j_\pm$ by identity (see (10.12))

$$\hbar\sigma\hat{\mathbf{L}} = \left[\hat{\mathbf{L}} + \frac{1}{2}\hbar\sigma\right]^2 - \hat{\mathbf{L}}^2 - \frac{1}{4}\hbar^2\sigma^2 = \hat{\mathbf{J}}^2 - \hat{\mathbf{L}}^2 - \frac{3}{4}\hbar^2, \tag{11.102}$$

since $j_\pm(j_\pm + 1) - l(l+1) - 3/4$ equals either l or $-(l+1)$. □

11.9 The Hydrogen Spectrum

The final step should consist in solving the coupled equations (11.88) for the components ψ_+, ψ_- in form (11.94), eliminating one of the components. The corresponding radial equation gives the energy levels by the method of Sect. 6.1.4.

Now, in contrast to nonrelativistic case, the energy levels depend on the angular momentum. This dependence perfectly explains the 'fine structure' of the spectrum.

11.9.1 The Radial Equation

We will obtain the radial equations by the method of [81], substituting (11.94) into (11.88). First, we need the expression of the operator $\sigma\mathbf{P}$ of (11.88) in terms of the orbital angular momentum and related operators.

Lemma 11.25 *The following relations hold*:

$$\sigma\mathbf{P} = |\mathbf{x}|^{-2}\sigma\mathbf{x}(\mathbf{x}\mathbf{P} + i\sigma\mathbf{L}), \tag{11.103}$$

$$\sigma\mathbf{x}(\sigma\mathbf{L} + \hbar) + (\sigma\mathbf{L} + \hbar)\sigma\mathbf{x} = 0. \tag{11.104}$$

Proof From the formulas for products of spin matrices we obtain

$$(\sigma\mathbf{x})(\sigma\mathbf{P}) = \mathbf{x}\mathbf{P} + i\sigma(\mathbf{x} \times \mathbf{P}) = \mathbf{x}\mathbf{P} + i\sigma\mathbf{L}. \tag{11.105}$$

Multiplying this equation on the left by $\sigma\mathbf{x}$ we arrive at (11.103). Now equation (11.104) follows on multiplying the commutation relations $[\mathbf{L}_j, \mathbf{x}^k] = i\hbar\epsilon_{jkl}\mathbf{x}^l$ by $\sigma_j\sigma_k = 2\delta_{jk} - \sigma_k\sigma_j$ and simplifying. ☐

Substituting (11.103) into (11.88) and using (11.104), this gives

$$\begin{cases} \big(E - \mathrm{m}c^2 - e\phi\big)\psi_+ = c|\mathbf{x}|^{-2}\sigma\mathbf{x}(\mathbf{x}\mathbf{P} + i\sigma\mathbf{L})\psi_-, \\ \qquad\qquad\qquad\qquad = c|\mathbf{x}|^{-2}\big(\mathbf{x}\mathbf{P} - i(\sigma\mathbf{L} + \hbar)\big)\sigma\mathbf{x}\psi_-, \\ \big(E + \mathrm{m}c^2 - e\phi\big)\psi_- = c|\mathbf{x}|^{-2}\sigma\mathbf{x}(\mathbf{x}\mathbf{P} + i\sigma\mathbf{L})\psi_+. \end{cases} \tag{11.106}$$

The last equation can be rewritten as

$$\big(E + \mathrm{m}c^2 - e\phi\big)\sigma\mathbf{x}\psi_- = c(\mathbf{x}\mathbf{P} + i\sigma\mathbf{L})\psi_+. \tag{11.107}$$

Together with the first equation of (11.106), this suggests the notation $\Psi_- := \frac{\sigma\mathbf{x}}{|\mathbf{x}|}\psi_-$ and $\Psi_+ := \psi_+$. Rewriting equations (11.106) in spherical coordinates, we get

$$\begin{cases} \big(E - \mathrm{m}c^2 - e\phi\big)\Psi_+ = c\Big(-i\hbar\dfrac{d}{dr} - ir^{-1}(\sigma\mathbf{L} + 2\hbar)\Big)\Psi_-, \\ \big(E + \mathrm{m}c^2 - e\phi\big)\Psi_- = c\Big(-i\hbar\dfrac{d}{dr} + ir^{-1}\sigma\mathbf{L}\Big)\Psi_+. \end{cases} \tag{11.108}$$

Let us construct all nonzero solutions of (11.88) in the form (11.94) for each fixed parameters j and k. For example, we can assume that

$$R_+^+(r) \not\equiv 0. \tag{11.109}$$

We denote by π_{jk}^+ the orthogonal projection in $L^2(S) \otimes \mathbb{C}^2$ onto the linear span of $\mathcal{S}_{jk}^+$, and by Π_{jk}^+, the corresponding projection in $L^2(\mathbb{R}^3) \otimes \mathbb{C}^2$. Then Π_{jk}^+ commutes with $\sigma \hat{\mathbf{L}}$, by Lemma 11.24. Hence, applying Π_{jk}^+ to both sides of (11.108), and using the fact that ϕ is spherically symmetric, it follows that

$$\begin{cases} \big(E - \mathrm{m}c^2 - e\phi\big)\Pi_{jk}^+\Psi_+ = c\Big(-i\hbar\dfrac{d}{dr} - ir^{-1}(\hbar l + 2\hbar)\Big)\Pi_{jk}^+\Psi_-, \\ \big(E + \mathrm{m}c^2 - e\phi\big)\Pi_{jk}^+\Psi_- = c\Big(-i\hbar\dfrac{d}{dr} + ir^{-1}\hbar l\Big)\Pi_{jk}^+\Psi_+, \end{cases} \tag{11.110}$$

by Lemma 11.24. Here

$$l = j - \frac{1}{2}, \tag{11.111}$$

since $\mathcal{S}_{jk}^+$ is defined by (11.101) with $j_+ = j$, and moreover, $j_+ := l + 1/2$. Further,

$$\Pi_{jk}^+\Psi_+ = R_1(r)\mathcal{S}_{jk}^+(\theta,\varphi), \qquad \Pi_{jk}^+\Psi_- = R_2(r)\mathcal{S}_{jk}^+(\theta,\varphi), \tag{11.112}$$

where, by (11.109),

$$R_1(r) \equiv R_+^+(r) \not\equiv 0. \tag{11.113}$$

Setting $R(r) := (R_1(r)\mathbb{R}_2(r))$ and substituting (11.112) into (11.110), it is found that (11.110) is equivalent to the following *radial equation* for the vector function $R(r)$

$$\big(E - e\phi - \mathrm{m}c^2\sigma_3\big)R(r) = -ic\hbar\left[\left(\frac{d}{dr} + \frac{1}{r}\right) + \frac{(l+1)}{r}\sigma_3\right]\sigma_1 R(r), \quad r > 0, \tag{11.114}$$

where σ_k are the Pauli matrices (10.11). Finally,

$$\big(E - e\phi - \mathrm{m}c^2\sigma_3\big)R(r) = -ic\hbar\left[\left(\frac{d}{dr} + \frac{1}{r}\right)\sigma_1 + i\frac{(l+1)}{r}\sigma_2\right]R(r), \quad r > 0, \tag{11.115}$$

because $\sigma_3\sigma_1 = i\sigma_2$.

11.9.2 Spectrum

As in the nonrelativistic case, we substitute $R(r) = e^{-\kappa r} P(r)$ with unknown parameter κ to be chosen later. Then equation (11.114) reduces to

$$\big(E - e\phi - \mathrm{mc}^2\sigma_3\big)P(r) = -ic\hbar\left[\left(\frac{d}{dr} + \frac{1}{r} - \kappa\right)\sigma_1 + i\frac{(l+1)}{r}\sigma_2\right]P(r),\quad r > 0 \tag{11.116}$$

or equivalently,

$$\big(E - \mathrm{mc}^2\sigma_3 - ic\hbar\kappa\sigma_1\big)P(r) = -ic\hbar\left[\left(\frac{d}{dr} + \frac{1}{r}\right)\sigma_1 + i\frac{(l+1)}{r}\sigma_2 + \frac{ie\phi}{c\hbar}\right]P(r),\quad r > 0. \tag{11.117}$$

Consider the matrix

$$M := \frac{i}{c\hbar}\big(E - \mathrm{mc}^2\sigma_3 - ic\hbar\kappa\sigma_1\big), \tag{11.118}$$

and rewrite the Coulombic potential as $e\phi = -c\hbar\alpha/r$, where

$$\alpha := \frac{e^2}{c\hbar} \approx \frac{1}{137}$$

is the dimensionless Sommerfeld *fine structure constant* (4.29).

Further, we suggest that $P(r)$ will be a 'polynomial'

$$P(r) = r^\delta \sum_0^n P_k r^k, \tag{11.119}$$

where we can assume that $P_0 \neq 0$ for a nontrivial solution. We should choose $\kappa > 0$ and $\delta > -3/2$ to ensure condition (11.86). To justify (11.119), we shall try to find a solution in the more general form

$$P(r) = r^\delta \sum_0^\infty P_k r^k. \tag{11.120}$$

Substituting this into (11.117), we get the equation

$$\sum_0^\infty r^{k+\delta} M P_k = \sum_0^\infty \big[(k+\delta+1)\sigma_1 + i(l+1)\sigma_2 + i\alpha\big] r^{k+\delta-1} P_k. \tag{11.121}$$

This gives the recurrence equation

$$M P_{k-1} = \big[(k+\delta+1)\sigma_1 + i(l+1)\sigma_2 + i\alpha\big] r^{k+\delta-1} P_k, \quad k = 0, 1, \ldots. \tag{11.122}$$

Taking $k = 0$, we obtain

$$\big[(\delta+1)\sigma_1 + i(l+1)\sigma_2 + i\alpha\big]P_0 = 0, \tag{11.123}$$

which implies that

$$\det\big[(\delta+1)\sigma_1 + i(l+1)\sigma_2 + i\alpha\big] = 0, \tag{11.124}$$

because $P_0 \neq 0$. This is equivalent to

$$(\delta+1)^2 = (l+1)^2 - \alpha^2. \tag{11.125}$$

Hence $|\delta+1| \approx l+1$, because α is small. Therefore, we should choose the positive root for $\delta+1$, since, for the negative root, we get $-\delta \approx l+2 \geq 2$, while $\delta > -3/2$.

Finally, examining the recurrence equation (11.122) it is found by condition (11.86) that the series (11.120) should terminate, as in the case of the Schrödinger equation. Hence, we arrive at (11.119) with an $n < \infty$.

Further, we may assume that $P_n \neq 0$ for a nontrivial solution. Therefore, (11.117) implies that $MP_n = 0$, and hence $\det M = 0$:

$$\det\big(E - \mathrm{m}c^2\sigma_3 - ic\hbar\kappa\sigma_1\big) = 0. \tag{11.126}$$

This is equivalent to

$$c^2\hbar^2\kappa^2 = \mathrm{m}^2c^4 - E^2 \tag{11.127}$$

and so, in particular, $E < \mathrm{m}c^2$. Equation (11.127) alone is not capable of determining the eigenvalues E, since we have the additional unknown parameter κ. Therefore, we need an additional equation, which we will derive from the recurrence equation (11.122) with $k = n$:

$$MP_{n-1} = \big[(n+\delta+1)\sigma_1 + i(l+1)\sigma_2 + i\alpha\big]r^{n+\delta-1}P_n. \tag{11.128}$$

Namely, the characteristic equation for the matrix M is as follows

$$(M - 2iE/c\hbar)M = 0 \tag{11.129}$$

since its determinant is zero and since its trace is $2iE/c\hbar$. Therefore, multiplying both sides of (11.122) by $M - 2iE/c\hbar$, it is found that

$$0 = (M - 2iE/c\hbar)\big[(n+\delta+1)\sigma_1 + i(l+1)\sigma_2 + i\alpha\big]P_n. \tag{11.130}$$

Multiplying by the Pauli matrices, we arrive at

$$0 = \big[2\kappa(n+\delta+1) - 2\alpha E/c\hbar\big]P_n. \tag{11.131}$$

This gives the quantization condition

$$\alpha E = c\hbar\kappa(n+\delta+1). \tag{11.132}$$

Hence $E > 0$, and so, substituting $c\hbar\kappa = \alpha E/(n + \delta + 1)$ into (11.127), we obtain the eigenvalues

$$E = E_{nj} = \frac{\mathrm{m}c^2}{\sqrt{1 + [\alpha^2/(n + \delta + 1)^2]}}, \tag{11.133}$$

where $\delta = \delta(l) > -1$ and $l = j - 1/2$ are given by (11.125) and (11.111), respectively.

Comparison with the Nonrelativistic Case Since α is small, we can approximate the eigenvalues by the binomial expansion:

$$E_{nj} \approx \mathrm{m}c^2 - \frac{\mathrm{m}c^2\alpha^2}{2(n + \delta + 1)^2}. \tag{11.134}$$

This approximation with $\delta = 0$ coincides with the nonrelativistic spectrum (6.3) of the hydrogen atom up to the (unessential) additive constant $\mathrm{m}c^2$. The relativistic formula depends on the angular momentum j through $\delta = \delta(l)$ by (11.125) and (11.111), while the nonrelativistic formula is independent of the angular momentum. This was another triumph of the Dirac theory, because this corresponds to the experimental observation of the *fine structure* of the spectrum.

Remark 11.26 The above analysis gives also the corresponding eigenfunctions

$$\psi_E = \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix} = \frac{\sigma\mathbf{x}}{|\mathbf{x}|}\begin{pmatrix} \Psi_+ \\ \Psi_- \end{pmatrix} = \frac{\sigma\mathbf{x}}{|\mathbf{x}|}\begin{pmatrix} R_1(r)\mathcal{S}^+_{jk}(\theta, \varphi) \\ R_2(r)\mathcal{S}^+_{jk}(\theta, \varphi) \end{pmatrix}. \tag{11.135}$$

Chapter 12
Electrodynamics and Special Relativity

Abstract The classical electrodynamics is a well established field theory, starting from formalization by Maxwell in 1855–1861 of previously known empirical facts. It serves up to now as a safe ground and a model for all subsequent field theories.

In 1862, Maxwell put forward the electromagnetic theory of light. In 1884, Heaviside recasted Maxwell's mathematical analysis from its original form to modern vector terminology, and in the same year Poynting discovered the energy propagation in a Maxwell field. In 1886, Hertz experimentally discovered electromagnetic waves and calculated the dipole radiation. Einstein in 1905 discovered the special relativity and the covariant electrodynamics, by postulating invariance of the Maxwell equations in all inertial frames—this was justified by the Michelson and Morley crucial experiments.

The next fundamental question arises on the interaction of the Maxwell field with matter. In 1890s Lorentz introduced the *electron theory* of polarization and magnetization of matter; this enabled one to avoid 'sharp questions', reducing the problem of the matter reaction to *constitutive equations*.

Alternatively, one should consider all details of the interaction of charged particles with a Maxwell field. However, this question cannot be solved in the classical theory, since for a point particle, the field energy is infinite—this effect is nowadays known as the 'ultraviolet divergence'.

To fix the situation, Abraham in 1905 introduced the model of 'extended electron'. However, this suggestion leads to next question on the shape of electron. This question was partially clarified after the discovery of quantum mechanics (1926); however a complete answer is still a matter for the future.

12.1 The Maxwell Equations and Potentials

In 1855–1861, Maxwell formulated equations of the electrodynamics by systematizing the empirical Coulomb, Biot–Savart–Laplace, and Faraday laws. He stated these equations in the MKS rationalized (or SI) system of units:

$$\begin{cases} \operatorname{div}\mathbf{E}(t,\mathbf{x}) = \dfrac{1}{\varepsilon_0}\rho(t,\mathbf{x}), & \operatorname{curl}\mathbf{E}(t,\mathbf{x}) = -\dot{\mathbf{B}}(t,\mathbf{x}), \\ \operatorname{div}\mathbf{B}(t,\mathbf{x}) = 0, & \dfrac{1}{\mu_0}\operatorname{curl}\mathbf{B}(t,\mathbf{x}) = \mathbf{j}(t,\mathbf{x}) + \varepsilon_0\dot{\mathbf{E}}(t,\mathbf{x}), \end{cases} \tag{12.1}$$

A. Komech, *Quantum Mechanics: Genesis and Achievements*,
DOI 10.1007/978-94-007-5542-0_12,

where $\rho(t,\mathbf{x})$ and $\mathbf{j}(t,\mathbf{x})$ stand for the charge and current density, and ε_0 and μ_0 are the dielectric permittivity and magnetic permeability of vacuum. The values of ε_0 and μ_0 were measured experimentally with high accuracy. In the MKS system,

$$\varepsilon_0 \approx \frac{1}{4\pi 9\times 10^9}\frac{\mathrm{A\,s}}{\mathrm{V\,m}}, \qquad \mu_0 \approx 4\pi\times 10^{-7}\frac{\mathrm{V\,s}}{\mathrm{A\,m}}. \tag{12.2}$$

It is instructive to note that the dimension of the product $\varepsilon_0\mu_0$ is s^2/m^2 (i.e., it is inversely proportional to the square of velocity). This observation was of crucial importance for the Maxwell theory of light and the Einstein's special theory of relativity.

The Maxwell equations imply the *continuity equation for charge and current*

$$\dot\rho(t,\mathbf{x}) + \operatorname{div}\mathbf{j}(t,\mathbf{x}) = 0; \tag{12.3}$$

this is equivalent to the charge conservation law. Of course, this equivalence cannot be considered as the proof of the (experimentally known) charge conservation since it was used in derivation of the Maxwell equations.

12.1.1 Synthesis of Maxwell Equations

Maxwell deduced the first and the last equations from the Coulomb and Biot–Savart–Laplace laws for the interaction of charges and currents. The second and the third Maxwell equations were conjectured and experimentally verified by Faraday.

The Coulomb law gives the force of the electrostatic interaction of two charges q_1, q_2 located at points $\mathbf{x}_1$ and $\mathbf{x}_2$, respectively:

$$\mathbf{F}_1 = \frac{1}{4\pi\varepsilon_0}\frac{q_1q_2(\mathbf{x}_1-\mathbf{x}_2)}{|\mathbf{x}_1-\mathbf{x}_2|^3}; \tag{12.4}$$

here $\mathbf{F}_1$ is the force acting onto the first elementary charge. Similarly, the Biot–Savart–Laplace law gives the force of the 'magnetostatic interaction' of two elementary currents $I_k d\mathbf{l}_k$, $k=1,2$, positioned at a distance r;

$$\mathbf{F}_1 = \frac{\mu_0}{4\pi}\frac{I_1 d\mathbf{l}_1\times[I_2 d\mathbf{l}_2\times(\mathbf{x}_1-\mathbf{x}_2)]}{|\mathbf{x}_1-\mathbf{x}_2|^3} \tag{12.5}$$

here $\mathbf{F}_1$ is the force acting onto the first elementary current.

The First Equation

From the Coulomb law (12.4) it follows that the electric field of the second elementary charge, at the position of the first one, is as follows

$$\mathbf{E}(\mathbf{x}_1) = \frac{1}{4\pi\varepsilon_0}\frac{q_2(\mathbf{x}_1-\mathbf{x}_2)}{|\mathbf{x}_1-\mathbf{x}_2|^3}. \tag{12.6}$$

Then, for the distribution of charges, $\rho(\mathbf{x}_2)\,d\mathbf{x}_2$, it follows by the principle of superposition,

$$\mathbf{E}(\mathbf{x}_1) = \frac{1}{4\pi\varepsilon_0}\int \frac{\rho(\mathbf{x}_2)(\mathbf{x}_1 - \mathbf{x}_2)}{|\mathbf{x}_1 - \mathbf{x}_2|^3}\,d\mathbf{x}_2. \tag{12.7}$$

Differentiating, we obtain

$$\operatorname{div}\mathbf{E}(\mathbf{x}_1) = \frac{1}{\varepsilon_0}\rho(\mathbf{x}_1); \tag{12.8}$$

this coincides with the first equation of (1.3).

Exercise 12.1 Verify (12.8) for $\rho(\mathbf{x}) \in C_0^\infty(\mathbb{R}^3)$. **Hint:** Verify geometrically that

$$\operatorname{div}_{\mathbf{x}_1} \frac{\mathbf{x}_1 - \mathbf{x}_2}{|\mathbf{x}_1 - \mathbf{x}_2|^3} = 4\pi\delta(\mathbf{x}_1 - \mathbf{x}_2).$$

Last Equation for Stationary Currents

The Biot–Savart–Laplace law (12.5) serves as the ground for definition of the magnetic field $\mathbf{B}(\mathbf{x}_1)$ at a point $\mathbf{x}_1$:

$$\mathbf{F}_1 = I_1 d\mathbf{l}_1 \times \mathbf{B}(\mathbf{x}_1), \tag{12.9}$$

which is the *Ampère law*. A comparison with (12.5) shows that

$$\mathbf{B}(\mathbf{x}_1) = \frac{\mu_0}{4\pi}\frac{I_2 d\mathbf{l}_2 \times (\mathbf{x}_1 - \mathbf{x}_2)}{|\mathbf{x}_1 - \mathbf{x}_2|^3}. \tag{12.10}$$

Integrating and using the principle of superposition, we obtain

$$\mathbf{B}(\mathbf{x}_1) = \frac{\mu_0}{4\pi}\int \frac{\mathbf{j}(\mathbf{x}_2) \times (\mathbf{x}_1 - \mathbf{x}_2)}{|\mathbf{x}_1 - \mathbf{x}_2|^3}\,d\mathbf{x}_2, \tag{12.11}$$

where $\mathbf{j}(\mathbf{x}_2)$ is the current density at the point $\mathbf{x}_2$. Differentiating, we get

$$\frac{1}{\mu_0}\operatorname{curl}\mathbf{B}(\mathbf{x}_1) = \mathbf{j}(\mathbf{x}_1), \tag{12.12}$$

which agrees with the last equation of (12.1) in the case of **stationary currents and fields**, because $\dot{\mathbf{E}}(t,\mathbf{x}) = 0$.

Exercise 12.2 Verify (12.12) when $\mathbf{j}(\mathbf{x}) \in C_0^\infty(\mathbb{R}^3)$. **Hint:** Use the continuity equation for charge and current, $\operatorname{div}\mathbf{j}(\mathbf{x}) \equiv 0$, for stationary currents.

Maxwell's Displacement Current

In Maxwell's time, the last equation of (12.1) was known in a form similar to (12.12),

$$\frac{1}{\mu_0}\operatorname{curl}\mathbf{B}(t,\mathbf{x}) = \mathbf{j}(t,\mathbf{x}). \tag{12.13}$$

The divergence of the LHS vanishes, because $\operatorname{div}\operatorname{curl} = 0$. On the other hand, $\operatorname{div}\mathbf{j}(t,\mathbf{x})$ generally does not vanish for nonstationary currents. This is obvious from the identity

$$\operatorname{div}\mathbf{j}(t,\mathbf{x}) + \operatorname{div}\varepsilon_0\dot{\mathbf{E}}(t,\mathbf{x}) = 0, \tag{12.14}$$

which follows from the continuity equation for charge and current (12.3) and the first Maxwell equation. To remove this contradiction, Maxwell introduced the *displacement current* $\varepsilon_0\dot{\mathbf{E}}(t,\mathbf{x})$ into the last equation (12.1). In this case, the divergence of the RHS of the equation vanishes by (12.14).

12.1.2 Maxwell Potentials

Maxwell reduced the system (12.1) to the wave equations by introducing potentials. Namely, $\operatorname{div}\mathbf{B}(t,\mathbf{x}) = 0$ implies that $\mathbf{B}(t,\mathbf{x}) = \operatorname{curl}\mathbf{A}(t,\mathbf{x})$. Then $\operatorname{curl}\mathbf{E}(t,\mathbf{x}) = -\dot{\mathbf{B}}(t,\mathbf{x})$ implies $\operatorname{curl}[\mathbf{E}(t,\mathbf{x}) + \dot{\mathbf{A}}(t,\mathbf{x})] = 0$, and hence $\mathbf{E}(t,\mathbf{x}) + \dot{\mathbf{A}}(t,\mathbf{x}) = -\nabla\phi(t,\mathbf{x})$. Finally,

$$\mathbf{B}(t,\mathbf{x}) = \operatorname{curl}\mathbf{A}(t,\mathbf{x}), \qquad \mathbf{E}(t,\mathbf{x}) = -\nabla\phi(t,\mathbf{x}) - \dot{\mathbf{A}}(t,\mathbf{x}). \tag{12.15}$$

Exercise 12.3 Prove the existence of potentials (12.15). **Hint:** Use the Fourier transform.

The choice of the potentials is not unique, since the *gauge transformation*

$$\phi(t,\mathbf{x}) \mapsto \phi(t,\mathbf{x}) + \dot{\chi}(t,\mathbf{x}), \qquad \mathbf{A}(t,\mathbf{x}) \mapsto \mathbf{A}(t,\mathbf{x}) - \nabla\chi(t,\mathbf{x}) \tag{12.16}$$

does not change the fields $\mathbf{E}(t,\mathbf{x})$, $\mathbf{B}(t,\mathbf{x})$ for any real function $\chi(t,\mathbf{x}) \in C^1(\mathbb{R}^4)$. Therefore, it is possible to satisfy an additional *gauge* condition. For example, we can provide

$$\phi(t,\mathbf{x}) = 0$$

choosing

$$\chi(t,\mathbf{x}) = -\int_0^t \phi\big(t',\mathbf{x}\big)\,dt' + \chi_0(\mathbf{x}).$$

Further we choose, the *Lorentz gauge*

$$\varepsilon_0 \dot{\phi}(t,\mathbf{x}) + \frac{1}{\mu_0}\operatorname{div}\mathbf{A}(t,\mathbf{x}) = 0. \tag{12.17}$$

Let us express the Maxwell equations (12.1) in terms of potentials. Substituting (12.15) into the first Maxwell equation, this gives $\frac{4\pi}{\varepsilon_0}\rho(t,\mathbf{x}) = \operatorname{div}\mathbf{E}(t,\mathbf{x}) = -\Delta\phi(t,\mathbf{x}) - \operatorname{div}\dot{\mathbf{A}}(t,\mathbf{x})$. Differentiating (12.17) in time, we get $\varepsilon_0\ddot{\phi}(t,\mathbf{x}) + \frac{1}{\mu_0}\operatorname{div}\dot{\mathbf{A}}(t,\mathbf{x}) = 0$. Hence, we can eliminate $\operatorname{div}\dot{\mathbf{A}}(t,\mathbf{x})$, obtaining

$$\big[\varepsilon_0\mu_0\partial_t^2 - \Delta\big]\phi(t,\mathbf{x}) = \frac{1}{\varepsilon_0}\rho(t,\mathbf{x}). \tag{12.18}$$

Similarly, substituting (12.15) into the last Maxwell equation, it is found that

$$\frac{1}{\mu_0}\operatorname{curl}\operatorname{curl}\mathbf{A}(t,\mathbf{x}) = \mathbf{j}(t,\mathbf{x}) + \varepsilon_0\dot{\mathbf{E}}(t,\mathbf{x}) = \mathbf{j}(t,\mathbf{x}) - \varepsilon_0\nabla\dot{\phi}(t,\mathbf{x}) - \varepsilon_0\ddot{\mathbf{A}}(t,\mathbf{x}). \tag{12.19}$$

Exercise 12.4 Prove the identity

$$\operatorname{curl}\operatorname{curl} = -\Delta + \nabla\operatorname{div}. \tag{12.20}$$

Substituting (12.20) into (12.19) and eliminating $\nabla\dot{\phi}(t,\mathbf{x})$ by the application of ∇ to (12.17), it follows that

$$\big[\varepsilon_0\mu_0\partial_t^2 - \Delta\big]\mathbf{A}(t,\mathbf{x}) = \mu_0\mathbf{j}(t,\mathbf{x}). \tag{12.21}$$

12.2 Maxwell's Theory of Light, and Einstein's Postulate

The special theory of relativity was discovered by Einstein through analysis of the covariance of the Maxwell equations under the Lorentz transformations.

12.2.1 Maxwell's Theory of Light

The Maxwell equations (12.1) contain the dielectric permittivity and magnetic permeability of vacuum, ε_0 and μ_0, but do not explicitly contain the speed of light c. Maxwell deduced the wave equations (12.18) and (12.21) with the coefficient $\varepsilon_0\mu_0$, and it was known at that time that this coefficient is $1/v^2$, where v is the speed of wave propagation: for instance, the dimension of $\varepsilon_0\mu_0$ is s^2/m^2!

Maxwell calculated $v = 1/\sqrt{\varepsilon_0\mu_0}$ by using values (12.2), obtaining $v \approx 3 \times 10^8$ m/s. His crucial observation was that v coincides with the speed of light c in vacuum with high precision! Hence Eqs. (12.18) and (12.21) become

$$\Box\phi(t,\mathbf{x}) = \frac{1}{\varepsilon_0}\rho(t,\mathbf{x}), \qquad \Box\mathbf{A}(t,\mathbf{x}) = \mu_0\mathbf{j}(t,\mathbf{x}), \quad (t,\mathbf{x}) \in \mathbb{R}^4, \tag{12.22}$$

where $\Box := \frac{1}{c^2}\partial_t^2 - \Delta$. Furthermore, Maxwell found that electromagnetic waves are transversal, like the light (see Sect. 14.3). This is why Maxwell identified light with electromagnetic waves (1862).

12.2.2 'Luminiferous Ether': The Michelson and Morley Experiment

The discovery of Maxwell had led to a very dramatic situation. Namely, according to classical physics, the wave propagation speed can be equal to c only in a **unique preferred reference frame**, in which the Maxwell equations have the form (12.1), (12.22). Such a reference frame was called the *luminiferous ether frame*. In all other frames, the propagation velocity depends on the vector $\mathbf{v}$ of relative velocity, because the space-time is transformed by the classical *Galilean transformations*: for $\mathbf{v} = (v, 0, 0)$,

$$\begin{pmatrix} t \\ \mathbf{x}^1 \\ \mathbf{x}^2 \\ \mathbf{x}^3 \end{pmatrix} \mapsto \begin{pmatrix} t_* \\ \mathbf{x}_*^1 \\ \mathbf{x}_*^2 \\ \mathbf{x}_*^3 \end{pmatrix} = \begin{pmatrix} t \\ \mathbf{x}^1 - vt \\ \mathbf{x}^2 \\ \mathbf{x}^3 \end{pmatrix}. \tag{12.23}$$

Therefore, *the Maxwell equations are not invariant under Galilean transformations*: otherwise, **the speed of light would be the same in any inertial reference frame**; this obviously contradicts the Galilean transformations!

This is why Michelson and Morley started around 1880 their famous experiment to identify the preferred *luminiferous ether frame* with the frame, in which the Sun is at rest. They tried to verify that the Earth moves with respect to the luminiferous ether, i.e., to check that the speed of light along and against the Earth orbital motion differs by twice speed of the Earth. More specifically, they compared the wavelengths of light along and against the motion of the Earth around the Sun. However, the result (1887) was negative and very discouraging: the wavelengths were identical with high accuracy, and hence the speed of propagation was found to be independent of the frame of reference! Astronomical observations of *double stars* performed by de Sitter (1908) confirmed the negative result of Michelson and Morley; this negative result was also confirmed by the Trouton and Noble experiment.

12.2.3 Einstein's Postulate

Various partial explanations of these negative results were put forward by Ritz, Fitzgerald, Lorentz and others. A complete explanation was provided in 1905 by

Einstein, who was able to culminate Maxwell's and Lorentz's ideas into a new complete theory. The main Einstein's postulate was the following:

$$\textbf{The laws of electrodynamics take identical form in all inertial frames}. \tag{12.24}$$

This postulate includes the invariance of the Maxwell equations, and hence the constancy of the speed of light. The required modification of the Galilean transformations was well known after Larmor (1897) and Lorentz (1899):

$$\begin{pmatrix} ct_* \\ \mathbf{x}_*^1 \\ \mathbf{x}_*^2 \\ \mathbf{x}_*^3 \end{pmatrix} = \begin{pmatrix} \frac{ct-\beta x^1}{\sqrt{1-\beta^2}} \\ \frac{x^1-\beta ct}{\sqrt{1-\beta^2}} \\ \mathbf{x}^2 \\ \mathbf{x}^3 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{1-\beta^2}} & -\frac{\beta}{\sqrt{1-\beta^2}} & 0 & 0 \\ -\frac{\beta}{\sqrt{1-\beta^2}} & \frac{1}{\sqrt{1-\beta^2}} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ \mathbf{x}^1 \\ \mathbf{x}^2 \\ \mathbf{x}^3 \end{pmatrix}, \tag{12.25}$$

where (t, x^1, x^2, x^3) are the time-space coordinates in the rest frame and $(t_*, x_*^1, x_*^2, x_*^3)$ corresponds to the moving frame, if the relative velocity is $(v, 0, 0)$ with $|v| < c$, and $\beta := v/c$.

Larmor and Lorentz discovered that the *Lorentz transformations* (12.25), leaving unchanged the form of the Maxwell equations (12.1) (see Sect. 12.5). This is why Einstein suggested to regard the Lorentz formulas (12.25) as the *true transformation of space and time* and rejected the Galilean transformation, because they do not preserve the form of the Maxwell equations. This suggestion includes, in particular, the following revolutionary claim:

Time in a moving frame is distinct from the time in the rest frame!

Exercise 12.5 Verify that the wave equation

$$\left[\frac{1}{c^2}\,\partial_t^2 - \Delta\right]\phi(t, \mathbf{x}) = 0 \tag{12.26}$$

is *invariant* with respect to the Lorentz transformations (12.25).
Hints:

(i) $\partial_2 = \partial_2^*$, $\partial_3 = \partial_3^*$, where $\partial_k := \frac{\partial}{\partial_{x^k}}$ and $\partial_k^* := \frac{\partial}{\partial_{x_*^k}}$.

(ii) Set $ct = x_0$ and check that the 1D equation $(\partial_0^2 - \partial_1^2)\phi(x^0, x^1) = 0$ is equivalent to

$$[(\partial_0^*)^2 - (\partial_1^*)^2)]\phi_*(x_*^0, x_*^1) = 0,$$

if $x_*^0 = ax^0 - bx^1$ and $x_*^1 = -bx^0 + ax^1$ with $a^2 - b^2 = 1$. By the definition, $\phi_*(x_*^0, x_*^1) = \phi(x^0, x^1)$.

Exercise 12.6 Verify that the wave equation (12.26) is *not invariant* with respect to the **Galilean transformation** (12.23). **Hint:** In the new variables, equation (12.26) takes the form

$$\left[\frac{1}{c^2}(\partial_{t_*} - v\nabla_{\mathbf{x}_*^1})^2 - \Delta_{\mathbf{x}_*}\right]\phi_*(t_*, \mathbf{x}_*) = 0. \tag{12.27}$$

Remark 12.7 For small velocities, $|v| \ll c$, the Galilean transformation (12.23) is close to the Lorentz one (12.25), and the coefficients of equation (12.27) are close to the ones of (12.26).

12.3 The Lorentz Group

The Lorentz group is the set of all transformations of the space-time, preserving the form of the wave equation.

12.3.1 The Lorentz Transformations

Einstein's postulate (12.24) implies the invariance of the wave equations (12.18) and (12.21). Let us calculate all possible transformations, which leave invariant the homogeneous wave equations. First, we introduce the new time variable $x^0 := ct$ and rewrite the homogeneous wave equation as follows:

$$\Box\phi(x) := g^{\alpha\beta}\partial_\alpha\partial_\beta\phi(x) = 0; \tag{12.28}$$

here $g^{\alpha,\beta} = \mathrm{diag}(1, -1, -1, -1)$, $\partial_\alpha := \frac{\partial}{\partial x^\alpha}$, and the summation in the repeated indexes is understood, $\alpha, \beta = 0, 1, 2, 3$.

We look for maps $M : \mathbb{R}^4 \to \mathbb{R}^4$ which are such that the equation preserves its form in the new variables $y = \Lambda x$; i.e., (12.28) is equivalent to

$$\Box\phi'(y) = g^{\alpha'\beta'}\partial_{\alpha'}\partial_{\beta'}\phi'(y) = 0, \tag{12.29}$$

where $\phi'(y) := \phi(x)$. The space-time is homogeneous, and hence the map M should be linear up to a translation:

$$y^{\alpha'} := (Mx)^{\alpha'} = \Lambda^{\alpha'}_{\alpha} x^\alpha + b^{\alpha'}, \quad \alpha' = 0, \ldots, 3. \tag{12.30}$$

Let us rewrite Eq. (12.28) in the new variables (12.30). By the chain rule,

$$\partial_\alpha := \frac{\partial}{\partial x^\alpha} = \frac{\partial y^{\alpha'}}{\partial x^\alpha}\frac{\partial}{\partial y^{\alpha'}} = \Lambda^{\alpha'}_{\alpha}\partial_{\alpha'}. \tag{12.31}$$

Hence, the equation in new variables reads

$$g^{\alpha,\beta}\Lambda^{\alpha'}_{\alpha}\Lambda^{\beta'}_{\beta}\partial_{\alpha'}\partial_{\beta'}\phi'(y) = 0, \quad y \in \mathbb{R}^4. \tag{12.32}$$

Comparing this with (12.29), we obtain the system of algebraic equations

$$g^{\alpha\beta}\Lambda_{\alpha}^{\alpha'}\Lambda_{\beta}^{\beta'} = g^{\alpha'\beta'}, \quad \alpha', \beta' = 0, 1, 2, 3. \tag{12.33}$$

In the matrix form, this system is equivalent to the equation

$$\Lambda g \Lambda^t = g, \tag{12.34}$$

in which Λ^t denotes the transposed matrix of Λ. Such matrices are called *Lorentz transformations*.

12.3.2 Properties of the Lorentz Transformations

Exercise 12.8 Verify that $|\det \Lambda| = 1$. **Hint:** Take the determinant of both sides of (12.34).

Thus Λ is invertible. Hence Eq. (12.34) is also equivalent to

$$\left(\Lambda^t\right)^{-1} g \Lambda^{-1} = g. \tag{12.35}$$

This matrix equation is equivalent to the invariance, with respect to the map Λ^{-1}, of the quadratic form

$$g(x, x) := (x, gx) = g_{\alpha\beta}x^{\alpha}x^{\beta}, \quad x \in \mathbb{R}^4, \tag{12.36}$$

which is known as the **Lorentz interval** (here $(g_{\alpha\beta}) := (g^{\alpha\beta})^{-1} = (g^{\alpha\beta})$). Then this form is also invariant with respect to Λ, and hence

$$\Lambda^t g \Lambda = g \tag{12.37}$$

or equivalently,

$$g_{\alpha\beta}\Lambda_{\alpha'}^{\alpha}\Lambda_{\beta'}^{\beta} = g_{\alpha'\beta'}, \quad \alpha', \beta' = 0, 1, 2, 3. \tag{12.38}$$

Definition 12.9 L is the set of all linear maps $\Lambda : \mathbb{R}^4 \to \mathbb{R}^4$ satisfying (12.37).

Exercise 12.10 Verify that the set L is a group. **Hint:** Use the characterization (12.37).

Definition 12.11

(i) The Minkowski space is the space $\mathbb{R}^4$ equipped with the quadratic form (12.36).
(ii) L is called the **Lorentz group**.

12.3.3 Examples

The simplest example of a Lorentz transform is given by a matrix

$$\Lambda = \hat{R} := \begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix}, \tag{12.39}$$

where $R \in \mathrm{SO}(3)$ is a rotation of the 3D space. The wave equation (12.26) is invariant with respect to transformation $x' = \hat{R}x$, because the Laplacian Δ is invariant with respect to rotations.

Exercise 12.12 Verify (12.34) for matrices (12.39).

Exercise 12.13 Construct all Lorentz transformations of the form

$$\Lambda = \begin{pmatrix} a & b & 0 & 0 \\ c & d & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{12.40}$$

Solution: Equation (12.37) is equivalent to the matrix equation

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{12.41}$$

which means that

$$a^2 - c^2 = 1, \qquad b^2 - d^2 = -1, \quad ab - cd = 0.$$

Equivalently,

$$\begin{aligned} &a = \pm\cosh\varphi, \qquad c = \sinh\varphi, \qquad d = \pm\cosh\chi, \qquad b = \sinh\chi, \\ &\pm\cosh\varphi\sinh\chi \pm \sinh\varphi\cosh\chi = 0, \end{aligned}$$

where the signs $\pm$ in the last formula are independent. Therefore, $\tanh\varphi = \pm\tanh\chi$; i.e., $\varphi = \pm\chi$. Hence, we get four one-parametric families of matrices

$$\begin{aligned} \Lambda_{\pm}^{+} &:= \begin{pmatrix} \cosh\varphi & \sinh\varphi & 0 & 0 \\ \pm\sinh\varphi & \pm\cosh\varphi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \\ \Lambda_{\pm}^{-} &:= \begin{pmatrix} -\cosh\varphi & \sinh\varphi & 0 & 0 \\ \pm\sinh\varphi & \mp\cosh\varphi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \end{aligned} \tag{12.42}$$

Here
Λ^+_+ are **hyperbolic rotations** (or **boosts**) by angle φ in the plane (x^0, x^1);
Λ^+_- are the hyperbolic rotations followed by reflection $(x^0, x^1) \mapsto (x^0, -x^1)$;
Λ^-_+ are the hyperbolic rotations followed by reflection $(x^0, x^1) \mapsto (-x^0, -x^1)$;
Λ^-_- are the hyperbolic rotations followed by reflection $(x^0, x^1) \mapsto (-x^0, x^1)$.

The Lorentz formulas (12.25) correspond to matrix Λ^+_+ with $\cosh\varphi = 1/\sqrt{1-\beta^2}$ and $\sinh\varphi = -\beta/\sqrt{1-\beta^2}$, so $\tanh\varphi = -\beta$.

12.4 The Lagrangian Formalism for Maxwell Field

The Maxwell equations can be represented as the Euler–Lagrange equations corresponding to a Lagrangian density, which is a quadratic polynomial in the potentials and their first derivatives.

12.4.1 The Maxwell Equations in Gaussian Units

Written in *unrationalized Gaussian units* (cgs), the Maxwell equations (12.1) read as

$$\begin{cases} \operatorname{div}\mathbf{E}(t,\mathbf{x}) = 4\pi\rho(t,\mathbf{x}), & \operatorname{curl}\mathbf{E}(t,\mathbf{x}) = -\dfrac{1}{c}\dot{\mathbf{B}}(t,\mathbf{x}), \\ \operatorname{div}\mathbf{B}(t,\mathbf{x}) = 0, & \operatorname{curl}\mathbf{B}(t,\mathbf{x}) = \dfrac{4\pi}{c}\mathbf{j}(t,\mathbf{x}) + \dfrac{1}{c}\dot{\mathbf{E}}(t,\mathbf{x}), \end{cases} \tag{12.43}$$

implying the same continuity equation for charge and current (12.3). Now the Maxwell potentials are defined by (cf. (12.15))

$$\mathbf{B}(t,\mathbf{x}) = \operatorname{curl}\mathbf{A}(t,\mathbf{x}), \qquad \mathbf{E}(t,\mathbf{x}) = -\nabla\phi(t,\mathbf{x}) - \frac{1}{c}\dot{\mathbf{A}}(t,\mathbf{x}). \tag{12.44}$$

The *Lorentz gauge* (12.17) now reads

$$\frac{1}{c}\dot{\phi}(t,\mathbf{x}) + \operatorname{div}\mathbf{A}(t,\mathbf{x}) = 0, \quad (t,\mathbf{x}) \in \mathbb{R}^4; \tag{12.45}$$

it can be satisfied by an appropriate gauge transformation of type (12.16)

$$\phi(t,\mathbf{x}) \mapsto \phi(t,\mathbf{x}) + \frac{1}{c}\dot{\chi}(t,\mathbf{x}), \qquad \mathbf{A}(t,\mathbf{x}) \mapsto \mathbf{A}(t,\mathbf{x}) - \nabla\chi(t,\mathbf{x}). \tag{12.46}$$

Consequently, the Maxwell equations (12.43) are equivalent to the wave equations for the potentials, similar to (12.22):

$$\Box\phi(t,\mathbf{x}) := \left[\frac{1}{c^2}\partial_t^2 - \Delta\right]\phi(t,\mathbf{x}) = 4\pi\rho(t,\mathbf{x}), \qquad \Box\mathbf{A}(t,\mathbf{x}) = \frac{4\pi}{c}\mathbf{j}(t,\mathbf{x}). \tag{12.47}$$

12.4.2 4D Vector Potential

Let us introduce four-dimensional vectors and derivatives

$$\begin{cases} x^0 = ct, \qquad x^\mu = \big(x^0, \dots, x^3\big), \\ \partial_\mu = \nabla_\mu = (\partial_0,\ \partial_1,\ \partial_2,\ \partial_3), \qquad \partial_\mu := \dfrac{\partial}{\partial x^\mu}, \\ x_\mu := g_{\mu\nu} x^\nu = \big(x^0, -x^1, -x^2, -x^3\big), \\ \partial^\mu := g^{\mu\nu}\partial_\nu = (\partial_0, -\partial_1, -\partial_2, -\partial_3), \end{cases} \tag{12.48}$$

where $g_{\mu\nu} = g^{\mu\nu} = \operatorname{diag}(1, -1, -1, -, 1)$ and the four-dimensional potentials and currents are

$$\begin{cases} \mathcal{A}^\mu(x) = \big(\phi(x), \mathbf{A}(x)\big), \qquad \mathcal{A}_\mu(x) := g_{\mu\nu}\mathcal{A}^\nu(x) = \big(\phi(x), -\mathbf{A}(x)\big), \\ \mathcal{J}^\mu(x) = \left(\rho(x), \dfrac{1}{c}\mathbf{j}(x)\right), \qquad \mathcal{J}_\mu(x) := g_{\mu\nu}\mathcal{J}^\nu(x) = \left(\rho(x), -\dfrac{1}{c}\mathbf{j}(x)\right). \end{cases} \tag{12.49}$$

In this notation, the Maxwell equations (12.47) read

$$\Box \mathcal{A}^\mu(x) = 4\pi \mathcal{J}^\mu(x). \tag{12.50}$$

Similarly, the continuity equation for charge and current (12.3), the gauge transformation (12.46) and the Lorentz gauge (12.45) are as follows:

$$\partial_\mu \mathcal{J}^\mu(x) = 0, \qquad \mathcal{A}^\mu(x) \mapsto \mathcal{A}^\mu(x) + \partial^\mu \chi(x), \qquad \partial_\mu \mathcal{A}^\mu(x) = 0. \tag{12.51}$$

12.4.3 Tensor Field

The Maxwell tensor is defined by

$$\mathcal{F}^{\mu\nu}(x) = \partial^\mu \mathcal{A}^\nu(x) - \partial^\nu \mathcal{A}^\mu(x). \tag{12.52}$$

In the matrix form,

$$\big(\mathcal{F}^{\mu\nu}(x)\big)_{\mu,\nu=0}^{3} = \begin{pmatrix} 0 & -\mathbf{E}_1(x) & -\mathbf{E}_2(x) & -\mathbf{E}_3(x) \\ \mathbf{E}_1(x) & 0 & -\mathbf{B}_3(x) & \mathbf{B}_2(x) \\ \mathbf{E}_2(x) & \mathbf{B}_3(x) & 0 & -\mathbf{B}_1(x) \\ \mathbf{E}_3(x) & -\mathbf{B}_2(x) & \mathbf{B}_1(x) & 0 \end{pmatrix}, \tag{12.53}$$

by formulas (12.44), (12.48), and (12.49).

Proposition 12.14 *The Maxwell equations* (12.50) *are equivalent to*

$$\partial_\mu \mathcal{F}^{\mu\nu}(x) = 4\pi \mathcal{J}^\nu(x), \quad x \in \mathbb{R}^4. \tag{12.54}$$

Proof The Maxwell tensor is invariant under the gauge transformations (12.51), since $\partial^\mu \partial^\nu \chi(x) - \partial^\nu \partial^\mu \chi(x) = 0$. Therefore, we can assume the Lorentz gauge (12.45) without loss of generality. Then by (12.50),

$$\partial_\mu \mathcal{F}^{\mu\nu}(x) = \partial_\mu \big(\partial^\mu \mathcal{A}^\nu(x) - \partial^\nu \mathcal{A}^\mu(x)\big) = \partial_\mu \partial^\mu \mathcal{A}^\nu(x) = \Box \mathcal{A}^\nu(x) = 4\pi \mathcal{J}^\nu(x). \tag{12.55}$$

□

12.4.4 The Lagrangian Density and Current

The Lagrangian density for the Maxwell equations (12.54) with a **given** external charge-current densities $\mathcal{J}^\nu(x)$ is defined by

$$\begin{gathered}\mathcal{L}(x, \mathcal{A}, \nabla\mathcal{A}) = -\frac{1}{16\pi} \mathcal{F}^{\mu\nu}\mathcal{F}_{\mu\nu} - \mathcal{J}^\nu(x)\mathcal{A}_\nu, \\ (x, \mathcal{A}, \nabla\mathcal{A}) \in \mathbb{R}^4 \times \mathbb{R}^4 \times \mathbb{R}^{16},\end{gathered} \tag{12.56}$$

where $\mathcal{F}^{\mu\nu} := \partial^\mu \mathcal{A}^\nu - \partial^\nu \mathcal{A}^\mu$ and $\mathcal{F}_{\mu\nu} := \partial_\mu \mathcal{A}_\nu - \partial_\nu \mathcal{A}_\mu$.

Proposition 12.15 *The Maxwell equations* (12.54) *with given charge-current densities* $\mathcal{J}^\nu(x)$ *are equivalent to the Euler–Lagrange equations* (4.15) *for the Lagrangian density* (12.56).

Proof The conjugate fields are canonically given by

$$\Pi^{\mu\nu} := \nabla_{\partial_\mu \mathcal{A}_\nu} \mathcal{L} = -\frac{1}{8\pi}\big(\mathcal{F}^{\mu\nu} - \mathcal{F}^{\nu\mu}\big) = -\frac{1}{4\pi}\mathcal{F}^{\mu\nu}. \tag{12.57}$$

Therefore,

$$\nabla_\mu \Pi^{\mu\nu}(x) = -\frac{1}{4\pi} \nabla_\mu \mathcal{F}^{\mu\nu}(x). \tag{12.58}$$

On the other hand,

$$\mathcal{L}_{\mathcal{A}_\nu} = -\mathcal{J}^\nu(x), \tag{12.59}$$

and so (12.54) is equivalent to equations of type (4.15). □

In variables $(\phi, \mathbf{A})$, Eq. (12.59) reads

$$\rho(x) = -\mathcal{L}_\phi, \qquad \frac{1}{c}\mathbf{j}_k(x) = \mathcal{L}_{\mathbf{A}^k}, \quad k = 1, 2, 3. \tag{12.60}$$

We note that $\mathcal{F}^{\mu\nu}\mathcal{F}_{\mu\nu} = 2(E^2 - B^2)$ by (12.53), and hence the Lagrangian density (12.56) reads

$$\mathcal{L}(x, \mathcal{A}, \nabla\mathcal{A}) = -\frac{1}{8\pi}\left(E^2 - B^2\right) - \rho\phi + \frac{\mathbf{j}\cdot\mathbf{A}}{c}. \tag{12.61}$$

Furthermore, the interaction term $\mathcal{J}^\nu(x)\mathcal{A}_\nu$ in (12.56) is a Lorentz invariant product of four vectors $\mathcal{J}^\nu(x)$ and $\mathcal{A}_\nu$.

12.5 Covariant Electrodynamics

We still need to complete the justification of the Einstein postulate (12.24) with the transformation of the Maxwell field by the Lorentz group.

The four-dimensional Maxwell equations (12.50) imply that the transformations for the potentials $\mathcal{A}^\mu(x)$ and 4-currents $\mathcal{J}^\mu(x)$ must be identical, because $\Box = \partial_0^2 - \partial_1^2 - \partial_2^2 - \partial_3^2$ is an invariant operator with respect to the Lorentz group. We will prove below that convection currents are transformed by the matrix Λ in the same way as a 4-vector x. Hence, this suggests that the same transformation also holds for all possible currents, and so

$$\mathcal{J}'(x') = \Lambda\mathcal{J}(x), \qquad \mathcal{A}'(x') = \Lambda\mathcal{A}(x), \quad x' = \Lambda x. \tag{12.62}$$

Further, as is shown in Sect. 12.4.4, the Maxwell equations are the canonical Euler–Lagrange equations corresponding to Lagrangian density

$$\mathcal{L}(x,\mathcal{A},\nabla\mathcal{A}) = -\frac{1}{16\pi}\mathcal{F}^{\mu\nu}\mathcal{F}_{\mu\nu} - g\big(\mathcal{J}(x),\mathcal{A}\big), \quad (x,\mathcal{A},\nabla\mathcal{A}) \in \mathbb{R}^4\times\mathbb{R}^4\times\mathbb{R}^{16}, \tag{12.63}$$

where $\mathcal{F}^{\mu\nu} := \partial^\mu\mathcal{A}^\nu - \partial^\nu\mathcal{A}^\mu$, $\mathcal{F}_{\mu\nu} := g_{\mu\mu'}g_{\nu\nu'}\mathcal{F}^{\mu'\nu'}$ and

$$g\big(\mathcal{J}(x),\mathcal{A}\big) := g_{\mu\nu}\mathcal{J}^\mu(x)\mathcal{A}^\nu. \tag{12.64}$$

Finally, it follows that the Einstein postulate (12.24) holds, because the Lagrangian density (12.63) is invariant under transformations (12.62).

Exercise 12.16 Verify that density (12.63) is invariant under the Lorentz transformations (12.62). **Hint:** Being symmetric, the bilinear form (12.64) is invariant, since the corresponding quadratic form is invariant by the definition of the Lorentz transformations.

12.5.1 Transformation of the Convection 4-Current

The transformation law (12.62) is one of basic postulates of the covariant electrodynamics. However, the validity of this postulate cannot be proved. Its efficiency is supported by the fruitful development and applications. We illustrate this transformation for a convection current.

Namely, we consider an electron at rest in the frame with coordinates x. Then the corresponding charge and current densities read

$$\rho_0(t,\mathbf{x}) = \rho_e(\mathbf{x}), \qquad \mathbf{j}_0(t,\mathbf{x}) = 0. \tag{12.65}$$

For a moving electron, we should take into account the Lorentz contraction. Hence

$$\rho(t,\mathbf{x}) = \gamma\rho_e\big(\gamma(x_1 - vt), x_2, x_3\big), \qquad \mathbf{j}(t,\mathbf{x}) = \gamma\rho_e\big(\gamma(x_1 - vt), x_2, x_3\big)(v,0,0), \tag{12.66}$$

where v is the speed is directed along Ox_1, and $\gamma = 1/\sqrt{1-\beta^2}$ is the Lorentz contraction, where $\beta := v/c$. The front factor γ provides the *charge invariance*

$$\int \rho(t,\mathbf{x})\,d\mathbf{x} = \int \rho_0(t,\mathbf{x})\,d\mathbf{x}. \tag{12.67}$$

Now let us demonstrate that (12.66) agrees with the postulated transformation law (12.62). Indeed, a moving electron is at rest in the frame with coordinates x_*, as defined by (12.25). Hence the Einstein postulate requires *the same formulas* (12.65) in the moving frame:

$$\rho_*(t_*,\mathbf{x}_*) = \rho_e(\mathbf{x}_*), \qquad \mathbf{j}_*(t_*,\mathbf{x}_*) = 0. \tag{12.68}$$

It remains to note that the 4-current $\mathcal{J}(x)$ (that corresponds to (12.66)) is expressed by (12.62) in terms of the 4-current $\mathcal{J}_*(x_*)$ (that corresponds to (12.68))):

$$\mathcal{J}(x) = \Lambda_*^{-1}\mathcal{J}_*(\Lambda_* x), \tag{12.69}$$

where Λ_* is the matrix (12.25).

12.5.2 Transformation of Maxwell Field

Formulas (12.62) imply the corresponding transformation for the Maxwell tensor (12.52):

$$\mathcal{F}'_{\mu'\nu'}(x) = \Lambda^{\mu}_{\mu'}\Lambda^{\nu}_{\nu'}\mathcal{F}_{\mu\nu}(x), \quad x'^{\mu'} = \Lambda^{\mu'}_{\mu}x^{\mu}. \tag{12.70}$$

In the matrix form, we have

$$\mathcal{F}'(x) = \Lambda\mathcal{F}(x)\Lambda^t. \tag{12.71}$$

Applying this formula to the Lorentz bust (12.25), we get the Lorentz transformation for the Maxwell fields:

$$\left\{\begin{array}{ll} \mathbf{E}^1_*(x_*) = \mathbf{E}^1(x), & \mathbf{B}^1_*(x_*) = \mathbf{B}^1(x), \\[2ex] \mathbf{E}^2_*(x_*) = \dfrac{\mathbf{E}^2(x) - \beta\mathbf{B}^3(x)}{\sqrt{1-\beta^2}}, & \mathbf{B}^2_*(x_*) = \dfrac{\mathbf{B}^2(x) + \beta\mathbf{E}^3(x)}{\sqrt{1-\beta^2}}, \\[2ex] \mathbf{E}^3_*(x_*) = \dfrac{\mathbf{E}^3(x) + \beta\mathbf{B}^2(x)}{\sqrt{1-\beta^2}}, & \mathbf{B}^3_*(x_*) = \dfrac{\mathbf{B}^3(x) - \beta\mathbf{E}^2(x)}{\sqrt{1-\beta^2}}. \end{array}\right. \tag{12.72}$$

In the vector form,

$$\begin{aligned} \mathbf{E}_*(x_*) &= \gamma\left[\mathbf{E}(x) + \frac{\mathbf{v}}{c} \times \mathbf{B}(x)\right] + \gamma' \mathbf{v}_1 \big\langle \mathbf{E}(x), \mathbf{v}_1 \big\rangle, \\ \mathbf{B}_*(x_*) &= \gamma\left[\mathbf{B}(x) - \frac{\mathbf{v}}{c} \times \mathbf{E}(x)\right] + \gamma' \mathbf{v}_1 \big\langle \mathbf{B}(x), \mathbf{v}_1 \big\rangle \end{aligned} \tag{12.73}$$

with $\gamma := 1/\sqrt{1-\beta^2}$, where $\beta := v/c$, $\gamma' := 1-\gamma$ and $\mathbf{v}_1 := \mathbf{v}/|\mathbf{v}|$.

Exercise 12.17 Verify formulas (12.72) and (12.73). **Hints:** Formulas (12.72) follow from (12.53) by (12.71) with the Lorentz matrix (12.25).

Neglecting the terms of order β^2, formulas (12.73) can be written as follows:

$$\mathbf{E}_*(x_*) = \mathbf{E}(x) + \frac{\mathbf{v}}{c} \times \mathbf{B}(x), \qquad \mathbf{B}_*(x_*) = \mathbf{B}(x) - \frac{\mathbf{v}}{c} \times \mathbf{E}(x). \tag{12.74}$$

12.6 The Lorentz Force

The Lorentz equation plays a crucial role in Quantum Mechanics, since its Hamilton function (3.30) provides the Schrödinger equation (3.53) by the canonical quantization (3.26). Historically, the Lorentz equation was suggested by the Coulomb and Ampère laws for stationary currents. On the other hand, the Lorentz force automatically follows from the variational principle applied to the Lagrangian density (12.56).

12.6.1 The Lorentz Equation

A Nonrelativistic Particle The motion of a charged particle with small velocity $|\dot{\mathbf{x}}(t)| \ll c$ in the Maxwell field is governed by the Lorentz equation (1.6),

$$\mathrm{m}\ddot{\mathbf{x}}(t) = e\left[\mathbf{E}\big(t, \mathbf{x}(t)\big) + \frac{1}{c}\dot{\mathbf{x}}(t) \times \mathbf{B}\big(t, \mathbf{x}(t)\big)\right], \tag{12.75}$$

where m is the mass of the particle and e is its charge.

A Relativistic Particle For large velocities $|\dot{\mathbf{x}}(t)| \sim c$, the Lorentz equation should be replaced by the following one

$$\dot{\mathbf{p}}(t) = e\left[\mathbf{E}\big(t, \mathbf{x}(t)\big) + \frac{1}{c}\dot{\mathbf{x}}(t) \times \mathbf{B}\big(t, \mathbf{x}(t)\big)\right], \tag{12.76}$$

where $\mathbf{p}(t) := \mathrm{m}\dot{\mathbf{x}}(t)/\sqrt{1 - \dot{\mathbf{x}}^2(t)/c^2}$ is the momentum of the *free relativistic particle* (see (13.11)).

12.6.2 The Lagrangian Formalism

The Lorentz equations (12.75) and (12.76) can be represented in the Euler–Lagrange form as follows:

$$\frac{d}{dt}L_{\dot{\mathbf{x}}}\big(t,\mathbf{x}(t),\dot{\mathbf{x}}(t)\big)=L_{\mathbf{x}}\big(t,\mathbf{x}(t),\dot{\mathbf{x}}(t)\big) \tag{12.77}$$

with $L=L_p+L_{\text{int}}$, where L_p is the Lagrangian of the free particle and L_{int} is the interaction term (i.e., the term depending on the Maxwell field). The Lagrangians for a free particle are well known: $L_p=\mathrm{m}\dot{\mathbf{x}}^2/2$ for a nonrelativistic particle and $L_p=-\mathrm{m}c^2\sqrt{1-\beta^2}$ for a relativistic one, where $\beta:=|\dot{\mathbf{x}}|/c$ (see (13.11)).

The Interaction Term

It remains to find the *interaction term* L_{int}. It is natural to obtain L_{int} by identifying its action with the one corresponding to the interaction term $-\mathcal{J}^\nu(x)\mathcal{A}_\nu$ in Lagrangian (12.56). The current densities $\mathcal{J}^\nu(t,\mathbf{x})$, which correspond to a particle trajectory $\mathbf{x}(t)$, are as follows:

$$\mathcal{J}^0(t,\mathbf{x})=e\delta\big(\mathbf{x}-\mathbf{x}(t)\big),\qquad \mathcal{J}^k(t,\mathbf{x})=\frac{e}{c}\dot{\mathbf{x}}^k(t)\delta\big(\mathbf{x}-\mathbf{x}(t)\big),\quad k=1,2,3. \tag{12.78}$$

We denote $\Omega=\{(t,\mathbf{x})\in\mathbb{R}^4: t_1<t<t_2\}$. Substituting (12.78) into (12.56), we get the corresponding action (4.8) in the form

$$\begin{aligned} S_\Omega &:= \int_\Omega \mathcal{L}(x,\mathcal{A}_\mu,\nabla\mathcal{A}_\mu)\,dx = S^0_\Omega(\mathcal{A})-\int_{t_1}^{t_2}\int_{\mathbb{R}^3}\mathcal{J}^\nu(t,\mathbf{x})\mathcal{A}_\nu(t,\mathbf{x})\,d\mathbf{x}\,dt\\ &= S^0_\Omega(\mathcal{A})-e\int_{t_1}^{t_2}\Big[\phi\big(t,\mathbf{x}(t)\big)-\frac{1}{c}\dot{\mathbf{x}}(t)\cdot\mathbf{A}\big(t,\mathbf{x}(t)\big)\Big]dt, \end{aligned} \tag{12.79}$$

where $S^0_\Omega(\mathcal{A})$ is action of the *free Maxwell field* in the absence of charged particles. Action (12.79) suggests considering the following *interaction term*

$$L_{\text{int}}(t,\mathbf{x},\mathbf{v})=-e\phi(t,\mathbf{x})+\frac{e}{c}\mathbf{v}\cdot\mathbf{A}(t,\mathbf{x}). \tag{12.80}$$

A Nonrelativistic Particle Let us check that the Euler–Lagrange equations corresponding to the Lagrangian

$$L(t,\mathbf{x},\mathbf{v})=\frac{\mathrm{m}\mathbf{v}^2}{2}-e\phi(t,\mathbf{x})+\frac{e}{c}\mathbf{v}\cdot\mathbf{A}(t,\mathbf{x}) \tag{12.81}$$

are equivalent to the Lorentz equations (12.75). First, we evaluate the momentum:

$$\mathbf{p}(t):=L_{\mathbf{v}}\big(t,\mathbf{x}(t),\dot{\mathbf{x}}(t)\big)=\mathrm{m}\dot{\mathbf{x}}(t)+\frac{e}{c}\mathbf{A}\big(t,\mathbf{x}(t)\big). \tag{12.82}$$

Now (12.77) becomes,

$$\dot{\mathbf{p}}_k(t) = L_{\mathbf{x}^k}\big(t, \mathbf{x}(t), \dot{\mathbf{x}}(t)\big) = -e\nabla_k \phi(t, \mathbf{x}) + \frac{e}{c}\dot{\mathbf{x}} \cdot \nabla_k \mathbf{A}(t, \mathbf{x}), \quad k = 1, 2, 3. \tag{12.83}$$

We calculate

$$\dot{\mathbf{p}}_k(t) = \frac{d}{dt}\left(\mathrm{m}\dot{\mathbf{x}}^k(t) + \frac{e}{c}\mathbf{A}_k\big(t, \mathbf{x}(t)\big)\right) = \mathrm{m}\ddot{\mathbf{x}}^k + \frac{e}{c}\left[\dot{\mathbf{A}}_k(t, \mathbf{x}) + \nabla_j \mathbf{A}_k(t, \mathbf{x})\dot{\mathbf{x}}^j\right]. \tag{12.84}$$

Substituting this expression into (12.83), it follows that

$$\mathrm{m}\ddot{\mathbf{x}}^k + \frac{e}{c}\left[\dot{\mathbf{A}}_k(t, \mathbf{x}) + \nabla_j \mathbf{A}_k(t, \mathbf{x})\dot{\mathbf{x}}^j\right] = -e\nabla_k \phi(t, \mathbf{x}) + \frac{e}{c}\dot{\mathbf{x}}^j \nabla_k \mathbf{A}_j(t, \mathbf{x}). \tag{12.85}$$

Equivalently,

$$\mathrm{m}\ddot{\mathbf{x}}^k = e\left[-\nabla_k \phi(t, \mathbf{x}) - \frac{1}{c}\dot{\mathbf{A}}_k(t, \mathbf{x})\right] + \frac{e}{c}\dot{\mathbf{x}}^j[\nabla_k \mathbf{A}_j - \nabla_j \mathbf{A}_k]. \tag{12.86}$$

The first square bracket on the RHS is $\mathbf{E}(t, \mathbf{x})$, by (12.44). Hence, it remains to verify that

$$\dot{\mathbf{x}}^j[\nabla_k \mathbf{A}_j - \nabla_j \mathbf{A}_k] = \dot{\mathbf{x}} \times \operatorname{curl} \mathbf{A}(t, \mathbf{x}).$$

We note that $\nabla_k \mathbf{A}_j - \nabla_j \mathbf{A}_k = (\operatorname{curl} \mathbf{A})_l \varepsilon_{kjl}$, where ε_{kjl} is the antisymmetric tensor. Therefore,

$$\dot{\mathbf{x}}^j[\nabla_k \mathbf{A}_j - \nabla_j \mathbf{A}_k] = \dot{\mathbf{x}}^j(\operatorname{curl} \mathbf{A})_l \varepsilon_{kjl} = \big[\dot{\mathbf{x}} \times \operatorname{curl} \mathbf{A}(t, \mathbf{x})\big]_k,$$

by the definition of vector product.

A Relativistic Particle For a relativistic particle, the interaction term (12.80) gives the Lagrangian function

$$L(t, \mathbf{x}, \mathbf{v}) = -\mathrm{m}c^2\sqrt{1-\beta^2} - e\phi(t, \mathbf{x}) + \frac{e}{c}\mathbf{v} \cdot \mathbf{A}(t, \mathbf{x}), \tag{12.87}$$

where $\beta := |\mathbf{v}|/c$. Now the momentum is given by

$$\mathbf{p} := L_{\mathbf{v}} = \mathrm{m}\mathbf{v}/\sqrt{1-\beta^2} + \frac{e}{c}\mathbf{A}(t, \mathbf{x}). \tag{12.88}$$

By the same calculation as above, we obtain (12.76).

12.6.3 The Hamiltonian Formalism

A Nonrelativistic Particle The Hamilton function is the Legendre transformation of the nonrelativistic Lagrangian (12.81):

$$\begin{aligned}\mathcal{H} &= \mathbf{p}\mathbf{v} - L = \mathbf{p}\mathbf{v} - \frac{\mathrm{m}\mathbf{v}^2}{2} + e\phi(t,\mathbf{x}) - \frac{e}{c}\mathbf{v}\cdot A(t,\mathbf{x}) \\ &= e\phi(t,\mathbf{x}) + \mathbf{v}\left(\mathbf{p} - \frac{e}{c}\mathbf{A}(t,\mathbf{x})\right) - \frac{\mathrm{m}\mathbf{v}^2}{2}. \qquad (12.89)\end{aligned}$$

Since $\mathbf{p} - \frac{e}{c}\mathbf{A}(t,\mathbf{x}) = \mathrm{m}\mathbf{v}$, we obtain

$$\mathcal{H} = e\phi(t,\mathbf{x}) + \frac{\mathrm{m}\mathbf{v}^2}{2} = e\phi(t,\mathbf{x}) + \frac{1}{2\mathrm{m}}\left[\mathbf{p} - \frac{e}{c}\mathbf{A}(t,\mathbf{x})\right]^2, \qquad (12.90)$$

which coincides with (3.30).

A Relativistic Particle Expression (12.88) implies that

$$\begin{aligned}\mathcal{H} &= \mathbf{p}\mathbf{v} - L = \mathbf{p}\mathbf{v} + \mathrm{m}c^2\sqrt{1-\beta^2} + e\phi(t,\mathbf{x}) - \frac{e}{c}\mathbf{v}\cdot\mathbf{A}(t,\mathbf{x}) \\ &= e\phi(t,\mathbf{x}) + \mathbf{v}\left(\mathbf{p} - \frac{e}{c}\mathbf{A}(t,\mathbf{x})\right) + \mathrm{m}c^2\sqrt{1-\beta^2}. \qquad (12.91)\end{aligned}$$

Since $\mathbf{p} - \frac{e}{c}\mathbf{A}(t,\mathbf{x}) = \mathrm{m}\mathbf{v}/\sqrt{1-\beta^2}$ it follows by (12.88) that

$$\begin{aligned}\mathcal{H} &= e\phi(t,\mathbf{x}) + \frac{\mathrm{m}\mathbf{v}^2}{\sqrt{1-\beta^2}} + \mathrm{m}c^2\sqrt{1-\beta^2} = e\phi(t,\mathbf{x}) + \frac{\mathrm{m}c^2}{\sqrt{1-\beta^2}} \\ &= e\phi(t,\mathbf{x}) + \mathrm{m}c^2\sqrt{1 + \left(\mathbf{p} - \frac{e}{c}\mathbf{A}(t,\mathbf{x})\right)^2/(\mathrm{m}c)^2}, \qquad (12.92)\end{aligned}$$

which is equivalent to (3.28):

$$\left(\mathcal{H}/c - \frac{e}{c}\phi(t,\mathbf{x})\right)^2 = \left[\mathbf{p} - \frac{e}{c}A(t,\mathbf{x})\right]^2 + \mathrm{m}^2c^2. \qquad (12.93)$$

This derivation of the Lorentz force from the Maxwell Lagrangian density is not too surprising, because the Maxwell equations were deduced from the Coulomb and Biot–Savart–Laplace laws, which imply (1.5) and (1.6).

12.7 Energy of the Maxwell Field: Poynting's Theorem

In 1884, Poynting discovered the energy flow in the Maxwell field. More precisely, he proved that

(i) $\mathcal{E}(t,\mathbf{x}) := \frac{1}{8\pi}(\mathbf{E}^2(t,\mathbf{x}) + \mathbf{B}^2(t,\mathbf{x}))$ is the energy density of the electromagnetic field, and

(ii) $\mathbf{S}(t,\mathbf{x}) := \dfrac{c}{4\pi}\mathbf{E}(t,\mathbf{x}) \times \mathbf{B}(t,\mathbf{x})$ is the density of energy flux.

This means that, for any bounded region $\Omega \subset \mathbb{R}^3$ with smooth boundary $\Sigma = \partial\Omega$,

$$\frac{d}{dt}\big[T_\Omega(t)+\mathcal{E}_\Omega(t)\big] = -\int_\Sigma \mathbf{S}(t,\mathbf{x})\cdot\mathbf{n}(\mathbf{x})\,d\Sigma(\mathbf{x}), \quad \mathcal{E}_\Omega(t) := \int_\Omega \mathcal{E}(t,\mathbf{x})\,d\mathbf{x}, \tag{12.94}$$

where $\mathbf{n}(\mathbf{x})$ is the unit outward normal vector to Σ, and $d\Sigma(\mathbf{x})$ is the Lebesgue measure on Σ; $T_\Omega(t)$ denotes the kinetic of charged particles in Ω:

$$T_\Omega(t) = \frac{1}{2}\sum_{\mathbf{x}_k(t)\in\Omega} \mathbf{m}_k \dot{\mathbf{x}}_k^2(t), \tag{12.95}$$

where $\mathbf{m}_k$ is mass of a particle and $\mathbf{x}_k(t)$ is its position.

The corresponding charge and current densities are given by

$$\rho(t,\mathbf{x}) = \sum e_k\delta\big(\mathbf{x}-\mathbf{x}_k(t)\big), \qquad \mathbf{j}(t,\mathbf{x}) = \sum e_k\dot{\mathbf{x}}_k(t)\delta\big(\mathbf{x}-\mathbf{x}_k(t)\big), \tag{12.96}$$

where e_k are the charges of the particles. Further, according to the Lorentz equation (12.75), we *formally* obtain

$$\begin{aligned}\dot{T}_\Omega(t) &= \sum_{\mathbf{x}_k(t)\in\Omega} \mathbf{m}_k\dot{\mathbf{x}}_k(t)\cdot\ddot{\mathbf{x}}_k(t)\\ &= \sum_{\mathbf{x}_k(t)\in\Omega} \dot{\mathbf{x}}_k(t)\cdot e_k\left[\mathbf{E}\big(\mathbf{x}_k(t),t\big) + \frac{1}{c}\dot{\mathbf{x}}_k(t)\times\mathbf{B}\big(\mathbf{x}_k(t),t\big)\right]\\ &= \sum_{\mathbf{x}_k(t)\in\Omega} e_k\dot{\mathbf{x}}_k(t)\cdot\mathbf{E}\big(\mathbf{x}_k(t),t\big) = \int_\Omega \mathbf{j}(t,\mathbf{x})\cdot\mathbf{E}(t,\mathbf{x})\,d\mathbf{x},\end{aligned} \tag{12.97}$$

provided that the field $\mathbf{E}(t,\mathbf{x})$ is continuous in $\mathbf{x}$ and if particles do not cross the boundary $\Sigma := \partial\Omega$ at time t. Now we can calculate

$$\dot{\mathcal{E}}_\Omega(t) = \frac{1}{4\pi}\int_\Omega \big[\mathbf{E}(t,\mathbf{x})\cdot\dot{\mathbf{E}}(t,\mathbf{x}) + \mathbf{B}(t,\mathbf{x})\cdot\dot{\mathbf{B}}(t,\mathbf{x})\big]\,d\mathbf{x}. \tag{12.98}$$

Substituting $\dot{\mathbf{E}}(t,\mathbf{x})$ and $\dot{\mathbf{B}}(t,\mathbf{x})$ from the Maxwell equations (12.43), we obtain

$$\begin{aligned}\dot{\mathcal{E}}_\Omega(t) &= \frac{1}{4\pi}\int_\Omega \big[\mathbf{E}\cdot(c\,\mathrm{curl}\,\mathbf{B} - 4\pi\mathbf{j}) - c\mathbf{B}\cdot\mathrm{curl}\,\mathbf{E}\big]\,d\mathbf{x}\\ &= -\dot{T}_\Omega(t) + \frac{c}{4\pi}\int_\Omega \big[\mathbf{E}\cdot(\nabla\times\mathbf{B}) - \mathbf{B}\cdot(\nabla\times\mathbf{E})\big]\,d\mathbf{x},\end{aligned} \tag{12.99}$$

according to (12.97). It remains to rewrite the last integrand as follows

$$(\nabla\cdot\mathbf{B})\times\mathbf{E} - (\nabla\cdot\mathbf{E})\times\mathbf{B} = -\nabla\cdot(\mathbf{E}\times\mathbf{B}), \tag{12.100}$$

since $\nabla \cdot \mathbf{B} = 0$, by the Maxwell equations (12.43). Then (12.99) implies (12.94) by the Stokes theorem.

Note that our proof of (12.97) is not satisfactory, because the field $\mathbf{E}(t, \mathbf{x})$ of point particles is singular at the particle positions. A rigorous proof for the Abraham model of *extended electron* [2, 111, 176] follows by the same arguments from the Maxwell equations (12.43) with the following charge and current densities

$$\rho(t,\mathbf{x}) := \sum \rho_k\big(\mathbf{x}-\mathbf{x}_k(t)\big), \qquad \mathbf{j}(t,\mathbf{x}) := \sum \dot{\mathbf{x}}_k(t)\rho_k\big(\mathbf{x}-\mathbf{x}_k(t)\big) \tag{12.101}$$

and with the suitably modified Lorentz equation

$$\mathrm{m}\ddot{\mathbf{x}}_k(t) = \int \Big[\mathbf{E}(t,\mathbf{x}) + \frac{1}{c}\dot{\mathbf{x}}_k(t) \times \mathbf{B}(t,\mathbf{x})\Big]\rho_k\big(\mathbf{x}-\mathbf{x}_k(t)\big)\,d\mathbf{x}, \quad t \in \mathbb{R}. \tag{12.102}$$

Here $\rho_k(\mathbf{x})$ stands for sufficiently smooth *extended charge density* of the corresponding particle. In the case of one relativistic particle, the relation of type (12.94) is proved in [176, (2.61)].

12.8 Momentum of the Maxwell Field

The momentum density of the Maxwell field is defined by

$$\mathbf{g}(t,\mathbf{x}) := \frac{1}{c^2}\mathbf{S}(t,\mathbf{x}).$$

The corresponding flux is the tensor field

$$\mathbf{T}_{\alpha\beta}(t,\mathbf{x}) := \frac{1}{4\pi}\Big[\mathbf{E}_\alpha\mathbf{E}_\beta + \mathbf{B}_\alpha\mathbf{B}_\beta - \frac{1}{2}\big(\mathbf{E}^2+\mathbf{B}^2\big)\delta_{\alpha\beta}\Big].$$

This means that for any fixed region $\Omega \subset \mathbb{R}^3$ with smooth boundary $\Sigma = \partial\Omega$,

$$\begin{aligned} \frac{d}{dt}\big[\mathbf{P}_\Omega(t) + \mathbf{G}_\Omega(t)\big] &= -\int_\Sigma \mathbf{T}(t,\mathbf{x}) \cdot \mathbf{n}(\mathbf{x})d\Sigma(\mathbf{x}), \\ \mathbf{G}_\Omega(t) &:= \int_\Omega \mathbf{g}(t,\mathbf{x})\,d\mathbf{x}, \end{aligned} \tag{12.103}$$

where $\mathbf{n}(\mathbf{x})$ is the unit outward normal vector to Σ, $\mathbf{P}_\Omega(t)$ is the momentum of charged particles in Ω,

$$\mathbf{P}_\Omega(t) = \sum_{\mathbf{x}_k(t)\in\Omega} \mathbf{m}_k\dot{\mathbf{x}}_k(t), \tag{12.104}$$

and $[\mathbf{T}(t,\mathbf{x}) \cdot \mathbf{n}(\mathbf{x})]_\alpha = \sum \mathbf{T}_{\alpha\beta}(t,\mathbf{x})\mathbf{n}_\beta(\mathbf{x})$.

Using the Lorentz equation (12.75) we *formally* obtain

$$\begin{aligned}\dot{\mathbf{P}}_\Omega(t) = \sum_{\mathbf{x}_k(t)\in\Omega} \mathbf{m}_k\ddot{\mathbf{x}}_k(t) &= \sum_{\mathbf{x}_k(t)\in\Omega} e_k\Big[\mathbf{E}\big(\mathbf{x}_k(t),t\big) + \frac{1}{c}\dot{\mathbf{x}}_k(t)\times\mathbf{B}\big(\mathbf{x}_k(t),t\big)\Big]\\ &= \int_\Omega\Big[\rho(t,\mathbf{x})\mathbf{E}(\mathbf{x},t) + \frac{1}{c}\mathbf{j}(t,\mathbf{x})\times\mathbf{B}(\mathbf{x},t)\Big]d\mathbf{x}\end{aligned} \tag{12.105}$$

in terms of the charge and current densities (12.96). Let us express $\rho(t,\mathbf{x})$ and $\mathbf{j}(t,\mathbf{x})$ from the Maxwell equations (12.43). Then the integrand of (12.105) reads

$$\begin{aligned}&\frac{1}{4\pi}\Big[\mathbf{E}\operatorname{div}\mathbf{E} + \Big(\operatorname{curl}\mathbf{B} - \frac{1}{c}\dot{\mathbf{E}}\Big)\times\mathbf{B}\Big]\\ &\quad= \frac{1}{4\pi}[\mathbf{E}\operatorname{div}\mathbf{E} + \operatorname{curl}\mathbf{B}\times\mathbf{B} + \operatorname{curl}\mathbf{E}\times\mathbf{E}] - \frac{1}{4\pi c}\partial_t[\mathbf{E}\times\mathbf{B}]\\ &\quad= \sum\nabla_\beta\mathbf{T}_{\alpha\beta} - \frac{\dot{\mathbf{S}}}{c^2}.\end{aligned} \tag{12.106}$$

Hence, by (12.105),

$$\dot{\mathbf{P}}_\Omega(t) = \int_\Omega\sum\nabla_\beta\mathbf{T}_{\alpha\beta}\,d\mathbf{x} - \dot{\mathbf{G}}(t). \tag{12.107}$$

Now (12.103) follows from the Stokes theorem.

As above, our proof of (12.103) is not satisfactory, because the field $\mathbf{E}(t,\mathbf{x})$ of point particles is singular at the particle positions. For the *Abraham model* (12.101), (12.102) a rigorous proof follows by the methods of [111, 112, 176].

12.9 Abraham: Electromagnetic Mass

In 1902, Abraham calculated the energy and momentum of a moving ‘extended electron’, and discovered the celebrated mass-energy equivalence. The electron is surrounded by the comoving Maxwell field, which contributes to the electron kinetic energy, and hence, to the electron mass.

The densities of charge and current of a moving extended electron are as follows (12.66)

$$\begin{aligned}\rho(t,\mathbf{x}) &= \gamma\rho_e\big(\gamma(x_1 - vt), x_2, x_3\big),\\ \mathbf{j}(t,\mathbf{x}) &= \gamma\rho_e\big(\gamma(x_1 - vt), x_2, x_3\big)(v,0,0).\end{aligned} \tag{12.108}$$

Here $\rho_e(\mathbf{x})$ is the charge density of the electron at rest, v is the speed directed along Ox_1, and $\gamma = 1/\sqrt{1-\beta^2}$ is the Lorentz contraction, $\beta := v/c$.

By (12.47), the corresponding wave equations for the potentials can be written, in the Lorentz gauge (12.45), as follows:

$$\Box\phi_v(t,\mathbf{x}) = 4\pi\gamma\rho_e\big(\gamma(x_1 - vt), x_2, x_3\big), \tag{12.109}$$

$$\Box\mathbf{A}_v(t,\mathbf{x}) = \frac{4\pi}{c}\gamma\rho_e\big(\gamma(x_1 - vt), x_2, x_3\big)(v,0,0), \tag{12.110}$$

where $\Box = \frac{1}{c^2}\partial_t^2 - \Delta$. Then the corresponding Maxwell fields are given by

$$\mathbf{B}_v(t,\mathbf{x}) = \operatorname{curl}\mathbf{A}_v(t,\mathbf{x}), \qquad \mathbf{E}_v(t,\mathbf{x}) = -\nabla\phi_v(t,\mathbf{x}) - \frac{1}{c}\dot{\mathbf{A}}(t,\mathbf{x}). \tag{12.111}$$

Finally, the electromagnetic energy and momentum read

$$\mathcal{E}(v) = \frac{1}{8\pi}\int\big[\mathbf{E}_v^2(t,\mathbf{x}) + \mathbf{B}_v^2(t,\mathbf{x})\big]\,d\mathbf{x}, \qquad \mathbf{G}(v) = \frac{1}{4\pi c}\int \mathbf{E}_v(t,\mathbf{x}) \times \mathbf{B}_v(t,\mathbf{x})\,d\mathbf{x}. \tag{12.112}$$

Exercise 12.18 Prove the Abraham relation (1.14)

$$\mathbf{G}(v) = \frac{\mathrm{m}_e\mathbf{v}}{\sqrt{1-\beta^2}} \quad \text{with } \mathrm{m}_e = \frac{4}{3}\frac{\mathcal{E}(0)}{c^2}. \tag{12.113}$$

Solution Substituting the scalar potential of the form $\phi_v(t,\mathbf{x}) = \gamma\phi(\gamma(x_1 - vt), x_2, x_3)$ into (12.109), we obtain the corresponding stationary equation $-\Delta\phi(\mathbf{x}) = 4\pi\rho_e(\mathbf{x})$. Hence, $\phi(\mathbf{x})$ is the spherically symmetric Coulombic potential

$$\phi(\mathbf{x}) = \int \frac{\rho_e(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}\,d\mathbf{y} \tag{12.114}$$

Respectively, the magnetic potential reads $A_v(t,\mathbf{x}) = \beta(\phi_v(t,\mathbf{x}), 0, 0)$. Now let us compute the Maxwell fields by substituting this potentials into (12.111). We have

$$\mathbf{B}_v(t,\mathbf{x}) = (0, \gamma\beta\nabla_3\phi, -\gamma\beta\nabla_2\phi), \qquad \mathbf{E}_v(t,\mathbf{x}) = -(\nabla_1\phi, \gamma\nabla_2\phi, \gamma\nabla_3\phi), \tag{12.115}$$

where all the derivatives of ϕ are taken at the point $(\gamma(x_1 - vt), x_2, x_3)$. Calculating the vector product, we obtain

$$\mathbf{G}(v) = \frac{1}{4\pi c}\big(2\beta\gamma^2, 0, 0\big)\int \big|\nabla_1\phi(x_1,x_2,x_3)\big|^2\frac{dx_1}{\gamma}\,dx_2\,dx_3 \tag{12.116}$$

by the spherical symmetry of $\phi(\mathbf{x})$. Similarly,

$$\mathcal{E}(v) = \frac{1}{8\pi}\big[1 + 2\gamma^2 + 2\gamma^2\beta^2\big]\int \big|\nabla_1\phi(x_1,x_2,x_3)\big|^2\frac{dx_1}{\gamma}\,dx_2\,dx_3. \tag{12.117}$$

Now (12.113) follows.

12.10 The Hertzian Dipole Radiation

In 1886 Hertz started working on his experiments on radiation and detection of electromagnetic waves. He elaborated the corresponding mathematical theory by solving the Maxwell equations [87] for the *electric dipole* and obtaining the power of radiation in all directions, as well as its polarization. The *total time-average radiation* is given by the famous Hertz formula

$$\overline{P} = \frac{1}{3c^3}\mathbf{p}^2\omega^4. \tag{12.118}$$

The electric dipole at point zero with moment $\mathbf{p}(t)$ is defined as the charge distribution $\rho(t,\mathbf{x}) = -\mathbf{p}(t)\cdot\nabla\delta^3(\mathbf{x})$. This distribution generates the electromagnetic wave radiating energy to infinity. Formula (12.118) holds for the harmonic dipole with $\mathbf{p}(t) = \mathbf{p}\cos\omega t$.

The Dipole Potentials First, with the help of the continuity equation for charge and current $\dot{\rho} + \nabla\cdot\mathbf{j} = 0$, the corresponding dipole current density may be calculated as follows:

$$\mathbf{j}(t,\mathbf{x}) = \dot{\mathbf{p}}(t)\delta^3(\mathbf{x}). \tag{12.119}$$

Then retarded vector potential (12.160), as generated by this current, is

$$\begin{aligned}\mathbf{A}(t,\mathbf{x}) &= \frac{1}{c}\int ds \int d^3\mathbf{y}\,\frac{\mathbf{j}(s,\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}\delta\big(s + |\mathbf{x}-\mathbf{y}|/c - t\big)\\ &= \frac{1}{c}\int ds\,\dot{\mathbf{p}}(s)\frac{1}{|\mathbf{x}|}\delta\big(s + |\mathbf{x}|/c - t\big) = \frac{1}{c}\frac{\dot{\mathbf{p}}(t-r/c)}{r},\end{aligned} \tag{12.120}$$

where $r = |\mathbf{x}|$. Similarly, the retarded scalar potential (12.160) equals

$$\begin{aligned}\phi(t,\mathbf{x}) &= \int ds \int d^3\mathbf{y}\,\frac{\rho(s,\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}\delta\big(s + |\mathbf{x}-\mathbf{y}|/c - t\big)\\ &= \int ds\,\mathbf{p}(s)\cdot\nabla\left[\frac{1}{|\mathbf{x}-\mathbf{y}|}\delta\big(s + |\mathbf{x}-\mathbf{y}|/c - t\big)\right]\Bigg|_{\mathbf{y}=0}\\ &= \int ds\,\mathbf{p}(s)\cdot\left[\frac{\mathbf{x}-\mathbf{y}}{|\mathbf{x}-\mathbf{y}|^3}\,\delta\big(s + |\mathbf{x}-\mathbf{y}|/c - t\big)\right.\\ &\qquad \left. - \frac{1}{|\mathbf{x}-\mathbf{y}|}\delta'(s + |\mathbf{x}-\mathbf{y}|/c - t)\frac{\mathbf{x}-\mathbf{y}}{c|\mathbf{x}-\mathbf{y}|}\right]\Bigg|_{\mathbf{y}=0}\\ &= \left[\frac{\mathbf{p}(t-r/c)\cdot\mathbf{x}}{r^3} + \frac{\dot{\mathbf{p}}(t-r/c)\cdot\mathbf{x}}{cr^2}\right] = \frac{\mathbf{x}\cdot\dot{\mathbf{p}}(t-r/c)}{cr^2} + \mathcal{O}\big(r^{-2}\big), \quad r\to\infty.\end{aligned}$$

The Dipole Fields Now we can calculate the magnetic field,

$$\begin{aligned}\mathbf{B}(t,\mathbf{x}) &= \nabla\times\mathbf{A}(t,\mathbf{x}) = \frac{1}{c}\nabla\times\frac{\dot{\mathbf{p}}(t-r/c)}{r}\\ &= \frac{1}{c}\dot{\mathbf{p}}(t-r/c)\times\frac{\mathbf{x}}{r^3} + \frac{1}{cr}\nabla(t-r/c)\times\ddot{\mathbf{p}}(t-r/c)\\ &= \frac{\ddot{\mathbf{p}}(t-r/c)\times\mathbf{x}}{c^2r^2} + \mathcal{O}\big(r^{-2}\big) \end{aligned}\tag{12.121}$$

and, in a similar fashion, the electric field

$$\begin{aligned}\mathbf{E}(t,\mathbf{x}) &= -\frac{1}{c}\dot{\mathbf{A}}(t,\mathbf{x}) - \nabla\phi(t,\mathbf{x}) = -\frac{1}{c^2}\frac{\ddot{\mathbf{p}}}{r} - \nabla\frac{\mathbf{x}\cdot\dot{\mathbf{p}}(t-r/c)}{cr^2} + \mathcal{O}\big(r^{-2}\big)\\ &= -\frac{1}{c^2}\frac{\ddot{\mathbf{p}}}{r} - \frac{1}{cr^2}\big(\dot{\mathbf{p}} + \mathbf{x}^j\nabla\big[\dot{\mathbf{p}}_j(t-r/c)\big]\big) - \frac{1}{c}\mathbf{x}\cdot\dot{\mathbf{p}}\nabla r^{-2} + \mathcal{O}\big(r^{-2}\big)\\ &= -\frac{1}{c^2}\frac{\ddot{\mathbf{p}}}{r} + \frac{\mathbf{x}(\mathbf{x}\cdot\ddot{\mathbf{p}})}{c^2r^3} + \mathcal{O}\big(r^{-2}\big) = \frac{1}{c^2r^3}\big(\ddot{\mathbf{p}}(t-r/c)\times\mathbf{x}\big)\times\mathbf{x} + \mathcal{O}\big(r^{-2}\big). \end{aligned}\tag{12.122}$$

We observe that $\mathbf{x}\perp\mathbf{E}\perp\mathbf{B}$; i.e., this is a radiation field.

The Dipole Radiation Now for the energy flow (the Poynting vector), we find

$$\begin{aligned}\mathbf{S}(t,\mathbf{x}) &= \frac{c}{4\pi}\mathbf{E}\times\mathbf{B} \sim \frac{1}{4\pi c^3r^5}\big[\ddot{\mathbf{p}}(t-r/c)\times\mathbf{x})\times\mathbf{x}\big]\times(\ddot{\mathbf{p}}\times\mathbf{x})\\ &= \frac{\mathbf{x}}{4\pi c^3r^3}\big|\big[1-|\mathbf{n}\rangle\langle\mathbf{n}|\big]\ddot{\mathbf{p}}\big|^2. \end{aligned}\tag{12.123}$$

Equivalently,

$$\mathbf{S}(t,\mathbf{x}) \sim \mathbf{n}\frac{\sin^2\chi}{4\pi c^3r^2}\ddot{\mathbf{p}}^2(t-r/c), \quad r\gg 1,\tag{12.124}$$

where $\mathbf{n}=\mathbf{x}/r$, and χ is the angle between $\mathbf{n}$ and $\ddot{\mathbf{p}}(t-r/c)$. For the fixed angle, averaging over time, gives

$$\overline{\mathbf{S}(t,\mathbf{x})} \sim \mathbf{n}\frac{\sin^2\chi}{4\pi c^3r^2}\overline{\ddot{\mathbf{p}}^2}, \quad r\gg 1.\tag{12.125}$$

Hence *mean total energy radiation* is given by

$$\overline{P} = \int_{|\mathbf{x}|=r}\overline{\mathbf{S}(\mathbf{x})}\cdot\mathbf{n}\,d\mathbf{x} = \frac{\overline{\ddot{\mathbf{p}}^2}}{4\pi c^3}\int_{|\mathbf{x}|=1}\sin^2\chi\,d\mathbf{x} = \frac{2}{3c^3}\overline{\ddot{\mathbf{p}}^2}.\tag{12.126}$$

Remark 12.19

(i) $\mathbf{S}$ behaves like r^{-2} for large distances, and hence the total energy flow through a large sphere is bounded in case $|\ddot{\mathbf{p}}(t)|$ is bounded.

(ii) All rejected terms with rapid decay at infinity do not contribute to the total energy flow through a large sphere.

(iii) There is no radiation (energy flow) in the direction of the dipole (for $\chi = 0, \pi$). In the orthogonal directions, for $\chi = \pi/2$, the electric field is parallel to $\mathbf{p}(t - r/c)$, while the magnetic field is orthogonal to $\mathbf{p}(t - r/c)$.

(iv) $\mathbf{S}$ is directed *outwards*—this means the energy radiation to infinity, which is a characteristic property of the retarded potentials.

Application to the Harmonic Oscillator For the harmonic oscillator we have $\mathbf{p}(t) = \mathbf{p}\cos\omega t$ with $\omega \neq 0$, and hence averaging over time $\overline{\ddot{\mathbf{p}}^2} = \overline{\mathbf{p}^2}\omega^4$ and using (12.126), it is found that

$$\overline{P} = \frac{2}{3c^3}\overline{\mathbf{p}^2}\omega^4. \tag{12.127}$$

Now (12.118) follows, because $\overline{\mathbf{p}^2} = \mathbf{p}^2/2$. Formula (12.125) holds with 2/3 instead of $\sin^2\chi$, provided that the orientation of the harmonic oscillator is *random isotropic*:

$$\overline{\mathbf{S}(t,\mathbf{x})} \sim \mathbf{n}\frac{\mathbf{p}^2\omega^4}{12\pi c^3 r^2}, \quad r \gg 1. \tag{12.128}$$

12.11 The Initial Problem for the Maxwell Equations

Solutions of the Maxwell equations (12.43) are uniquely determined by the initial conditions

$$\mathbf{E}|_{t=0} = \mathbf{E}_0(\mathbf{x}), \qquad \mathbf{B}|_{t=0} = \mathbf{B}_0(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^3. \tag{12.129}$$

By (12.103), the total energy of the Maxwell field is given by

$$W(t) = \frac{1}{8\pi}\int_{\mathbb{R}^3}\left[\mathbf{E}^2(t,\mathbf{x}) + \mathbf{B}^2(t,\mathbf{x})\right]d\mathbf{x}. \tag{12.130}$$

Hence for finite energy solutions, we have $(\mathbf{E}(t,\cdot), \mathbf{B}(t,\cdot)) \in L^2 \oplus L^2$ for all $t \in \mathbb{R}$, where $L^2 := L^2(\mathbb{R}^3) \otimes \mathbb{R}^3$. Therefore, it is natural to look for a solution $(\mathbf{E}(t,\cdot), \mathbf{B}(t,\cdot)) \in C(\mathbb{R}, L^2 \oplus L^2)$ of the *initial problem* (12.43), (12.129).

For the existence of such a solution it is necessary that $(\mathbf{E}_0(\mathbf{x}), \mathbf{B}_0(\mathbf{x})) \in L^2 \oplus L^2$. Moreover, from the Maxwell equations it follows that

$$\operatorname{div}\mathbf{E}_0(\mathbf{x}) = 4\pi\rho(0,\mathbf{x}), \qquad \operatorname{div}\mathbf{B}_0(\mathbf{x}) = 0, \quad \mathbf{x} \in \mathbb{R}^3; \tag{12.131}$$

these are necessary constraints for the existence of solutions of the overdetermined system (12.43).

Finally, it is natural to assume that

$$\rho(t,\mathbf{x}) \in C\big(\mathbb{R}, L^2(\mathbb{R}^3)\big), \qquad \mathbf{j}(t,\mathbf{x}) \in C\big(\mathbb{R}, L^2\big). \tag{12.132}$$

Lemma 12.20 *Let* $\mathbf{E}_0(x), \mathbf{B}_0(x)$ *and* $\rho(t,\mathbf{x}), \mathbf{j}(t,\mathbf{x})$ *satisfy all the above conditions. Then*

(i) *The initial problem* (12.43), (12.129) *has a unique solution* $(\mathbf{E}(t,\cdot),\mathbf{B}(t,\cdot))\in C(\mathbb{R},L^2\oplus L^2)$;

(ii) *Let* $\mathbf{j}(t,\mathbf{x})\equiv 0$. *Then the corresponding field energy* (12.130) *is conserved*:

$$\int_{\mathbb{R}^3}\big[\mathbf{E}^2(t,\mathbf{x})+\mathbf{B}^2(t,\mathbf{x})\big]\,d\mathbf{x}=\text{const},\qquad t\in\mathbb{R}^3. \tag{12.133}$$

(iii) *The convolution representation holds*

$$\begin{pmatrix}\mathbf{E}(t)\\ \mathbf{B}(t)\end{pmatrix}=\mathbf{M}(t)*\begin{pmatrix}\mathbf{E}_0\\ \mathbf{B}_0\end{pmatrix}+4\pi\int_0^t\mathbf{G}(t-s)*\begin{pmatrix}\rho(s)\\ \frac{1}{c}\mathbf{j}(s)\end{pmatrix}ds,\qquad t\in\mathbb{R}, \tag{12.134}$$

where $\mathbf{E}(t):=\mathbf{E}(t,\cdot)$, *etc*., *and* $\mathbf{M}(t)$ *and* $\mathbf{G}(t)$ *are*, *respectively*, *the* 6×6 *and the* 6×4 *matrix-valued distributions concentrated on the sphere* $|\mathbf{x}|=|t|$ *for every fixed* $t\in\mathbb{R}$:

$$\mathbf{M}(t)(\mathbf{x})=0,\qquad \mathbf{G}(t,\mathbf{x})=0\quad \textit{for } |\mathbf{x}|\neq|t|. \tag{12.135}$$

Proof Step (i) We introduce the complex field $\mathbf{C}(t,\mathbf{x})=\mathbf{E}(t,\mathbf{x})+i\mathbf{B}(t,\mathbf{x})$ and rewrite (12.43) as

$$\frac{1}{c}\dot{\mathbf{C}}(t,\mathbf{x})=-i\operatorname{curl}\mathbf{C}(t,\mathbf{x})-\frac{4\pi}{c}\mathbf{j}(t,\mathbf{x}),\qquad \operatorname{div}\mathbf{C}(t,\mathbf{x})=4\pi\rho(t,\mathbf{x}), \tag{12.136}$$

$$\mathbf{C}|_{t=0}=\mathbf{C}_0(\mathbf{x}), \tag{12.137}$$

where $\mathbf{C}_0(\mathbf{x})=\mathbf{E}_0(\mathbf{x})+i\mathbf{B}_0(\mathbf{x})$.

Step (ii) Equations (12.136) and (12.137) in the Fourier transform $\hat{\mathbf{C}}(t,\mathbf{k})=\int\exp(i\mathbf{k}\cdot x)\mathbf{C}(t,\mathbf{x})\,d\mathbf{x}$ read as

$$\dot{\hat{\mathbf{C}}}(t,\mathbf{k})=c\hat{m}(\mathbf{k})\hat{\mathbf{C}}(t,\mathbf{k})-4\pi\hat{\mathbf{j}}(t,\mathbf{k}),\qquad -i\mathbf{k}\cdot\hat{\mathbf{C}}(t,\mathbf{k})=4\pi\hat{\rho}(t,\mathbf{k}), \tag{12.138}$$

$$\hat{\mathbf{C}}|_{t=0}=\hat{\mathbf{C}}_0(\mathbf{k}), \tag{12.139}$$

where $\hat{m}(\mathbf{k})$ denotes the 3×3 skew-adjoint matrix of the operator $-\mathbf{k}\times$ in $\mathbb{C}^3$.

The solution $\hat{C}(t,\mathbf{k})$ is defined uniquely from the first equation of (12.138) of the overdetermined system (12.138), (12.139),

$$\hat{\mathbf{C}}(t,\mathbf{k})=\exp\big(c\hat{m}(\mathbf{k})t\big)\hat{\mathbf{C}}_0(k)-4\pi\int_0^t\exp\big(c\hat{m}(\mathbf{k})(t-s)\big)\hat{\mathbf{j}}(s,\mathbf{k})\,ds. \tag{12.140}$$

We still have to show that (12.140) satisfies the constraint from (12.138):

$$S(t,\mathbf{k}):=4\pi\hat{\rho}(t,\mathbf{k})+i\mathbf{k}\cdot\hat{\mathbf{C}}(t,\mathbf{k})\equiv 0. \tag{12.141}$$

Indeed, the Fourier transformed equations (12.131), (12.3) are

$$-i\mathbf{k}\cdot\hat{\mathbf{C}}_0(\mathbf{k}) = 4\pi\hat{\rho}(0,\mathbf{k}), \tag{12.142}$$

$$\dot{\hat{\rho}}(t,\mathbf{k}) - i\mathbf{k}\cdot\hat{\mathbf{j}}(t,\mathbf{k}) = 0. \tag{12.143}$$

Hence

$$\begin{aligned} S(0,\mathbf{k}) &= 4\pi\hat{\rho}(0,\mathbf{k}) + i\mathbf{k}\cdot\hat{C}_0(\mathbf{k}) = 0,\\ \dot{S}(t,\mathbf{k}) &= 4\pi\dot{\hat{\rho}}(t,\mathbf{k}) - 4\pi i\mathbf{k}\cdot\hat{\mathbf{j}}(t,\mathbf{k}) = 0, \end{aligned} \tag{12.144}$$

where we employed the first equation of (12.138) and $\mathbf{k}\cdot\hat{m}(\mathbf{k})\hat{\mathbf{C}}(t,\mathbf{k}) = 0$. Now (12.141) follows from (12.144). Finally, (12.140) implies that $\hat{C}(t,\cdot) \in C(\mathbb{R}, L^2\otimes\mathbb{C})$, and hence $C(t,\cdot) \in C(\mathbb{R}, L^2\otimes\mathbb{C})$, by Plancherel's theorem. This proves Lemma 12.20(i).

Step (iii) From the Plancherel identity it follows that

$$\int_{\mathbb{R}^3}\left[\mathbf{E}^2(t,\mathbf{x}) + \mathbf{B}^2(t,\mathbf{x})\right]d\mathbf{x} = \int_{\mathbb{R}^3}\left|\mathbf{C}(t,\mathbf{x})\right|^2 d\mathbf{x} = (2\pi)^{-3}\int_{\mathbb{R}^3}\left|\hat{\mathbf{C}}(t,\mathbf{k})\right|^2 d\mathbf{k}. \tag{12.145}$$

As a result, (12.133) follows from (12.140), because $\mathbf{j}(t,\mathbf{k}) \equiv 0$ and $\exp(c\hat{m}(\mathbf{k})t)$ is a unitary matrix.

Step (iv) Finally, we need to transform (12.140) back to position space in order to check (12.135). We have $\hat{m} = \hat{m}(\mathbf{k}) = -\mathbf{k}\times$, $\hat{m}^2 = -\mathbf{k}^2 + |\mathbf{k}\rangle\langle\mathbf{k}|$, $\hat{m}^3 = -|\mathbf{k}|^2\hat{m}, \ldots$. Hence,

$$\hat{m}^{2j+1} = (-1)^j|\mathbf{k}|^{2j}\hat{m} = (-1)^j\frac{\hat{m}}{|\mathbf{k}|}|\mathbf{k}|^{2j+1} \quad \text{for } j \geq 0,$$

$$\hat{m}^{2j} = \hat{m}^{2j-1}\hat{m} = (-1)^{j-1}|\mathbf{k}|^{2j-2}\hat{m}^2 = -(-1)^j\left(\frac{\hat{m}}{|\mathbf{k}|}\right)^2|\mathbf{k}|^{2j}, \quad j \geq 1,$$

and further, using Euler's trick,

$$\begin{aligned} \exp\bigl(\hat{m}(\mathbf{k})t\bigr) &= \sum_0^\infty(\hat{m}t)^n/n! = \sum_0^\infty(\hat{m}t)^{2j}/(2j)! + \sum_0^\infty(\hat{m}t)^{2j+1}/(2j+1)\\ &= 1 + \left(\frac{\hat{m}}{|\mathbf{k}|}\right)^2\bigl(1-\cos|\mathbf{k}|t\bigr) + \frac{\hat{m}}{|\mathbf{k}|}\sin|\mathbf{k}|t\\ &= \cos|\mathbf{k}|t + \hat{m}\frac{\sin|\mathbf{k}|t}{|\mathbf{k}|} + \bigl(1-\cos|\mathbf{k}|t\bigr)\frac{|\mathbf{k}\rangle\langle\mathbf{k}|}{|\mathbf{k}|^2}. \end{aligned} \tag{12.146}$$

We set $\hat{K}(t,\mathbf{k}) = \sin|\mathbf{k}|t/|\mathbf{k}|$, $\hat{\mathbf{m}}(t,\mathbf{k}) = \partial_t\hat{K}(t,\mathbf{k}) + \hat{m}\hat{K}(t,\mathbf{k})$, and $\hat{D}(t,\mathbf{k}) = 1-\cos|\mathbf{k}|t$. Then we finally obtain,

$$\exp\bigl(c\hat{m}(\mathbf{k})t\bigr) = \hat{\mathbf{m}}(ct,\mathbf{k}) + |\mathbf{k}\rangle\frac{\hat{D}(ct,\mathbf{k})}{|\mathbf{k}|^2}\langle\mathbf{k}|. \tag{12.147}$$

Inserting this into (12.140) and using (12.142) and (12.143), it follows that

$$\begin{aligned}\hat{\mathbf{C}}(t,\mathbf{k}) &= \hat{\mathbf{m}}(ct,\mathbf{k})\hat{\mathbf{C}}_0(\mathbf{k}) + 4\pi i|\mathbf{k}\rangle\frac{\hat{D}(ct,\mathbf{k})}{|\mathbf{k}|^2}\hat{\rho}(\mathbf{k},0)\\ &\quad - 4\pi\int_0^t\left[\hat{\mathbf{m}}\bigl(c(t-s),\mathbf{k}\bigr)\hat{\mathbf{j}}(s,\mathbf{k}) - i|\mathbf{k}\rangle\frac{\hat{D}(c(t-s),\mathbf{k})}{|\mathbf{k}|^2}\dot{\hat{\rho}}(s,\mathbf{k})\right]ds.\end{aligned} \tag{12.148}$$

Further, integrating by parts, this becomes

$$\begin{aligned}\hat{\mathbf{C}}(t,\mathbf{k}) &= \hat{\mathbf{m}}(ct,\mathbf{k})\hat{\mathbf{C}}_0(\mathbf{k})\\ &\quad - 4\pi\int_0^t\left[\hat{\mathbf{m}}\bigl(c(t-s),\mathbf{k}\bigr)\hat{\mathbf{j}}(s,\mathbf{k}) - i|\mathbf{k}\rangle\frac{\partial_t\hat{D}(c(t-s),\mathbf{k})}{|\mathbf{k}|^2}\hat{\rho}(s,\mathbf{k})\right]ds.\end{aligned} \tag{12.149}$$

Step (v) Using $\partial_t\hat{D}(t,\mathbf{k}) = |\mathbf{k}|\sin|\mathbf{k}|t = |\mathbf{k}|^2\hat{K}(t,\mathbf{k})$, this gives the inverse Fourier transforms,

$$\mathbf{m}(t,\mathbf{x}) := F^{-1}_{\mathbf{k}\mapsto\mathbf{x}}\hat{\mathbf{m}}(t,\mathbf{k}) = \partial_t K(t,\mathbf{x}) - i\,\mathrm{curl}\circ K(t,\mathbf{x}), \tag{12.150}$$

$$\mathbf{g}(t,\mathbf{x}) := F^{-1}_{\mathbf{k}\mapsto\mathbf{x}}\bigl(i|\mathbf{k}\rangle\hat{K}(t,\mathbf{k}),\ -\hat{\mathbf{m}}(t,\mathbf{k})\bigr) = \bigl(-\nabla K(t,\mathbf{x}), -\mathbf{m}(t,\mathbf{x})\bigr), \tag{12.151}$$

where $K(t,x)$ denotes the Kirchhoff kernel

$$K(t,\mathbf{x}) := F^{-1}\hat{K}(t,\mathbf{k}) = \frac{1}{4\pi t}\delta\bigl(|\mathbf{x}|-|t|\bigr). \tag{12.152}$$

In this notation, (12.149) implies (12.134) in the ‘complex’ form

$$\mathbf{C}(t) = \mathbf{m}(ct)*\mathbf{C}_0 + 4\pi\int_0^t\mathbf{g}\bigl(c(t-s)\bigr)*\begin{pmatrix}c\rho(s)\\ \mathbf{j}(s)\end{pmatrix}ds, \quad t\in\mathbb{R}. \tag{12.153}$$

Separating into real and imaginary parts, we obtain

$$\mathbf{E}(t,\mathbf{x}) = \mathbf{E}_{(r)}(t,\mathbf{x}) + \mathbf{E}_{(0)}(t,\mathbf{x}), \qquad \mathbf{B}(t,\mathbf{x}) = \mathbf{B}_{(r)}(t,\mathbf{x}) + \mathbf{B}_{(0)}(t,\mathbf{x}), \tag{12.154}$$

where we denote

$$\begin{pmatrix}\mathbf{E}_{(0)}(t)\\ \mathbf{B}_{(0)}(t)\end{pmatrix} = \begin{pmatrix}\partial_t K(ct) & \mathrm{curl}\circ K(ct)\\ -\,\mathrm{curl}\circ K(ct) & \partial_t K(ct)\end{pmatrix} * \begin{pmatrix}\mathbf{E}_0\\ \mathbf{B}_0\end{pmatrix} \tag{12.155}$$

and

$$\begin{pmatrix} \mathbf{E}_{(r)}(t) \\ \mathbf{B}_{(r)}(t) \end{pmatrix} = 4\pi \int_0^t \begin{pmatrix} -\nabla K(c(t-s)) & -\partial_t K(c(t-s)) \\ 0 & \operatorname{curl} \circ K(c(t-s)) \end{pmatrix} * \begin{pmatrix} c\rho(s) \\ \mathbf{j}(s) \end{pmatrix} ds. \tag{12.156}$$

Now (12.134) and (12.135) immediately follow. □

12.12 The Long Time Asymptotics: Retarded Potentials

The *retarded potentials* are particular solutions of the wave and Maxwell equations. We explain the outstanding role of these potentials as regards the long time asymptotics.

Formula (12.156) may be rewritten as follows

$$\mathbf{E}_{(r)}(t,\mathbf{x}) = -\nabla\phi_{(r)}(t,\mathbf{x}) - \frac{1}{c}\dot{\mathbf{A}}_{(r)}(t,\mathbf{x}), \qquad \mathbf{B}_{(r)}(t,\mathbf{x}) = \operatorname{curl}\mathbf{A}_{(r)}(t,\mathbf{x}), \tag{12.157}$$

where the potentials are given by

$$\phi_{(r)}(t,\mathbf{x}) = \int \frac{\Theta(t_{\text{ret}})}{|\mathbf{x}-\mathbf{y}|}\rho(t_{\text{ret}},\mathbf{y})\,d\mathbf{y}, \qquad \mathbf{A}_{(r)}(t,\mathbf{x}) = \frac{1}{c}\int \frac{\Theta(t_{\text{ret}})}{|\mathbf{x}-\mathbf{y}|}\mathbf{j}(t_{\text{ret}},\mathbf{y})\,d\mathbf{y}, \tag{12.158}$$

where $t_{\text{ret}} = t - |\mathbf{x}-\mathbf{y}|/c$. Let us assume that the charge and current densities are continuous and localized in space,

$$\rho(t,\mathbf{x}) = 0, \qquad \mathbf{j}(t,\mathbf{x}) = 0, \quad |\mathbf{x}| > R,\ t \in \mathbb{R}. \tag{12.159}$$

Then potentials (12.158) for $ct > R + |\mathbf{x}|$ become the *retarded potentials* [93],

$$\left\{\begin{aligned} \phi_{(r)}(t,\mathbf{x}) &= \phi_{\text{ret}}(t,\mathbf{x}) := \int \frac{\rho(t-|\mathbf{x}-\mathbf{y}|/c,\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}\,d\mathbf{y}, \\ \mathbf{A}_{(r)}(t,\mathbf{x}) &= \mathbf{A}_{\text{ret}}(t,\mathbf{x}) := \frac{1}{c}\int \frac{\mathbf{j}(t-|\mathbf{x}-\mathbf{y}|/c,\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}\,d\mathbf{y}, \end{aligned}\right| \quad ct > R + |\mathbf{x}|. \tag{12.160}$$

Similarly, fields (12.157) for $ct > R + |\mathbf{x}|$ become the *retarded fields*

$$\left\{\begin{aligned} \mathbf{E}_{(r)}(t,\mathbf{x}) &= -\nabla\phi_{\text{ret}}(t,\mathbf{x}) - \dot{A}_{\text{ret}}(t,\mathbf{x}), \\ \mathbf{B}_{(r)}(t,\mathbf{x}) &= \operatorname{curl}\mathbf{A}_{\text{ret}}(t,\mathbf{x}), \end{aligned}\right| \quad ct > R + |\mathbf{x}|. \tag{12.161}$$

A distinguished role of the particular retarded solutions (12.160) of the wave equations (12.47) is justified in *scattering theory*. Namely, as we have proved above, the solutions to the wave equations (12.47) are defined uniquely by the initial conditions at time zero:

$$\phi|_{t=0} = \phi_0(\mathbf{x}), \qquad \dot{\phi}|_{t=0} = \pi_0(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^3, \tag{12.162}$$

$$\mathbf{A}|_{t=0} = \mathbf{A}_0(\mathbf{x}), \qquad \dot{\mathbf{A}}|_{t=0} = \Pi_0(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^3. \tag{12.163}$$

However, the **long time asymptotics** of solutions as $t \to +\infty$ **do not depend** on the initial data for **space-localized** functions $\phi_0, \pi_0, \mathbf{A}_0, \Pi_0$, and coincide with the retarded potentials (12.160). For example, let us consider the initial functions with compact supports.

Lemma 12.21 *Let* (12.159) *hold and let the initial functions* $\phi_0(x), \pi_0(x), \mathbf{A}_0(x)$, $\Pi_0(x)$ *be continuous and localized in space*,

$$\phi_0(\mathbf{x}) = \pi_0(\mathbf{x}) = 0, \qquad \mathbf{A}_0(\mathbf{x}) = \Pi_0(\mathbf{x}) = 0, \quad |\mathbf{x}| > R_0. \tag{12.164}$$

Then for the large time the solutions of the initial problems (12.47), (12.162) *coincide with the retarded potentials* (12.160):

$$\phi(t,\mathbf{x}) = \phi_{\mathrm{ret}}(t,\mathbf{x}), \quad \mathbf{A}(t,\mathbf{x}) = \mathbf{A}_{\mathrm{ret}}(t,\mathbf{x}), \quad ct > R_0 + |\mathbf{x}|. \tag{12.165}$$

Proof Let us prove this lemma for the scalar potential $\phi(t,\mathbf{x})$. The Kirchhoff formula for the solution reads

$$\phi(t,\mathbf{x}) = \frac{1}{4\pi ct}\int_{S_{ct}(\mathbf{x})} \pi_0(\mathbf{y})\,dS(\mathbf{y}) + \partial_t\left(\frac{1}{4\pi t}\int_{S_{ct}(\mathbf{x})} \phi_0(\mathbf{y})\,dS(\mathbf{y})\right) + \phi_{\mathrm{ret}}(t,\mathbf{x}),$$
$$t > 0, \tag{12.166}$$

where $S_{ct}(\mathbf{x})$ denotes the sphere $\{\mathbf{y} \in \mathbb{R}^3 : |\mathbf{x}-\mathbf{y}| = ct\}$, and $dS(\mathbf{y})$ is the Lebesgue measure on the sphere. Now (12.165) follows from (12.164). □

Similar theorem holds for the Maxwell equations.

Lemma 12.22 *Let* (12.159) *and the conditions of Theorem* 12.20 *hold, and*

$$\mathbf{E}_0(\mathbf{x}) = \mathbf{B}_0(\mathbf{x}) = 0, \quad |\mathbf{x}| > R_0. \tag{12.167}$$

Then

$$\mathbf{E}(t,\mathbf{x}) = \mathbf{E}_{\mathrm{ret}}(t,\mathbf{x}), \qquad \mathbf{B}(t,\mathbf{x}) = \mathbf{B}_{\mathrm{ret}}(t,\mathbf{x}), \quad ct > R_0 + |\mathbf{x}|. \tag{12.168}$$

Proof (12.168) follows from (12.154) and (12.161) by (12.167), (12.155) and (12.152) since $\mathbf{E}_{(0)}, \mathbf{B}_{(0)}$ vanish for $ct > R_0 + |\mathbf{x}|$. □

Finite Energy Initial Data For the Maxwell equations with $\mathbf{j} = 0$, the energy is conserved by (12.133). For the wave equations (12.47) with $\rho = 0$ and $\mathbf{j} = 0$, the energy conservations read,

$$\begin{cases} \displaystyle\int \left[\frac{1}{c^2} \left| \dot{\phi}(t, \mathbf{x}) \right|^2 + \left| \nabla \phi(t, \mathbf{x}) \right|^2 \right] d\mathbf{x} = \text{const}, \\ \displaystyle\int \left[\frac{1}{c^2} \left| \dot{\mathbf{A}}(t, \mathbf{x}) \right|^2 + \left| \nabla \mathbf{A}(t, \mathbf{x}) \right|^2 \right] d\mathbf{x} = \text{const}. \end{cases} \tag{12.169}$$

For solutions with finite initial energy, (12.165) and (12.168) hold asymptotically as $t \to +\infty$ in the local energy semi-norms.

Lemma 12.23 *Suppose that* (12.159) *holds. Let* $\phi(t, \mathbf{x})$, $\mathbf{A}(t, \mathbf{x})$ *and* $\mathbf{E}(t, \mathbf{x})$, $\mathbf{B}(t, \mathbf{x})$, *respectively, be finite energy solutions of the wave equations* (12.47) *and of the Maxwell equations* (12.43). *Then, for any* $R > 0$,

$$\left\{ \begin{array}{l} \displaystyle\int_{|\mathbf{x}|<R} \left[\left| \dot{\phi}(t, \mathbf{x}) - \dot{\phi}_{\text{ret}}(t, \mathbf{x}) \right|^2 + \left| \nabla \phi(t, \mathbf{x}) - \nabla \phi_{\text{ret}}(t, \mathbf{x}) \right|^2 \right] d\mathbf{x} \to 0, \\ \displaystyle\int_{|\mathbf{x}|<R} \left[\left| \dot{\mathbf{A}}(t, \mathbf{x}) - \dot{\mathbf{A}}_{\text{ret}}(t, \mathbf{x}) \right|^2 + \left| \nabla \mathbf{A}(t, \mathbf{x}) - \nabla \mathbf{A}_{\text{ret}}(t, \mathbf{x}) \right|^2 \right] d\mathbf{x} \to 0, \\ \displaystyle\int_{|\mathbf{x}|<R} \left[\left| \mathbf{E}(t, \mathbf{x}) - \mathbf{E}_{\text{ret}}(t, \mathbf{x}) \right|^2 + \left| \mathbf{B}(t, \mathbf{x}) - \mathbf{B}_{\text{ret}}(t, \mathbf{x}) \right|^2 \right] d\mathbf{x} \to 0, \end{array} \right| \quad t \to +\infty. \tag{12.170}$$

Proof We split the initial functions in two components: the first, space-localized component, similarly to (12.164) and (12.167), and the remainder, which is small in the energy norm. We denote by $\phi', \mathbf{A}', \mathbf{E}', \mathbf{B}'$ the solutions corresponding to the space-localized initial states, with given charge and current densities. On the other hand, let $\phi'', \mathbf{A}'', \mathbf{E}'', \mathbf{B}''$ be the solutions corresponding to the remainder initial states and to zero charge and current densities.

Now asymptotics (12.170) hold for $\phi', \mathbf{A}', \mathbf{E}', \mathbf{B}'$, instead of $\phi, \mathbf{A}, \mathbf{E}, \mathbf{B}$, by (12.165) and (12.168).

Finally, the contributions of ϕ'', $\mathbf{A}''$, $\mathbf{E}''$, $\mathbf{B}''$ into integrals (12.170) are uniformly small in time, by the energy conservation law (12.169) for ϕ'' and $\mathbf{A}''$, and by (12.133), for $\mathbf{E}''$ and $\mathbf{B}''$. □

12.13 The Lorentz Theory of Polarization and Magnetization

In 1890s, Lorentz developed a phenomenological theory of matter in the Maxwell field in the context of his *electron theory*. This approach works equally well in the classical and quantum contexts.

The state of a neutral molecule is characterized by its *dipole moment and magnetic moment*. The parameters allow one to describe with high precision the field, generated by the molecule, at 'large distances' from the molecule. Hence, the parameters give a complete description of the molecular field for any macroscopic observation, which can be made only at distances much larger than the size of a molecule. Hence, it suffices to specify the influence of external fields onto these parameters by the corresponding *constitutive equations*.

Suppose that a molecule is located near the origin; i.e.,

$$\rho(t,\mathbf{y})=0, \qquad \mathbf{j}(t,\mathbf{y})=0, \quad |x|>a,\ t\in\mathbb{R}. \tag{12.171}$$

where $a>0$ denotes the size of a molecule, and $\rho,\mathbf{j}$ denote the corresponding molecular charge and current densities.

12.13.1 Stationary Molecular Fields in Dipole Approximation

First, we consider the static case when densities are independent of time. In this case, equations (12.47) become stationary Poisson equations and their solutions are the Coulomb potentials

$$\phi(\mathbf{x})=\int_{|\mathbf{y}|\le a}\frac{\rho(\mathbf{y})\,d\mathbf{y}}{|\mathbf{x}-\mathbf{y}|}, \qquad \mathbf{A}(\mathbf{x})=\frac{1}{c}\int_{|\mathbf{y}|\le a}\frac{\mathbf{j}(\mathbf{y})\,d\mathbf{y}}{|\mathbf{x}-\mathbf{y}|}, \quad \mathbf{x}\in\mathbb{R}^3. \tag{12.172}$$

Expanding $1/|\mathbf{x}-\mathbf{y}|$ in a Taylor series for small $|\mathbf{y}|\le a$ and large $|\mathbf{x}|\gg a$

$$\frac{1}{|\mathbf{x}-\mathbf{y}|}=\frac{1}{\sqrt{\mathbf{x}^2+\mathbf{y}^2-2\mathbf{y}\mathbf{x}}}=\frac{1}{|\mathbf{x}|}+\frac{\mathbf{y}\mathbf{x}}{|\mathbf{x}|^3}+\mathcal{O}\Big(\frac{1}{|\mathbf{x}|^3}\Big), \tag{12.173}$$

transforms (12.172) into

$$\left\{\begin{aligned}\phi(\mathbf{x})&=\frac{Q}{|\mathbf{x}|}+\frac{\mathbf{p}\mathbf{x}}{|\mathbf{x}|^3}+\mathcal{O}\Big(\frac{a^2}{|\mathbf{x}|^3}\Big),\\ \mathbf{A}(\mathbf{x})&=\frac{\mathbf{J}}{c|\mathbf{x}|}+\frac{\mathcal{M}\mathbf{x}}{|\mathbf{x}|^3}+\mathcal{O}\Big(\frac{a^2}{|\mathbf{x}|^3}\Big),\end{aligned}\right| \quad |\mathbf{x}|\gg a, \tag{12.174}$$

where we set

$$\left\{\begin{aligned}Q&=\int\rho(\mathbf{y})\,d\mathbf{y}, \qquad &\mathbf{p}&=\int\mathbf{y}\rho(\mathbf{y})\,d\mathbf{y},\\ \mathbf{J}&=\int\mathbf{j}(\mathbf{y})\,d\mathbf{y}, \qquad &\mathcal{M}_{kl}&=\frac{1}{c}\int\mathbf{j}_k(\mathbf{y})\mathbf{y}_l\,d\mathbf{y}.\end{aligned}\right. \tag{12.175}$$

We will identify molecular fields with the first two terms in (12.174), because $|\mathbf{x}|/a\gg 1$ in all *macroscopic observations*. This corresponds to the following *dipole approximations* for $\rho(\mathbf{y})$ and $\mathbf{j}(\mathbf{y})$:

$$\rho_d(\mathbf{y})=Q\delta(\mathbf{y})-\mathbf{p}\cdot\nabla\delta(\mathbf{y}), \qquad \mathbf{j}_d(\mathbf{y})=\mathbf{J}\delta(\mathbf{y})-c\mathcal{M}\nabla\delta(\mathbf{y}). \tag{12.176}$$

12.13.2 Multipole Expansions of Non-Stationary Fields

Asymptotic formulas of type (12.174) can be extended to non-stationary localized densities satisfying (12.171). In this case, the corresponding Maxwell field is represented by the retarded potentials (12.160):

$$\left\{\begin{aligned} \phi(t,\mathbf{x}) &= \int_{|\mathbf{y}|\le a} \frac{\rho(t-|\mathbf{x}-\mathbf{y}|/c,\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}\,d\mathbf{y} \approx \int_{|\mathbf{y}|\le a} \frac{\rho(t-|\mathbf{x}|/c,\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}\,d\mathbf{y},\\ A(t,\mathbf{x}) &= \frac{1}{c}\int_{|\mathbf{y}|\le a} \frac{\mathbf{j}(t-|\mathbf{x}-\mathbf{y}|/c,\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}\,d\mathbf{y} \approx \frac{1}{c}\int_{|\mathbf{y}|\le a} \frac{\mathbf{j}(t-|\mathbf{x}|/c,\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}\,d\mathbf{y}, \end{aligned}\right| \quad |\mathbf{x}| \gg a.$$

Continuing the Taylor expansion (12.173), we obtain a complete series of the type (12.174), including all negative powers of $|\mathbf{x}|$:

$$\frac{1}{|\mathbf{x}-\mathbf{y}|} = \frac{1}{|\mathbf{x}|} + \frac{\mathbf{y}\mathbf{x}}{|\mathbf{x}|^3} + \sum_{|\alpha|\ge 2} \frac{\mathbf{y}^\alpha P_\alpha(\mathbf{n})}{|\mathbf{x}|^{|\alpha|+1}}, \quad |\mathbf{y}| \le a < |\mathbf{x}|, \tag{12.177}$$

where $\mathbf{n} := \mathbf{x}/|\mathbf{x}|$ and $|P_\alpha(\mathbf{n})| \le C(1+|\alpha|)^2$. The series converge, provided that $|\mathbf{y}| < a$ and $|\mathbf{x}| > a$. Hence, the retarded potentials may be expressed by the converging series

$$\left\{\begin{aligned} \phi(t,\mathbf{x}) &= \frac{Q(t-|\mathbf{x}|/c)}{|\mathbf{x}|} + \frac{\mathbf{p}(t-|\mathbf{x}|/c)\mathbf{x}}{|\mathbf{x}|^3} + \sum_{|\alpha|\ge 2} \frac{\phi_\alpha(t-|\mathbf{x}|/c)P_\alpha(\mathbf{n})}{|\mathbf{x}|^{|\alpha|+1}},\\ \mathbf{A}(t,\mathbf{x}) &= \frac{\mathbf{J}(t-|\mathbf{x}|/c)}{c|\mathbf{x}|} + \frac{\mathcal{M}(t-|\mathbf{x}|/c)\mathbf{x}}{|\mathbf{x}|^3} + \sum_{|\alpha|\ge 2} \frac{A_\alpha(t-|\mathbf{x}|/c)P_\alpha(\mathbf{n})}{|\mathbf{x}|^{|\alpha|+1}}, \end{aligned}\right| \quad |\mathbf{x}| \gg a, \tag{12.178}$$

where in particular

$$\left\{\begin{aligned} Q(t) &= \int_{|\mathbf{y}|\le a} \rho(t,\mathbf{y})\,d\mathbf{y}, \qquad & \mathbf{p}(t) &= \int_{|\mathbf{y}|\le a} \mathbf{y}\rho(t,\mathbf{y})\,d\mathbf{y},\\ \mathbf{J}(t) &= \int_{|\mathbf{y}|\le a} \mathbf{j}(t,\mathbf{y})\,d\mathbf{y}, \qquad & \mathcal{M}_{kl}(t) &= \frac{1}{c}\int_{|\mathbf{y}|\le a} \mathbf{j}_k(t,\mathbf{y})\mathbf{y}_l\,d\mathbf{y}. \end{aligned}\right. \tag{12.179}$$

These series correspond to the following *multipole expansions* for $\rho(t,\mathbf{y})$ and $\mathbf{j}(t,\mathbf{y})$:

$$\left\{\begin{aligned} \rho(t,\mathbf{y}) &= Q(t)\delta(\mathbf{y}) - \mathbf{p}(t)\cdot\nabla\delta(\mathbf{y}) + \sum_{|\alpha|\ge 2} \rho_\alpha(t)\nabla^\alpha\delta(\mathbf{y}),\\ \mathbf{j}(t,\mathbf{y}) &= \mathbf{J}(t)\delta(\mathbf{y}) - c\mathcal{M}(t)\nabla\delta(\mathbf{y}) + \sum_{|\alpha|\ge 2} \mathbf{j}_\alpha(t)\nabla^\alpha\delta(\mathbf{y}), \end{aligned}\right. \tag{12.180}$$

where

$$\rho_\alpha(t) = \int_{|\mathbf{y}|\le a} \frac{(-\mathbf{y})^\alpha}{\alpha!}\rho(t,\mathbf{y})\,d\mathbf{y}, \qquad \mathbf{j}_\alpha(t) = \int_{|\mathbf{y}|\le a} \frac{(-\mathbf{y})^\alpha}{\alpha!}\mathbf{j}(t,\mathbf{y})\,d\mathbf{y}, \tag{12.181}$$

by (12.171). The series (12.180) converge in the sense of distributions dual to the space of test functions $\varphi(\mathbf{x})$ with

$$\frac{|\nabla^\alpha \varphi(0)|}{\alpha!} \le C b^{-|\alpha|}, \quad 0 \le |\alpha| < \infty, \tag{12.182}$$

where $b > a$. For instance, formula (12.177) with $|\mathbf{x}| > a$ gives a complete set of such test functions. In particular, the continuity equation for charge and current

$$\dot\rho(t,\mathbf{y}) + \nabla \cdot \mathbf{j}(t,\mathbf{y}) = 0 \tag{12.183}$$

holds in this sense. Substituting here the series (12.180), we obtain

$$\dot Q(t)\delta(\mathbf{y}) - \dot{\mathbf{p}}(t)\cdot\nabla\delta(\mathbf{y}) + \mathbf{J}(t)\cdot\nabla\delta(\mathbf{y}) + \sum_{|\alpha|\ge 2} C_\alpha(t)\nabla^\alpha\delta(\mathbf{y}) = 0, \tag{12.184}$$

and hence

$$\dot Q(t) = 0, \qquad \mathbf{J}(t) = \dot{\mathbf{p}}(t). \tag{12.185}$$

12.13.3 Molecule in a Stationary State: Magnetic Moment

Let us consider a molecule in a stationary state; i.e., $\rho(t,\mathbf{y}) \equiv \rho(\mathbf{y})$ and $\mathbf{j}(t,\mathbf{y}) \equiv \mathbf{j}(\mathbf{y})$. Then the multipole expansions (12.180) do not depend on time,

$$\begin{cases} \rho(t,\mathbf{y}) \equiv \rho(\mathbf{y}) = Q\delta(\mathbf{y}) - \mathbf{p}\cdot\nabla\delta(\mathbf{y}) + \sum\limits_{|\alpha|\ge 2} \rho_\alpha \nabla^\alpha\delta(\mathbf{y}), \\ \mathbf{j}(t,\mathbf{y}) \equiv \mathbf{j}(\mathbf{y}) = \mathbf{J}\delta(\mathbf{y}) - c\mathcal{M}\nabla\delta(\mathbf{y}) + \sum\limits_{|\alpha|\ge 2} \mathbf{j}_\alpha \nabla^\alpha\delta(\mathbf{y}). \end{cases} \tag{12.186}$$

Now the second identity of (12.185) gives that

$$\mathbf{J} = 0. \tag{12.187}$$

Lemma 12.24 *Let the multipole charge-current densities (12.186) correspond to a stationary state of a molecule. Then the matrix $\mathcal{M}$ of (12.174), (12.175) is skew-symmetric and*

$$\mathcal{M}\nabla = \mathbf{m} \times \nabla, \tag{12.188}$$

*where $\mathbf{m} \in \mathbb{R}^3$ is the **magnetic moment** of the molecule.*

Proof Substituting (12.186) into (12.183), it is found that

$$\nabla \cdot \mathbf{j}(\mathbf{y}) = 0. \tag{12.189}$$

Therefore, in particular

$$\nabla \cdot \left[\mathcal{M}\nabla\delta(\mathbf{y})\right] = 0. \tag{12.190}$$

Hence, $\mathcal{M}_{kl} + \mathcal{M}_{lk} = 0$, and so the matrix $\mathcal{M}$ admits the representation

$$\mathcal{M} = \begin{pmatrix} 0 & -\mathbf{m}_3 & \mathbf{m}_2 \\ \mathbf{m}_3 & 0 & -\mathbf{m}_1 \\ -\mathbf{m}_2 & \mathbf{m}_1 & 0 \end{pmatrix} \tag{12.191}$$

with a vector $\mathbf{m} \in \mathbb{R}^3$. □

Further, by (12.175),

$$\mathbf{m} = \frac{1}{2c}\int \mathbf{y} \times \mathbf{j}(\mathbf{y})\, d\mathbf{y}, \tag{12.192}$$

and so the integral is independent of the choice of the origin, because $\int \mathbf{j}(\mathbf{y})\, d\mathbf{y} = 0$ by (12.187).

From now on we will assume the following:

The Adiabatic Hypothesis Molecular dynamics can be described as an *adiabatic evolution* of stationary states with corresponding dipole electric moment $\mathbf{p}(t)$ and magnetic moment $\mathbf{m}(t)$.

12.13.4 The Maxwell Equations in Matter

Polarization and Magnetization Let us consider a neutral molecule, so $Q = 0$. Now (12.185)–(12.188) imply, in the dipole approximation,

$$\rho(t,\mathbf{x}) \approx -\mathbf{p}(t)\cdot\nabla\delta(\mathbf{x}), \qquad \mathbf{j}(t,\mathbf{x}) \approx \dot{\mathbf{p}}(t)\,\delta(\mathbf{x}) + c\nabla\times\mathbf{m}(t)\delta(\mathbf{x}) \tag{12.193}$$

for the charge and current densities of a molecule. The total molecular charge–current densities $\rho_{\text{mol}}(t,\mathbf{x})$ and $\mathbf{j}_{\text{mol}}(t,\mathbf{x})$ in matter are sums of contributions of all molecules. In the dipole approximation (12.193), the densities are given by

$$\rho_{\text{mol}}(t,\mathbf{x}) \approx -\nabla\cdot\mathbf{P}_d(t,\mathbf{x}), \qquad \mathbf{j}_{\text{mol}}(t,\mathbf{x}) \approx \dot{\mathbf{P}}_d(t,\mathbf{x}) + c\nabla\times\mathbf{M}_d(t,\mathbf{x}), \tag{12.194}$$

where

$$\mathbf{P}_d(t,\mathbf{x}) = \sum_n \mathbf{p}_n(t)\cdot\nabla\delta(\mathbf{x}-\mathbf{x}_n), \qquad \mathbf{M}_d(t,\mathbf{x}) = \sum_n \mathbf{m}_n(t)\delta(\mathbf{x}-\mathbf{x}_n). \tag{12.195}$$

These singular distributions and equations (12.194) were obtained by dipole approximation of molecular charge and current densities for $|\mathbf{x}| \gg a$.

Now we consider the limit $a \to 0$: the corresponding limiting **electric polarization** $\mathbf{P}(t, \mathbf{x})$ and **magnetization** $\mathbf{M}(t, \mathbf{x})$ are the solutions to the system of type (12.194),

$$\rho_{\text{mol}}(t, \mathbf{x}) = -\nabla \cdot \mathbf{P}(t, \mathbf{x}), \qquad \mathbf{j}_{\text{mol}}(t, \mathbf{x}) = \dot{\mathbf{P}}(t, \mathbf{x}) + c\nabla \times \mathbf{M}(t, \mathbf{x}). \tag{12.196}$$

For instance, a solution $\mathbf{M}(t, \mathbf{x})$ exists, since $\nabla \cdot [\mathbf{j}_{\text{mol}} - \dot{\mathbf{P}}] = \nabla \cdot \mathbf{j}_{\text{mol}} + \dot{\rho}_{\text{mol}} = 0$. Moreover, the equations can be written as

$$\nabla \cdot \mathcal{B}(t, \mathbf{x}) = 4\pi \rho_{\text{mol}}, \qquad \nabla \times \mathcal{B}(t, \mathbf{x}) = \frac{4\pi}{c} \mathbf{j}_{\text{mol}}(t, \mathbf{x}) + \frac{1}{c} \dot{\mathcal{E}}(t, \mathbf{x}), \tag{12.197}$$

where

$$\mathcal{E}(t, \mathbf{x}) := -4\pi \mathbf{P}(t, \mathbf{x}), \qquad \mathcal{B}(t, \mathbf{x}) := 4\pi \mathbf{M}(t, \mathbf{x}). \tag{12.198}$$

Equations (12.197) coincide with the first and last Maxwell equations (12.43), corresponding to molecular charges and currents. However, to uniquely identify the fields, we need two additional vector equations.

Below we will assume that $\mathcal{E}$ and $\mathcal{B}$ are proportional to $\mathbf{E}$ and $\mathbf{B}$ respectively (see (12.200) and (12.202)). Then the field $\mathcal{B}$ also satisfies the third Maxwell equation. Finally, $\mathcal{E}$ and $\mathcal{B}$ also satisfy the second Maxwell equation (which is the Faraday induction equation), if $\mathbf{E}$, $\mathbf{B}$ do not depend on time as well as $\mathcal{E}$, $\mathcal{B}$. Thus, **in the static case**, $\mathcal{E}$ and $\mathcal{B}$ coincide with the electric and magnetic fields generated by the molecular charges and currents.

The Dielectric Displacement and Magnetic Intensity Let us separate the macroscopic and molecular charge and current densities in the Maxwell equations (12.43):

$$\rho(t, \mathbf{x}) = \rho_{\text{mac}}(t, \mathbf{x}) + \rho_{\text{mol}}(t, \mathbf{x}), \qquad \mathbf{j}(t, \mathbf{x}) = \mathbf{j}_{\text{mac}}(t, \mathbf{x}) + \mathbf{j}_{\text{mol}}(t, \mathbf{x}). \tag{12.199}$$

Further, we introduce new fields

$$\mathbf{D}(t, \mathbf{x}) := \mathbf{E}(t, \mathbf{x}) - \mathcal{E}(t, \mathbf{x}), \qquad \mathbf{H}(t, \mathbf{x}) := \mathbf{B}(t, \mathbf{x}) - \mathcal{B}(t, \mathbf{x}), \tag{12.200}$$

which are called the *dielectric displacement* and *magnetic field intensity*. Finally, substituting expressions (12.199) and (12.200) into the Maxwell equations (12.43), and using (12.197), we obtain the *Maxwell equations in matter*:

$$\begin{cases} \operatorname{div} \mathbf{D}(t, \mathbf{x}) = 4\pi \rho_{\text{mac}}(t, \mathbf{x}), \qquad \operatorname{curl} \mathbf{E}(t, \mathbf{x}) = -\dfrac{1}{c} \dot{\mathbf{B}}(t, \mathbf{x}), \\ \operatorname{div} \mathbf{B}(t, \mathbf{x}) = 0, \qquad \operatorname{curl} \mathbf{H}(t, \mathbf{x}) = \dfrac{1}{c} \dot{\mathbf{D}}(t, \mathbf{x}) + \dfrac{4\pi}{c} \mathbf{j}_{\text{mac}}(t, \mathbf{x}). \end{cases} \tag{12.201}$$

The Constitutive Equations Equations (12.201) contain two additional unknown vector fields $\mathbf{D}$, $\mathbf{H}$. Therefore, we need two additional vector equations to make the theory self-consistent. For **isotropic** materials, the simplest hypothesis are linear **constitutive equations**

$$\mathbf{D}(t, \mathbf{x}) = \varepsilon \mathbf{E}(t, \mathbf{x}), \qquad \mathbf{B}(t, \mathbf{x}) = \mu \mathbf{H}(t, \mathbf{x}), \tag{12.202}$$

where ε is called the **permittivity** and μ is called the **permeability** of matter. The **constitutive equations** can be rewritten also as follows

$$\mathbf{P}(t,\mathbf{x}) = \chi_e \mathbf{E}(t,\mathbf{x}), \qquad \mathbf{M}(t,\mathbf{x}) = \chi_m \mathbf{H}(t,\mathbf{x}), \tag{12.203}$$

where χ_e is called the **electric susceptibility** and χ_m is called the **magnetic susceptibility** of matter, and in this case

$$\varepsilon = 1 + 4\pi\chi_e, \qquad \mu = 1 + 4\chi_m, \tag{12.204}$$

by (12.200). With constitutive equations (12.202), the choice of fields $\mathbf{P}, \mathbf{M}$ and $\mathbf{D}, \mathbf{H}$ is uniquely fixed. In practice, the fields are measured by observations in the corresponding cylindrical cavities [93]. It is worth noting that Eqs. (12.196), (12.201) were suggested by the dipole approximations (12.194).

12.13.5 The Refraction Coefficient

We use (12.202) to eliminate **D** and **H** from (12.201). Introducing potentials as in (12.44), and using the 'Lorentz' gauge condition $\frac{\varepsilon\mu}{c}\dot{\phi}(t,\mathbf{x}) + \operatorname{div}\mathbf{A}(t,\mathbf{x}) = 0$, we obtain, as in (12.47),

$$\Box_{\text{mac}}\phi(t,\mathbf{x}) = \frac{4\pi}{\varepsilon}\rho_{\text{mac}}(t,\mathbf{x}), \qquad \Box_{\text{mac}}\mathbf{A}(t,\mathbf{x}) = \frac{4\pi\mu}{c}\mathbf{j}_{\text{mac}}(t,\mathbf{x}), \tag{12.205}$$

where $\Box_{\text{mac}} := \frac{\varepsilon\mu}{c^2}\partial_t^2 - \Delta$.

Corollary 12.25 *Let $\varepsilon\mu > 0$. Then the speed of propagation of the electromagnetic field in matter is $c/\sqrt{\varepsilon\mu}$, and hence the refraction coefficient with respect to vacuum is given by*

$$n = \sqrt{\varepsilon\mu}. \tag{12.206}$$

Chapter 13
Mathematical Appendices

Abstract The Lagrangian and Hamiltonian formalism provide a unifying language for modern field theory.

The Hamilton–Jacobi theory and geometrical optics are considered as one of the main issues for the introduction of the Schrödinger equation and quantum observables.

We give an updated version of Noether's theorem on currents and give its applications to the conservation laws for the Schrödinger equation and to the charge continuity.

The limiting amplitude principle provides an explanation of the photoelectric effect.

13.1 The Lagrangian and Hamiltonian Mechanics

13.1.1 The Lagrangian Mechanics

The Newton equations for n particles

$$m_i \ddot{\mathbf{x}}_i(t) = -\nabla_{\mathbf{x}_i} V\bigl(x(t), t\bigr), \quad x(t) := \bigl(\mathbf{x}_1(t), \dots, \mathbf{x}_n(t)\bigr) \tag{13.1}$$

can be represented as a Lagrangian system. The corresponding Lagrangian $L(x, v, t)$ is the function on the *extended phase space* $\mathcal{E}^+ = \mathbb{R}^{3n} \times \mathbb{R}^{3n} \times \mathbb{R}$,

$$L(x, v, t) = \sum_i \frac{m_i \mathbf{v}_i^2}{2} - V(x, t), \quad (x, v, t) \in \mathcal{E}^+, \tag{13.2}$$

where $v = (\mathbf{v}_1, \dots, \mathbf{v}_n)$. The corresponding **Euler–Lagrange** equations read as

$$\frac{d}{dt} L_v\bigl(x(t), \dot{x}(t), t\bigr) = L_x\bigl(x(t), \dot{x}(t), t\bigr), \quad t \in \mathbb{R}. \tag{13.3}$$

We consider more general systems with extended phase space $\mathcal{E}^+ := \mathbb{R}^N \times \mathbb{R}^N \times \mathbb{R}$, where $N = 1, 2, \dots$ and a Lagrangian function $L(x, v, t) \in C^2(\mathcal{E}^+)$. The momentum and energy of the Lagrangian system are defined, respectively, by

$$p = L_v(x, v, t), \tag{13.4}$$

A. Komech, *Quantum Mechanics: Genesis and Achievements*,

DOI 10.1007/978-94-007-5542-0_13,

$$E(x, v, t) = pv - L(x, v, t). \tag{13.5}$$

Then $p = (\mathbf{p}_1, \ldots, \mathbf{p}_n)$, where $\mathbf{p}_i = m_i \mathbf{v}_i$, and

$$E = \sum_i \frac{m_i \mathbf{v}_i^2}{2} + V(x, t). \tag{13.6}$$

Lemma 13.1 *Suppose that the Lagrangian functional is independent of time; i.e., $L(x, v, t) = L(x, v)$. Then the energy is conserved: $E(t) := E(x(t), v(t)) = \text{const}$.*

Proof Differentiating, we obtain $\dot{E}(t) = \dot{p}v + p\dot{v} - L_x\dot{x} - L_v\dot{v} = 0$ by (13.4) and (13.3). □

13.1.2 The Legendre Transformation and Hamiltonian Mechanics

The Legendre transformation can be viewed as a change of variables in the extended phase space $\mathcal{E}^+ := \mathbb{R}^N \times \mathbb{R}^N \times \mathbb{R}$, translating the Euler–Lagrange equations into the Hamiltonian form.

A function $f(v)$ on $\mathbb{R}^N$ defines the map of the extended phase space $\mathbb{R}^N$ into itself defined by $\lambda : v \mapsto p := f_v(v)$. The *Legendre* transform of $f(v)$ is the function on $\mathbb{R}^N$ defined by $(\Lambda f)(p) := pv - f(v)$ for $p = \lambda(v)$, provided that the Legendre map λ is a diffeomorphism $\mathbb{R}^N \to \lambda(\mathbb{R}^N)$.

By the inverse function theorem, λ is the (local) diffeomorphism if and only if $\det |f_{vv}| \neq 0$ for $v \in \mathbb{R}^N$. For example, the last inequality holds for $f(v) = \sum_1^N m_k v_k^2$ if all $m_k > 0$.

Exercise 13.2 Calculate $\Lambda v^2 = p^2/4$, $\Lambda v^4 = 3v^4 = 3(p/4)^{4/3}, \ldots$.

Exercise 13.3 Prove that $\Lambda(\Lambda f) = f$ if $\lambda : \mathbb{R}^N \to \lambda(\mathbb{R}^N)$ is a diffeomorphism.

We extend the Legendre transformation to Lagrangian functions $L(x, v, t)$ as

$$H(x, p, t) = pv - L(x, v, t), \quad p := L_v(x, v, t) \tag{13.7}$$

for every fixed x, t. The function $H(x, p, t)$ coincides with the energy (13.5), where v is expressed in p.

Lemma 13.4 *Suppose that $v \mapsto p := L_v(x, v, t)$ is a diffeomorphism $\mathbb{R}^N \to \lambda(\mathbb{R}^N)$ for all fixed $(x, t) \in \mathbb{R}^N \times \mathbb{R}$. Then the Euler–Lagrange equations (13.3) are equivalent to the Hamiltonian ones,*

$$\dot{x}(t) = H_p\big(x(t), p(t), t\big), \qquad \dot{p}(t) = -H_x\big(x(t), p(t), t\big). \tag{13.8}$$

Proof The first equation of (13.8) follows by differentiation of the first identity (13.7) with fixed x, t:

$$H_p = v + pv_p - L_v v_p = v = \dot{x}, \tag{13.9}$$

because, by definition, $p = L_v$. The second equation of (13.8) follows by differentiation of the first identity (13.7) with fixed p, t: by (13.3), we have $H_x = -L_x = -\dot{p}$. □

Exercise 13.5 Verify that $H(x, p, t) = \sum_i \frac{\mathbf{p}_i^2}{2m_i} + V(x, t)$ for Lagrangian (13.2).

Exercise 13.6 Calculate the momentum, energy and the Hamilton function for the Lagrangian of a relativistic particle

$$L(\mathbf{x}, \mathbf{v}) = -\mathrm{m}c^2\sqrt{1 - \left(\frac{\mathbf{v}}{c}\right)^2}, \qquad (\mathbf{x}, \mathbf{v}) \in \mathbb{R}^3 \times \mathbb{R}^3. \tag{13.10}$$

Solution:

$$\begin{aligned} \mathbf{p} &= \frac{\mathrm{m}\mathbf{v}}{\sqrt{1 - (\frac{\mathbf{v}}{c})^2}}, \qquad & \mathbf{v} &= \frac{\mathbf{p}}{\sqrt{\mathrm{m}^2 + (\frac{\mathbf{p}}{c})^2}}, \\ E &= \frac{\mathrm{m}c^2}{\sqrt{1 - (\frac{\mathbf{v}}{c})^2}}, \qquad & H &= c^2\sqrt{\mathrm{m}^2 + \left(\frac{\mathbf{p}}{c}\right)^2}. \end{aligned} \tag{13.11}$$

13.1.3 The Hamilton–Jacobi Equation

Let $L(x, v, t)$ be a Lagrangian function on $\mathcal{E}^+ = \mathbb{R}^N \times \mathbb{R}^N \times \mathbb{R}$ and let $H(x, p, t)$ be the corresponding Hamilton function. Consider the initial problem

$$-\dot{S}(x, t) = H\big(x, \nabla S(x, t), t\big), \tag{13.12}$$

$$S(x, 0) = S_0(x), \quad x \in \mathbb{R}^N, \tag{13.13}$$

where $S_0(x) \in C^1(\mathbb{R}^N)$ is a given function. Our aim is to describe the Hamilton–Jacobi method for construction of the solution (see [4, 73, 129]).

Namely, consider the corresponding initial problem for the Hamilton system,

$$\begin{cases} \dot{x}(t) = H_p\big(x(t), p(t), t\big), & \dot{p}(t) = -H_x\big(x(t), p(t), t\big) \\ x|_{t=0} = x_0, & p|_{t=0} = \nabla S_0(x_0) \end{cases} \tag{13.14}$$

with $x_0 \in \mathbb{R}^N$. Let $(x(x_0, t), p(x_0, t))$ be the solution of this system, which is known to exist and be C^1-smooth for small $|t|$ depending on x_0. Consider the function $\mathcal{S}$

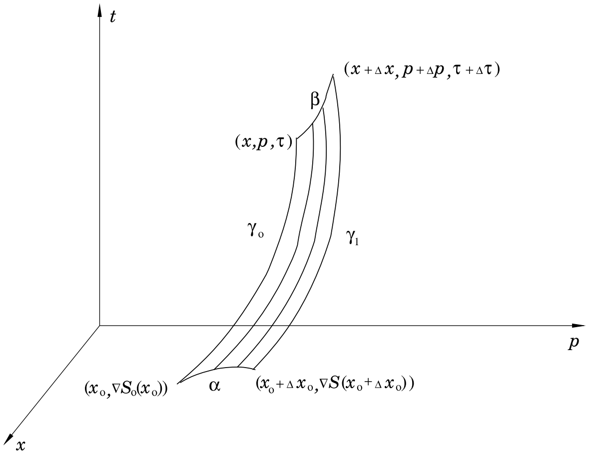


Fig. 13.1 The Lagrangian manifold

defined by the action integral

$$\mathcal{S}(x_0, t) = S_0(x_0) + \int_0^t L\big(x(x_0, s), \dot{x}(x_0, s), s\big)\, ds \tag{13.15}$$

for $x_0 \in \mathbb{R}^N$ and small $|t|$. At last, we express x_0 in terms of $x(x_0, t)$ for small $|t|$: this is possible, because the Jacobian matrix $\frac{\partial x(x_0,t)}{\partial x_0} = E$ for $t = 0$. We have $x_0 = x_0(x, t)$, where $x_0(x, t) \in C^1(\mathbb{R}^N \times \mathbb{R})$, and so we can define, for small $|t|$,

$$S(x, t) = \mathcal{S}\big(x_0(x, t), t\big), \quad x \in \mathbb{R}^N. \tag{13.16}$$

Theorem 13.7 *Function* (13.16) *is the* (*unique*) *solution of the initial problem* (13.12) *for small* $|t|$.

Proof This result follows from suitable properties of the differential 1-form $\omega = p\,dx - H\,dt$ known as the *Poincaré–Cartan integral invariant* [4].

Step (i). Consider $x_0, x_0 + \Delta x_0 \in \mathbb{R}^N$ and small $|\tau|$, $|\tau + \Delta\tau|$. By $\mathcal{M}_\tau$ we denote the following two-dimensional submanifold of the extended phase space $\mathcal{E}^+$ (see Fig. 13.1),

$$\mathcal{M}_\tau = \big\{x(x_0 + s\Delta x_0, t), p\big((x_0 + s\Delta x_0, t), t\big) : s \in [0, 1], t \in [0, \tau + s\Delta\tau]\big\}. \tag{13.17}$$

We have $\partial\mathcal{M}_\tau = \alpha \cup \gamma_1 \cup \beta \cup \gamma_0$, where

$$\left|\begin{aligned}
&\alpha := \big\{\big(x_0 + s\Delta x_0, \nabla S_0(x_0 + s\Delta x_0), 0\big) : s \in [0, 1]\big\},\\
&\beta := \big\{\big(x(x_0 + s\Delta x_0, t + s\Delta\tau), p(x_0 + s\Delta x_0, t + s\Delta\tau), t + s\Delta\tau\big) : s \in [0, 1]\big\},\\
&\gamma_0 := \big\{\big(x(x_0, t), p(x_0, t), t\big) : t \in [0, \tau]\big\},\\
&\gamma_1 := \big\{\big(x(x_0 + \Delta x_0, t), p(x_0 + \Delta x_0, t)\big) : t \in [0, \tau + \Delta\tau]\big\}.
\end{aligned}\right. \tag{13.18}$$

Let us orient the intervals according to increment of parameters s, t. Then, by the Stokes theorem,

$$\int_{\mathcal{M}_\tau} d\omega = \int_\alpha \omega + \int_{\gamma_1} \omega - \int_\beta \omega - \int_{\gamma_0} \omega. \tag{13.19}$$

Step (ii). The key observation is that the restriction of $d\omega$ onto $\mathcal{M}_\tau$ vanishes,

$$d\omega|_{\mathcal{M}_\tau} = 0. \tag{13.20}$$

In other words, $\mathcal{M}_\tau$ is the *Lagrangian manifold* by the following two facts:

(i) $\mathcal{H} := (H_p, -H_x, 1)$ is a tangent vector field to $\mathcal{M}_\tau$ at every point, and
(ii) For every vector field $\mathcal{V}$ in a neighborhood $\mathcal{O}(\mathcal{M}_\tau)$ of $\mathcal{M}_\tau$, we have

$$d\omega(\mathcal{H}, \mathcal{V})|_{\mathcal{O}(\mathcal{M}_\tau)} \equiv 0. \tag{13.21}$$

The last identity holds by the following arguments. First,

$$d\omega = dp \wedge dx - dH \wedge dt = dp \wedge dx - (H_p\, dp + H_x\, dx) \wedge dt$$

is antisymmetric bilinear form with matrix

$$A = \begin{pmatrix} 0 & E & H_x \\ -E & 0 & H_p \\ -H_x & -H_p & 0 \end{pmatrix}, \tag{13.22}$$

where E is the $N \times N$ identity matrix. Second, $A\mathcal{H} \equiv 0$, and hence

$$d\omega(\mathcal{H}, \mathcal{V}) = \langle A\mathcal{H}, \mathcal{V} \rangle \equiv 0. \tag{13.23}$$

Step (iii) Now (13.19) reads as

$$0 = \int_\alpha \nabla S_0(x)\, dx + \int_{\gamma^1} L\, dt - \int_\beta (p\, dx - H\, dt) - \int_{\gamma^0} L\, dt, \tag{13.24}$$

because $dt|_\alpha = 0$, $p|_\alpha = \nabla S_0(x)$ and $\omega|_{\gamma^i} = L\, dt$. The first term on the RHS of (13.24) is equal to $S_0(x_0 + \Delta x_0) - S_0(x_0)$. Therefore, by (13.15) and (13.16),

$$\begin{aligned} \int_\beta (p\, dx - H\, dt) &= \left[S_0(x_0 + \Delta x_0) + \int_{\gamma^1} L\, dt \right] - \left[S_0(x_0) + \int_{\gamma^0} L\, dt \right] \\ &= S(x + \Delta x, \tau + \Delta\tau) - S(x, \tau), \end{aligned} \tag{13.25}$$

where $x + \Delta x = x(x_0 + \Delta x_0, \tau + \Delta\tau)$ and $x = x(x_0, \tau)$. Finally, (13.25) implies

$$\dot{S}(x, t) = -H(x, p, t), \quad \nabla S(x, t) = p; \tag{13.26}$$

Eq. (13.12) now follows. □

13.2 Geometrical Optics

The short wavelength solutions of the Schrödinger equation propagate along *rays*, which are trajectories of the corresponding classical Lorentz equation. A mathematical approach to justification of this *geometrical optics* was discovered by Fresnel in 1815, and continued by Rayleigh in 1877, Debye in 1909, Jeffreys in 1923, and by Schrödinger, Wentzel, Kramers and Brillouin in 1926.

The Lorentz equation (1.6) for a classical particle in the Maxwell field admits the Hamilton form (3.31),

$$\begin{cases}\dot{\mathbf{x}}(t)=\nabla_{\mathbf{p}}\mathcal{H}(\mathbf{x},\mathbf{p},t)=\frac{1}{\mathrm{m}}\Big(\mathbf{p}-\frac{e}{c}\mathbf{A}(t,\mathbf{x})\Big),\\ \dot{\mathbf{p}}(t)=-\nabla_{\mathbf{x}}\mathcal{H}(\mathbf{x},\mathbf{p},t)=-e\nabla\phi(t,\mathbf{x})+\frac{e}{\mathrm{m}c}\nabla\Big[\mathbf{A}(t,\mathbf{x})\cdot\Big(\mathbf{p}-\frac{e}{c}\mathbf{A}(t,\mathbf{x})\Big)\Big]\end{cases} \tag{13.27}$$

with the Hamiltonian function (3.30):

$$E=\mathcal{H}(t,\mathbf{x},\mathbf{p})=\frac{1}{2\mathrm{m}}\Big[\mathbf{p}-\frac{e}{c}\mathbf{A}(t,\mathbf{x})\Big]^2+e\phi(t,\mathbf{x}). \tag{13.28}$$

Schrödinger associated the wave equation

$$i\hbar\partial_t\psi(t,\mathbf{x})=\frac{1}{2\mathrm{m}}\Big(-i\hbar\nabla-\frac{e}{c}A(t,\mathbf{x})\Big)^2\psi(t,\mathbf{x})+e\phi(t,\mathbf{x})\,\psi(t,\mathbf{x}) \tag{13.29}$$

to the Hamilton system (13.27). We shall demonstrate that the short-wavelength solutions to (13.29) are governed by the Hamilton equations (13.27). More precisely, we consider the Cauchy problem for (13.29) with the initial condition

$$\psi|_{t=0}=a_0(\mathbf{x})e^{iS_0(\mathbf{x})/\hbar},\quad \mathbf{x}\in\mathbb{R}^3, \tag{13.30}$$

where $S_0(\mathbf{x})$ is a real function. We denote by $(\mathbf{x}(\mathbf{x}_0,t),\mathbf{p}(\mathbf{x}_0,t))$ the solution of the Hamilton equations (13.27) with the initial condition (cf. (13.14)):

$$\mathbf{x}|_{t=0}=\mathbf{x}_0,\qquad \mathbf{p}|_{t=0}=\nabla S_0(\mathbf{x}_0). \tag{13.31}$$

The solution exists for $|t|<T(\mathbf{x}_0)$, where $T(\mathbf{x}_0)>0$.

Definition 13.8

(i) The curve $\mathbf{x}=\mathbf{x}(t,\mathbf{x}_0)$ is the **ray** of the Cauchy problem (13.29), (13.30) starting at the point $\mathbf{x}_0$.

(ii) The **ray tube** or the **ray beam emanating from the initial function** (13.30) is the set

$$\mathcal{T}=\big\{\big(t,\mathbf{x}(t,\mathbf{x}_0)\big)\in\mathbb{R}^4:|t|<T(\mathbf{x}_0),\ \mathbf{x}_0\in\operatorname{supp}a_0\big\}. \tag{13.32}$$

The following lemma means, roughly speaking, that the set $\mathcal{T}$ is the support of the solution of the Cauchy problem (13.29), (13.30) if $\hbar \ll 1$. Namely, we construct the formal WKB–Debye expansion

$$\psi(t,\mathbf{x}) \sim \left(\sum_{k=0}^{\infty} \hbar^k a_k(t,\mathbf{x})\right) e^{iS(t,\mathbf{x})/\hbar}, \quad \hbar \to 0. \tag{13.33}$$

Assume that the potentials $\phi(t,\mathbf{x})$, $A(t,\mathbf{x}) \in C^\infty(\mathbb{R}^4)$ and a_0, $S_0 \in C^\infty(\mathbb{R}^3)$. Then the map $\mathbf{x}_0 \mapsto \mathbf{x}(\mathbf{x}_0, t)$ is a local C^∞-diffeomorphism of $\mathbb{R}^3$ for small $|t|$.

Lemma 13.9 *Suppose that the map* $\mathbf{x}_0 \mapsto \mathbf{x}(t,\mathbf{x}_0)$ *is a diffeomorphism of* $\operatorname{supp} a_0$ *for* $|t| < T$. *Then the formal expansion* (13.33) *holds for* $|t| < T$ *and is identically zero outside* $\mathcal{T}$; *i.e.*,

$$a_k(t,\mathbf{x}) = 0, \quad (t,\mathbf{x}) \notin \mathcal{T},\ |t| < T,\ k = 0, 1, 2, \ldots. \tag{13.34}$$

Proof Substituting the asymptotic expansion (13.33) into (13.29) and equating the leading term with $\hbar^0 = 1$, we obtain the Hamilton–Jacobi equation (3.49),

$$-\dot{S}(t,\mathbf{x}) = H\big(\mathbf{x}, \nabla S(t,\mathbf{x}), t\big), \quad \mathbf{x} \in \mathbb{R}^3,\ |t| < T, \tag{13.35}$$

whose solution is integral (13.15) over the trajectories of the *classical equations* (13.27), according to Hamilton–Jacobi's Theorem 13.7.

Further, substitute $\psi(t,\mathbf{x}) = a(t,\mathbf{x})e^{iS(t,\mathbf{x})/\hbar}$ into the Schrödinger equation (13.29). Then Eq. (13.35) implies the following *transport equation* for the amplitude $a(t,\mathbf{x})$:

$$\begin{cases} \dot{a}(t,\mathbf{x}) = -\dfrac{1}{m}\left[\nabla S(t,\mathbf{x}) - \dfrac{e}{c}A(t,\mathbf{x})\right] \cdot \nabla a(t,\mathbf{x}) \\ \qquad\qquad + \dfrac{1}{2\mathrm{m}}\big[\Delta S(t,\mathbf{x}) + \nabla \cdot A(t,\mathbf{x})\big]a(t,\mathbf{x}) + \dfrac{i\hbar}{2\mathrm{m}}\Delta a(t,\mathbf{x}) \\ \qquad =: -Da(t,\mathbf{x}) + M(t,\mathbf{x})a(t,\mathbf{x}) + \dfrac{i\hbar}{2\mathrm{m}}d(t,\mathbf{x}), \quad |t| < T, \\ a|_{t=0} = a_0(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^3, \end{cases} \tag{13.36}$$

where D is the first order differential operator $Da(t,\mathbf{x}) := \frac{1}{m}[\nabla S(t,\mathbf{x}) - \frac{e}{c}A(t,\mathbf{x})] \cdot \nabla a(t,\mathbf{x})$, and where we set

$$M(t,\mathbf{x}) := \frac{1}{2\mathrm{m}}\big[\Delta S(t,\mathbf{x}) + \nabla \cdot A(t,\mathbf{x})\big], \qquad d(t,\mathbf{x}) := \Delta a(t,\mathbf{x}).$$

The key observation is that

$$\dot{a}(t,\mathbf{x}) + Da(t,\mathbf{x}) = \frac{d}{dt}a\big(t,\mathbf{x}(t,\mathbf{x}_0)\big), \tag{13.37}$$

because $\dot{\mathbf{x}}(t,\mathbf{x}_0) = \frac{1}{m}[\mathbf{p}(t,\mathbf{x}_0) - \frac{e}{c}A(t,\mathbf{x}(t,\mathbf{x}_0))]$ by (13.27), and since

$$\mathbf{p}(t,\mathbf{x}_0) = \nabla S\big(t,\mathbf{x}(t,\mathbf{x}_0)\big), \tag{13.38}$$

by (13.26). Hence (13.36) can be rewritten as

$$\begin{cases} \dfrac{d}{dt}\tilde{a}(t,\mathbf{x}_0) = \tilde{M}(t,\mathbf{x}_0)\tilde{a}(t,\mathbf{x}_0) + \dfrac{i\hbar}{2\mathrm{m}}\tilde{d}(t,\mathbf{x}_0), \quad |t| < T, \\ \tilde{a}(0,\mathbf{x}_0) = a_0(\mathbf{x}_0), \quad \mathbf{x}_0 \in \mathbb{R}^3 \end{cases} \tag{13.39}$$

where all functions are expressed in the *ray coordinates* $(t,\mathbf{x}_0)$:

$$\begin{aligned} &\tilde{a}(t,\mathbf{x}_0) := a\big(t,\mathbf{x}(t,\mathbf{x}_0)\big), \qquad \tilde{M}(t,\mathbf{x}_0) := M\big(t,\mathbf{x}(t,\mathbf{x}_0)\big), \\ &\tilde{d}(t,\mathbf{x}_0) := d\big(t,\mathbf{x}(t,\mathbf{x}_0)\big). \end{aligned}$$

We substitute the formal expansion $\tilde{a}(t,\mathbf{x}_0) \sim \sum_{k=0}^{\infty} \hbar^k \tilde{a}_k(t,\mathbf{x}_0)$ into the first equation (13.39). Equating the terms with identical powers of $\hbar$, we arrive at the recursive *transport equations*

$$\left.\begin{aligned} &\frac{d}{dt}\tilde{a}_0(t,\mathbf{x}_0) = \tilde{M}(t,\mathbf{x}_0)\tilde{a}_0(t,\mathbf{x}_0), \\ &\frac{d}{dt}\tilde{a}_1(t,\mathbf{x}_0) = \tilde{M}(t,\mathbf{x}_0)\tilde{a}_1(t,\mathbf{x}_0) + \frac{i}{2\mathrm{m}}\tilde{d}_0(t,\mathbf{x}_0), \\ &\dots \\ &\frac{d}{dt}\tilde{a}_k(t,\mathbf{x}_0) = \tilde{M}(t,\mathbf{x}_0)\tilde{a}_k(t,\mathbf{x}_0) + \frac{i}{2\mathrm{m}}\tilde{d}_{k-1}(t,\mathbf{x}_0), \\ &\dots \end{aligned}\right| \quad |t| < T, \tag{13.40}$$

where $\tilde{d}_0(t,\mathbf{x}_0)$ is the function $d_0(t,\mathbf{x}) := \Delta a_0(t,\mathbf{x})$ expressed in the ray coordinates, etc. It remains to substitute the same expansion into the initial conditions (13.39), which gives

$$\tilde{a}_0(0,\mathbf{x}_0) = a_0(\mathbf{x}_0), \quad \tilde{a}_1(0,\mathbf{x}_0) = 0, \quad \dots, \quad a_k(0,\mathbf{x}_0) = 0, \dots. \tag{13.41}$$

Now (13.40) and (13.41) imply that $\tilde{a}_0(0,\mathbf{x}) = 0$, $|t| < T$, if $\mathbf{x}_0 \notin \operatorname{supp} a_0$. Hence, we also have $\tilde{d}_0(0,\mathbf{x}) = 0$ for $|t| < T$, if $\mathbf{x}_0 \notin \operatorname{supp} a_0$. Consequently, from (13.40) and (13.41) it follows that $\tilde{a}_1(0,\mathbf{x}) = 0$ for $|t| < T$, if $\mathbf{x}_0 \notin \operatorname{supp} a_0$, etc. ☐

Corollary 13.10 *Let $a_0(\mathbf{x}) = 0$ for $|\mathbf{x} - \mathbf{x}_0| \geq \varepsilon$. Then, for any $N > 0$,*

$$\psi(t,\mathbf{x}) = \mathcal{O}\big(\hbar^N\big) \quad \textit{for } \big|\mathbf{x} - \mathbf{x}(t,\mathbf{x}_0)\big| \geq 2\varepsilon, \ |t| \leq T_\varepsilon \tag{13.42}$$

with sufficiently small $T_\varepsilon > 0$.

Similar *geometrical optics* also holds for short wavelength solutions to relativistic Klein–Gordon equation. The theory was extended to the Dirac equation by Pauli [149]. General hyperbolic systems were considered by Lax [132]. Further development had culminated in the Maslov–Hörmander theory of the Fourier integral operators, as developed by Duistermaat, Helffer, Sjöstrand and others (see [78, 164, 186]). For nonlinear wave equations, the theory was developed by Joly, Métivier, Rauch and others.

13.3 The Noether Symmetry Theory

Time evolution of Lagrangian fields $\psi(x) \in \mathbb{R}^N$ is governed by the Euler–Lagrange equations (4.15)

$$\nabla_\alpha \pi^\alpha(x) = \mathcal{L}_\psi\big(x, \psi(x), \nabla\psi(x)\big), \quad x \in \mathbb{R}^4, \tag{13.43}$$

where π^α are the canonically conjugate fields defined by (4.13).

Assume that the corresponding Lagrangian density is invariant with respect to a Lie symmetry group. Then it is possible to construct the Noether free divergent 4-currents and the corresponding invariants.

13.3.1 Field Symmetry

Consider the one-parametric group of transformations $g_s : \mathbb{R}^4 \times \mathbb{R}^N \to \mathbb{R}^4 \times \mathbb{R}^N$ of the form

$$g_s : \begin{pmatrix} x \\ \psi \end{pmatrix} \mapsto \begin{pmatrix} y_s \\ \psi_s \end{pmatrix} := \begin{pmatrix} a_s(x) \\ b_s(\psi) \end{pmatrix}, \quad s \in \mathbb{R}, \tag{13.44}$$

where a_s and b_s are some diffeomorphisms $a_s : \mathbb{R}^4 \to \mathbb{R}^4$ and $b_s : \mathbb{R}^N \to \mathbb{R}^N$, respectively. We define the corresponding transformations for *the fields*

$$\psi_s(y_s) := b_s\big(\psi(x)\big), \quad \text{where } y_s = a_s(x). \tag{13.45}$$

This definition implies the corresponding transformation for the derivatives: by the chain rule,

$$\nabla\psi_s(y_s) := Db_s\big(\psi(x)\big)\nabla\psi(x)\frac{\partial x(y_s)}{\partial y_s}, \tag{13.46}$$

where Db_s is the differential of the map b_s.

Definition 13.11 The transformation group g_s, $s \in \mathbb{R}$, is a symmetry of the Lagrangian field with Lagrangian density $\mathcal{L}(x, \psi, \nabla\psi)$, provided that the following

algebraic identity holds:

$$\mathcal{L}(x,\psi,\xi)=\mathcal{L}(y_s,\psi_s,\xi_s)\left|\frac{\partial y_s(x)}{\partial x}\right|,\quad x\in\mathbb{R}^4,\ \psi\in\mathbb{R}^N,\ \xi\in\mathbb{R}^{4N},\ s\in\mathbb{R}, \tag{13.47}$$

where $\psi_s := b_s(\psi)$ and $\xi_s := Db_s(\psi)\xi\frac{\partial x(y_s)}{\partial y_s}$, in accordance with (13.45) and (13.46), respectively.

Note that algebraic identity (13.47) is equivalent to the **functional identity**

$$\mathcal{L}\big(x,\psi(x),\nabla\psi(x)\big)=\mathcal{L}\big(y_s,\psi_s(y_s),\nabla\psi_s(y_s)\big)\left|\frac{\partial y_s(x)}{\partial x}\right|,\quad x\in\mathbb{R}^4,\ s\in\mathbb{R} \tag{13.48}$$

for any differentiable field $\psi(x)$.

Example 13.12 (Time-Translations) Consider the time translation along $e_0 = (1,0,0,0)$,

$$g_s:\begin{pmatrix}x\\ \psi\end{pmatrix}\mapsto\begin{pmatrix}y_s\\ \psi_s\end{pmatrix}=\begin{pmatrix}x-se_0\\ \psi\end{pmatrix},\quad s\in\mathbb{R}. \tag{13.49}$$

Then

$$\xi_s=\xi,\qquad \left|\frac{\partial y_s(x)}{\partial x}\right|\equiv 1. \tag{13.50}$$

Now (13.47) is equivalent to the independence of the Lagrangian density $\mathcal{L}(x,\psi,\xi)$ on time $t := x^0$.

Example 13.13 (Space-Translations) Consider the space-translation along $e_1 = (0,1,0,0)$,

$$g_s:\begin{pmatrix}x\\ \psi\end{pmatrix}\mapsto\begin{pmatrix}y_s\\ \psi_s\end{pmatrix}=\begin{pmatrix}x-se_1\\ \psi\end{pmatrix},\quad s\in\mathbb{R}. \tag{13.51}$$

Then

$$\xi_s=\xi,\qquad \left|\frac{\partial y_s(x)}{\partial x}\right|\equiv 1. \tag{13.52}$$

Now (13.47) means that the Lagrangian density $\mathcal{L}(x,\psi,\xi)$ does not depend on x^1.

Example 13.14 (Space-Rotations) Consider the group

$$g_s:\begin{pmatrix}(x^0,\mathbf{x})\\ \psi\end{pmatrix}\mapsto\begin{pmatrix}y_s\\ \psi_s\end{pmatrix}=\begin{pmatrix}(x^0,R_n(s)\mathbf{x})\\ \psi\end{pmatrix},\quad s\in\mathbb{R}, \tag{13.53}$$

where $\psi \in \mathbb{C}$ is a scalar complex field, and $R_n(s)$ is rotation of $\mathbb{R}^3$ around $\mathbf{e}_n$ by an angle of s *radian*. Then

$$\xi_s = \begin{pmatrix} 1 & 0 \\ 0 & R_n(s) \end{pmatrix} \xi, \qquad \left| \frac{\partial y_s(x)}{\partial x} \right| \equiv 1. \tag{13.54}$$

Now (13.47) is equivalent to the identity

$$\mathcal{L}(x, \psi, \xi) \equiv \mathcal{L}(y_s, \psi, \xi_s), \quad s \in \mathbb{R},\ x \in \mathbb{R}^4,\ \psi \in \mathbb{C},\ \xi \in \mathbb{C}^4. \tag{13.55}$$

Example 13.15 (Phase Rotations) For $\psi \in \mathbb{C}^N = \mathbb{R}^{2N}$, we define

$$g_s \begin{pmatrix} x \\ \psi \end{pmatrix} \mapsto \begin{pmatrix} y \\ \psi_s \end{pmatrix} = \begin{pmatrix} x \\ e^{is}\psi \end{pmatrix}, \quad s \in \mathbb{R}. \tag{13.56}$$

Then

$$\xi_s = e^{is}\xi, \qquad \left| \frac{\partial y_s(x)}{\partial x} \right| \equiv 1. \tag{13.57}$$

Now (13.47) is equivalent to the identity

$$\mathcal{L}\big(x, e^{is}\psi, e^{is}\xi\big) \equiv \mathcal{L}(x, \psi, \xi), \quad s \in \mathbb{R},\ x \in \mathbb{R}^4,\ \psi \in \mathbb{C}^N,\ \xi \in \mathbb{C}^{4N}. \tag{13.58}$$

13.3.2 The Noether Current and the Continuity Equation

Definition 13.16 For a given one-parametric group g_s of transformations (13.44) and a given field $\psi(x)$, we define the vector fields

$$v(x) = \frac{d}{ds}\bigg|_{s=0} a_s(x), \qquad w(x) = \frac{d}{ds}\bigg|_{s=0} \psi_s(x), \quad x \in \mathbb{R}^4, \tag{13.59}$$

where $\psi_s(x) = b_s(\psi(a_s^{-1}(x)))$, according to definition (13.45).

Definition 13.17 The Noether 4-current corresponding to the field $\psi(x)$ and the symmetry group g_s is the vector field

$$S^\alpha(x) = \pi^\alpha(x) \cdot w(x) + \mathcal{L}\big(x, \psi(x), \nabla\psi(x)\big) v^\alpha(x), \quad x \in \mathbb{R}^4,\ \alpha = 0, \ldots, 3, \tag{13.60}$$

where $\pi^\alpha(x) \cdot w(x) := \sum_1^N \pi_j^\alpha(x) \cdot w^j(x)$.

Theorem 13.18 (E. Noether [146, 194]) *Let g_s be a one-parametric symmetry group, and let $\psi(x) \in C^2(\mathbb{R}^4, \mathbb{R}^N)$ be a solution of Eqs.* (13.43). *Then the following continuity equation holds,*

$$\nabla_\alpha S^\alpha(x) = 0, \quad x \in \mathbb{R}^4. \tag{13.61}$$

Corollary 13.19 *Under the hypotheses of Theorem 13.18 assume that* $S^\alpha(t,\mathbf{x})|\mathbf{x}|^2 \to 0$ *as* $|\mathbf{x}| \to \infty$ *for* $\alpha = 1,2,3$. *Then the following conservation law holds,*

$$\mathcal{S}^0(t) := \int_{\mathbb{R}^3} S^0(t,\mathbf{x})\, d\mathbf{x} = \text{const}, \qquad t := x^0 \in \mathbb{R}. \tag{13.62}$$

Proof We have $\mathcal{S}^0(t) = \lim_{R\to\infty} \mathcal{S}^0_R(t)$, where

$$\mathcal{S}^0_R(t) := \int_{|\mathbf{x}|\le R} S^0(t,\mathbf{x})\, d\mathbf{x}, \qquad t \in \mathbb{R}. \tag{13.63}$$

Differentiating and using (13.61) and Stokes's theorem, it follows that

$$\dot{\mathcal{S}}^0_R(t) := -\int_{|\mathbf{x}|\le R} \nabla_\alpha S^\alpha(t,\mathbf{x})\, d\mathbf{x} = -\int_{|\mathbf{x}|=R} n_\alpha(\mathbf{x}) S^\alpha(t,\mathbf{x})\, d\Sigma, \qquad t \in \mathbb{R}, \tag{13.64}$$

where $n^\alpha(\mathbf{x}) := \mathbf{x}^\alpha/|\mathbf{x}|$ and $d\Sigma$ is the Lebesgue measure on the sphere $|\mathbf{x}| = R$. Therefore, $\dot{\mathcal{S}}^0_R(t) \to 0$ as $R \to \infty$. Hence (13.62) follows. □

The integral identity (13.64) means that the vector field $S^\alpha(t,\mathbf{x})$ is the current density of the field $S^0(t,\mathbf{x})$.

Proof of Theorem 13.18 (cf. [70, 163, 194]) Differentiating (13.48) in s, this gives

$$\begin{aligned} 0 = {} & \mathcal{L}_x\big(x, \psi(x), \nabla\psi(x)\big) \cdot \frac{d}{ds}\bigg|_{s=0} y_s + \mathcal{L}_\psi\big(x, \psi(x), \nabla\psi(x)\big) \cdot \frac{d}{ds}\bigg|_{s=0} \psi_s(y_s) \\ & + \mathcal{L}_{\nabla_\alpha\psi}\big(x, \psi(x), \nabla\psi(x)\big) \cdot \frac{d}{ds}\bigg|_{s=0} \nabla_\alpha\psi_s(y_s) \\ & + \mathcal{L}\big(x, \psi(x), \nabla\psi(x)\big) \frac{d}{ds}\bigg|_{s=0} \left|\frac{\partial y_s(x)}{\partial x}\right|. \end{aligned} \tag{13.65}$$

We calculate these four derivatives in s.

I. By definition (13.59),

$$\frac{d}{ds}\bigg|_{s=0} y_s = v(x). \tag{13.66}$$

II. From definition (13.45), the chain rule, and (13.59), it follows that

$$\begin{aligned} \frac{d}{ds}\bigg|_{s=0} \psi_s(y_s) &= \frac{d}{ds}\bigg|_{s=0} \big[\psi_s(y_0) + \psi_0(y_s)\big] \\ &= \frac{d}{ds}\bigg|_{s=0} \big[\psi_s(x) + \psi(y_s)\big] = w(x) + \nabla\psi(x)\cdot v(x). \end{aligned} \tag{13.67}$$

III. The same arguments imply that

$$\begin{aligned}\frac{d}{ds}\Big|_{s=0}\big[\nabla_\alpha\psi_s(y_s)\big] &= \frac{d}{ds}\Big|_{s=0}\frac{\partial\psi_s(y_s)}{\partial y_s^\alpha}\\ &= \frac{d}{ds}\Big|_{s=0}\left[\frac{\partial\psi_s(y_0)}{\partial y_0^\alpha}+\frac{\partial\psi_0(y_s)}{\partial y_0^\alpha}+\frac{\partial\psi_0(y_0)}{\partial y_s^\alpha}\right]\\ &= \nabla_\alpha w(x)+\nabla_\alpha\big(\nabla\psi(x)\cdot v(x)\big)+\frac{d}{ds}\Big|_{s=0}\frac{\partial\psi}{\partial y_s^\alpha}. \end{aligned}\tag{13.68}$$

To calculate the last derivative, we employ the matrix identity

$$\frac{\partial\psi}{\partial y_s^\beta}\frac{\partial y_s^\beta}{\partial x^\alpha}=\frac{\partial\psi}{\partial x^\alpha}.\tag{13.69}$$

Differentiating in s, it follows that

$$\frac{d}{ds}\Big|_{s=0}\frac{\partial\psi}{\partial y_s^\alpha}+\frac{\partial\psi}{\partial x^\beta}\frac{\partial v_\beta}{\partial x^\alpha}=0.\tag{13.70}$$

Therefore, the last derivative in (13.68) equals $-\nabla\psi\cdot\nabla_\alpha v$. Hence, (13.68) becomes,

$$\frac{d}{ds}\Big|_{s=0}\nabla_\alpha\psi_s(y_s)=\nabla_\alpha w(x)+\nabla\big(\nabla_\alpha\psi(x)\big)\cdot v(x).\tag{13.71}$$

IV. Finally, the derivative of the determinant $|\frac{\partial y_s(x)}{\partial x}|$ is the trace of the matrix derivative,

$$\frac{d}{ds}\Big|_{s=0}\left|\frac{\partial y_s(x)}{\partial x}\right|=\operatorname{tr}\frac{d}{ds}\Big|_{s=0}\frac{\partial a_s(x)}{\partial x}=\operatorname{tr}\frac{\partial v(x)}{\partial x}=\nabla\cdot v(x).\tag{13.72}$$

Collecting all calculations **I–IV** in (13.65), we obtain

$$\begin{aligned}&\mathcal{L}_x\big(x,\psi(x),\nabla\psi(x)\big)\cdot v(x)+\mathcal{L}_\psi\big(x,\psi(x),\nabla\psi(x)\big)\cdot\big(w(x)+\nabla\psi(x)\cdot v(x)\big)\\ &\quad+\pi^\alpha(x)\cdot\big(\nabla_\alpha w(x)+\nabla\big(\nabla_\alpha\psi(x)\big)\cdot v(x)\big)\\ &\quad+\mathcal{L}\big(x,\psi(x),\nabla\psi(x)\big)\nabla\cdot v(x)=0.\end{aligned}\tag{13.73}$$

We rewrite it as follows:

$$\mathcal{L}_\psi\cdot w(x)+\pi^\alpha(x)\cdot\nabla_\alpha w(x)+\nabla\cdot\big[\mathcal{L}\big(x,\psi(x),\nabla\psi(x)\big)v(x)\big]=0.\tag{13.74}$$

Finally, we substitute $\mathcal{L}_\psi=\nabla^\alpha\pi_\alpha(x)$ from the Euler–Lagrange equations (13.43). Then (13.74) becomes

$$\nabla_\alpha\big[\pi^\alpha(x)\cdot w(x)\big]+\nabla_\alpha\big[\mathcal{L}\big(x,\psi(x),\nabla\psi(x)\big)v^\alpha(x)\big]=0,\tag{13.75}$$

which coincides with (13.61) by (13.60). □

13.4 Application of Noether's Theorem

Noether's theorem implies, in particular, the four conservation laws: conservation of total energy, momentum, angular momentum and charge. Namely, the energy conservation follows from the invariance under time translations, the momentum conservation follows from the translation invariance in space, the angular momentum conservation follows from the rotational invariance, and the charge conservation is a consequence of the phase invariance.

Below we assume an appropriate decay of the wave function $\psi(t, \mathbf{x})$ and of its derivatives of the first and second order for $|\mathbf{x}| \to \infty$, providing the convergence of all spatial integrals defining the conservation laws.

13.4.1 The Energy, Momentum, and Angular Momentum Conservation

We denote $t = x^0$, and so $x = (x^0, x^1, x^2, x^3) = (t, \mathbf{x})$.

Definition 13.20

(i) The energy of a Lagrangian field at time t is defined by

$$E(t) = \int_{\mathbb{R}^3} \big[\pi^0(x) \cdot \nabla_0 \psi(x) - \mathcal{L}\big(x, \psi(x), \nabla\psi(x)\big)\big]\, d\mathbf{x}, \qquad (13.76)$$

where $\pi^0(x) \cdot \nabla_0 \psi(x) := \pi^0_j(x)\, \nabla_0 \psi^j(x)$.

(ii) The momentum of a Lagrangian field at time t is the vector

$$\mathbf{p}(t) = -\int_{\mathbb{R}^3} \pi^0(x) \cdot \nabla\psi(x)\, d\mathbf{x}. \qquad (13.77)$$

(iii) The angular momentum of a scalar complex Lagrangian field at time t is the vector

$$\mathbf{J}(t) = \int_{\mathbb{R}^d} \pi^0(x) \cdot \big[\mathbf{x} \times \boldsymbol{\nabla}\psi(x)\big]\, d\mathbf{x}. \qquad (13.78)$$

Exercise 13.21 Verify that for Lagrangian density (4.7) of the Schrödinger equation, definitions (13.76), (13.77) and (13.78) coincide with (3.72), (3.57) and (3.58) up to a factor. **Hints:**

(i) Express (13.76)–(13.78) in Ψ and Π^0 instead of ψ and π^0 (see (4.14) and Sect. 3.4).

(ii) $\Pi^0 = -J\hbar\Psi$ by (4.14).

Theorem 13.22

(i) *Let Lagrangian density* $\mathcal{L}(x,\psi,\nabla\psi)$ *be independent of time* $t=x^0$. *Then, for any solution* $\psi(x)$ *of* (13.43), *the energy is conserved,*

$$E(t)=\text{const},\qquad t\in\mathbb{R}. \tag{13.79}$$

(ii) *Let Lagrangian density* $\mathcal{L}(x,\psi,\nabla\psi)$ *be independent of* $\mathbf{x}^n$. *Then, for any solution* $\psi(x)$ *of* (13.43), *the corresponding component of the momentum is conserved,*

$$\mathbf{p}_n(t)=\text{const},\qquad t\in\mathbb{R}. \tag{13.80}$$

Proof The theorem follows from Corollary 13.19 applied to the conserved current (13.60) corresponding to the symmetry groups of translations in time and space:

(i) The Lagrangian density satisfies identity (13.47) for the symmetry group (13.49). Definitions (13.59) imply

$$v(x)=(-1,0,0,0),\qquad w(x)=\nabla_0\psi(x),\quad x\in\mathbb{R}^4, \tag{13.81}$$

because $\psi_s(x)=\psi(x_0+s,\mathbf{x})$. Hence, the Noether 4-current (13.60) reads

$$\begin{aligned}S_0(x)=\big(&\pi^0(x)\nabla_0\psi(x)-\mathcal{L}\big(x,\psi(x),\nabla\psi(x)\big),\pi^1(x)\nabla_0\psi(x),\\&\pi^2(x)\nabla_0\psi(x),\pi^3(x)\nabla_0\psi(x)\big).\end{aligned} \tag{13.82}$$

Now (13.79) follows from Corollary 13.19.

(ii) The Lagrangian density satisfies identity (13.47) for the symmetry group (13.51). Definitions (13.59) imply

$$v(x)=(0,-1,0,0),\qquad w(x)=\nabla_1\psi(x),\quad x\in\mathbb{R}^4, \tag{13.83}$$

because $\psi_s(x)=\psi(x_0,x_1+s,x_2,x_3)$. Hence, the Noether 4-current (13.60) reads

$$\begin{aligned}S_1(x)=\big(&\pi^0(x)\nabla_1\psi(x),\pi^1(x)\nabla_1\psi(x)-\mathcal{L}\big(x,\psi(x),\nabla\psi(x)\big),\\&\pi^2(x)\nabla_1\psi(x),\pi^3(x)\nabla_1\psi(x)\big).\end{aligned} \tag{13.84}$$

Now (13.80) follows from Corollary 13.19. □

Now consider Lagrangian densities for a complex scalar field $\psi(x)$, which are invariant with respect to rotations (13.53), (13.54).

Theorem 13.23 *Let the Lagrangian density of a complex scalar field satisfy* (13.55). *Then, for any solution of* (13.43), *the corresponding component of the angular momentum is conserved,*

$$\mathbf{J}_n(t)=\text{const},\qquad t\in\mathbb{R}. \tag{13.85}$$

This theorem follows from Corollary 13.19 applied to the conserved current (13.59) corresponding to the symmetry group of space rotations. □

Exercise 13.24 Prove Theorem 13.23. **Hints:** (13.53) means that $a_s x = (x_0, R_n(s)\mathbf{x})$ and $b_s\psi = \psi$. Hence, definitions (13.59) give

$$v(x) = (0, \mathbf{e}_n \times \mathbf{x}), \qquad w(x) = \frac{d}{ds}\bigg|_{s=0} \psi\big(R_n(-s)x\big). \tag{13.86}$$

Differentiating, we obtain

$$w(x) = \nabla\psi \cdot [\mathbf{x} \times \mathbf{e}_n] = \mathbf{e}_n \cdot [\nabla\psi \times \mathbf{x}] = [\nabla\psi \times \mathbf{x}]_n. \tag{13.87}$$

Now definition (13.60) gives

$$S^0(x) = \pi^0(x) \cdot [\nabla\psi \times \mathbf{x}]_n. \tag{13.88}$$

Therefore, Corollary 13.19 implies (13.85).

Theorems 13.22 and 13.23 imply the conservation laws (3.61), (3.64) and (3.67) for the case of the Schrödinger equation (see Exercise 13.21).

13.4.2 The Energy-Momentum Tensor

Suppose that the Lagrangian density $\mathcal{L}(\psi, \nabla\psi)$ is independent of x. Then the continuity equation (13.61) holds for currents (13.82) and (13.84), as well as for the 4-currents S_2^α and S_3^α, which are defined similarly to S_1^α from (13.84). The four continuity equations read

$$\partial_\alpha S_\beta^\alpha(x) = 0, \quad \beta = 0, 1, 2, 3, \tag{13.89}$$

where $S_\beta^\alpha(x) := \pi^\alpha(x)\nabla_\beta\psi(x) - \delta_\beta^\alpha\mathcal{L}(x, \psi(x), \nabla\psi(x))$. We recall that $g_{\alpha\beta} = \text{diag}(1, -1, -1, -1)$.

Definition 13.25 Energy-momentum tensor is defined by

$$T_{\alpha\beta}(x) := g_{\alpha\alpha'}S_\beta^{\alpha'}(x) = \pi_\alpha(x)\nabla_\beta\psi(x) - g_{\alpha\beta}\mathcal{L}\big(x, \psi(x), \nabla\psi(x)\big).$$

Note that $T_{00}(x) = S_0^0(x)$ is the energy density and $T_{0k}(x) = -S_k^0(x)$ is the momentum density, according to definitions (13.76) and (13.77), respectively. Now identities (13.89) read

$$\partial^\alpha T_{\alpha\beta}(x) = 0, \quad \beta = 0, 1, 2, 3, \tag{13.90}$$

where $\partial^\alpha := g^{\alpha\alpha'}\partial_{\alpha'} = (\partial_0, -\partial_1, -\partial_2, -\partial_3)$, because $(g^{\alpha\alpha'}) := (g_{\alpha\alpha'})^{-1} = (g_{\alpha\alpha'})$. In the three dimensional form,

$$\dot{T}_{0\beta}(t,\mathbf{x}) = -\operatorname{div}\mathbf{T}_\beta(t,\mathbf{x}), \quad \beta = 0,1,2,3, \tag{13.91}$$

where $\mathbf{T}_\beta(t,\mathbf{x}) := (T_{1\beta}(t,\mathbf{x}), T_{2\beta}(t,\mathbf{x}), T_{3\beta}(t,\mathbf{x}))$ and now $\mathbf{x} := (x_1, x_2, x_3) = -(x^1, x^2, x^3)$.

Let us consider the field energy and momentum located in a bounded open region $\Omega \subset \mathbb{R}^3$ at time t: by definition,

$$E(\Omega,t) := \int_\Omega T_{00}(t,\mathbf{x})\,d\mathbf{x}, \qquad \mathbf{p}_k(\Omega,t) := \int_\Omega T_{0k}(t,\mathbf{x})\,d\mathbf{x}, \quad k = 1,2,3. \tag{13.92}$$

Suppose that the boundary $\partial\Omega$ is smooth. Hence from identities (13.91) it follows by Stokes's theorem that

$$\begin{aligned} \dot{E}(\Omega,t) &:= -\int_{\partial\Omega} \mathbf{T}_0(t,\mathbf{x})\cdot\mathbf{n}(\mathbf{x})\,d\mathbf{x}, \\ \dot{\mathbf{p}}_k(\Omega,t) &:= -\int_{\partial\Omega} \mathbf{T}_k(t,\mathbf{x})\cdot\mathbf{n}(\mathbf{x})\,d\mathbf{x}, \quad k = 1,2,3, \end{aligned} \tag{13.93}$$

where $\mathbf{n}(\mathbf{x})$ is the unit outward normal vector to the boundary $\partial\Omega$ at the point $\mathbf{x}$. The identities mean that $\mathbf{T}_0(t,\mathbf{x})$ is the energy flux, while $\mathbf{T}_k(t,\mathbf{x})$ is the flux of the kth component of the momentum.

13.4.3 Phase Invariance and the Charge Continuity Equation

Consider a Lagrangian field $\psi \in \mathbb{C}^N = \mathbb{R}^{2N}$ with Lagrangian density $\mathcal{L}(x,\psi,\nabla\psi)$. For phase rotations (13.56) formulas (13.59) give

$$v(x) = 0, \qquad w(x) = i\psi(x). \tag{13.94}$$

Respectively,

$$S^\alpha(x) = \pi^\alpha(x)\cdot\big[i\psi(x)\big], \quad \alpha = 0,\dots,3, \tag{13.95}$$

where $\pi^\alpha(x)$ and $\psi(x)$ are identified with the corresponding real vector from $\mathbb{R}^{2N}$, the dot '$\cdot$' denotes the inner product on $\mathbb{R}^{2N}$, and the multiplication by i is identified with the application of the corresponding skew-symmetric matrix. Now Noether's theorem implies the following corollary.

Corollary 13.26 *Let a Lagrangian density $\mathcal{L}(x,\psi,\nabla\psi)$ be invariant with respect to the phase rotations* (13.58). *Then the continuity equation* (13.61) *holds for the* 4-*current* (13.95).

In particular, for $N = 1$ formula (13.95) is equivalent to

$$S^\alpha(x) = \mathrm{Im}\big[\pi^\alpha(x)\overline{\psi}(x)\big] = \frac{1}{2}\big[\pi^\alpha(x)\overline{\psi}(x) - \psi(x)\overline{\pi}^\alpha(x)\big]. \tag{13.96}$$

Application to the Schrödinger and Maxwell–Schrödinger Equations

Let us apply formula (13.96) to Lagrangian densities of the Schrödinger equation (4.7) and Maxwell–Schrödinger equations (4.21). Both densities are invariant with respect to the phase rotations (13.58). The canonically conjugate momenta $\pi^\alpha(x)$ for both densities are identical and given by (4.14). Substituting this into (13.96), we obtain the electric charge and current (3.34) up to the factor $e/\hbar$:

$$S^0(x) = \frac{\hbar}{e}\,\rho(t,\mathbf{x}), \qquad S^k(x) = \frac{\hbar}{e}\mathbf{j}^k(t,\mathbf{x}), \quad k = 1, 2, 3. \tag{13.97}$$

Therefore, the Noether theorem (13.61) implies the charge continuity equation (3.78) for electric 4-current $(\rho, \mathbf{j})$.

13.4.4 Gauge Invariance

For a *gauge invariant* Lagrangian density $\mathcal{L}$, the current (13.95) can be represented in terms of variational derivatives in the Maxwell potentials. Namely, consider a field $\psi \in \mathbb{C}^N$ interacting with the Maxwell vector potential $\mathcal{A}(x) = (\mathcal{A}_0(x), \mathcal{A}_1(x), \mathcal{A}_2(x), \mathcal{A}_3(x)) = (\phi_0, -\mathbf{A})$ (see (12.49)). This means that the corresponding Lagrangian density $\mathcal{L}$ depends also on $\mathcal{A}(x)$ and $\nabla\mathcal{A}(x)$. By definition, the Lagrangian density is *gauge invariant* if it depends on $\nabla\psi$ and $\mathcal{A}(x)$ only through combination $i\hbar\nabla\psi - \frac{e}{c}\mathcal{A}(x)\psi$; i.e.,

$$\mathcal{L}\big(x, \psi, \nabla\psi, \mathcal{A}(x), \nabla\mathcal{A}(x)\big) = \tilde{\mathcal{L}}\Big(x, \psi, i\hbar\nabla\psi - \frac{e}{c}\mathcal{A}(x)\psi, \nabla\mathcal{A}(x)\Big). \tag{13.98}$$

For example, the Lagrangian densities (4.7), (4.34), (11.57) and (11.59) are gauge invariant.

We next extend the definition of charge and current densities (12.59).

Definition 13.27 The current $\mathcal{J}^\alpha(x)$ is defined by

$$\mathcal{J}^\alpha(x) = -\mathcal{L}_{\mathcal{A}_\alpha}\big(x, \psi, \nabla\psi, \mathcal{A}_\mu(x), \nabla\mathcal{A}_\mu(x)\big), \quad \alpha = 0, 1, 2, 3. \tag{13.99}$$

Lemma 13.28 *Suppose that a Lagrangian density $\mathcal{L}$ is gauge invariant. Then the Noether current* (13.95) *coincides with* (13.99) *up to a factor*:

$$\mathcal{J}^\alpha(x) = \frac{e}{\hbar c}\pi^\alpha(x)\cdot\big[i\psi(x)\big]. \tag{13.100}$$

Proof By definition (4.13),

$$\pi^\alpha(x) := \nabla_{\nabla_\alpha \psi}\Big[\tilde{\mathcal{L}}\Big(x, \psi(x), i\hbar\nabla\psi(x) - \frac{e}{c}\mathcal{A}(x)\psi(x), \nabla\mathcal{A}(x)\Big)\Big], \qquad (13.101)$$

where $\nabla_\alpha \psi$ is identified with the corresponding vector in $\mathbb{R}^{2N}$. Secondly,

$$\mathcal{J}^\alpha(x) := -\nabla_{\mathcal{A}_\alpha(x)}\Big[\tilde{\mathcal{L}}\Big(x, \psi(x), i\hbar\nabla\psi(x) - \frac{e}{c}\mathcal{A}(x)\psi(x), \nabla\mathcal{A}(x)\Big)\Big], \qquad (13.102)$$

by definition (13.99). Hence, (13.100) follows. □

Example 13.29 For the coupled Maxwell–Schrödinger and Maxwell–Dirac equations (4.19), (4.22) and (11.58), (11.60), it follows from formulas (13.99) and (13.97) in variables $(\phi, \mathbf{A})$ that

$$\mathcal{J}^0 = \rho(x) = -\mathcal{L}_\phi, \qquad \mathcal{J}^k = \frac{\mathbf{j}_k(x)}{c} = \mathcal{L}_{\mathbf{A}^k}, \quad k = 1, 2, 3, \qquad (13.103)$$

similarly to (12.60).

13.5 The Limiting Amplitude Principle

For a finite dimensional Hamiltonian system forced by an external harmonic source, in nonresonant case the motion reduces to a quasiperiodic regime, which is a mixing of eigenmodes and particular periodic solution. In the resonant case, the forced oscillations runaway.

A similar situation holds in the long time limit for linear wave fields that are infinite dimensional Hamiltonian systems. However, in this case, the quasiperiodic limit behavior holds even in the resonant case, provided the frequency belongs to the *continuous spectrum*, which newer occurs for finite dimensional systems. This peculiarity of wave fields is caused by the *radiation of energy to infinity in continuous spectrum*. This radiation prevents the runaway behavior of the field in any bounded region though the total energy growth linearly in time.

This *limiting amplitude principle* plays a crucial role in the theory of long time behavior of wave processes. In particular, it explains wave propagation along wave guides and justifies the quantum Kramers–Kronig dispersion theory (Sect. 8.3), and Einstein's rules for photoelectric effect (Sect. 8.4).

For the first time, this principle was justified mathematically in 1960s for the wave and Klein–Gordon equations by Lax, Morawetz, Phillips, Vainberg and others. For the Schrödinger equation, this principle was justified by Eidus [55, 56] relying on the Agmon–Jensen–Kato spectral theory [3, 97].

13.5.1 Harmonic Source and Spectrum

We consider the Schrödinger equation with harmonic source (cf. (8.56) and (8.57))

$$i\hbar\dot{\psi}(t,\mathbf{x}) = H\psi(t,\mathbf{x}) + b(\mathbf{x})e^{-i\omega t}, \tag{13.104}$$

where $\omega \in \mathbb{R}$ and the amplitude $b(\mathbf{x})$ decays rapidly as $|\mathbf{x}| \to \infty$. Here H is the Schrödinger operator corresponding to a *static* Maxwell field with potentials $\phi(t,\mathbf{x}) \equiv \phi(\mathbf{x})$ and $\mathbf{A}(t,\mathbf{x}) \equiv \mathbf{A}(\mathbf{x})$ (see (3.37)),

$$H := \frac{1}{2m}\left[-i\hbar\nabla - \frac{e}{c}\mathbf{A}(\mathbf{x})\right]^2 + e\phi(\mathbf{x}). \tag{13.105}$$

The corresponding 'unperturbed equation' reads

$$i\dot{\psi}(t,\mathbf{x}) = H\psi(t,\mathbf{x}). \tag{13.106}$$

We consider *finite energy solutions* of (13.104), (13.106). In particular, $\psi(t,\cdot) \in C(\mathbb{R}, E)$, by the charge conservation, where $E := L^2(\mathbb{R}^3)$. Let us denote by $U(t)$ the corresponding dynamical group:

$$\psi(t) = U(t)\psi(0), \qquad \psi(t) := \psi(t,\cdot). \tag{13.107}$$

The charge conservation means that $U(t)$ is a unitary operator on E.

We shall look for a solution of the type $a_\omega(\mathbf{x})e^{-i\omega t}$. Substituting into (13.104), we get the *Helmholtz stationary equation*

$$(H-\omega)a_\omega(\mathbf{x}) = -b(\mathbf{x}). \tag{13.108}$$

The *spectrum*, Spec$H$, is the set of all $\omega \in \mathbb{C}$ at which the operator $H-\omega$ is not invertible in E. The *resolvent* of H is defined by

$$R(\omega) := (H-\omega)^{-1}, \quad \omega \in \mathbb{C}\setminus \operatorname{Spec} H. \tag{13.109}$$

Equation (13.108) admits a unique solution $a_\omega(\mathbf{x}) \in E$ for every $b \in E$ if $\omega \notin \operatorname{Spec} H$:

$$a_\omega = -R(\omega)b, \quad \omega \in \mathbb{C}\setminus \operatorname{Spec} H. \tag{13.110}$$

For example, let us calculate the spectrum and the resolvent for the free Schrödinger operator $H_0 = -\Delta$. In the Fourier transform operator, H_0 becomes the multiplication by $\hat{H}_0(\mathbf{k}) = \mathbf{k}^2$ on the space $\hat{E} = L^2(\mathbb{R}^3) = E$, by Plancherel's theorem. Therefore, $R_0(\omega)$, in the Fourier transform, is the multiplication by $(\mathbf{k}^2-\omega)^{-1}$. This operator is bounded in $\hat{E}$ if and only if the function $(\mathbf{k}^2-\omega)^{-1}$ is bounded in $\mathbb{R}^3$. Hence, Spec$H_0 = \overline{\mathbb{R}}_+ = \{\omega \in \mathbb{R} : \omega \geq 0\}$. In the coordinate representation, $R_0(\omega)$ is the convolution with tempered fundamental solution $\mathcal{E}_\omega(\mathbf{x})$ of the operator $-\Delta - \omega$. The tempered fundamental solution is unique,

$$\mathcal{E}_\omega(\mathbf{x}) = \mathcal{E}^+_\omega(\mathbf{x}) := \frac{e^{i\sqrt{\omega}|\mathbf{x}|}}{4\pi|\mathbf{x}|}, \quad \omega \in \mathbb{C}\setminus\overline{\mathbb{R}}_+, \tag{13.111}$$

where we choose $\mathrm{Im}\sqrt{\omega} > 0$ for all $\omega \in \mathbb{C} \setminus \overline{\mathbb{R}}_+$. Here $\mathcal{E}^+_\omega(\mathbf{x})$ is a tempered distribution, because its Fourier transform is a bounded function. Hence, the tempered solution of Eq. (13.108) is unique for $\omega \in \mathbb{C} \setminus \overline{\mathbb{R}}_+$, and is given by

$$a_\omega = -\mathcal{E}^+_\omega * b. \tag{13.112}$$

The distribution $\mathcal{E}^-_\omega(\mathbf{x}) := \frac{e^{-i\sqrt{\omega}|\mathbf{x}|}}{4\pi|\mathbf{x}|}$ is also a fundamental solution; however, for every $\omega \in \mathbb{C} \setminus \overline{\mathbb{R}}_+$, it is not tempered, inasmuch as it grows exponentially as $|\mathbf{x}| \to \infty$.

In the case $\omega > 0$, both fundamental solutions $\mathcal{E}^\pm_\omega(\mathbf{x})$ are tempered; hence, the tempered solution of (13.108) is not unique. Moreover, in this case, the space of tempered solutions is infinite dimensional, because any distribution $a(\mathbf{k}/|\mathbf{k}|)\delta(|\mathbf{k}| - \sqrt{\omega})$ is a solution of the homogeneous equation

$$(\mathbf{k}^2 - \omega)a(\mathbf{k}/|\mathbf{k}|)\delta(|\mathbf{k}| - \sqrt{\omega}) = 0.$$

A similar nonuniqueness also holds for the perturbed Schrödinger operator (13.105); this is the well-known problem in *diffraction theory*. The uniqueness will be provided using either the *Sommerfeld radiation conditions* or the *limiting absorption principle* (see below).

13.5.2 The Limiting Absorption Principle

The *limiting amplitude principle* yields the following long-time asymptotics (see [55, 56]):

$$\psi(t,\mathbf{x}) = a_\omega(\mathbf{x})e^{-i\omega t} + \sum_l C_l\psi_l(\mathbf{x})e^{-i\omega_l t} + r(t,\mathbf{x}), \tag{13.113}$$

where $a_\omega(\mathbf{x})$ is a *limiting amplitude*, $\psi_l(\mathbf{x})$ are the eigenfunctions of the discrete spectrum of the operator H, and $\|r(t,\cdot)\| \to 0$ as $t \to \infty$ in an appropriate norm.

The asymptotics (13.113) hold for the initial states $\psi(\mathbf{x},0)$ from the *Agmon–Sobolev weighted spaces*:

$$\mathcal{H}^2_\sigma = \left\{\psi(\mathbf{x}) : \|\psi\|_{2,\sigma} = \left\|\left(1+|\mathbf{x}|\right)^\sigma(1-\Delta)\psi\right\|_{L^2(\mathbb{R}^3)} < \infty\right\}. \tag{13.114}$$

The asymptotics (13.113) are proved for a restricted class of potentials: for example, if

$$\phi(\mathbf{x}) = \mathcal{O}\left(|\mathbf{x}|^{-5-\varepsilon}\right), \qquad \mathbf{A}(\mathbf{x}) \equiv 0, \tag{13.115}$$

where $\varepsilon > 0$.

The limiting amplitude $a_\omega(\mathbf{x})$ obviously satisfies the stationary Helmholtz equation (13.108). Generally, $a_\omega(\mathbf{x}) \notin E$ for $\omega \geq 0$, but anyhow it is a tempered distribution. Now the problem of nonuniqueness for $\omega > 0$ becomes significant: we should decide which of the solutions is 'true'. The answer is given by the *limiting absorption principle*.

Lemma 13.30 *Let* $\omega \in \mathbb{R}$, $\omega \neq 0$ *and* $\omega \neq \omega_l$ *for all* l. *Let moreover, the potential satisfy* (13.115), *and* $\psi(\mathbf{x}, 0), b(\mathbf{x}) \in H^2_\sigma$ *with* $\sigma > 5/2$. *Then*

(i) *The asymptotics* (13.113) *hold with*

$$\left\| r(t, \cdot) \right\|_{H^2_{-\sigma}} \to 0, \quad t \to \infty. \tag{13.116}$$

(ii) *The limiting amplitude is given by* $\psi_\omega = -R(\omega)b$, *if* $\omega < 0$, *and*

$$a_\omega = -\lim_{\varepsilon \to 0+} R(\omega + i\varepsilon)b, \quad \text{if } \omega > 0, \tag{13.117}$$

where the limit holds in the space $H^2_{-\sigma}$ *with* $\sigma > 1/2$.

(iii) *The coefficients* C_l *depend on* $\psi(0, x)$ *but do not depend on* $b(x)$. *More precisely,*

$$\sum_l C_l \psi_l = P_d \psi(0, \cdot), \tag{13.118}$$

where P_d *is the orthogonal projection onto the discrete spectral space of* H.

In particular, for the free Schrödinger operator $H_0 = -\Delta$,

$$a_\omega = \lim_{\varepsilon \to 0+} \mathcal{E}^+_{\omega + i\varepsilon} * b \quad \text{if } \omega > 0. \tag{13.119}$$

Formula (13.117) can be proved by an appropriate development of methods [97] (see also [56, 116]). Let us sketch a *formal* proof using the results of [97]. Namely, the Duhamel representation for solutions of the inhomogeneous equation (13.104) reads

$$\psi(t) = U(t)\psi(0) - i \int_0^t U(t - s) b e^{-i\omega s}\, ds, \tag{13.120}$$

where $U(t)$ is the dynamical group of Eq. (13.106).

From results of [97] it follows that the first term on the RHS of (13.120) admits the asymptotic expansion (13.113) with $a_\omega = 0$ and the coefficients C_l defined by (13.118). This is the central achievement of the scattering theory: the contribution of the continuous spectrum to the solution $U(t)\psi(0)$ of the homogeneous equation vanishes as $t \to \infty$.

It remains to examine the integral term. We can rewrite it as follows:

$$I(t) = -i e^{-i\omega t} \int_0^t U(s) b e^{i\omega s}\, ds. \tag{13.121}$$

Therefore, writing formally $U(s) = e^{-iHs}$, we obtain

$$I(t) = -i e^{-i\omega t} \int_0^t e^{-iHs} b e^{i\omega s}\, ds \sim e^{-i\omega t} a_\omega, \quad t \to \infty, \tag{13.122}$$

where

$$
\begin{aligned}
a_\omega &= -i\int_0^\infty e^{-iHs}be^{i\omega s}\,ds = -i\lim_{\varepsilon\to 0+}\int_0^\infty e^{-iHs}be^{i(\omega+i\varepsilon)s}\,ds \\
&= -\lim_{\varepsilon\to 0+}(H-\omega-i\varepsilon)^{-1}b,
\end{aligned} \tag{13.123}
$$

which coincides with (13.117). In particular, (13.122) shows that contribution of $I(t)$ into the coefficients C_l vanishes. The details can be found in [119].

Remark 13.31

(i) Lemma 13.30(iii) implies that the sum over the discrete spectrum on the RHS of (13.113), does not appear if the eigenfunctions were not presented in the initial state $\psi(0,x)$.

(ii) Formula (13.117) is called the '*limiting absorption principle*', because $R(\omega+i\varepsilon) = (H-\omega-i\varepsilon)^{-1}$ is the resolvent corresponding to the 'damped' equation

$$
i\dot\psi(t,\mathbf{x}) = H\psi(t,\mathbf{x}) - i\varepsilon\psi(t,\mathbf{x}) + b(\mathbf{x})e^{-i\omega t}, \tag{13.124}
$$

where the term $-i\varepsilon\psi(t,\mathbf{x})$ describes an *absorption of energy*.

Chapter 14
Exercises

Abstract Here we collect classical calculations lying in the ground of the 'old quantum mechanics': the Kepler problem, the Bohr–Sommerfeld quantization, electromagnetic plane waves, the Lorentz theory of polarization and dispersion, the normal Zeeman effect, diamagnetism and paramagnetism, and the Landé factor. Finally, we present the Heisenberg quantization of harmonic oscillator.

14.1 The Kepler Problem

The Kepler problem is defined by the two-particle Lagrangian

$$L = \frac{m_1}{2}\dot{x}_1^2 + \frac{m_2}{2}\dot{x}_2^2 + \frac{\gamma m_1 m_2}{|x_1 - x_2|}.$$

We shall solve it through the following steps:

(a) Reduction to the one-particle problem (coordinate x) via separation of the center-of-mass coordinate X.
(b) Reduction to the two-dimensional problem by observing that the plane spanned by $(x(0), \dot{x}(0))$ is invariant.
(c) Reduction to the one-dimensional problem (coordinate r) by the introduction of polar coordinates (r, φ). Show that Kepler's third law holds, find the effective potential of the one-dimensional problem, and check for energy conservation.
(d) Investigate the orbits of the one-dimensional problem.
(e) Integrate the orbit equation to find the possible orbits. Show, depending on the values of the integration constants (angular momentum, energy), that the orbits are ellipses, parabolas or hyperbolas.

Solution

(a) Denote by

$$X(t) := \frac{m_1}{M}x_1(t) + \frac{m_2}{M}x_2(t)$$

A. Komech, *Quantum Mechanics: Genesis and Achievements*,
DOI 10.1007/978-94-007-5542-0_14,

the center of mass, where the total mass is $M := m_1 + m_2$, and $x := x_1 - x_2$. Then

$$\sum_{i=1}^{2} m_i \dot{x}_i^2 = \frac{m_1 m_2 \dot{x}^2(t)}{M} + M\dot{X}^2.$$

Hence we can rewrite the Lagrangian in the form

$$L = \frac{m_1 m_2 \dot{x}^2}{2M} + \frac{M\dot{X}^2}{2} + \frac{\gamma m_1 m_2}{|x_1 - x_2|} = L_1(x, \dot{x}, \dot{X}).$$

Applying the Euler–Lagrange equation to the coordinate X,

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{X}} - \frac{\partial L}{\partial X} = 0,$$

we get $\frac{d}{dt} M\dot{X} = 0$. Hence,

$$\dot{X} = \text{const}, \tag{14.1}$$

and so $X(t) = at + b$, where a, b are some constants.

(b) It remains to consider the Lagrangian

$$L \equiv L(x, \dot{x}) := \frac{\dot{x}^2}{2M} + \frac{\gamma}{|x|}.$$

Each trajectory lies in a plane. By rotational symmetry, we may assume that the considered trajectory lies in the plane $x_3 = 0$. As a result, $x(t) = (u(t), v(t), 0)$ and the Lagrangian for the variables $u(t), v(t)$ is as follow:

$$L(u, v, \dot{u}, \dot{v}) := \frac{\dot{u}^2 + \dot{v}^2}{2M} + \frac{\gamma}{\sqrt{u^2 + v^2}}. \tag{14.2}$$

(c) In polar coordinates (r, φ),

$$u = r\cos\varphi, \qquad v = r\sin\varphi, \tag{14.3}$$

we have $u^2 + v^2 = r^2$, $\dot{u}^2 + \dot{v}^2 = \dot{r}^2 + r^2\dot{\varphi}^2$. Substituting this into the Lagrangian, this gives

$$L(r, \varphi, \dot{r}, \dot{\varphi}) := \frac{\dot{r}^2 + r^2\dot{\varphi}^2}{2M} + \frac{\gamma}{r}. \tag{14.4}$$

Applying the Euler–Lagrange equations read

$$\ddot{r} = r\dot{\varphi}^2 - \frac{\gamma M}{r^2}, \quad \frac{d}{dt}\left[\frac{r^2\dot{\varphi}}{M}\right] = 0.$$

The last equation implies Kepler's third law:

$$r^2\dot{\varphi} = \text{const} =: I. \tag{14.5}$$

Hence, the first equation can be written as follows:

$$\ddot{r} = \frac{I^2}{r^3} - \frac{\gamma M}{r^2} = -V'(r), \quad V(r) := \frac{I^2}{2r^2} - \frac{2\gamma M}{r}.$$

Multiplying this equation by $\dot{r}$ and integrating, we arrive at the reduced energy conservation law

$$\dot{r}^2 + V(r) = E. \tag{14.6}$$

(d) Now we find $r(\varphi)$. Since $\frac{dr}{d\varphi} = \frac{\dot{r}}{\dot{\varphi}}$, from (14.6) and (14.5) it is found that

$$\frac{dr}{d\varphi} = \frac{\pm\sqrt{-\frac{I^2}{r^2} + \frac{2\gamma M}{r} + 2E}}{\frac{I}{r^2}}.$$

Hence

$$\int \frac{I\,dr}{r^2\sqrt{-\frac{I^2}{r^2} + \frac{2\gamma M}{r} + 2E}} = \pm \int d\varphi.$$

Introducing the **Clairaut substitution** $\rho := 1/r$, we have $dr = -\rho^{-2}\,d\rho$. Hence,

$$\int \frac{I\,dr}{r^2\sqrt{-\frac{I^2}{r^2} + \frac{2\gamma M}{r} + 2E}} = -\int \frac{I\,d\rho}{\sqrt{-I^2\rho^2 + 2\gamma M\rho + 2E}}$$
$$= \int \frac{I\,d\rho}{\sqrt{D - I^2(\rho - B)^2}}.$$

Here $D := 2E + (\gamma M)^2 I^{-2}$, $B := \gamma M I^{-2}$. Note that the constant D must be positive for any non-empty trajectory. Further,

$$\int \frac{I\,d\rho}{\sqrt{D - I^2(\rho - B)^2}} = \arcsin\left\{\frac{I(\rho - B)}{\sqrt{D}}\right\} = \pm\varphi + \varphi_0.$$

Finally, we get $\frac{I(\rho-B)}{\sqrt{D}} = \sin(\varphi + \varphi_0)$. Since $\rho := 1/r$, it is found that

$$\frac{A_1}{r} = A_2 + \sin(\varphi + \varphi_0), \tag{14.7}$$

where

$$A_1 := \frac{I}{\sqrt{D}}, \qquad A_2 := \frac{IB}{\sqrt{D}} = \frac{\gamma M}{\sqrt{2EI^2 + (\gamma M)^2}}. \tag{14.8}$$

Note that r is bounded away from zero by (14.7).

(e) For simplicity, we assume that $\varphi_0 = 0$. Then (14.7) becomes

$$A_2 r = A_1 - r\cos\varphi. \tag{14.9}$$

We recall that $r\sin\varphi = v$ and $r = \sqrt{u^2+v^2}$. Now we can rewrite (14.9) in the form $(u^2+v^2)A_2^2 = (A_1 - v)^2$. Hence,

$$u^2 + v^2\left(1 - \frac{1}{A_2^2}\right) + 2\frac{A_1}{A_2^2}v = \frac{A_1^2}{A_2^2}.$$

Obviously, $A_1, A_2 \geq 0$. First, consider the case $A_1 > 0$. Then, we have an ellipse for $A_2 > 1$, and a hyperbola, for $A_2 < 1$; for $A_2 = 1$ we have a parabola. From (14.8) it follows that: if $E < 0$ we have an ellipse, if $E > 0$ we have a hyperbola, and if $E = 0$ we have a parabola.

Finally, in the case $A_1 = 0$ we have $I = 0$. Then either $r = 0$ and φ are arbitrary (the point mass is at the origin) or $\sin\varphi = -A_2 = \text{const}$ (the point mass moves along a line through the origin).

14.2 The Bohr–Sommerfeld Quantization

Problem Use the Bohr–Sommerfeld quantization rules (1.62):

$$\oint p_r\, dr = hk, \qquad \int p_\varphi\, d\varphi = hm \tag{14.10}$$

to derive the quantization of angular momentum and the Schrödinger formula (6.3) for the energy levels of the electron in the Hydrogen atom.

Solution The Hamiltonian for an electron with charge $e < 0$ and mass m in the Coulomb field of an infinitely heavy nucleus with charge $|e| > 0$ is

$$H = \frac{1}{2\mathrm{m}}\left(p_r^2 + \frac{p_\varphi^2}{r^2}\right) - \frac{e^2}{r}, \quad p_r = \mathrm{m}\dot{r}, \;\; p_\varphi = \mathrm{m}r^2\dot{\varphi}. \tag{14.11}$$

Hence, the corresponding Hamilton equations read

$$\dot{\varphi} = H_{p_\varphi} = \frac{p_\varphi}{\mathrm{m}r^2}, \qquad \dot{p}_\varphi = -H_\varphi = 0, \tag{14.12}$$

$$\dot{r} = H_{p_r} = \frac{p_r}{\mathrm{m}}, \qquad \dot{p}_r = -H_r = \frac{p_\varphi^2}{\mathrm{m}r^3} - \frac{e^2}{r^2}. \tag{14.13}$$

From (14.12), we obtain (14.5):

$$p_\varphi = J = \text{const} > 0. \tag{14.14}$$

Then, using (14.13),

$$\dot{p}_r = \frac{J^2}{\mathrm{m}r^3} - \frac{e^2}{r^2}. \tag{14.15}$$

Since $\frac{dp_r}{dr} = \frac{\dot{p}_r}{\dot{r}}$, it follows from (14.13) and (14.15) that

$$\frac{dp_r}{dr} = \frac{\frac{J^2}{\mathrm{m}r^3} - \frac{e^2}{r^2}}{\frac{p_r}{\mathrm{m}}} \quad \text{or} \quad \frac{p_r\, dp_r}{\mathrm{m}} = \left(\frac{J^2}{\mathrm{m}r^3} - \frac{e^2}{r^2}\right) dr. \tag{14.16}$$

Hence,

$$\frac{p_r^2}{2\mathrm{m}} + \frac{J^2}{2\mathrm{m}r^2} - \frac{e^2}{r} = \mathcal{E} = \text{const}. \tag{14.17}$$

To have bounded orbits (ellipses), we should consider $\mathcal{E} < 0$. Then

$$p_r = \sqrt{2\mathrm{m}\left(\frac{e^2}{r} - E\right) - \frac{J^2}{r^2}}, \tag{14.18}$$

where $E \equiv |\mathcal{E}| > 0$ was introduced for convenience.

Now we find the extrema of r. From Eq. (14.13) it follows that $\dot{r} = 0$ if and only if $p_r = 0$. From (14.18), we have

$$2\mathrm{m}\left(\frac{e^2}{r} - E\right) - \frac{J^2}{r^2} = 0,$$

which can be rewritten as

$$r^2 - \frac{e^2 r}{E} + \frac{J^2}{2\mathrm{m}E} = 0.$$

Therefore,

$$r = \frac{e^2}{2E} \pm \sqrt{\frac{e^4}{4E^2} - \frac{J^2}{2\mathrm{m}E}} = \frac{e^2}{2E} \pm \frac{1}{2E}\sqrt{e^4 - \frac{2J^2 E}{\mathrm{m}}} =: r_\pm.$$

Hence, $r_{\max} = r_+$ and $r_{\min} = r_-$.

Now we employ (14.10). At first, note that from (14.10) and (14.14) it follows that

$$\int p_\varphi\, d\varphi = \int J\, d\varphi = 2\pi J = hm. \tag{14.19}$$

As a result,

$$J = \hbar m, \quad m \in \mathbb{Z},\ m > 0. \tag{14.20}$$

Second, (14.18) implies

$$\begin{aligned}\int p_r\,dr &= 2\int_{r_{\min}}^{r_{\max}} p_r\,dr = 2\int_{r_{\min}}^{r_{\max}} \sqrt{2\mathrm{m}\left(\frac{e^2}{r}-E\right)-\frac{J^2}{r^2}}\,dr\\ &= 2\int_{r_{\min}}^{r_{\max}} \sqrt{-2\mathrm{m}Er^2+2\mathrm{m}e^2r-J^2}\,\frac{dr}{r}.\end{aligned} \tag{14.21}$$

Applying here the formula

$$\begin{aligned}\int \sqrt{ax^2+bx+c}\,\frac{dx}{x} &= \sqrt{ax^2+bx+c}-\frac{b}{2\sqrt{-a}}\arcsin\frac{2ax+b}{b^2-4ac}\\ &\quad -\sqrt{-c}\arcsin\frac{bx+2c}{x\sqrt{b^2-4ac}}, \quad a<0,\ c<0\end{aligned}$$

with $a=-2\mathrm{m}E<0$, $b=2\mathrm{m}e^2$, $c=-J^2<0$, we obtain

$$\begin{aligned}\int p_r\,dr &= 2\Bigg[\sqrt{2\mathrm{m}\left(\frac{e^2}{r}-E\right)-\frac{J^2}{r^2}}\,\Bigg|_{r_{\min}}^{r_{\max}}\\ &\quad -\frac{2\mathrm{m}e^2}{2\sqrt{2\mathrm{m}E}}\arcsin\frac{-4\mathrm{m}Er+2\mathrm{m}e^2}{\sqrt{4\mathrm{m}^2e^4-8J^2\mathrm{m}E}}\Bigg|_{r_{\min}}^{r_{\max}}\\ &\quad -J\arcsin\frac{2\mathrm{m}e^2r-2J^2}{r\sqrt{4\mathrm{m}^2e^4-8J^2\mathrm{m}E}}\Bigg|_{r_{\min}}^{r_{\max}}\Bigg]\\ &= 2\left[0+\pi\frac{\mathrm{m}e^2}{\sqrt{2\mathrm{m}E}}-J\pi\right]\\ &= 2\pi\left[\sqrt{\frac{\mathrm{m}e^4}{2E}}-J\right].\end{aligned} \tag{14.22}$$

Now (14.10) gives that

$$2\pi\left[\sqrt{\frac{\mathrm{m}e^4}{2E}}-J\right]=hk, \quad k\in\mathbb{Z}.$$

Applying (14.20), we get

$$\sqrt{\frac{\mathrm{m}e^4}{2E}}=\hbar(k+m).$$

We denote $k+m=n\geq 1$, and now

$$\frac{\mathrm{m}e^4}{2E}=\hbar^2n^2, \quad n\geq 1.$$

Hence,

$$\mathcal{E} = -E = -\frac{\mathrm{m}e^4}{2\hbar^2 n^2} = -\frac{2\pi\hbar c R}{n^2}, \quad R := \frac{\mathrm{m}e^4}{4\pi\hbar^3 c}$$

which coincides with the Schrödinger formula (6.3).

14.3 Electromagnetic Plane Waves

Problem Find the electric and magnetic fields of a plane electromagnetic wave by inserting the complex plane wave ansatz

$$\mathbf{E} = \mathbf{E}_0 \exp i(\mathbf{k} \cdot \mathbf{x} - \omega t), \quad \mathbf{B} = \mathbf{B}_0 \exp i(\mathbf{k} \cdot \mathbf{x} - \omega t) \tag{14.23}$$

into the free Maxwell equations. Find two linearly independent solutions and construct a linearly and circularly polarized waves. Calculate the corresponding energy flow per time unit.

Solution The Maxwell equations in vacuum are

$$\nabla \times \mathbf{B} = \frac{1}{c}\dot{\mathbf{E}}, \qquad \nabla \times \mathbf{E} = -\frac{1}{c}\dot{\mathbf{B}}, \qquad \nabla \cdot \mathbf{B} = 0, \qquad \nabla \cdot \mathbf{E} = 0.$$

Substituting the ansatz (14.23), we get

$$\mathbf{k} \times \mathbf{B}_0 = -\frac{\omega}{c}\mathbf{E}_0, \qquad \mathbf{k} \times \mathbf{E}_0 = \frac{\omega}{c}\mathbf{B}_0, \qquad \mathbf{k} \cdot \mathbf{B}_0 = 0, \qquad \mathbf{k} \cdot \mathbf{E}_0 = 0. \tag{14.24}$$

A general solution is obvious: vector $\mathbf{k} \in \mathbb{R}^3$ is arbitrary, $\mathbf{E}_0 \perp \mathbf{k}$, and $\mathbf{B}_0 = \frac{c}{\omega}\mathbf{k} \times \mathbf{E}_0$, where ω is any solution of the dispersion relation

$$\omega^2 = \mathbf{k}^2 c^2. \tag{14.25}$$

The plane wave allows the representation (12.44)

$$\mathbf{E} = -\frac{1}{c}\dot{\mathbf{A}}, \qquad \mathbf{B} = \nabla \times \mathbf{A} \tag{14.26}$$

with the magnetic potential

$$\mathbf{A}(x,t) = \mathbf{A}_0 \exp i(\mathbf{k} \cdot \mathbf{x} - \omega t), \quad \mathbf{A}_0 = \frac{c}{i\omega}\mathbf{E}_0. \tag{14.27}$$

In particular, $\mathbf{k} \cdot \mathbf{A}_0 = 0$, and hence

$$\nabla \cdot \mathbf{A} = 0. \tag{14.28}$$

Here we allow for a complex notation for the plane wave field for convenience, but one should keep in mind that only the real part of the field is physical. Splitting $\mathbf{E}_0 = \mathbf{E}_0' + i\mathbf{E}_0''$ with $\mathbf{E}_0', \mathbf{E}_0'' \in \mathbb{R}^3$, we obtain

$$\mathbf{E}_{\text{ph}} = \text{Re}\big(\mathbf{E}_0 e^{i\mathbf{k}\cdot\mathbf{x}-i\omega t}\big) = \mathbf{E}_0' \cos(\mathbf{k}\cdot\mathbf{x} - \omega t) - \mathbf{E}_0'' \sin(\mathbf{k}\cdot\mathbf{x} - \omega t). \tag{14.29}$$

This field is linearly polarized if $\mathbf{E}_0' \| \mathbf{E}_0''$. Otherwise, it is 'elliptically polarized' ('circularly polarized' if $\mathbf{E}_0' \perp \mathbf{E}_0''$ and $|\mathbf{E}_0'| = |\mathbf{E}_0''|$).

The corresponding magnetic field is given by (14.24):

$$\mathbf{B}_{\text{ph}} = \frac{c}{\omega}\mathbf{k} \times \mathbf{E}_{\text{ph}} = \mathbf{B}_0' \cos(\mathbf{k}\cdot\mathbf{x} - \omega t) + \mathbf{B}_0'' \sin(\mathbf{k}\cdot\mathbf{x} - \omega t), \tag{14.30}$$

where

$$\mathbf{B}_0' = \frac{c}{\omega}\mathbf{k} \times \mathbf{E}_0' = \frac{\mathbf{k}}{|\mathbf{k}|} \times \mathbf{E}_0', \qquad \mathbf{B}_0'' = -\frac{c}{\omega}\mathbf{k} \times \mathbf{E}_0'' = -\frac{\mathbf{k}}{|\mathbf{k}|}\mathbf{E}_0''$$

since $\omega = |\mathbf{k}|c$ by (14.25). Therefore, the energy flow density $\mathbf{S} = \frac{c}{4\pi}\mathbf{E} \times \mathbf{B}$ oscillates in time, and points to the direction of the wave vector $\mathbf{k}$. Its **time average** is given by

$$\overline{\mathbf{S}} = \frac{c}{4\pi}\overline{\mathbf{E} \times \mathbf{B}} = \frac{c}{8\pi}\big[\left|\mathbf{E}_0'\right|^2 + \left|\mathbf{E}_0''\right|^2\big]\frac{\mathbf{k}}{|\mathbf{k}|}. \tag{14.31}$$

14.4 Polarization and Dispersion

In many instances, a classical point-like electron in matter may be modeled by a damped harmonic oscillator, which oscillates around a fixed, positively charged center. In an external electromagnetic field, the corresponding equation of motion for the position of the electron is

$$\mathrm{m}\big[\ddot{\mathbf{x}}(t) + \gamma\dot{\mathbf{x}}(t) + \omega_0^2\mathbf{x}(t)\big] = e\left[\mathbf{E} + \frac{1}{c}\dot{\mathbf{x}} \times \mathbf{B}\right]. \tag{14.32}$$

Here m is the mass and e the charge of the electron, γ is a damping constant and $\mathbf{x}$ is the distance vector from the center.

The model was suggested by Lorentz. The friction is introduced to provide stability under external perturbation $\mathbf{E} = \mathbf{E}_0 e^{-i\omega t}$ for all frequencies $\omega \in \mathbb{R}$. This friction should model a radiation of energy to infinity by the atom.

Problem Calculate the electric dipole moment $\mathbf{p}_0 = e\mathbf{x}$ of the electron in an oscillating electric field $\mathbf{E} = \mathbf{E}_0 \exp(-i\omega t)$. Discuss the frequency dependence of the resulting polarizability.

Solution Inserting $\mathbf{x} = \frac{\mathbf{p}_0}{e}\exp(-i\omega t)$ into the above equation, we get

$$\big[-\omega^2 - i\omega\gamma + \omega_0^2\big]\mathbf{p}_0 = \frac{e^2}{\mathrm{m}}\mathbf{E}_0 \tag{14.33}$$

or

$$\mathbf{p}_0 = \frac{e^2}{\mathrm{m}}\frac{1}{\omega_0^2 - \omega^2 - i\omega\gamma}\mathbf{E}_0 \equiv \alpha(\omega)\mathbf{E}_0. \tag{14.34}$$

with the 'atomic polarizability' α. The absolute value of the polarizability $|\alpha(\omega)| = |\mathbf{E}_0|/|\mathbf{p}_0|$ equals

$$\big|\alpha(\omega)\big| = \frac{e^2}{\mathrm{m}}\frac{1}{[(\omega_0^2 - \omega^2)^2 + \omega^2\gamma^2]^{1/2}}. \tag{14.35}$$

It has local maxima at $\omega = \pm\omega(\gamma)$ with $\omega(\gamma) = \sqrt{\omega_0^2 + \frac{\gamma^2}{2}}$ and a local minimum at $\omega = 0$ (we assume that $\omega_0 > 0$). Hence there is anomalous dispersion (weaker polarizability coefficient $|\alpha(\omega)|$ for higher frequency $|\omega|$) for values of $|\omega| \geq \omega(\gamma)$.

The complex nature of polarizability means that there exists a phase difference between $\mathbf{E}_0$ and $\mathbf{p}_0$. Away from the resonance, i.e., for $|\omega| - \omega_0 \gg \gamma$, the polarizability $\alpha(\omega)$ is approximately real.

If there are N atoms per volume and if f_k denotes the fraction of electrons per atom that oscillate with eigenfrequency $\omega_k > 0$, then the total polarization per volume is as follows:

$$\mathbf{p} = \chi_e\mathbf{E}, \quad \chi_e = \frac{Ne^2}{\mathrm{m}}\sum_k \frac{f_k}{\omega_k^2 - \omega^2 - i\omega\gamma}. \tag{14.36}$$

Hence, we find the electric permittivity

$$\varepsilon = 1 + 4\pi\chi_e = 1 + \frac{4\pi Ne^2}{\mathrm{m}}\sum_k \frac{f_k}{\omega_k^2 - \omega^2 - i\omega\gamma}, \tag{14.37}$$

which is the Drude formula for permittivity. This famous formula was established in the framework of 'Old Quantum Mechanics' by Kramers and Heisenberg [121–123]. It is similar to its quantum analogue (8.83) and implies many important conclusions which agree with experimental observations (cf. with the comments below (8.83)). See also the detailed discussions in [7]):

I. The formula implies an important approximation to the refraction coefficient n. Namely, for many materials $\mu \sim 1$, and hence $n = \sqrt{\varepsilon\mu} \sim \sqrt{\varepsilon}$ by (12.206) if $\varepsilon > 0$ (here we should set $\gamma = 0$).
II. It implies the sharp peak of $|\varepsilon(\omega)|$ near the resonances ω_k for small γ, and the form of the peaks is in a good agreement with experimental observations.
III. It explains the anomalous dispersion near the eigenfrequencies ω_k for small γ.

IV. The poles of $\chi_e(\omega)$ and $\varepsilon(\omega)$ lie in the low complex half plane $\mathrm{Im}\,\omega < 0$ since $\gamma > 0$. Hence, $\chi_e(\omega)$ and $\varepsilon(\omega)$ are holomorphic in the upper complex half plane $\mathrm{Im}\,\omega > 0$. This fact implies the integral *dispersion relations* between real and imaginary parts of $\chi_e(\omega)$ and $\varepsilon(\omega)$ discovered by Kramers and Kronig [124–126]. This theory was developed later in the framework of Quantum Field Theory, see the survey in [156, v. III].

14.5 The Normal Zeeman Effect

Electrons in matter are again described by spherical harmonic oscillators, which oscillate around a fixed center (nucleus), as in Sect. 14.4. Here we assume that the electrons are excited with their eigenfrequency by some unspecified mechanism (e.g., heat or radiation); therefore we ignore a possible damping by assuming $\gamma = 0$ in Eq. (14.32). These harmonic oscillators are now exposed to a constant, external magnetic field, therefore the equation of motion for the position of one electron is

$$\mathrm{m}\big[\ddot{\mathbf{x}}(t) + \omega_0^2\mathbf{x}(t)\big] = \frac{e}{c}\dot{\mathbf{x}} \times \mathbf{B}. \tag{14.38}$$

Problem

(i) Find the three eigenfrequencies and eigenmodes of this equation of motion.
(ii) Use the resulting dipole moments $\mathbf{p} = e\mathbf{x}$ to calculate the dipole radiation of all three modes and their polarization patterns, both parallel and perpendicular to the constant, external magnetic field.

Solution We will follow Lorentz's arguments who was awarded the Nobel Prize in 1902 for this theory. Without loss of generality we assume that the constant external magnetic field is directed along Ox_3-axis, i.e., $\mathbf{B} = B\mathbf{e}_3$. Then the equations of motion read,

$$\ddot{x}_1 + \omega_0^2 x_1 = 2\omega_{\mathcal{L}}\dot{x}_2, \qquad \ddot{x}_2 + \omega_0^2 x_2 = -2\omega_{\mathcal{L}}\dot{x}_1, \tag{14.39}$$

$$\ddot{x}_3 + \omega_0^2 x_3 = 0, \tag{14.40}$$

where $\omega_{\mathcal{L}} = \frac{eB}{2\mathrm{mc}}$ is the *Larmor frequency*. Obviously, the last (decoupled) equation is solved by $x_3 = r_0 \sin(\omega_0 t - \phi_0)$, where ϕ_0 is an irrelevant integration constant, which will be set to zero in the sequel. Therefore the parallel (to the magnetic field) eigenmode is as follows

$$\mathbf{x}_{\|} = r_0\mathbf{e}_3 \sin\omega_0 t \tag{14.41}$$

with eigenfrequency ω_0 (the eigenfrequency is not altered by the magnetic field).

To solve the system (14.39), we write it in the complex form. Introducing $z = x_1 + ix_2$, we obtain

$$\ddot{z} + \omega_0^2 z = -2i\omega_{\mathcal{L}}\dot{z}. \tag{14.42}$$

Substituting $z = r_0 e^{i\omega t}$, we arrive at the *characteristic equation* $-\omega^2 + \omega_0^2 = 2\omega_{\mathcal{L}}\omega$. The solution is given by

$$\omega_\pm = -\omega_{\mathcal{L}} \pm \sqrt{\omega_0^2 + \omega_{\mathcal{L}}^2}, \tag{14.43}$$

where $\omega_+ > 0$, while $\omega_- < 0$. Therefore, we find the following two perpendicular (to the magnetic field) eigenmodes

$$\mathbf{x}_\pm = r_0(\mathbf{e}_1 \cos\omega_\pm t + \mathbf{e}_2 \sin\omega_\pm t), \tag{14.44}$$

with the eigenfrequencies $\omega_\pm$ altered by the magnetic field.

Our next aim is to calculate the dipole radiation emitted by these three eigenmodes. Here we shall use the formulas for Hertzian dipole radiation (see Sect. 12.10) with $\mathbf{p} = e\mathbf{x}$. From formulas (12.121), (12.122), and (12.124), it follows that for $r = |\mathbf{x}| \gg r_0$ we can assume approximately

$$\mathbf{B}(t,\mathbf{x}) = \frac{e}{c^2 r}\ddot{\mathbf{x}}(t - r/c) \times \mathbf{n}, \tag{14.45}$$

$$\mathbf{E}(t,\mathbf{x}) = \frac{e}{c^2 r}\big[-\ddot{\mathbf{x}}(t - r/c) + \mathbf{n}\big(\mathbf{n}\cdot\ddot{\mathbf{x}}(t - r/c)\big)\big] = \frac{e}{c^2 r}\big[\ddot{\mathbf{x}}(t - r/c) \times \mathbf{n}\big] \times \mathbf{n}, \tag{14.46}$$

$$\mathbf{S} = \mathbf{n}\frac{e^2 \sin^2\chi}{4\pi c^3 r^2}\ddot{\mathbf{x}}^2(t - r/c), \tag{14.47}$$

where $\mathbf{n} := \mathbf{x}/r$, and χ is the angle between $\mathbf{x}$ and $\ddot{\mathbf{x}}(t - r/c)$. Let us consider the radiation of the parallel and perpendicular eigenmodes separately:

I. For the parallel eigenmode (14.41) we have $\ddot{\mathbf{x}} \| \mathbf{e}_3$, hence $\mathbf{B}(t,\mathbf{x}) \perp \mathbf{e}_3$ by (14.45). Hence, $\mathbf{B}(t,\mathbf{x})$ is linearly polarized since it is also perpendicular to $\mathbf{x}$. Therefore, the electric field $\mathbf{E}(t,\mathbf{x})$ is also linearly polarized since it is perpendicular to $\mathbf{B}(t,\mathbf{x})$ and $\mathbf{x}$. Finally, the Poynting vector

$$\mathbf{S} = \mathbf{n}\frac{e^2\omega_0^4 |r_0|^2}{4\pi c^3 r^2}\sin^2\theta \sin^2\omega_0 t', \tag{14.48}$$

where $t' := t - r/c$, and θ is the angle between $\mathbf{x}$ and $\mathbf{e}_3$. Hence, there is no radiation in the $\mathbf{e}_3$ direction. The maximal radiation is emitted in the orthogonal directions to $\mathbf{e}_3$, along the 'median' plane $x_3 = 0$. At the points of this median plane we have $\mathbf{E}(t,\mathbf{x}) \| \mathbf{e}_3$ by (14.46).

II. For the perpendicular eigenmodes (14.44), the accelerations

$$\ddot{\mathbf{x}}_\pm(t - r/c) = -r_0\omega_\pm^2\big(\mathbf{e}_1 \cos\omega_\pm t' + \mathbf{e}_2 \sin\omega_\pm t'\big), \tag{14.49}$$

rotate with angular velocities $\omega_\pm$. Hence, the directions of the radiation fields $\mathbf{E}(t,\mathbf{x})$ and $\mathbf{B}(t,\mathbf{x})$ also rotate according to (14.45) and (14.46). Finally, the Poynting vector (14.47) reads

$$\mathbf{S} = \mathbf{n}\frac{e^2\omega_\pm^4|r_0|^2}{4\pi c^3 r^2}\big[1-\cos^2\chi\big] = \mathbf{n}\frac{e^2\omega_\pm^4|r_0|^2}{4\pi c^3 r^2}\big[1-\sin^2\theta\cos^2\big(\omega_\pm t'-\varphi\big)\big] \tag{14.50}$$

since $\mathbf{n} = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$, and hence

$$\begin{aligned}\cos\chi &= -\mathbf{n}\cdot\big(\mathbf{e}_1\cos\omega_\pm t' + \mathbf{e}_2\sin\omega_\pm t'\big)\\ &= -\sin\theta\big(\cos\varphi\cos\omega_\pm t' + \sin\varphi\sin\omega_\pm t'\big).\end{aligned}$$

Now the radiation is maximal in the $\mathbf{e}_3$ direction, and it does not depend on time. The radiation is minimal in the orthogonal directions to $\mathbf{e}_3$, along the 'median' plane. At the points of this median plane, we have $\mathbf{B}(t,\mathbf{x})\|\mathbf{e}_3$ by (14.45).

14.6 Diamagnetism and Paramagnetism

The relation between the magnetic induction **B**, the magnetic field intensity **H** and the magnetization **M** of matter is as follows (for isotropic matter)

$$\mathbf{B} = \mathbf{H} + 4\pi\mathbf{M}, \qquad \mathbf{B} = \mu\mathbf{H}, \qquad \mathbf{M} = \chi_m\mathbf{H}, \quad \mu = 1 + 4\pi\chi_m.$$

Here, a substance is called diamagnetic if $\mu < 1$.

Problem Find the diamagnetic behavior of electrons moving on circular orbits.

Solution Use again the harmonic oscillator model in a constant external magnetic field (like in Sect. 14.5). Electrons moving on circular orbits have to be interpreted as currents inducing the magnetic field.

We again assume that the external magnetic field is along the $\mathbf{e}_3$ direction and $\mathbf{B} = B\mathbf{e}_3$. The parallel eigenmode (14.41) does not contribute to the magnetic moment since the integral (12.192) for the corresponding current vanishes. Hence, we are interested only in the perpendicular eigenmodes (14.44) with frequencies

$$\omega_\pm = -\omega_\mathcal{L} \pm \sqrt{\omega_0^2 + \omega_\mathcal{L}^2}, \tag{14.51}$$

according to (14.43). Without external field these are exactly two modes rotating with the same angular velocity ω_0, one (the + mode) in the counter-clockwise direction, and the other, in the clockwise direction. The two modes induce the magnetic moments of equal strengths, but of opposite orientations. Therefore, macroscopically, their net contribution to the magnetization is zero, since the eigenmodes are equiprobable.

With external field, however, the additional contribution is in the same direction in both cases, and a net contribution remains. The magnetic moment of each eigenmode is given by (12.192):

$$\mathbf{m}_\pm = \frac{1}{2c}\int \mathbf{x} \times \mathbf{j}_\pm(t,\mathbf{x})\, d^3\mathbf{x}, \tag{14.52}$$

where $\mathbf{j}_\pm(t,\mathbf{x})$ is the corresponding current density. In our case, the electron is in a circular orbit in the x_1, x_2 plane, therefore

$$\mathbf{j}_\pm(t,\mathbf{x}) = e\delta\big(\mathbf{x} - \mathbf{x}_\pm(t)\big)\dot{\mathbf{x}}_\pm(t). \tag{14.53}$$

By (14.44), we have

$$\mathbf{m}_\pm = \frac{1}{2c} e\mathbf{x}_\pm(t) \times \dot{\mathbf{x}}_\pm(t) = \frac{1}{2c} e\omega_\pm r_0^2 \mathbf{e}_3. \tag{14.54}$$

Hence, the mean magnetic moment reads

$$\mathbf{m} = \frac{\mathbf{m}_+ + \mathbf{m}_-}{2} = -\frac{1}{2c} e\omega_{\mathcal{L}} r_0^2 \mathbf{e}_3 = -\frac{e^2 B r_0^2}{4\mathrm{m}c^2}\mathbf{e}_3 = -\frac{e^2 r_0^2}{4\mathrm{m}c^2}\mathbf{B} \tag{14.55}$$

since $(\omega_+ + \omega_-)/2 = -\omega_{\mathcal{L}}$ according to (14.51). Hence, the magnetic moment is *opposite to the external magnetic field* which means the *diamagnetism*.

For general orbits of electrons, r_0^2 must be replaced by the orbit average $\overline{|x^1|^2 + |x^2|^2}$; if the force law for the electron is spherically symmetric, as in our case, we have $\overline{|x^1|^2 + |x^2|^2} = (2/3)\overline{\mathbf{x}^2}$. If there are several electrons per atom with different average orbit radii and if there are N atoms per volume, then the magnetization is

$$\mathbf{M} = -N\frac{e^2 B}{6\mathrm{m}c^2}\sum_j \overline{\mathbf{x}_j^2}\mathbf{e}_3 = -N\frac{e^2}{6\mathrm{m}c^2}\sum_j \overline{\mathbf{x}_j^2}\mathbf{B}, \tag{14.56}$$

and the susceptibility is approximately given (if it is small) by the Langevin formula (cf. (9.26))

$$\chi_m = -N\frac{e^2}{6\mathrm{m}c^2}\sum_j \overline{\mathbf{x}_j^2}. \tag{14.57}$$

The assumptions here were rather general (electrons orbiting around nuclei) and, in fact, all materials are diamagnetic, but the diamagnetism may be over-compensated by other effects with positive magnetic susceptibility (like paramagnetism).

Paramagnetism is present when the following conditions hold: (1) the atoms or molecules of the material already have a fixed magnetic moment $\mathbf{m}_0$ (even without external magnetic field), and (2) the orientations of these magnetic moments are randomly distributed in the absence of an external magnetic field. Without external magnetic field there is, therefore, no macroscopic magnetization of the paramagnetic material. However, in an external magnetic field, the magnetic moments are partially aligned along the direction of the external field, because this aligned position is energetically favorable. This alignment is thwarted by thermal fluctuations, and the average magnetic moment per molecule is

$$\overline{\mathbf{m}} = \eta \frac{|\mathbf{m}_0|^2}{kT} \mathbf{B}_{\text{ext}}, \tag{14.58}$$

where η is some constant depending on the molecule type, T is the temperature, and k is the Boltzmann constant. Both quantum mechanical and thermodynamic considerations are needed to derive equation (14.58), which is beyond the scope of the present discussion.

14.7 The Vector Model

The phenomenological *vector model* has been used in the *old quantum theory* by Landé [128] for *an addition of the magnetic moments* in many-electron atoms and molecules. We explain the Landé ideas in the framework of the Schrödinger–Pauli theory (see Sect. 10.2).

14.7.1 Precession of the Angular Momentum

We consider the orbital and spin angular momenta $\mathbf{L}$ and $\mathbf{s}$ introduced in (3.58) and (10.13) respectively.

Problem Verify that, in the uniform magnetic field $\mathbf{B} = (0, 0, B)$, the components $\mathbf{L}_3(t)$ and $\mathbf{s}_3(t)$ are conserved, while the vectors $(\mathbf{L}_1(t), \mathbf{L}_2(t)) \in \mathbb{R}^2$ and $(\mathbf{s}_1(t), \mathbf{s}_2(t)) \in \mathbb{R}^2$ rotate with angular velocity $\omega_{\mathcal{L}}$ and $2\omega_{\mathcal{L}}$, respectively, where $\omega_{\mathcal{L}} = -\frac{eB}{2\mathrm{m}c}$ is the Larmor frequency.

Solution For concreteness, we consider the orbital momentum. The conservation of $\mathbf{L}_3$ follows from (3.67), because the magnetic field $\mathbf{B} = (0, 0, B)$ is axially symmetric (see (3.66), (3.68)). Since $\hat{\mathbf{L}}_k = -\hbar\mathbf{H}_k$, from (6.39) we have

$$[\hat{\mathbf{L}}_1, \hat{\mathbf{L}}_2] = i\hbar\hat{\mathbf{L}}_3, \qquad [\hat{\mathbf{L}}_2, \hat{\mathbf{L}}_3] = i\hbar\hat{\mathbf{L}}_1, \qquad [\hat{\mathbf{L}}_3, \hat{\mathbf{L}}_1] = i\hbar\hat{\mathbf{L}}_2. \tag{14.59}$$

Therefore, (10.17) implies that

$$[\mathcal{P}, \hat{\mathbf{L}}_1] = -i\omega_{\mathcal{L}}\hbar\hat{\mathbf{L}}_2, \qquad [\mathcal{P}, \hat{\mathbf{L}}_2] = i\omega_{\mathcal{L}}\hbar\hat{\mathbf{L}}_1 \tag{14.60}$$

since each of $\hat{\mathbf{L}}_k$ commutes with any of $\hat{\mathbf{s}}_j$. Hence, by the Heisenberg equation (3.93),

$$\dot{\mathbf{L}}_1(t) = \langle \Psi(t), i\hbar^{-1}[\mathcal{P}\hat{\mathbf{L}}_1 - \hat{\mathbf{L}}_1\mathcal{P}]\Psi(t)\rangle = \langle \Psi(t), \omega_{\mathcal{L}}\hat{\mathbf{L}}_2\Psi(t)\rangle = \omega_{\mathcal{L}}\mathbf{L}_2(t). \tag{14.61}$$

Similarly, we have $\dot{\mathbf{L}}_2(t) = -\omega_{\mathcal{L}}\mathbf{L}_1(t)$, and hence

$$\frac{d}{dt}\big(\mathbf{L}_1(t) + i\mathbf{L}_2(t)\big) = -i\omega_{\mathcal{L}}\big(\mathbf{L}_1(t) + i\mathbf{L}_2(t)\big), \tag{14.62}$$

as required.

The result just proved implies the *precession* of the vectors $\mathbf{L}(t)$ and $\mathbf{s}(t)$ with angular velocities $\omega_{\mathcal{L}}$ and $2\omega_{\mathcal{L}}$, respectively, around the magnetic field $\mathbf{B}$.

Remark 14.1 A similar precession can be proved for a classical system of electrons rotating as rigid bodies in a uniform magnetic field [93].

14.7.2 The Vector Model

Let us apply the idea of precession to the addition of the orbital and spin magnetic moments. This addition explains the Einstein–de Haas and the anomalous Zeeman experiments.

The model explains the magnetization in the Einstein–de Haas experiment by the classical mechanism of reorientation in a magnetic field of a total magnetic moment $\mathbf{m}$, which exists even in the absence of the magnetic field. Namely, the total magnetic moment $\mathbf{m}$ is the sum of the orbital and spin magnetic moments $\mathbf{m}_o$ and $\mathbf{m}_s$:

$$\mathbf{m} = \mathbf{m}_o + \mathbf{m}_s = \frac{e}{2\mathrm{m}c}\mathbf{L} + \frac{e}{\mathrm{m}c}\mathbf{s}, \tag{14.63}$$

where the moments are defined by the orbital and spin angular momenta $\mathbf{L}$ and $\mathbf{s}$, respectively, with corresponding distinct gyromagnetic ratios $\frac{e}{2\mathrm{m}c}$ and $\frac{e}{\mathrm{m}c}$.

In absence of an external magnetic field, the total angular momentum $\mathbf{J} = \mathbf{L} + \mathbf{s}$ is conserved, while $\mathbf{L}$ and $\mathbf{s}$ are generally not conserved. For example, the spin angular momentum $\mathbf{s}$ should precess in the magnetic field generated by the orbital angular momentum. Similarly, the orbital angular momentum $\mathbf{L}$ precesses in the magnetic field generated by the spin angular momentum. Thus the lengths of the vectors $\mathbf{L}$ and $\mathbf{s}$ should be constant. Hence the conservation of the sum $\mathbf{J} = \mathbf{L} + \mathbf{s}$ implies that the vectors $\mathbf{L}$ and $\mathbf{s}$ rotate around $\mathbf{J}$. Then the total magnetic moment $\mathbf{m} = \frac{e}{2\mathrm{m}c}\mathbf{L} + \frac{e}{\mathrm{m}c}\mathbf{s}$ also rotates around $\mathbf{J}$. Since the angular velocity of the rotation is very high, we have to take into account only the effective value of the magnetic moment, $\mathbf{m}_{\mathrm{eff}}$, which is the projection of the total magnetic moment onto $\mathbf{J}$. Let us calculate this projection.

The angle α between the vectors $\mathbf{J}$ and $\mathbf{L}$ is conserved, as well as the angle β between the vectors $\mathbf{J}$ and $\mathbf{s}$, and so

$$\cos\alpha = \frac{|\mathbf{J}|^2 + |\mathbf{L}|^2 - |\mathbf{s}|^2}{2|\mathbf{J}||\mathbf{L}|}, \qquad \cos\beta = \frac{|\mathbf{J}|^2 + |\mathbf{s}|^2 - |\mathbf{L}|^2}{2|\mathbf{J}||\mathbf{s}|}. \tag{14.64}$$

Then the projection is as follows:

$$\mathbf{m}_{\mathrm{eff}} = \frac{e}{2\mathrm{m}c}|\mathbf{L}|\cos\alpha + \frac{e}{\mathrm{m}c}|\mathbf{s}|\cos\beta. \tag{14.65}$$

Now the effective gyromagnetic ratio reads

$$\begin{aligned} g_{\text{eff}} := \frac{\mathbf{m}_{\text{eff}}}{|\mathbf{J}|\frac{e}{2\text{m}c}} &= \frac{|\mathbf{J}|^2 + |\mathbf{L}|^2 - |\mathbf{s}|^2}{2|\mathbf{J}|^2} + 2\frac{|\mathbf{J}|^2 + |\mathbf{s}|^2 - |\mathbf{L}|^2}{2|\mathbf{J}|^2} \\ &= \frac{3}{2} + \frac{|\mathbf{s}|^2 - |\mathbf{L}|^2}{2|\mathbf{J}|^2}, \end{aligned} \tag{14.66}$$

which is the Landé formula [128].

A final novelty arises from the Bohr Correspondence Principle (see [20]): we redefine the lengths of the vectors **J**, **L** and **s** as follows:

$$\begin{aligned} &|\mathbf{J}|^2 := \langle \Psi, (\hat{\mathbf{J}}_1^2 + \hat{\mathbf{J}}_2^2 + \hat{\mathbf{J}}_3^2)\Psi \rangle, \qquad |\mathbf{L}|^2 := \langle \Psi, (\hat{\mathbf{L}}_1^2 + \hat{\mathbf{L}}_2^2 + \hat{\mathbf{L}}_3^2)\Psi \rangle, \\ &|\mathbf{s}|^2 := \langle \Psi, (\hat{\mathbf{s}}_1^2 + \hat{\mathbf{s}}_2^2 + \hat{\mathbf{s}}_3^2)\Psi \rangle. \end{aligned} \tag{14.67}$$

The operators $\hat{\mathbf{J}}_1^2 + \hat{\mathbf{J}}_2^2 + \hat{\mathbf{J}}_3^2$, $\hat{\mathbf{L}}_1^2 + \hat{\mathbf{L}}_2^2 + \hat{\mathbf{L}}_3^2$ and $\hat{\mathbf{s}}_1^2 + \hat{\mathbf{s}}_2^2 + \hat{\mathbf{s}}_3^2 = 3/4$ are known to commute. Hence, the quantum stationary states can be classified by the eigenvalues of the operators that are equal to $J(J+1)$, $L(L+1)$, and $3/4$, where $J, L = 0, 1, 2, \ldots$. For the states, we have $|\mathbf{J}|^2 = J(J+1)$, $|\mathbf{L}|^2 = L(L+1)$, and $|\mathbf{s}|^2 = 3/4$. Substituting this into (14.66), we obtain exactly formula (10.31):

$$g_{\text{eff}} = \frac{3}{2} + \frac{3/4 - L(L+1)}{2J(J+1)}. \tag{14.68}$$

This formula is confirmed experimentally by the Einstein–de Haas and the anomalous Zeeman effects.

14.8 Quantization of Harmonic Oscillator

Here we prove Proposition 2.1 on eigenvalues of the energy operator

$$\hat{E} = \frac{1}{2}\hat{p}^2 + \frac{1}{2}\omega^2\hat{x}^2 \tag{14.69}$$

corresponding to the Hamilton function of the harmonic oscillator $H = \frac{1}{2}p^2 + \frac{1}{2}\omega^2 x^2$. Here $\hat{x}$ and $\hat{p}$ satisfy the commutation relations (2.30):

$$[\hat{x}, \hat{p}] = i\hbar. \tag{14.70}$$

We should prove that eigenvalues of $\hat{E}$ are given by $E_n = \hbar\omega(n + \frac{1}{2})$, $n = 0, 1, \ldots$. For the proof, we factorize the Hamilton function as $H = \frac{1}{2}(p - i\omega x)(p + i\omega x)$, and respectively,

$$\hat{E} = \frac{1}{2}(\hat{p} + i\omega\hat{x})(\hat{p} - i\omega\hat{x}) + \frac{1}{2}\hbar\omega \tag{14.71}$$

since $[\hat{x}, \hat{p}] = i\hbar$. Equivalently,

$$\hat{E} = \hbar\omega\left(a^*a + \frac{1}{2}\right),$$

$$a := (\hat{p} - i\omega\hat{x})/\sqrt{2\hbar\omega}, \ \left[a, a^*\right] = \frac{-i\omega[\hat{x}, \hat{p}] + i\omega[\hat{p}, \hat{x}]}{2\hbar\omega} = 1, \tag{14.72}$$

because $\hat{x}^* = \hat{x}$ and $\hat{p}^* = \hat{p}$.

Problem Let e be an eigenvector of $\hat{E}$, i.e., $\hat{E}e = Ee$. Verify that

$$\hat{E}ae = (E - \hbar\omega)ae, \qquad \hat{E}a^*e = (E + \hbar\omega)a^*e. \tag{14.73}$$

Solution From the commutation relation (14.72) it follows that

$$\begin{aligned}\hat{E}ae &= \hbar\omega\left(a^*a + \frac{1}{2}\right)ae = \hbar\omega\left(aa^* - \frac{1}{2}\right)ae = a\hbar\omega\left(a^*a - \frac{1}{2}\right)e \\ &= a(\hat{E} - \hbar\omega)e = (E - \hbar\omega)ae, \\ \hat{E}a^*e &= \hbar\omega\left(a^*a + \frac{1}{2}\right)a^*e = a^*\hbar\omega\left(aa^* + \frac{1}{2}\right)e = a^*\hbar\omega\left(a^*a + \frac{3}{2}\right)e \\ &= a^*(\hat{E} + \hbar\omega)e = (E + \hbar\omega)a^*e,\end{aligned}$$

as required. The operators a, a^* are called, respectively, the ‘annihilation’ and ‘creation’ *ladder* operators. The final crucial observation is that $\frac{1}{2}\hat{E} \geq \hbar\omega$, giving $a^Ne = 0$ if $E - N\hbar\omega < \frac{1}{2}\hbar\omega$. Let $a^{n+1}e = 0$, while $a^ne \neq 0$. Hence

$$\begin{aligned}\left\langle a^ne, (E - n\hbar\omega)a^ne\right\rangle = \left\langle a^ne, \hat{E}a^ne\right\rangle &= \hbar\omega\left\langle a^{n+1}e, a^{n+1}e\right\rangle + \frac{1}{2}\hbar\omega\left\langle a^ne, a^ne\right\rangle \\ &= \frac{1}{2}\hbar\omega\left\langle a^ne, a^ne\right\rangle.\end{aligned}$$

As a result, $E - n\hbar\omega = \frac{1}{2}\hbar\omega$.

Let Ω be the unit eigenvector of H corresponding to the minimal eigenvalue $E_0 = \frac{1}{2}\hbar\omega$.

Problem Verify that $e_n := \frac{(a*)^n\Omega}{\sqrt{n!}}$ is the unit eigenvector of H corresponding to $E_n = \hbar\omega(n + \frac{1}{2})$, $n = 0, 1, \ldots$.

Solution It suffices to check that

$$\left\langle(a*)^n\Omega, (a*)^n\Omega\right\rangle = n! \tag{14.74}$$

First, we have

$$\left\langle(a*)^n\Omega, (a*)^n\Omega\right\rangle = \left\langle a^n(a*)^n\Omega, \Omega\right\rangle. \tag{14.75}$$

Second, from the commutation relation (14.72) it follows by induction that

$$a(a*)^n - (a*)^n a = n(a*)^{n-1}. \tag{14.76}$$

Indeed, multiplying from the left by a, we obtain the same identity (with $n+1$ instead of n). Therefore,

$$a(a*)^n \Omega = n(a*)^{n-1}\Omega \tag{14.77}$$

because $a\Omega = 0$. Hence,

$$\big\langle (a*)^n \Omega, (a*)^n \Omega \big\rangle = n \big\langle (a*)^{n-1}\Omega, (a*)^{n-1}\Omega \big\rangle = \cdots = n! \langle \Omega, \Omega \rangle = n!, \tag{14.78}$$

as required.

Bibliography

1. M. Abraham, Dynamik des Elektrons. Phys. Z. **4**, 57–63 (1902); G tt. Nachr., 20-41, 1902
2. M. Abraham, *Theorie der Elektrizität, Bd. 2: Elektromagnetische Theorie der Strahlung* (Teubner, Leipzig, 1905)
3. S. Agmon, Spectral properties of Schrödinger operator and scattering theory. Ann. Sc. Norm. Super. Pisa, Ser. IV **2**, 151–218 (1975)
4. V. Arnold, *Mathematical Methods of Classical Mechanics* (Springer, Berlin, 1978)
5. V.M. Babich, V.S. Buldirev, *Short-Wavelength Diffraction Theory. Asymptotic Methods* (Springer, Berlin, 1991)
6. V. Bach, F. Klopp, H. Zenk, Mathematical analysis of the photoelectric effect. Adv. Theor. Math. Phys. **5**(6), 969–999 (2001)
7. R. Becker, *Electromagnetic Fields and Interactions, Vol. I, II: Quantum Theory of Atoms and Radiation* (Blaisdell, Boston, 1964)
8. A. Bensoussan, C. Iliine, A. Komech, Breathers for a relativistic nonlinear wave equation. Arch. Ration. Mech. Anal. **165**, 317–345 (2002)
9. H. Berestycki, P.L. Lions, Arch. Ration. Mech. Anal. **82**(4), 313–375 (1983)
10. F.A. Berezin, M.A. Shubin, *The Schrödinger Equation* (Kluwer Academic, Dordrecht, 1991)
11. H. Bethe, *Intermediate Quantum Mechanics* (Benjamin, New York, 1964)
12. H. Bethe, E. Salpeter, *Quantum Mechanics of One- and Two-Electron Atoms* (Springer, Berlin, 1957)
13. J.D. Bjorken, S.D. Drell, *Relativistic Quantum Mechanics* (McGraw-Hill, New York, 1964)
14. J.D. Bjorken, S.D. Drell, *Relativistic Quantum Fields* (McGraw-Hill, New York, 1965)
15. N. Bohr, Philos. Mag. **26**, 1; 476; 857 (1913)
16. N. Bohr, K. Dan. Vidensk. Selsk. Skr., Nat.vidensk. Math. Afd. 8. Raekke **IV.1**, 1–3 (1918) [English translation: On the quantum theory of line spectra, pp. 95–138 in: Sources in Quantum Mechanics, ed. B.L. van der Waerden, North-Holland, Amsterdam, 1967]
17. L. Boltzman, *Vorlesungen über Gastheory* (Barth, Leipzig, 1896) (Part I), 1898 (Part II)
18. M. Born, P. Jordan, Zur Quantenmechanik. Z. Phys. **34**, 858–888 (1925) [English translation: On quantum mechanics, pp. 277–306 in: Sources in Quantum Mechanics, ed. B.L. van der Waerden, North-Holland, Amsterdam, 1967]
19. M. Born, W. Heisenberg, P. Jordan, Zur Quantenmechanik. Z. Phys. **34**, 858–888 (1925) [English translation: On quantum mechanics II, pp. 321–385 in: Sources in Quantum Mechanics, ed. B.L. van der Waerden, North-Holland, Amsterdam, 1967]
20. M. Born, *Atomic Physics* (Blackie, London, 1951)
21. A.B. Bhatia, P.C. Clemmow, D. Gabor, A.R. Stokes, A.M. Taylor, P.A. Wayman, W.L. Wilcock, in *Principles of Optics, Electromagnetic Theory of Propagation, Interference and Diffraction of Light*, ed. by M. Born, E. Wolf (Pergamon, London, 1959)

A. Komech, *Quantum Mechanics: Genesis and Achievements*,
DOI 10.1007/978-94-007-5542-0,

22. N. Bournaveas, Local existence for the Maxwell–Dirac equations in three space dimensions. Commun. Partial Differ. Equ. **21**(5–6), 693–720 (1996)
23. J.J. Brehm, W.J. Mullins, *Introduction to the Structure of Matter: A Course in Modern Physics* (Wiley, New York, 1989)
24. L. de Broglie, Ondes et quanta. C. R. Math. Acad. Sci. Paris **177**, 507–510 (1923)
25. L. de Broglie, Quanta de lumi‘re, diffraction et interférences. C. R. Math. Acad. Sci. Paris **177**, 548–550 (1923)
26. L. de Broglie, Les quanta, la théorie cinétique des gaz et le principe de Fermat. C. R. Math. Acad. Sci. Paris **177**, 630–632 (1923)
27. L. de Broglie, *Physicien et Penseur* (Editions Albin Michel, Paris, 1953)
28. V.S. Buslaev, G.S. Perelman, Scattering for the nonlinear Schrödinger equation: states close to a soliton. St. Petersburg Math. J. **4**, 1111–1142 (1993)
29. V.S. Buslaev, G.S. Perelman, On the stability of solitary waves for nonlinear Schrödinger equations. Am. Math. Soc. Trans. (2) **164**, 75–98 (1995)
30. V.S. Buslaev, C. Sulem, On asymptotic stability of solitary waves for nonlinear Schrödinger equations. Ann. Inst. Henri Poincaré, Anal. Non Linéaire **20**(3), 419–475 (2003)
31. L. Catto, C. Le Bris, P.L. Lions, *The Mathematical Theory of Thermodynamic Limits: Thomas-Fermi Type Models* (Clarendon, Oxford, 1998)
32. R.G. Chambers, Shift of an electron interference pattern by enclosed magnetic flux. Phys. Rev. Lett. **5**, 3–5 (1960)
33. S.J. Chang, *Introduction to Quantum Field Theory* (World Scientific, Singapore, 1990)
34. E.U. Condon, G.H. Shortley, *The Theory of Atomic Spectra* (Cambridge University Press, Cambridge, 1963)
35. O. Costin, R.D. Costin, J.L. Lebowitz, A. Rokhlenko, Evolution of a model quantum system under time periodic forcing: conditions for complete ionization. Commun. Math. Phys. **221**(1), 1–26 (2001)
36. O. Costin, R.D. Costin, J.L. Lebowitz, Time asymptotics of the Schrödinger wave function in time-periodic potentials. J. Stat. Phys. **116**(1–4), 283–310 (2004)
37. O. Costin, J.L. Lebowitz, C. Stucchio, Ionization in a 1-dimensional dipole model. Rev. Math. Phys. **20**(7), 835–872 (2008)
38. O. Costin, J.L. Lebowitz, C. Stucchio, S. Tanveer, Exact results for ionization of model atomic systems. J. Math. Phys. **51**(1), 015211 (2010)
39. C.G. Darwin, The wave equation of the electron. Proc. R. Soc. Lond. A **118**, 654–680 (1928)
40. C. Davisson, L. Germer, Nature **119**, 558 (1927)
41. M. Defranceschi, C. Le Bris, *Mathematical Models and Methods for Ab Initio Quantum Chemistry*. Lecture Notes in Chemistry, vol. 74 (Springer, Berlin, 2000)
42. P. Debye, Zur Theorie der spezifischen Waerme. Ann. Phys. (Leipz.) **39**(4), 789 (1912)
43. P.A.M. Dirac, Proc. R. Soc. A **112**, 661 (1926)
44. P.A.M. Dirac, The quantum theory of the emission and absorption of radiation. Proc. R. Soc. A **114**, 243 (1927)
45. P.A.M. Dirac, The quantum theory of the electron, I. Proc. R. Soc. Lond. A **117**, 610–624 (1928)
46. P.A.M. Dirac, *The Principles of Quantum Mechanics* (Oxford University Press, Oxford, 1999)
47. T. Dudnikova, A. Komech, E.A. Kopylova, Yu.M. Suhov, On convergence to equilibrium distribution, I: Klein–Gordon equation with mixing. Commun. Math. Phys. **225**(1), 1–32 (2002)
48. T. Dudnikova, A. Komech, N. Mauser, On the convergence to a statistical equilibrium for the Dirac equation. Russ. J. Math. Phys. **10**(4), 399–410 (2003)
49. T. Dudnikova, A. Komech, N. Mauser, On two-temperature problem for harmonic crystals. J. Stat. Phys. **114**(3/4), 1035–1083 (2004)
50. T. Dudnikova, A. Komech, N.E. Ratanov, Yu.M. Suhov, On convergence to equilibrium distribution, II: wave equation with mixing. J. Stat. Phys. **108**(4), 1219–1253 (2002)

51. T. Dudnikova, A. Komech, H. Spohn, On a two-temperature problem for wave equation with mixing. Markov Process. Relat. Fields **8**(1), 43–80 (2002)
52. T. Dudnikova, A. Komech, H. Spohn, On convergence to statistical equilibrium for harmonic crystals. J. Math. Phys. **44**(6), 2596–2620 (2003)
53. T. Dudnikova, A. Komech, H. Spohn, Energy-momentum relation for solitary waves of relativistic wave equation. Russ. J. Math. Phys. **9**(2), 153–160 (2002)
54. W. Eckhaus, A. van Harten, *The Inverse Scattering Transformation and the Theory of Solitons. An Introduction* (North-Holland, Amsterdam, 1981)
55. D.M. Eidus, The principle of limit amplitude. Russ. Math. Surv. **24**(3), 97–167 (1969)
56. D.M. Eidus, The limiting amplitude principle for the Schrödinger equation in domains with unbounded boundaries. Asymptot. Anal. **2**(2), 95–99 (1989)
57. A. Einstein, Die Plancksche Theorie der Strahlung und die Theorie der spezifischen Wärme. Ann. Phys. (4) **22**, 180–190 (1907) [English translation: Planck's theory and the theory of specific heat. CPAE, vol. 2, Doc. 38]
58. A. Einstein, Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt (Concerning the generation and transformation of light as seen from a heuristic point of view), in *Annalen der Physik*, March 18, 1905
59. A. Einstein, Planck's theory of radiation and the theory of specific heat. Ann. Phys. (4) **22**, 180–190 (1906)
60. A. Einstein, Quantentheorie der Strahlung (On the quantum theory of radiation). Phys. Z. **18**, 121–128 (1917)
61. M. Esteban, V. Georgiev, E. Sere, Stationary solutions of the Maxwell–Dirac and the Klein–Gordon–Dirac equations. Calc. Var. Partial Differ. Equ. **4**(3), 265–281 (1996)
62. M. Fedoriuk, *Asymptotic Methods for Partial Differential Equations*. Encyclopaedia of Mathematical Sciences, vol. 34 (Springer, Berlin, 1999)
63. R. Feynman, R.B. Leighton, M. Sands, *The Feynman Lectures on Physics, Vol. I, II, III* (Addison-Wesley, Reading, 1964)
64. R. Feynman, *Quantum Electrodynamics* (Addison-Wesley, Reading, 1998)
65. M. Fowler. http://galileo.phys.virginia.edu/classes/252/spectra.html
66. Y. Frenkel, Z. Phys. **37**, 43 (1926)
67. S. Frabboni, G.C. Gazzadi, G. Pozzi, Young's double-slit interference experiment with electrons. Am. J. Phys. **75**(11), 1053–1055 (2007)
68. J. Fröhlich, T.P. Tsai, H.T. Yau, On the point-particle (Newtonian) limit of the nonlinear Hartree equation. Commun. Math. Phys. **225**(2), 223–274 (2002)
69. M. Gardner, *The New Ambidextrous Universe. Symmetry and Asymmetry from Mirror Reflections to Superstrings* (Freeman, New York, 1990)
70. I. Gelfand, S. Fomin, *Calculus of Variations* (Dover, Mineola, 2000)
71. I.M. Gelfand, R.A. Minlos, Z.Ya. Shapiro, *Representations of the Rotation and Lorentz Groups and Their Applications* (Pergamon, Oxford, 1963)
72. C. Gérard, Scattering theory for Klein–Gordon equations with non-positive energy. arXiv:1101.2145
73. H. Goldstein, *Classical Mechanics* (Addison-Wesley, Reading, 1980)
74. W. Gordon, Die Energieniveaus der Wassenstoffatoms nach der Diracschen Quantentheorie des Elektrons. Z. Phys. **48**, 11–14 (1928)
75. W. Greiner, *Quantum Mechanics: An Introduction* (Springer, Berlin, 2001)
76. M. Griesemer, H. Zenk, On the atomic photoeffect in non-relativistic QED. arXiv:0910.1809
77. F. Gross, *Relativistic Quantum Mechanics and Field Theory* (Wiley, New York, 1999)
78. V. Guillemin, S. Sternberg, *Geometric Asymptotics* (American Mathematical Society, Providence, 1977)
79. S.J. Gustafson, I.M. Sigal, *Mathematical Concepts of Quantum Mechanics* (Springer, Berlin, 2003)
80. Y. Guo, K. Nakamitsu, W. Strauss, Global finite-energy solutions of the Maxwell–Schrödinger system. Commun. Math. Phys. **170**(1), 181–196 (1995)
81. K. Hannabus, *An Introduction to Quantum Theory* (Clarendon, Oxford, 1997)

82. W. Heisenberg, Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen. Z. Phys. **33**, 879–893 (1925) [English translation: Quantum-mechanical reinterpretation of kinematic and mechanical relations, pp. 261–276 in: Sources in Quantum Mechanics, ed. B.L. van der Waerden, North-Holland, Amsterdam, 1967]
83. W. Heisenberg, Der derzeitige Stand der nichtlinearen Spinortheorie der Elementarteilchen. Acta Phys. Austriaca **14**, 328–339 (1961)
84. W. Heisenberg, *Introduction to the Unified Field Theory of Elementary Particles* (Interscience, London, 1966)
85. W. Heitler, *The Quantum Theory of Radiation* (Oxford University Press, Oxford, 1954)
86. E. Hellinger, Neue Begründung der Theorie quadratischer Formen von unendlichvielen Veränderlichen. J. Math. **136**, 210–271 (1910)
87. H. Hertz, *Electric Waves: Being Researches on the Propagation of Electric Action with Finite Velocity Through Space* (Dover, New York, 1893)
88. D. Hilbert, *Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen* (Teubner, Leipzig, 1912)
89. P.D. Hislop, I.M. Sigal, *Introduction to Spectral Theory. With Applications to Schrödinger Operators* (Springer, New York, 1996)
90. L. Houllevigue, *L'Évolutlon des Sciences* (Collin, Paris, 1908)
91. V. Imaikin, A. Komech, N. Mauser, Soliton-type asymptotics for the coupled Maxwell–Lorentz equations. Ann. Inst. Poincaré, Phys. Theor. **5**, 1117–1135 (2004)
92. C. Itzykson, J.B. Zuber, *Quantum Field Theory* (McGraw-Hill, New York, 1980)
93. J.D. Jackson, *Classical Electrodynamics*, 3rd edn. (Wiley, New York, 1999)
94. V. Jaksic, C.-A. Pillet, Ergodic properties of classical dissipative systems, I. Acta Math. **181**, 245–282 (1998)
95. M. Jammer, *The Conceptual Development of Quantum Mechanics* (McGraw-Hill, New York, 1966)
96. C. Jönsson, Electron diffraction at multiple slits. Am. J. Phys. **42**, 4–11 (1974)
97. A. Jensen, T. Kato, Spectral properties of Schrödinger operators and time-decay of the wave functions. Duke Math. J. **46**, 583–611 (1979)
98. W. Kauffmann, Ann. Phys. **61**, 544 (1897)
99. G. Kirchhoff, Poggendorff's Ann. **59**, 275 (1860) [English translation: On the relation between the radiating and absorbing powers of different bodies for light and heat, Philos. Mag. Ser. 4 **20**(130), 1–21 (1860)]
100. A. Komech, On stabilization of string-nonlinear oscillator interaction. J. Math. Anal. Appl. **196**, 384–409 (1995)
101. A. Komech, On transitions to stationary states in Hamiltonian nonlinear wave equations. Phys. Lett. A **241**, 311–322 (1998)
102. A. Komech, On transitions to stationary states in one-dimensional nonlinear wave equations. Arch. Ration. Mech. Anal. **149**(3), 213–228 (1999)
103. A. Komech, Attractors of nonlinear Hamiltonian one-dimensional wave equations. Russ. Math. Surv. **55**(1), 43–92 (2000)
104. A.I. Komech, A.A. Komech, On global attraction to solitary waves for the Klein–Gordon equation coupled to nonlinear oscillator. C. R. Math. Acad. Sci. Paris **343**(2), 111–114 (2006)
105. A.I. Komech, A.A. Komech, Global attractor for a nonlinear oscillator coupled to the Klein–Gordon field. Arch. Ration. Mech. Anal. **185**, 105–142 (2007)
106. A.I. Komech, A.A. Komech, Global attraction to solitary waves in models based on the Klein–Gordon equation SIGMA, symmetry integrability. Geom. Methods Appl. **4**, 010 (2008). Electronic only. http://www.emis.de/journals/SIGMA/2008/, arXiv:0711.0041
107. A.I. Komech, A.A. Komech, Global attraction to solitary waves for Klein–Gordon equation with mean field interaction. Ann. Inst. Henri Poincaré, Anal. Non Linéaire **26**(3), 855–868 (2009)
108. A.I. Komech, A.A. Komech, On global attraction to solitary waves for the Klein–Gordon field coupled to several nonlinear oscillators. J. Math. Pures Appl. **93**, 91–111 (2010). arXiv:

math/0702660

109. A. Komech, *On Global Attractors of Hamilton Nonlinear Wave Equations*. Lecture Notes of the Max Planck Institute for Mathematics in the Sciences, vol. 24 (2005). http://www.mis.mpg.de/preprints/ln/lecturenote-2405-abstr.html
110. A. Komech, H. Spohn, Soliton-like asymptotics for a classical particle interacting with a scalar wave field. Nonlinear Anal. **33**(1), 13–24 (1998)
111. A. Komech, H. Spohn, Long-time asymptotics for the coupled Maxwell–Lorentz equations. Commun. Partial Differ. Equ. **25**, 558–585 (2000)
112. A. Komech, H. Spohn, M. Kunze, Long-time asymptotics for a classical particle interacting with a scalar wave field. Commun. Partial Differ. Equ. **22**, 307–335 (1997)
113. A. Komech, Lectures on classical fields and quantum mechanics, part I. Research Report INRIA No. RR-4736, INRIA Rocquencourt, 2003. http://hal.inria.fr/inria-00071851/fr/
114. A. Komech, *Lectures on Quantum Mechanics. Nonlinear PDE Point of View*. Lecture Notes of the Max Planck Institute for Mathematics in the Sciences, vol. 25/2005 (2005). http://www.mis.mpg.de/preprints/ln/lecturenote-2505-abstr.html, arXiv:math-ph/0505059
115. A. Komech, N. Mauser, A. Vinnichenko, On attraction to solitons in relativistic nonlinear wave equations. Russ. J. Math. Phys. **11**(3), 289–307 (2004)
116. A. Komech, E. Kopylova, *Dispersion Decay* (Wiley, New York, 2012). http://www.mat.univie.ac.at/~komech/
117. A. Komech, E. Kopylova, On asymptotic stability of moving kink for relativistic Ginzburg–Landau equation. Commun. Math. Phys. **302**(1), 225–252 (2011). arXiv:0910.5538
118. A. Komech, E. Kopylova, On asymptotic stability of kink for relativistic Ginzburg–Landau equation. Arch. Rat. Mech. Anal. **202**, 213–245 (2011). arXiv:0910.5539
119. A. Komech, E. Kopylova, *Dispersion Decay and Scattering Theory* (Wiley, Hoboken, 2012)
120. A. Komech, M. Kunze, H. Spohn, Effective dynamics for a mechanical particle coupled to a wave field. Commun. Math. Phys. **203**, 1–19 (1999)
121. H.A. Kramers, The quantum theory of dispersion. Nature **113**, 673–676 (1924) [pp. 177–180 in: Sources in Quantum Mechanics, ed. B.L. van der Waerden, North-Holland, Amsterdam, 1967]
122. H.A. Kramers, The law of dispersion and Bohr's theory of spectra. Nature **114**, 310–311 (1924) [pp. 199–202 in: Sources in Quantum Mechanics, ed. B.L. van der Waerden, North-Holland, Amsterdam, 1967]
123. H.A. Kramers, W. Heisenberg, Über die Streuung von Strahlen durch Atome. Z. Phys. **31**, 681–708 (1925) [English translation: On the dispersion of radiation by atoms, pp. 223–252 in: Sources in Quantum Mechanics, ed. B.L. van der Waerden, North-Holland, Amsterdam, 1967]
124. H.A. Kramers, in *La diffusion de la lumiere par les atomes, Atti Cong. Intern. Fisica (Transactions of Volta Centenary Congress)* (1927), pp. 545–557
125. R.L. Kronig, On the theory of the dispersion of X-rays. J. Opt. Soc. Am. **12**, 547–557 (1926). doi:10.1364/JOSA.12.000547
126. http://en.wikipedia.org/wiki/Kramers-Kronig_relation
127. R. Ladenburg, Die quantentheoretische Zahl der Dispersionelektronen. Z. Phys. **4**, 451–468 (1921) [English translation: The quantum-theoretical interpretation of the number of dispersion electrons, pp. 139–158 in: Sources in Quantum Mechanics, ed. B.L. van der Waerden, North-Holland, Amsterdam, 1967]
128. A. Landé, Z. Phys. **5**, 5 (1921)
129. C. Lanczos, *The Variational Principles of Mechanics* (University of Toronto Press, Toronto, 1970)
130. L.D. Landau, E.M. Lifshitz, *Quantum Mechanics. Non-Relativistic Theory* (Pergamon, London, 1965)
131. D.F. Lawden, *The Mathematical Principles of Quantum Mechanics* (Methuen, London, 1967)
132. P.D. Lax, Asymptotic solutions of oscillatory initial value problems. Duke Math. J. **24**, 627–646 (1957)

133. V.G. Levich, Yu.A. Vdovin, V.A. Mamlin, *Course of Theoretical Physics*, vol. II (Nauka, Moscow, 1971). In Russian
134. E.H. Lieb, R. Seiringer, *The Stability of Matter* (Cambridge University Press, Cambridge, 2010)
135. E. Long, D. Stuart, Effective dynamics for solitons in the nonlinear Klein–Gordon–Maxwell system and the Lorentz force law. Rev. Math. Phys. **21**(4), 459–510 (2009)
136. H.A. Lorentz, The theory of electrons and its applications to the phenomena of light and radiant heat. Lectures from a course held at Columbia University, New York, NY, USA, March and April 1906, Repr. of the 2nd ed., Sceaux: Éditions Jacques Gabay (1992)
137. M. Maggiore, *A Modern Introduction to Quantum Field Theory* (Oxford University Press, Oxford, 2005)
138. F. Mandl, G. Shaw, *Quantum Field Theory* (Wiley, Chichester, 1999)
139. http://www.discoveriesinmedicine.com/Bar-Cod/Cathode-Ray-Tube-CRT.html
140. J. Miller, M. Weinstein, Asymptotic stability of solitary waves for the regularized long-wave equation. Commun. Pure Appl. Math. **49**(4), 399–441 (1996)
141. http://en.wikipedia.org/wiki/Oil_drop_experiment, http://library.thinkquest.org/19662/low/eng/exp-millikan.html
142. J. von Neumann, Allgemeine Eigenwerttheorie Hermitescher Funktionaloparatoren. Math. Ann. **102**, 49–131 (1929)
143. J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton, 1955)
144. R. Newburgh, J. Peidle, W. Rueckner, Einstein, Perrin, and the reality of atoms: 1905 revisited. Am. J. Phys. **74**(6), 478–481 (2006). www.ias.u-psud.fr/medoc-OLD/BOUDINE/presentations/EPA%205.doc
145. R. Newton, *Quantum Physics* (Springer, New York, 2002)
146. E. Noether, *Gesammelte Abhandlungen. Collected Papers* (Springer, Berlin, 1983)
147. W. Pauli, Quantentheorie, in *Handbuch der Physik*, vol. 23, ed. by H. Geiger, K. Scheel (Springer, Berlin, 1926), pp. 1–278
148. W. Pauli, Über das Wasserstoffspektrum vom Standpunkt der neuen Quantenmechanik. Z. Phys. **36**, 336–363 (1926) [English translation: On the hydrogen spectrum from the standpoint of the new quantum mechanics, pp. 387–427 in: Sources in Quantum Mechanics, ed. B.L. van der Waerden, North-Holland, Amsterdam, 1967]
149. W. Pauli, Diracs Wellengleichungen des Elektrons und geometrische Optik. Helv. Phys. Acta **5**, 179–199 (1932)
150. R.L. Pego, M.I. Weinstein, On asymptotic stability of solitary waves. Phys. Lett. A **162**, 263–268 (1992)
151. R.L. Pego, M.I. Weinstein, Asymptotic stability of solitary waves. Commun. Math. Phys. **164**, 305–349 (1994)
152. J. Perrin, New experiments on the cathode rays. Nature **53**, 298–299 (1896). http://www.chemteam.info/Chem-History/Perrin-1895.html
153. M. Peshkin, A. Tonomura, *The Aharonov–Bohm's Effect* (Springer, New York, 1989)
154. M. Planck, Über eine Verbesserung der wienschen Spektralgleichung. Verh. Dtsch. Phys. Ges. **2**, 202–204 (1967) [English translation: On an improvement of Wien's equation for the spectrum, pp. 79–81 in: D. ter-Haar, The Old Quantum Theory, Pergamon, Oxford, 1967]. http://www.ffn.ub.es/luisnavarro/nuevo_maletin/Planck%20%281900%29,%20Improvement%20of%20Wien%27s.pdf
155. M. Planck, *The Theory of Heat Radiation* (Blakiston's, London, 1914)
156. M. Reed, B. Simon, *Methods of Modern Mathematical Physics* (Academic Press, New York, 1980), II (1975), III (1979), IV (1978)
157. A. Rokhlenko, O. Costin, J.L. Lebowitz, Decay versus survival of a localized state subjected to harmonic forcing: exact results. J. Phys. A, Math. Gen. **35**(42), 8943–8951 (2002)
158. J.J. Sakurai, *Advanced Quantum Mechanics* (Addison-Wesley, Reading, 1967)
159. G. Scharf, *Finite Quantum Electrodynamics. The Causal Approach* (Springer, Berlin, 1995)
160. L. Schiff, *Quantum Mechanics* (McGraw-Hill, New York, 1955)

161. E. Schrödinger, Quantisierung als Eigenwertproblem, I. Ann. Phys. **79**, 361 (1926); II **79**, 489; III **80**, 437 (1926); IV **81**, 109 (1926)
162. E. Schrödinger, Über das Verhältnis der Heisenberg–Born–Jordanschen Quantenmechanik zu der Meinen. Ann. Phys. **79**, 734–756 (1926)
163. F. Schwabl, *Advanced Quantum Mechanics* (Springer, Berlin, 1997)
164. M.A. Shubin, *Pseudodifferential Operators and Spectral Theory* (Springer, Berlin, 2001)
165. W.H. Slabaugh, Determination of Avogadro's number by Perrin's law. J. Chem. Educ. **42**(9), 471–472 (1965). http://scitation.aip.org/journals/doc/AJPIAS-ft/vol_74/iss_6/478_1.html
166. A. Soffer, M.I. Weinstein, Multichannel nonlinear scattering in nonintegrable systems. Commun. Math. Phys. **133**, 119–146 (1990)
167. A. Soffer, M.I. Weinstein, Multichannel nonlinear scattering and stability, II: the case of anisotropic and potential and data. J. Differ. Equ. **98**, 376–390 (1992)
168. A. Soffer, M.I. Weinstein, Resonances, radiation damping and instability in Hamiltonian nonlinear wave equations. Invent. Math. **136**(1), 9–74 (1999)
169. A. Soffer, M.I. Weinstein, Selection of the ground state for nonlinear Schrödinger equations. Rev. Math. Phys. **16**(8), 977–1071 (2004)
170. A. Soffer, M.I. Weinstein, Theory of nonlinear dispersive waves and selection of the ground state. Phys. Rev. Lett. **95**, 213905 (2005)
171. A. Sommerfeld, *Atombau und Spektrallinien, Vol. I and II* (Vieweg, Braunschweig, 1951)
172. A. Sommerfeld, *Thermodynamics and Statistical Mechanics*. Lectures on Theoretical Physics, vol. V (Academic Press, New York, 1956)
173. A. Sommerfeld, *Optics* (Academic Press, New York, 1954)
174. A. Sommerfeld, G. Schur, Ann. Phys. **4**, 409 (1930)
175. A. Sommerfeld, H. Bethe, *Elektronentheorie der Metalle* (Springer, Berlin, 1967)
176. H. Spohn, *Dynamics of Charged Particles and Their Radiation Field* (Cambridge University Press, Cambridge, 2004)
177. N. Straumann, On the first solvay congress in 1911. arXiv:1109.3785
178. A. Szabo, N.S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory* (Dover, New York, 1996)
179. B. Thaller, *The Dirac Equation* (Springer, Berlin, 1991)
180. L.H. Thomas, Nature **117**, 514 (1926)
181. G.P. Thomson, The diffraction of cathode rays by thin films of platinum. Nature **120**, 802 (1927)
182. J.J. Thomson, Philos. Mag. **44**, 298 (1897)
183. http://www.youtube.com/watch?v=IdTxGJjA4Jw
184. A. Tonomura, J. Endo, T. Matsuda, T. Kawasaki, H. Ezawa, Demonstration of single-electron buildup of an interference pattern. Am. J. Phys. **57**(2), 117–120 (1989)
185. J. VanderLinde, *Classical Electromagnetic Theory, Fundamental Theories of Physics*, vol. 145 (Kluwer Academic, Dordrecht, 2004)
186. B.R. Vainberg, *Asymptotic Methods in Equations of Mathematical Physics* (Gordon and Breach, New York, 1989)
187. B.L. van der Waerden, *Group Theory and Quantum Mechanics* (Springer, New York, 1974)
188. B.L. van der Waerden (ed.), *Sources in Quantum Mechanics* (North-Holland, Amsterdam, 1967)
189. S. Weinberg, *The Quantum Theory of Fields. Vol. 1. Foundations* (Cambridge University Press, Cambridge, 2005)
190. G. Wentzel, Z. Phys. **43**(1), 779 (1927)
191. H. Weyl, *The Theory of Groups and Quantum Mechanics* (Dover, New York, 1949)
192. E. Whittaker, *A History of the Theories of Aether and Electricity*, vol. 1 (Nelson, London, 1951)
193. http://www.physics.rutgers.edu/ugrad/387/zeeman.pdf, hyperphysics.phy-astr.gsu.edu/hbase/quantum/zeeman.html, http://www.thphys.de/docs/F44_Normal_Zeeman_Effect.pdf
194. E. Zeidler, *Applied Functional Analysis. Main Principles and Their Applications*. Applied Mathematical Sciences, vol. 109 (Springer, Berlin, 1995)

195. E. Zeidler, *Quantum Field Theory. I: Basics in Mathematics and Physics. A Bridge Between Mathematicians and Physicists* (Springer, Berlin, 2006)
196. E. Zeidler, *Quantum Field Theory. II: Quantum Electrodynamics. A Bridge Between Mathematicians and Physicists* (Springer, Berlin, 2009)
197. H. Zenk, Ionisation by quantised electromagnetic fields: the photoelectric effect. Rev. Math. Phys. **20**, 367–406 (2008)

Index

A. Komech, *Quantum Mechanics: Genesis and Achievements*,
DOI 10.1007/978-94-007-5542-0,