Stony Brook University
Academic Commons

Essential Graduate Physics | Department of Physics and Astronomy

2024

Part QM: Quantum Mechanics

Konstantin Likharev
SUNY Stony Brook, konstantin.likharev@stonybrook.edu

Follow this and additional works at: https://commons.library.stonybrook.edu/egp

 Part of the Physics Commons

Recommended Citation

Likharev, Konstantin, "Part QM: Quantum Mechanics" (2024). *Essential Graduate Physics*. 4.
https://commons.library.stonybrook.edu/egp/4

This Book is brought to you for free and open access by the Department of Physics and Astronomy at Academic Commons. It has been accepted for inclusion in Essential Graduate Physics by an authorized administrator of Academic Commons. For more information, please contact mona.ramonetti@stonybrook.edu, hu.wang.2@stonybrook.edu.



Konstantin K. Likharev

Essential Graduate Physics

Lecture Notes and Problems

Open online access at mirror sites

http://commons.library.stonybrook.edu/egp/

https://essentialgraduatephysics.org/

https://sites.google.com/site/likharevegp/

under the Creative Commons http://creativecommons.org/licenses/by-nc-sa/4.0/ license

Part QM:
Quantum Mechanics

Last edit: July 2, 2024

B/W paperback copies of this volume are also available on *Amazon.com*:
https://www.amazon.com/gp/product/B0D7S1GXDG

About the author:
https://you.stonybrook.edu/likharev/

© K. Likharev

Table of Contents

Chapter 1. Introduction (26 pp.)
- 1.1. Experimental motivations
- 1.2. Wave mechanics postulates
- 1.3. Postulates' discussion
- 1.4. Continuity equation
- 1.5. Eigenstates and eigenvalues
- 1.6. Time evolution
- 1.7. Space dependence
- 1.8. Exercise problems (18)

Chapter 2. 1D wave mechanics (76 pp.)
- 2.1. Basic relations
- 2.2. Free particle: Wave packets
- 2.3. Particle reflection and tunneling
- 2.4. Motion in soft potentials
- 2.5. Resonant tunneling, and metastable states
- 2.6. Localized state coupling, and quantum oscillations
- 2.7. Periodic systems: Energy bands and gaps
- 2.8. Periodic systems: Particle dynamics
- 2.9. Harmonic oscillator: Brute force approach
- 2.10. Exercise problems (47)

Chapter 3. Higher Dimensionality Effects (64 pp.)
- 3.1. Quantum interference and the AB effect
- 3.2. Landau levels and the quantum Hall effect
- 3.3. Scattering and diffraction
- 3.4. Energy bands in higher dimensions
- 3.5. Axially symmetric systems
- 3.6. Spherically symmetric systems: Brute force approach
- 3.7. Atoms
- 3.8. Spherically symmetric scatterers
- 3.9. Exercise problems (49)

Chapter 4. Bra-ket Formalism (52 pp.)
- 4.1. Motivation
- 4.2. States, state vectors, and linear operators
- 4.3. State basis and matrix representation
- 4.4. Change of basis, and matrix diagonalization
- 4.5. Observables: Expectation values and uncertainties
- 4.6. Quantum dynamics: Three pictures
- 4.7. Coordinate and momentum representations
- 4.8. Exercise problems (34)

Chapter 5. Some Exactly Solvable Problems (48 pp.)
5.1. Two-level systems
5.2. The Ehrenfest theorem
5.3. The Feynman path integral
5.4. Revisiting harmonic oscillator
5.5. Glauber states and squeezed states
5.6. Revisiting spherically-symmetric problems
5.7. Spin and its addition to orbital angular momentum
5.8. Exercise problems (55)

Chapter 6. Perturbative Approaches (36 pp.)
6.1. Time-independent perturbations
6.2. The linear Stark effect
6.3. Fine structure of atomic levels
6.4. The Zeeman effect
6.5. Time-dependent perturbations
6.6. Quantum-mechanical Golden Rule
6.7. Golden Rule for step-like perturbations
6.8. Exercise problems (31)

Chapter 7. Open Quantum Systems (54 pp.)
7.1. Open systems, and the density matrix
7.2. Coordinate representation and the Wigner function
7.3. Open system dynamics: Dephasing
7.4. Fluctuation-dissipation theorem
7.5. The Heisenberg-Langevin approach
7.6. Density matrix approach
7.7. Application to two-level systems
7.8. Damped harmonic oscillator
7.9. Continuous-spectrum systems
7.10. Exercise problems (17)

Chapter 8. Multiparticle Systems (52 pp.)
8.1. Distinguishable and indistinguishable particles
8.2. Singlets, triplets, and the exchange interaction
8.3. Multiparticle systems
8.4. Perturbative approaches
8.5. Quantum computation and cryptography
8.6. Exercise problems (35)

Chapter 9. Elements of Relativistic Quantum Mechanics (36 pp.)
9.1. Electromagnetic field quantization
9.2. Photon absorption and counting
9.3. Photon emission: spontaneous and stimulated
9.4. Cavity QED

9.5. The Klien-Gordon and relativistic Schrödinger equations
9.6. Dirac's theory
9.7. Low energy limit
9.8. Exercise problems (22)

Chapter 10. Making Sense of Quantum Mechanics (16 pp.)
10.1. Quantum measurements
10.2. QND measurements
10.3. Hidden variables, the Bell theorem, and local reality
10.4. Interpretations of quantum mechanics
10.5. Exercise problem (1)

* * *

Supplemental file **Exercise Problems with Model Solutions** (311 problems, 552 pp.)
is available online:
https://essentialgraduatephysics.org/Files/QM%20exercises.pdf .

B/W paperback copies of these materials are available on *Amazon.com*:
https://www.amazon.com/gp/product/B0D7S393GS .

Additional file **Test Problems with Model Solutions** (68 problems, 61 pp.)
is available for course instructors from the author upon request – see *Front Matter*.

* * *

Introductory Remarks

The structure of this course is more or less traditional for graduate physics education, with most attention paid to the non-relativistic quantum mechanics, and only Chapter 9 reviewing the relativistic effects.

As in many (though not all) textbooks on this level, the discussion of Dirac's bra-ket formalism is postponed until after the discussion of numerous quantum-mechanical effects in Chapters 1-3 by using the conceptually simpler wave-mechanics approach. One reason for that decision was the author's serious commitment to the *Occam Razor* principle, in particular to the analysis of each physical effect by using the simplest suitable theoretical tools.

A really distinguishing feature of the course is Chapter 7 on open quantum systems, with a focus on the decoherence ('*dephasing*') and energy dissipation ('*relaxation*') effects. These effects are frequently discussed in statistical physics courses, but their understanding is necessary for any informed discussion of quantum measurements and quantum effects in macroscopic systems, with their substantial coupling to the environment.

Chapter 1. Introduction

This introductory chapter briefly reviews the major experimental motivations for quantum mechanics and then discusses its simplest formalism – Schrödinger's wave mechanics. Much of this material may be found in undergraduate textbooks,[1] *so the discussion is rather brief and focused on the most important conceptual issues.*

1.1. Experimental motivations

By the beginning of the 1900s, physics (which by that time included what we now call non-relativistic classical mechanics, classical thermodynamics and statistics, and classical electrodynamics including geometric and wave optics) looked an almost completed discipline, with most human-scale phenomena reasonably explained, and just a couple of mysterious "dark clouds"[2] on the horizon. However, rapid technological progress and the resulting development of more refined scientific instruments have led to a fast multiplication of observed phenomena that could not be explained on a classical basis. Let me list the most consequential of those experimental findings.

(i) The blackbody radiation measurements, pioneered by G. Kirchhoff in 1859, have shown that in thermal equilibrium, the power of electromagnetic radiation by a fully absorbing ("black") surface, per a unit frequency interval, drops exponentially at high frequencies. This is not what could be expected from the combination of classical electrodynamics and statistics, which predicted an infinite growth of the radiation density with frequency.

Indeed, classical electrodynamics shows[3] that electromagnetic field modes evolve in time just as harmonic oscillators, and that the number dN of these modes in a relatively large free-space volume $V >> \lambda^3$, within a small frequency interval $d\omega << \omega$ near some frequency ω, is

$$dN = 2V\frac{d^3k}{(2\pi)^3} = 2V\frac{4\pi k^2 dk}{(2\pi)^3} = V\frac{\omega^2}{\pi^2 c^3}d\omega, \tag{1.1}$$

where $c \approx 3\times10^8$ m/s is the free-space speed of light, $k = \omega/c$ is the free-space wave number, and $\lambda = 2\pi/k$ is the radiation wavelength. On the other hand, classical statistics[4] predicts that in thermal equilibrium at temperature T, the average energy E of each 1D harmonic oscillator should be equal to $k_B T$, where k_B is the Boltzmann constant.[5] Combining these two results, we readily get the so-called *Rayleigh-Jeans formula* for the average electromagnetic wave energy per unit volume:

[1] See, for example, S. Gasiorowicz, *Quantum Physics*, 3rd ed., Wiley, 2003; D. Griffith, *Quantum Mechanics*, 2nd ed., Cambridge U. Press, 2016.

[2] This famous expression was used in a 1900 talk by Lord Kelvin (born William Thomson), in reference to the results of blackbody radiation measurements and the Michelson-Morley experiments, i.e. the precursors of quantum mechanics and relativity theory.

[3] See, e.g., EM Sec. 7.8, in particular Eq. (7.211).

[4] See, e.g., SM Sec. 2.2.

[5] In the SI units, used throughout this series, $k_B \approx 1.38\times10^{-23}$ J/K – see *Appendix UCA: Selected Units and Constants* for the exact value.

© K. Likharev

$$u \equiv \frac{1}{V}\frac{dE}{d\omega} = \frac{k_B T}{V}\frac{dN}{d\omega} = \frac{\omega^2}{\pi^2 c^3} k_B T\,, \tag{1.2}$$

which diverges at $\omega \to \infty$ (Fig. 1) – the so-called *ultraviolet catastrophe*. On the other hand, the blackbody radiation measurements, improved by O. Lummer and E. Pringsheim, and by H. Rubens and F. Kurlbaum to reach a 1%-scale accuracy, were compatible with the law suggested in 1900 by Max Planck:

$$u = \frac{\omega^2}{\pi^2 c^3}\frac{\hbar\omega}{\exp\{\hbar\omega / k_B T\} - 1}\,. \tag{1.3a}$$

Planck's radiation law

This law may be reconciled with the fundamental Eq. (1) if the following replacement is made for the average energy of each field oscillator:

$$k_B T \to \frac{\hbar\omega}{\exp(\hbar\omega / k_B T) - 1}\,, \tag{1.3b}$$

with the factor

$$\hbar \approx 1.055 \times 10^{-34}\ \text{J}\cdot\text{s}\,, \tag{1.4}$$

Planck's constant

now called *Planck's constant*.[6] At low frequencies ($\hbar\omega << k_B T$), the denominator in Eqs. (3) may be approximated as $\hbar\omega/k_B T$, so the average energy (3b) tends to its classical value $k_B T$, and the Planck law (3a) reduces to the Rayleigh-Jeans formula (2). However, at higher frequencies ($\hbar\omega >> k_B T$), Eq. (3) describes the experimentally observed rapid decrease of the radiation density – see Fig. 1.

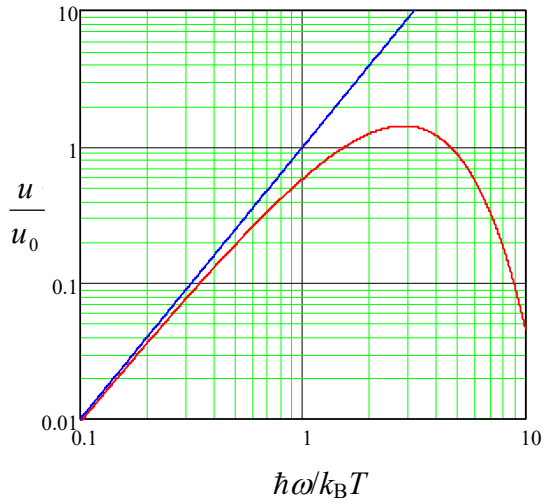


Fig. 1.1. The blackbody radiation density u, in units of $u_0 \equiv (k_B T)^3/\pi^2\hbar^2 c^3$, as a function of frequency, according to the Rayleigh-Jeans formula (blue line) and Planck's law (red line).

M. Plank derived Eq. (3b) from basic statistics, by assuming (just to fit the experimental results for u) that the energy of a field oscillator of frequency ω can only take values that differ by

$$E = \hbar\omega\,. \tag{1.5}$$

Energy vs frequency

[6] M. Planck himself wrote $\hbar\omega$ as $h\nu$, where $\nu = \omega/2\pi$ is the "cyclic" frequency, so in early texts on quantum mechanics the term "Planck's constant" referred to $h \equiv 2\pi\hbar$, while $\hbar$ was called "the Dirac constant" for a while. I will use contemporary terminology and abstain from using the constant h at all, to avoid confusion.

(ii) The photoelectric effect, discovered in 1887 by H. Hertz and studied quantitatively by A. Stoletov in 1888-89, shows a sharp lower boundary ("*red border*") for the frequency of the incident light that may kick electrons out from metallic surfaces, independent of its intensity. Albert Einstein, in one of his three famous 1905 papers, noticed that this threshold ω_{min} could be explained by assuming that light consisted of certain particles (later called *photons*) with the same energy (5).[7] Indeed, with this assumption, at the photon absorption by an electron, its energy $E = \hbar\omega$ is divided between a fixed energy U_0 (nowadays called the *workfunction*) of the electron's binding inside the metal, and the excess kinetic energy $m_e v^2/2 > 0$ of the freed electron – see Fig. 2. In this picture, the red border finds a natural explanation as $\omega_{min} = U_0/\hbar$.[8]

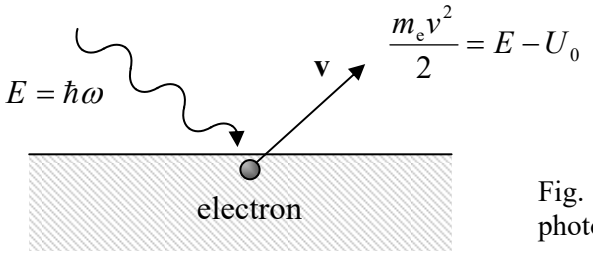


Fig. 1.2. Einstein's explanation of the photoelectric effect's frequency threshold.

(iii) The discrete frequency spectra of the electromagnetic radiation by excited atomic gases could not be explained by classical physics. (Applied to the planetary model of atoms, proposed by Ernst Rutherford, classical electrodynamics predicts the collapse of electrons on nuclei in ~10^{-10}s, due to the electric-dipole radiation of electromagnetic waves.[9]) Especially challenging was the observation by Johann Jacob Balmer (in 1885) that the radiation frequencies of simple atoms may be well described by simple formulas. For example, for the lightest, hydrogen atom, all radiation frequencies may be numbered with just two positive integers n and $n' > n$:

$$\omega_{n,n'} = \omega_0\left(\frac{1}{n^2} - \frac{1}{n'^2}\right), \tag{1.6}$$

with $\omega_0 \equiv \omega_{1,\infty} \approx 2.07\times10^{16}$ s^{-1}. This observation, and the experimental value of ω_0, have found its first explanation in the famous 1913 theory by Niels Henrik David Bohr, which was a phenomenological precursor of present-day quantum mechanics. In this theory, $\omega_{n,n'}$ was interpreted as the frequency of a photon that obeys Eq. (5), with its energy $E_{n,n'} = \hbar\omega_{n,n'}$ being the difference between two *quantized* (discrete) energy levels of the atom (Fig. 3):

$$E_{n,n'} = E_{n'} - E_n > 0\,. \tag{1.7}$$

Bohr showed that Eq. (6) may be obtained from Eqs. (5) and (7), and classical mechanics, augmented with just one additional postulate[10] equivalent to the assumption that the angular momentum

[7] As a reminder, A. Einstein received his only Nobel Prize (in 1921) for exactly this work rather than for his relativity theory.

[8] For most metals, U_0 is between 4 and 5 electron-volts (eV), so the threshold corresponds to $\lambda_{max} = 2\pi c/\omega_{min} = 2\pi c/(U_0/\hbar) \approx 300$ nm – approximately at the border between the visible light and the ultraviolet radiation.

[9] See, e.g., EM Sec. 8.2.

[10] For more on his actual postulate, see Problem 1.

$L = m_e v r$ of an electron moving with velocity v on a circular orbit of radius r about the hydrogen's nucleus (the proton, assumed to be at rest because of its much higher mass), is quantized as

$$L = \hbar n, \tag{1.8}$$

Angular momentum quantization

where $\hbar$ is again the same Planck's constant (4), and n is an integer. (In Bohr's theory, n could not be equal to zero, though in genuine quantum mechanics, it can.)

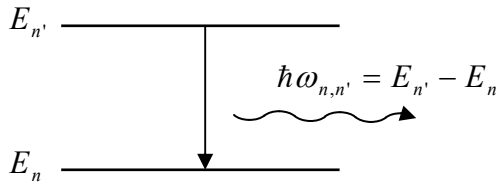


Fig. 1.3. The electromagnetic radiation of a system as a result of the transition between its quantized energy levels.

Indeed, it is sufficient to solve Eq. (8), $m_e v r = \hbar n$, together with the equation

$$m_e \frac{v^2}{r} = \frac{e^2}{4\pi\varepsilon_0 r^2}, \tag{1.9}$$

which expresses the 2nd Newton's law for an electron rotating in the Coulomb field of the nucleus. (Here $e \approx 1.6\times10^{-19}$C is the fundamental electric charge, and $m_e \approx 0.91\times10^{-30}$ kg is the electron's rest mass.) The result for r is

$$r = n^2 r_B, \qquad \text{where } r_B \equiv \frac{\hbar^2/m_e}{e^2/4\pi\varepsilon_0} \approx 0.0529\ \text{nm}. \tag{1.10}$$

Bohr radius

The constant r_B, called the *Bohr radius*, is the most important spatial scale of phenomena in atomic, molecular, and condensed-matter physics – and hence in all chemistry and biochemistry.

Now plugging these results into the non-relativistic expression for the full electron energy (with its rest energy taken for reference),

$$E = \frac{m_e v^2}{2} - \frac{e^2}{4\pi\varepsilon_0 r}, \tag{1.11}$$

we get the following simple expression for the electron's energy levels:

$$E_n = -\frac{E_H}{2n^2} < 0, \tag{1.12}$$

Hydrogen atom's energy levels

which, together with Eqs. (5) and (7), immediately gives Eq. (6) for the radiation frequencies. Here E_H is called the so-called *Hartree energy constant* (or just the "Hartree energy")[11]

$$E_H \equiv \frac{\left(e^2/4\pi\varepsilon_0\right)^2}{\hbar^2/m_e} \approx 4.360\times10^{-18}\ \text{J} \approx 27.21\ \text{eV}. \tag{1.13a}$$

Hartree energy constant

(Please note the useful relations that follow from Eqs. (10) and (13a):

[11] Unfortunately, another name, the "Rydberg constant", is sometimes used for either this energy unit or its half, $E_H/2 \approx 13.6$ eV. To add to the confusion, the same term "Rydberg constant" is used, in some subfields of physics, for the reciprocal free-space wavelength ($1/\lambda_0 = \omega_0/2\pi c$) corresponding to the frequency $\omega_0 \equiv E_H/2\hbar$.

$$E_{\rm H} = \frac{e^2}{4\pi\varepsilon_0 r_{\rm B}} = \frac{\hbar^2}{m_{\rm e} r_{\rm B}^2}, \qquad \text{i.e. } r_{\rm B} = \frac{e^2/4\pi\varepsilon_0}{E_{\rm H}} = \left(\frac{\hbar^2/m_{\rm e}}{E_{\rm H}}\right)^{1/2}; \tag{1.13b}$$

the first of them shows, in particular, that $r_{\rm B}$ is the distance at which the natural scales of the electron's potential and kinetic energies are equal.)

Note also that Eq. (8), in the form $pr = \hbar n$, where $p = m_{\rm e}v$ is the electron momentum's magnitude, may be rewritten as the condition that an integer number (n) of wavelengths λ of certain (before the late 1920s, hypothetic) waves[12] fits the circular orbit's perimeter: $2\pi r \equiv 2\pi\hbar n/p = n\lambda$. Dividing both parts of the last equality by n, we see that for this statement to be true, the wave number $k \equiv 2\pi/\lambda$ of the de Broglie waves should be proportional to the electron's momentum $p = mv$:

Momentum vs wave number

$$\boxed{p = \hbar k}, \tag{1.14}$$

again with the same Planck's constant as in Eq. (5).

(iv) The Compton effect[13] is the reduction of frequency of X-rays at their scattering on free (or nearly free) electrons – see Fig. 4.

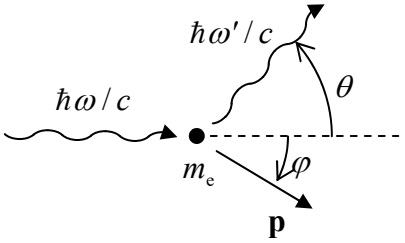


Fig. 1.4. The Compton effect.

The effect may be explained by assuming that the X-ray photon also has a definite momentum that obeys the vector-generalized version of Eq. (14):

$$\mathbf{p}_{\rm photon} = \hbar\mathbf{k} = \frac{\hbar\omega}{c}\mathbf{n}, \tag{1.15}$$

where $\mathbf{k}$ is the wavevector (whose magnitude is equal to the wave number k, and whose direction coincides with the unit vector $\mathbf{n}$ directed along the wave propagation[14]), and that the momenta of both the photon and the electron are related to their energies E by the classical relativistic formula[15]

$$E^2 = (cp)^2 + (mc^2)^2. \tag{1.16}$$

(For a photon, the rest energy m is zero, and this relation is reduced to Eq. (5): $E = cp = c\hbar k = \hbar\omega$.) Indeed, a straightforward solution of the following system of three equations,

$$\hbar\omega + m_{\rm e}c^2 = \hbar\omega' + \left[(cp)^2 + (m_{\rm e}c^2)^2\right]^{1/2}, \tag{1.17}$$

[12] This concept was first proposed in 1924 by Louis Victor Pierre Raymond de Broglie (in his PhD thesis!), so instead of speaking of wavefunctions, we are still frequently speaking of the *de Broglie waves*, especially when free particles are discussed. (In some subfields of physics, the term "matter waves" is used for the same notion.)

[13] This effect was observed in 1922, and explained a year later by Arthur Holly Compton, using Eqs. (5) and (15).

[14] See, e.g., EM Sec. 7.1.

[15] See, e.g., EM Sec. 9.3, in particular Eq. (9.78).

$$\frac{\hbar\omega}{c} = \frac{\hbar\omega'}{c}\cos\theta + p\cos\varphi, \tag{1.18}$$

$$0 = \frac{\hbar\omega'}{c}\sin\theta - p\sin\varphi, \tag{1.19}$$

(that express the conservation of, respectively, the full energy of the system and of the two relevant Cartesian components of its full momentum, at the scattering event – see Fig. 4), yields the result:

$$\frac{1}{\hbar\omega'} = \frac{1}{\hbar\omega} + \frac{1}{m_e c^2}(1-\cos\theta), \tag{1.20a}$$

which is traditionally represented as the relation between the initial and final values of the photon's wavelength $\lambda = 2\pi/k = 2\pi/(\omega/c)$: [16]

$$\lambda' = \lambda + \frac{2\pi\hbar}{m_e c}(1-\cos\theta) \equiv \lambda + \lambda_C(1-\cos\theta), \quad \text{with } \lambda_C \equiv \frac{2\pi\hbar}{m_e c}, \tag{1.20b}$$

Compton effect

and agrees with experiment.

(v) De Broglie wave diffraction. In 1927, Clinton Joseph Davisson and Lester Germer, and independently George Paget Thomson succeeded in observing the diffraction of electrons on solid crystals (Fig. 5). Specifically, they found that the intensity of the elastic reflection of electrons from a crystal increases sharply when the angle α between the incident beam of electrons and the crystal's atomic planes, separated by distance d, satisfies the following relation:

$$2d\sin\alpha = n\lambda, \tag{1.21}$$

Bragg condition

where $\lambda = 2\pi/k = 2\pi\hbar/p$ is the de Broglie wavelength of the electrons, and n is an integer. As Fig. 5 shows, this is just the well-known condition[17] that the path difference $\Delta l = 2d\sin\alpha$ between the de Broglie waves reflected from two adjacent crystal planes coincides with an integer number of λ, i.e. of the constructive interference of the waves.[18]

To summarize, all the listed experimental observations could be explained starting from two very simple (and similarly looking) formulas: Eq. (5) (at that stage, for photons only), and Eq. (15) for both photons and electrons – both relations involving the same Planck's constant. This fact might give an impression of experimental evidence sufficient to declare the light consisting of discrete particles (photons), and, on the contrary, electrons being the de Broglie waves rather than particles. However, by

[16] The constant $\lambda_C \approx 2.426\times10^{-12}$ m that participates in this relation, is called the electron's *Compton wavelength*. This term is somewhat misleading: as the reader can see from Eqs. (17)-(19), no wave in the Compton problem has such a wavelength – either before or after the scattering.

[17] See, e.g., EM Sec. 8.4, in particular Fig. 8.9 and Eq. (8.82). Frequently, Eq. (21) is called the *Bragg condition*, due to the pioneering experiments by W. Bragg on X-ray scattering from crystals, which were started in 1912.

[18] Later, spectacular experiments on diffraction and interference of heavier particles (with the correspondingly smaller de Broglie wavelength), e.g., neutrons, whole atoms, and even large organic molecules, have also been carried out – see, e.g., the review by A. Zeilinger *et al., Rev. Mod. Phys*. **60**, 1067 (1988) and a later publication Y. Fein *et al*., *Nature Physics* **15**, 1242 (2019) and references therein. Nowadays, such interference of heavy particles is used for ultrasensitive measurements of gravity, rotation, and tilt – see, e.g., the reviews by A. Cronin *et al*., *Rev. Mod. Phys*. **81**, 1051 (2009) and M. Arndt, *Phys. Today* **67**, 30 (May 2014).

that time (the mid-1920s), physics has accumulated overwhelming evidence of *wave* properties of light, such as interference and diffraction.[19] In addition, there was also strong evidence for the lumped-particle ("corpuscular") behavior of electrons. It is sufficient to mention the famous oil-drop experiments by Robert Andrew Millikan and Harvey Fletcher (1909-1913), in which only whole electrons could be added to an oil drop, changing its total electric charge by multiples of the electron's charge (-*e*) – and never by its fraction. It was apparently impossible to reconcile these observations with a purely wave picture, in which an electron and hence its charge needed to be spread over the de Broglie wave's extension, so an arbitrary part of it could be cut off using an appropriate experimental setup.

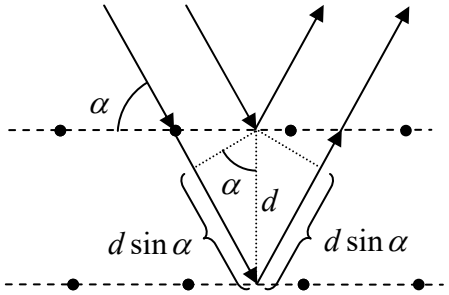


Fig. 1.5. The De Broglie wave interference at electron scattering from a crystal lattice.

Thus the founding fathers of quantum mechanics faced the formidable task of reconciling the wave and corpuscular properties of electrons and photons – and other particles. The decisive breakthrough in that task was achieved in 1926 by Ervin Schrödinger and Max Born, who formulated what is now known either formally as the *Schrödinger picture of non-relativistic quantum mechanics of the orbital motion*[20] *in the coordinate representation* (this term will be explained later in the course) or informally just as the *wave mechanics*. I will now formulate the main postulates of this theory.

1.2. Wave mechanics postulates

Let us consider a spinless,[21] non-relativistic point-like particle, whose classical dynamics may be described by a certain Hamiltonian function $H(\mathbf{r}, \mathbf{p}, t)$,[22] where $\mathbf{r}$ is the particle's radius vector and $\mathbf{p}$ is its momentum. (This condition is important because it excludes from our current discussion the systems whose interaction with their environment results in *irreversible* effects, in particular the energy's decay. Such "open" systems need a more general description, which will be discussed in Chapter 7.) Wave mechanics of such *Hamiltonian particles* may be based on the following set of postulates that are comfortingly elegant – though their final justification is given only by the agreement of all their corollaries with experiment.[23]

[19] See, e.g., EM Sec. 8.4.

[20] *Orbital motion* is the historic (and rather misleading) term used for *any* translational motion of the particle.

[21] Actually, in wave mechanics, the spin of the described particle has not to be equal to zero. Rather, it is assumed that the particle spin's *effects* on its orbital motion are negligible.

[22] As a reminder, for many systems (including all those whose kinetic energy is a quadratic-homogeneous function of generalized velocities, like $mv^2/2$), H coincides with the total energy E – see, e.g., CM Sec. 2.3. In what follows, I will assume that $H = E$ until noticed otherwise.

[23] Quantum mechanics, like any theory, may be built on different sets of postulates/axioms leading to the same results. In this text, I will not try to beat down the number of postulates to the absolute possible minimum, not

(i) Wavefunction and probability. Not always such variables as **r** or **p** can be measured exactly, even at "perfect conditions" when all external uncertainties, including measurement instrument imperfection, varieties of the initial state preparation, and unintended particle interactions with its environment, have been removed.[24] Moreover, the **r** and **p** of the same particle can *never* be measured exactly simultaneously. Instead, the *most detailed* description of the particle's state allowed by Nature is given by a certain complex function $\Psi(\mathbf{r}, t)$, called the *wavefunction* (or "wave function"), which generally enables only *probabilistic* predictions of the measured values of **r**, **p**, and other directly measurable variables – in quantum mechanics, usually called *observables*.

Specifically, the probability dW of finding a particle inside an elementary volume $dV \equiv d^3r$ is proportional to this volume and hence may be characterized by a volume-independent *probability density* $w \equiv dW/d^3r$, which in turn is related to the wavefunction as

$$w = \left|\Psi(\mathbf{r},t)\right|^2 \equiv \Psi^*(\mathbf{r},t)\Psi(\mathbf{r},t), \tag{1.22a}$$

Probability via wavefunction

where the sign * denotes the usual complex conjugation. As a result, the total probability of finding the particle somewhere inside a volume V may be calculated as

$$W = \int_V w d^3r = \int_V \Psi^*\Psi d^3r. \tag{1.22b}$$

In particular, if volume V contains the particle *definitely* (i.e. with the 100% probability, $W = 1$), Eq. (22b) is reduced to the so-called *wavefunction normalization condition*

$$\int_V \Psi^*\Psi d^3r = 1. \tag{1.22c}$$

Normalization condition

(ii) Observables and operators. With each observable A, quantum mechanics associates a certain *linear operator* $\hat{A}$, such that the average measured value of A (usually called the *expectation value*) is expressed as[25]

$$\langle A \rangle = \int_V \Psi^* \hat{A} \Psi d^3r, \tag{1.23}$$

Observable's expectation value

where $\langle \ldots \rangle$ means the statistical average, i.e. the result of averaging the measurement results over a large *ensemble* (set) of macroscopically similar experiments, and Ψ is the normalized wavefunction that satisfies Eq. (22c). Note immediately that for Eqs. (22) and (23) to be compatible, the *identity* (or "unit") *operator* defined by the relation

$$\hat{I}\Psi = \Psi, \tag{1.24}$$

Identity operator

has to be associated with a particular type of measurement, namely with the particle's *detection*, i.e. the observation of its presence.

---

only because that would require longer argumentation, but chiefly because such attempts typically result in making certain implicit assumptions hidden from the reader – the practice as common as regrettable.

[24] I will imply such perfect conditions in the further narrative, until the discussion of the system's interaction with its environment in Chapter 7.

[25] This key measurement postulate is sometimes called the *Born rule*, though sometimes this term is used for the (less general) Eqs. (22).

(iii) The Hamiltonian operator and the Schrödinger equation. Another particular operator, *Hamiltonian* $\hat{H}$ whose observable is the particle's energy E, also plays in wave mechanics a very special role, because it participates in the *Schrödinger equation*,

Schrödinger equation

$$i\hbar\frac{\partial\Psi}{\partial t}=\hat{H}\Psi\,, \tag{1.25}$$

that determines the wavefunction's dynamics, i.e. its time evolution.

(iv) The radius-vector and momentum operators. In wave mechanics (i.e. in the coordinate representation), the vector operator of the particle's radius vector $\mathbf{r}$ just multiples the wavefunction by this vector, while the operator of the particle's momentum is proportional to the spatial derivative:

$$\hat{\mathbf{r}}=\mathbf{r},\qquad \hat{\mathbf{p}}=-i\hbar\nabla\,, \tag{1.26a}$$

Operators of coordinate and momentum

where ∇ is the *del* (or "nabla") vector operator.[26] Thus in the Cartesian coordinates,

$$\hat{\mathbf{r}}=\mathbf{r}=\{x,y,z\},\qquad \hat{\mathbf{p}}=-i\hbar\left\{\frac{\partial}{\partial x},\frac{\partial}{\partial y},\frac{\partial}{\partial z}\right\}. \tag{1.26b}$$

(v) The correspondence principle. In the limit when quantum effects are insignificant, e.g., when the characteristic scale of action[27] (i.e. the product of the relevant energy and time scales of the problem) is much larger than Planck's constant $\hbar$, all wave mechanics results have to tend to those given by classical mechanics. Mathematically, this correspondence is achieved by duplicating the classical relations between various observables with similar relations between the corresponding operators. For example, for a free particle, the Hamiltonian (which in this particular case corresponds to the kinetic energy $T=p^2/2m$ alone) has the form

Free particle's Hamiltonian

$$\hat{H}=\hat{T}=\frac{\hat{p}^2}{2m}=-\frac{\hbar^2}{2m}\nabla^2. \tag{1.27}$$

Now, even before a deeper discussion of the postulates' physics (offered in the next section), we may immediately see that they indeed provide a formal way toward a resolution of the apparent contradiction between the wave and corpuscular properties of particles. Indeed, for a free particle, the Schrödinger equation (25), with the substitution of Eq. (27), takes the form

Free particle's Schrödinger equation

$$i\hbar\frac{\partial\Psi}{\partial t}=-\frac{\hbar^2}{2m}\nabla^2\Psi\,, \tag{1.28}$$

whose particular but most important solution is a plane, single-frequency ("monochromatic") traveling wave,[28]

Plane wave

$$\Psi(\mathbf{r},t)=ae^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}, \tag{1.29}$$

[26] If you need, see, e.g., Secs. 8-10 of the *Selected Mathematical Formulas* appendix – below, referred to as MA. Note that according to those formulas, the del operator follows all the rules of the usual (geometric) vectors. This is, by definition, true for other quantum-mechanical vector operators to be discussed below.

[27] See, e.g., CM Sec. 10.3.

[28] See, e.g., CM Sec. 6.4 and/or EM Sec. 7.1.

where a, $\mathbf{k}$, and ω are constants. Indeed, plugging Eq. (29) into Eq. (28), we immediately see that such a plane wave, with an arbitrary complex amplitude a, is indeed a solution of this Schrödinger equation, provided a specific *dispersion relation* between the wave number $k \equiv |\mathbf{k}|$ and the frequency ω:

$$\hbar\omega = \frac{(\hbar k)^2}{2m}. \tag{1.30}$$

Free particle's dispersion relation

The constant a may be calculated, for example, assuming that the wave (29) is extended over a certain volume V, while beyond it, $\Psi = 0$. Then from the normalization condition (22c) and Eq. (29), we get[29]

$$|a|^2 V = 1. \tag{1.31}$$

Let us use Eqs. (23), (26), and (27) to calculate the expectation values of the particle's momentum $\mathbf{p}$ and energy $E = H$ in the state (29). The result is

$$\langle \mathbf{p} \rangle = \hbar\mathbf{k}, \qquad \langle E \rangle = \langle H \rangle = \frac{(\hbar k)^2}{2m}; \tag{1.32}$$

according to Eq. (30), the last equality may be rewritten as $\langle E \rangle = \hbar\omega$.

Next, Eq. (23) enables calculation of not only the average (in math speak, the *first moment*) of an observable but also its higher moments, notably the *second moment* – in physics, usually called *variance*:

$$\langle \tilde{A}^2 \rangle \equiv \langle (A - \langle A \rangle)^2 \rangle = \langle A^2 \rangle - \langle A \rangle^2, \tag{1.33}$$

Observable's variance

and hence its *uncertainty* alternatively called the "root-mean-square (r.m.s.) fluctuation",

$$\delta A \equiv \langle \tilde{A}^2 \rangle^{1/2}. \tag{1.34}$$

Observable's uncertainty

The uncertainty is the scale of deviations $\tilde{A} \equiv A - \langle A \rangle$ of measurement results from their average. In the particular case when the uncertainty δA equals zero, every measurement of the observable A will give the same value $\langle A \rangle$; such a state is said to have a *definite value* of the variable. For example, in application to the state with wavefunction (29), these relations yield $\delta E = 0$, $\delta\mathbf{p} = 0$. This means that in this plane-wave, monochromatic state, the energy and momentum of the particle have definite values, so the statistical average signs in Eqs. (32) might be removed. Thus, these relations are reduced to the experimentally inferred Eqs. (5) and (15).

Hence the wave mechanics postulates indeed may describe the observed wave properties of non-relativistic particles. (For photons, we would need its relativistic generalization – see Chapter 9 below.) On the other hand, due to the linearity of the Schrödinger equation (25), any sum of its solutions is also a solution – the so-called *linear superposition principle*. For a free particle, this means that any set of plane waves (29) is also a solution to this equation. Such sets, with close values of $\mathbf{k}$ and hence $\mathbf{p} = \hbar\mathbf{k}$ (and, according to Eq. (30), of ω as well), may be used to describe spatially localized "pulses", called *wave packets* – see Fig. 6. In Sec. 2.1, I will prove (or rather reproduce H. Weyl's proof :-) that the wave packet's extension δx in any direction (say, x) is related to the width δk_x of the distribution of the

[29] For infinite space ($V \to \infty$), Eq. (31) yields $a \to 0$, i.e. wavefunction (29) vanishes. This formal problem may be readily resolved considering sufficiently long wave packets – see Sec. 2.2 below.

corresponding component of its wave vector as $\delta x \delta k_x \geq ½$, and hence, according to Eq. (15), to the width δp_x of the momentum component distribution as

Heisenberg's uncertainty relation

$$\delta x \cdot \delta p_x \geq \frac{\hbar}{2}. \tag{1.35}$$

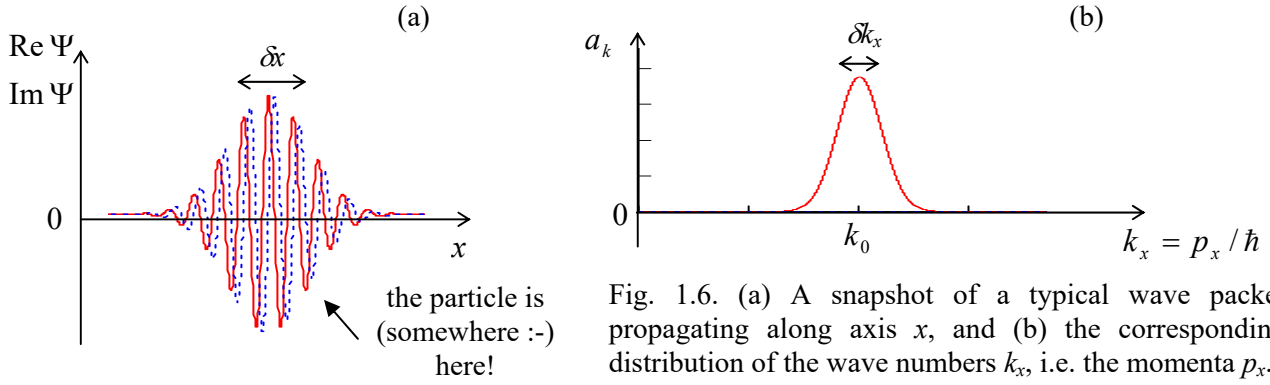


Fig. 1.6. (a) A snapshot of a typical wave packet propagating along axis x, and (b) the corresponding distribution of the wave numbers k_x, i.e. the momenta p_x.

This is the famous *Heisenberg's uncertainty principle*, which quantifies the first postulate's point that the coordinate and the momentum cannot be defined exactly simultaneously. However, since Planck's constant, $\hbar \sim 10^{-34}$ J·s, is extremely small on the human scale of things, it still allows for particle localization in a very small volume even if the momentum spread in a wave packet is also small on that scale. For example, according to Eq. (35), a 0.1% spread of momentum of a 1 keV electron ($p \sim 1.7\times10^{-24}$ kg·m/s) allows its wave packet to be as small as $\sim3\times10^{-10}$ m. (For a heavier particle such as a proton, the packet would be even tighter.) As a result, wave packets may be used to describe the particles that are quite point-like from the macroscopic point of view.

In a nutshell, this is the main idea of wave mechanics, and the first part of this course (Chapters 1-3) will be essentially a discussion of the various effects described by this approach. During this discussion, however, we will not only witness the wave mechanics' many triumphs within its applicability domain but also gradually accumulate evidence for its handicaps, which will force an eventual transfer to a more general formalism – to be discussed in Chapter 4 and beyond.

1.3. Postulates' discussion

The wave mechanics' postulates listed in the previous section (hopefully, familiar to the reader from their undergraduate studies) may look very simple. However, the physics of these axioms is very deep, leading to some counter-intuitive conclusions, and their in-depth discussion requires solutions of several key problems of wave mechanics. This is why in this section I will give only an initial, admittedly superficial discussion of the postulates, and will be repeatedly returning to the conceptual foundations of quantum mechanics throughout the course, especially in the concluding Chapter 10.

First of all, the fundamental uncertainty of observables, which is in the core of the first postulate, is very foreign to the basic ideas of classical mechanics, and historically has made quantum mechanics so hard to swallow for many star physicists, notably including Albert Einstein – despite his 1905 work, which advanced the field so much. However, this fact has been confirmed by numerous experiments,

and (more importantly) there has not been a single confirmed experiment that would contradict this postulate, so quantum mechanics was long ago promoted from a theoretical hypothesis to the rank of a reliable scientific theory.

One more remark in this context is that Eq. (25) itself is *deterministic*, i.e. conceptually enables an *exact* calculation of the wavefunction's distribution in space at any instant t, provided that its initial distribution, and the particle's Hamiltonian, are known exactly. Note that in classical statistical mechanics, the probability density distribution $w(\mathbf{r}, t)$ may be also calculated from deterministic differential equations, for example, the Liouville equation.[30] The quantum-mechanical description differs from that situation in two important aspects. First, in the perfect conditions outlined above (the best possible initial state preparation and measurements), the Liouville equation is reduced to the 2nd Newton law of classical mechanics, i.e. the statistical uncertainty of its results disappears. In quantum mechanics this is not true: the quantum uncertainly, such as that described by Eq. (35), persists even in this limit. Second, the wavefunction $\Psi(\mathbf{r}, t)$ gives more information than just $w(\mathbf{r}, t)$ because, besides the modulus of Ψ involved in Eq. (22), this complex function also has the *phase* $\varphi \equiv \arg\Psi$, which may affect some observables, describing, in particular, interference of the de Broglie waves.

Next, it is very important to understand that the relation between the quantum mechanics and experiment, given by the second postulate, necessarily involves another key notion: that of the corresponding *statistical ensemble*, in this case, a set of many experiments carried out at apparently (*macroscopically*) similar settings including the initial conditions. Indeed, the probability of a certain (n^{th}) result (*outcome*) of an experiment may be only defined for a certain statistical ensemble, as the limit

$$W_n \equiv \lim_{M\to\infty} \frac{M_n}{M}, \qquad \text{with } M \equiv \sum_{n=1}^{N} M_n, \tag{1.36}$$

Probability: definition

where M is the total number of experiments, M_n is the number of outcomes of the n^{th} type, and N is the number of different outcomes.

Note that a particular choice of statistical ensemble may affect probabilities W_n very significantly. For example, if we pull out playing cards at random from a standard pack of 52 different cards of 4 suits, the probability W_n of getting a certain card (e.g., the queen of spades) is 1/52. However, if the cards of a certain suit (say, hearts) had been taken out from the pack in advance, the probability of getting the queen of spades is higher, 1/39. It is important that we would also get the last number for the probability even if we had used the full 52-card pack, but for some reason discarded results of all experiments giving us any rank of hearts. Hence, the ensemble definition (or its *redefinition* in the middle of the game) may change outcome probabilities.

In wave mechanics, with its fundamental relation (22) between w and Ψ, this means that not only the outcome probabilities but the wavefunction itself may also depend on the statistical ensemble we are using, i.e. not only on the preparation of the system and the experimental setup, but also on the subset of outcomes taken into account. This is why an attribution of the wavefunction to a single experiment, both before and after the measurement, may lead to very unphysical interpretations of the results, including some wavefunction's evolution stages not described by the Schrödinger equation (the so-called *wave packet reduction*), superluminal action on distance, etc. Later in the course, we will see that minding the fundamentally statistical nature of quantum mechanics, and in particular the dependence of

[30] See, e.g., SM Sec. 6.1.

wavefunctions on the statistical ensemble's definition (or redefinition), readily resolves most, though not all, paradoxes of quantum measurements.

Note, however, again that the standard quantum mechanics, as discussed in Chapters 1-6 and 9 of this course, is limited to statistical ensembles with *the least possible uncertainty* of the considered systems, i.e. with the best possible knowledge of their state.[31] This condition requires, first, the least uncertain initial preparation of the system, and second, its total isolation from the rest of the world, or at least from its disordered part (the "environment"), in the course of its evolution. Only such ensembles may be described by certain wavefunctions. A detailed discussion of more general ensembles, which are necessary if these conditions are not satisfied, will be given in Chapters 7, 8, and 10.

Finally, regarding Eq. (23): a better feeling of this expression may be obtained by its comparison with the general definition of the expectation value (i.e. of the statistical average) in the probability theory. Namely, let each of N possible outcomes in a set of M experiments give a certain value A_n of an observable A; then

Definition of statistical average

$$\langle A \rangle \equiv \lim_{M\to\infty} \frac{1}{M}\sum_{n=1}^{N} A_n M_n = \sum_{n=1}^{N} A_n W_n . \tag{1.37}$$

Taking into account Eq. (22), which relates W and Ψ, the structures of Eq. (23) and the final form of Eq. (37) are similar. Their exact relation will be further discussed in Sec. 4.1.

1.4. Continuity equation

The wave mechanics postulates survive one more sanity check: they satisfy the natural requirement that the particle does not appear or vanish in the course of the quantum evolution.[32] Indeed, let us use Eq. (22b) to calculate the rate of change of the probability W to find a particle within a certain volume V:

$$\frac{dW}{dt} = \frac{d}{dt}\int_V \Psi\Psi^* d^3 r . \tag{1.38}$$

Assuming for simplicity that the boundaries of this volume V do not move, it is sufficient to carry out the partial differentiation of the product $\Psi\Psi^*$ inside the integral. Using the Schrödinger equation (25), together with its complex conjugate,

$$-i\hbar\frac{\partial\Psi^*}{\partial t} = (\hat{H}\Psi)^* , \tag{1.39}$$

we readily get

$$\frac{dW}{dt} = \int_V \frac{\partial}{\partial t}\left(\Psi\Psi^*\right) d^3 r = \int_V \left(\Psi^*\frac{\partial\Psi}{\partial t} + \Psi\frac{\partial\Psi^*}{\partial t}\right) d^3 r = \frac{1}{i\hbar}\int_V \left[\Psi^*\left(\hat{H}\Psi\right) - \Psi\left(\hat{H}\Psi\right)^*\right] d^3 r. \tag{1.40}$$

[31] The reader should not be surprised by the use of the notion of "knowledge" (or "information") in this context. Indeed, due to the statistical character of experiment outcomes, quantum mechanics (or at least its relation to experiment) is intimately related to information theory. In contrast to much of classical physics, which may be discussed without any reference to information, in quantum mechanics, as in classical statistical physics, such abstraction is possible only in some very special (and not the most interesting) cases.

[32] Note that this requirement may be violated in the *relativistic* quantum theory – see Chapter 9.

Let the particle move in a field of external forces (not necessarily constant in time), so its classical Hamiltonian function H is the sum of the particle's kinetic energy $T = p^2/2m$ and its potential energy $U(\mathbf{r}, t)$.[33] According to the correspondence principle and Eq. (27), the Hamiltonian operator may be represented as the sum[34]

$$\hat{H} = \hat{T} + \hat{U} = \frac{\hat{p}^2}{2m} + U(\mathbf{r},t) = -\frac{\hbar^2}{2m}\nabla^2 + U(\mathbf{r},t)\,. \tag{1.41}$$

Potential field: Hamiltonian

At this stage, we should notice that this operator, when acting on a real function, gives a real function.[35] Hence, the result of its action on an arbitrary complex function $\Psi = a + ib$ (where a and b are real) is

$$\hat{H}\Psi = \hat{H}(a+ib) = \hat{H}a + i\hat{H}b\,, \tag{1.42}$$

where $\hat{H}a$ and $\hat{H}b$ are also real, while

$$\left(\hat{H}\Psi\right)^* = \left(\hat{H}a + i\hat{H}b\right)^* = \hat{H}a - i\hat{H}b = \hat{H}(a-ib) = \hat{H}\Psi^*\,. \tag{1.43}$$

This means that Eq. (40) may be rewritten as

$$\frac{dW}{dt} = \frac{1}{i\hbar}\left(\Psi^*\hat{H}\Psi - \Psi\hat{H}\Psi^*\right)d^3r = -\frac{\hbar^2}{2m}\frac{1}{i\hbar}\int_V\left(\Psi^*\nabla^2\Psi - \Psi\nabla^2\Psi^*\right)d^3r\,. \tag{1.44}$$

Now let us use the general rules of vector calculus[36] to write the following identity:

$$\nabla\cdot\left(\Psi^*\nabla\Psi - \Psi\nabla\Psi^*\right) = \Psi^*\nabla^2\Psi - \Psi\nabla^2\Psi^*, \tag{1.45}$$

A comparison of Eqs. (44) and (45) shows that we may write

$$\frac{dW}{dt} = -\int_V(\nabla\cdot\mathbf{j})\,d^3r, \tag{1.46}$$

where the vector $\mathbf{j}$ is defined as

$$\mathbf{j} \equiv \frac{i\hbar}{2m}\left(\Psi\nabla\Psi^* - \text{c.c.}\right) \equiv \frac{\hbar}{m}\text{Im}\left(\Psi^*\nabla\Psi\right), \tag{1.47}$$

Probability current density

where c.c. means the complex conjugate of the previous expression – in this case, $(\Psi\nabla\Psi^*)^*$, i.e. $\Psi^*\nabla\Psi$. Now using the well-known divergence theorem,[37] Eq. (46) may be rewritten as the *continuity equation*

$$\frac{dW}{dt} + I = 0, \quad \text{with } I \equiv \int_S j_n\,d^2r\,, \tag{1.48}$$

Continuity equation: integral form

[33] As a reminder, such a description is valid not only for conservative forces (in that case U has to be time-independent) but also for any force $\mathbf{F}(\mathbf{r}, t)$ that may be expressed via the gradient of $U(\mathbf{r}, t)$ – see, e.g., CM Chapters 2 and 10. (A good counter-example when such a description is *impossible* is given by the magnetic component of the Lorentz force – see, e.g., EM Sec. 9.7 and also Sec. 3.1 below.)

[34] Historically, this was the main step made (in 1926) by E. Schrödinger on the background of L. de Broglie's idea. The probabilistic interpretation of the wavefunction was put forward, almost simultaneously, by M. Born.

[35] In Chapter 4, we will discuss a more general family of *Hermitian operators*, which have this property.

[36] See, e.g., MA Eq. (11.4a) combined with the del operator's definition $\nabla^2 \equiv \nabla\cdot\nabla$.

[37] See, e.g., MA Eq. (12.2).

where j_n is the component of the vector **j** along the outwardly directed normal to the closed surface S that limits the volume V, i.e. the scalar product $\mathbf{j}\cdot\mathbf{n}$, where **n** is the unit vector along this normal.

Formulas (47) and (48) show that if the wavefunction on the surface vanishes, the total probability W of finding the particle within the volume does not change, providing the intended sanity check. In the general case, Eq. (48) says that dW/dt equals the flux I of the vector **j** through the surface, with the minus sign. It is clear that this vector may be interpreted as the *probability current density* – and I, as the total *probability current* through the surface S. This interpretation may be further supported by applying Eq. (47) to any wavefunction represented in the polar form $\Psi = ae^{i\varphi}$, with real a and φ:

$$\mathbf{j} = a^2 \frac{\hbar}{m}\nabla\varphi\,. \tag{1.49}$$

Note that for a real wavefunction, or even for a wavefunction with an arbitrary but space-constant phase φ, the probability current density vanishes. On the contrary, for the traveling wave (29), with a constant probability density (1.22a), $w = a^2$, Eq. (49) yields a non-zero (and physically very transparent) result:

$$\mathbf{j} = w\frac{\hbar}{m}\mathbf{k} = w\frac{\mathbf{p}}{m} = w\mathbf{v}\,, \tag{1.50}$$

where $\mathbf{v} = \mathbf{p}/m$ is the particle's velocity. If multiplied by the particle's mass m, the probability density w turns into the (average) mass density ρ, and the probability current density, into the *mass flux density* $\rho\mathbf{v}$. Similarly, if multiplied by the total electric charge q of the particle, with w turning into the *charge density* σ, **j** becomes the electric current density. As the reader (hopefully :-) knows, both these currents satisfy classical continuity equations similar to Eq. (48).[38]

Finally, let us recast the continuity equation, rewriting Eq. (46) as

$$\int_V \left(\frac{\partial w}{\partial t} + \nabla\cdot\mathbf{j}\right) d^3r = 0\,. \tag{1.51}$$

Now we may argue that this equality may be true for any choice of volume V only if the expression under the integral vanishes everywhere, i.e. if

Continuity equation: differential form

$$\boxed{\frac{\partial w}{\partial t} + \nabla\cdot\mathbf{j} = 0.} \tag{1.52}$$

This differential form of the continuity equation may be more convenient than its integral form (48).

1.5. Eigenstates and eigenvalues

Now let us discuss the most important corollaries of wave mechanics' *linearity*. First of all, it uses only *linear operators*. This term means that the operators must obey the following two rules:[39]

$$\left(\hat{A}_1 + \hat{A}_2\right)\Psi = \hat{A}_1\Psi + \hat{A}_2\Psi, \tag{1.53}$$

[38] See, e.g., respectively, CM 8.3 and EM Sec. 4.1.

[39] By the way, if any equality involving operators is valid for an arbitrary wavefunction, the latter is frequently dropped from the notation, resulting in *operator equality*. In particular, Eq. (53) may be readily used to prove that the linear operators are *commutative*: $\hat{A}_2 + \hat{A}_1 = \hat{A}_1 + \hat{A}_2$, and *associative*: $\left(\hat{A}_1 + \hat{A}_2\right) + \hat{A}_3 = \hat{A}_1 + \left(\hat{A}_2 + \hat{A}_3\right)$.

$$\hat{A}\left(c_1\Psi_1 + c_2\Psi_2\right) = \hat{A}\left(c_1\Psi_1\right) + \hat{A}\left(c_2\Psi_2\right) = c_1\hat{A}\Psi_1 + c_2\hat{A}\Psi_2, \tag{1.54}$$

where Ψ_n are arbitrary wavefunctions and c_n are arbitrary constants (in quantum mechanics, frequently called *c-numbers*, to distinguish them from operators and wavefunctions). The most important examples of linear operators are given by:

(i) the multiplication by a function, such as for the operator $\hat{\mathbf{r}}$ given by Eq. (26), and
(ii) the spatial or temporal differentiation, such as in Eqs. (25)-(27).

Next, it is of key importance that the Schrödinger equation (25) is also linear. (This fact was already used in the discussion of wave packets in Sec. 2.) This means that if each of several functions Ψ_n are particular solutions of Eq. (25) with a certain Hamiltonian, then their arbitrary linear combination,

$$\Psi = \sum_n c_n \Psi_n, \tag{1.55}$$

is also a solution of the same equation.[40]

Let us use the linearity to accomplish an apparently impossible feat: immediately find the *general* solution of the Schrödinger equation for the important case when the system's Hamiltonian does not depend on time explicitly – for example, is given by Eq. (41) with time-independent potential energy $U = U(\mathbf{r})$, so the corresponding Schrödinger equation has the form

$$i\hbar\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Psi + U(\mathbf{r})\Psi. \tag{1.56}$$

Static field: Schrödinger equation

First of all, let us prove that the following product,

$$\Psi_n = a_n(t)\psi_n(\mathbf{r}), \tag{1.57}$$

Variable separation

qualifies as a particular solution of this equation. Indeed, plugging Eq. (57) into Eq. (25) with *any* time-independent Hamiltonian, using the fact that in this case

$$\hat{H}a_n(t)\psi_n(\mathbf{r}) = a_n(t)\hat{H}\psi_n(\mathbf{r}), \tag{1.58}$$

and dividing both parts of the equation by $a_n\psi_n$, we get

$$\frac{i\hbar}{a_n}\frac{da_n}{dt} = \frac{\hat{H}\psi_n}{\psi_n}. \tag{1.59}$$

The left-hand side of this equation may depend only on time, while the right-hand side, only on coordinates. This may be true for all $\mathbf{r}$ and t only if we assume that each of these parts is equal to (the same) constant of the dimension of energy, which I will denote as E_n.[41] As a result, we are getting two separate equations for the temporal and spatial parts of the wavefunction:

[40] At first glance, it may seem strange that the *linear* Schrödinger equation correctly describes quantum properties of systems whose classical dynamics is described by *nonlinear* equations of motion, e.g., an anharmonic oscillator – see, e.g., CM Sec. 5.2. Note, however, that statistical equations of classical dynamics (see, e.g., SM Chapters 5 and 6) also have this property, so it is not specific to quantum mechanics.

[41] This argumentation, leading to *variable separation*, is very common in mathematical physics – see, e.g., its discussion in EM Sec. 2.5.

Stationary Schrödinger equation

$$\hat{H}\psi_n = E_n\psi_n, \tag{1.60}$$

$$i\hbar\frac{da_n}{dt} = E_n a_n . \tag{1.61a}$$

The latter of these equations, rewritten in the form

$$\frac{da_n}{a_n} = -i\frac{E_n}{\hbar}dt, \tag{1.61b}$$

is readily integrable, giving

Stationary state: time evolution

$$\ln a_n = -i\omega_n t + \text{const}, \qquad \text{so } a_n = \text{const}\times\exp\{-i\omega_n t\}, \qquad \text{with } \omega_n \equiv \frac{E_n}{\hbar}. \tag{1.62}$$

Now plugging Eqs. (57) and (62) into Eq. (22), we see that in the quantum state described by Eqs. (57)-(62), the probability w of finding the particle at a certain location does not depend on time:

$$w \equiv \psi_n^*(\mathbf{r})\psi_n(\mathbf{r}) = w(\mathbf{r}). \tag{1.63}$$

With the same substitution, Eq. (23) shows that the expectation value of any operator that does not depend on time explicitly is also time-independent:

$$\langle A\rangle \equiv \int \psi_n^*(\mathbf{r})\hat{A}\psi_n(\mathbf{r})d^3r = \text{const}. \tag{1.64}$$

Due to this property, the states described by Eqs. (57)-(62) are called *stationary*; they are fully defined by the possible solutions of the *stationary* (or "time-independent") *Schrödinger equation* (60).[42] Note that for the Hamiltonian (41), the stationary Schrödinger equation (60),

Static field: stationary Schrödinger equation

$$-\frac{\hbar^2}{2m}\nabla^2\psi_n + U(\mathbf{r})\psi_n = E_n\psi_n, \tag{1.65}$$

is a linear, homogeneous differential equation for the function ψ_n, with *a priory* unknown parameter E_n. Such equations fall into the mathematical category of *eigenproblems*,[43] whose *eigenfunctions* ψ_n and *eigenvalues* E_n should be found simultaneously, i.e. self-consistently.[44]

Mathematics[45] tells us that for such equations with space-confined eigenfunctions ψ_n, tending to zero at $r \to \infty$, the spectrum of eigenvalues is *discrete*. It also proves that the eigenfunctions corresponding to different eigenvalues are *orthogonal,* i.e. that space integrals of the products $\psi_n\psi_{n'}^*$ vanish for all pairs with $n \neq n'$. Due to the Schrödinger equation's linearity, each of these functions may be multiplied by a proper constant coefficient to make their set *orthonormal*:

$$\int \psi_n^*\psi_{n'}d^3r = \delta_{n,n'} \equiv \begin{cases} 1, & \text{for } n = n', \\ 0, & \text{for } n \neq n'. \end{cases} \tag{1.66}$$

[42] In contrast, the *full* Schrödinger equation (25) is frequently called *time-dependent* or *non-stationary*.

[43] From the German root *eigen*, meaning "particular" or "characteristic".

[44] Eigenvalues of energy are frequently called *eigenenergies*, and it is often said that the eigenfunction ψ_n and the corresponding eigenenergy E_n together determine the n^{th} *stationary eigenstate* of the system.

[45] See, e.g., Sec. 9.3 of the handbook by G. Korn and T. Korn, listed in MA Sec. 16(ii).

Moreover, the eigenfunctions $\psi_n(\mathbf{r})$ form a *full set*, meaning that an arbitrary function $\psi(\mathbf{r})$, in particular the actual wavefunction Ψ of the system at the initial moment of its evolution (which I will always, with a few clearly marked exceptions, take for $t = 0$), may be represented as a unique expansion over the eigenfunction set:

$$\Psi(\mathbf{r},0) = \sum_n c_n \psi_n(\mathbf{r}) . \tag{1.67}$$

The expansion coefficients c_n may be readily found by multiplying both sides of Eq. (67) by $\psi^*_{n'}$, integrating the results over the space, and using Eq. (66). The result is

$$c_n = \int \psi_n^*(\mathbf{r})\Psi(\mathbf{r},0)d^3r . \tag{1.68}$$

Now let us consider the following wavefunction[46]

$$\Psi(\mathbf{r},t) = \sum_n c_n a_k(t)\psi_k(\mathbf{r}) = \sum_n c_n \exp\left\{-i\frac{E_n}{\hbar}t\right\}\psi_n(\mathbf{r}) . \tag{1.69}$$

General solution of Eq. (56)

Since each term of the sum has the form (57) and satisfies the Schrödinger equation, so does the sum as the whole. Moreover, if the coefficients c_n are derived in accordance with Eq. (68), then the solution (69) satisfies the initial conditions as well. At this moment we can use one more bit of help from mathematicians, who tell us that the linear, partial differential equation (56), with fixed initial conditions, may have only one (*unique*) solution. This means that in our case of time-independent potential Hamiltonian, Eq. (69) gives the *general* solution of the Schrödinger equation (25).

So, we have succeeded in our apparently over-ambitious goal. Now let us pause this mad mathematical dash for a minute, and discuss this key result.

1.6. Time evolution

For the time-dependent factor $a_n(t)$ of each component (57) of the general solution (69), our procedure gave a very simple and universal result (62), describing a linear change of the phase $\varphi_n \equiv \arg(a_n)$ of this complex function in time, with a constant rate

$$\frac{d\varphi_n}{dt} = -\omega_n = -\frac{E_n}{\hbar} , \tag{1.70}$$

so the real and imaginary parts of a_n oscillate sinusoidally with this frequency. The relation (70) coincides with the Planck-Einstein conjecture (5), but could these oscillations of the wavefunctions represent a physical reality? Indeed, for photons, described by Eq. (5), E may be (and as we will see in Chapter 9, is) the actual, well-defined energy of one photon, and ω is the frequency of the radiation so quantized. However, for non-relativistic particles described by wave mechanics, the potential energy U and hence the full energy E are defined up to an arbitrary constant because we may measure them from an arbitrary reference level. How can such a change of the energy reference level (which may be made just in our mind) alter the frequency of oscillations of a variable?

According to Eqs. (22)-(23), this time evolution of a wavefunction does not affect the particle's probability distribution, or even any observable (including the energy E, provided that it is always

[46] Note that according to Eq. (22b), the probability of finding the system in the k^{th} state equals $|c_k|^2$. Because of that, the complex coefficients c_k (or sometimes the products $c_k a_k$) are called *probability amplitudes*.

referred to the same origin as U), in any stationary state. However, let us combine Eq. (5) with Bohr's assumption (7):

$$\hbar\omega_{nn'} = E_{n'} - E_n . \tag{1.71}$$

The *difference* $\omega_{nn'}$ of the eigenfrequencies ω_n and $\omega_{n'}$, participating in this formula, is evidently independent of the energy reference, and as will be proved later in the course, determines the measurable frequency of the electromagnetic radiation (or possibly of a wave of a different physical nature) emitted or absorbed at the quantum transition between the states.

As another but related example, consider two similar particles 1 and 2, each in the same (say, the lowest-energy) eigenstate, but with their potential energies (and hence the ground state energies $E_{1,2}$) different by a constant $\Delta U \equiv U_1 - U_2$. Then, according to Eq. (70), the difference $\varphi \equiv \varphi_1 - \varphi_2$ of their wavefunction phases evolves in time with the reference-independent rate

$$\frac{d\varphi}{dt} = -\frac{\Delta U}{\hbar} . \tag{1.72}$$

Certain measurement instruments, weakly coupled to the particles, may allow observation of this evolution, while keeping the particle's quantum dynamics virtually unperturbed, i.e. Eq. (70) intact. Perhaps the most spectacular measurement of this type is possible using the *Josephson effect* in weak links between two superconductors – see Fig. 7.[47]

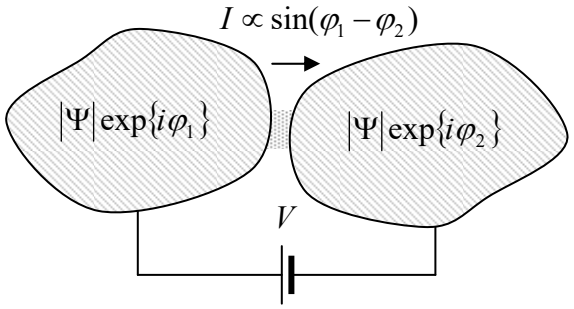


Fig. 1.7. The Josephson effect in a weak link between two bulk superconductor electrodes.

As a brief reminder,[48] superconductivity may be explained by a specific coupling between conduction electrons in solids, that leads, at low temperatures, to the formation of the so-called *Cooper pairs*. Such pairs behave as Bose particles and form a coherent *Bose-Einstein condensate*.[49] Most properties of such a condensate may be described by a single, common wavefunction Ψ, evolving in time just as that of a free particle, with the effective potential energy $U = q\phi = -2e\phi$, where ϕ is the electrochemical potential,[50] and $q = -2e$ is the electric charge of a Cooper pair. As a result, for the system shown in Fig. 7, in which externally applied voltage V fixes the difference $\phi_1 - \phi_2$ between the electrochemical potentials of the superconductors, Eq. (72) takes the form

Eq. (72) for Josephson effect

$$\boxed{\frac{d\varphi}{dt} = \frac{2e}{\hbar}V} . \tag{1.73}$$

[47] The effect was predicted in 1962 by Brian Josephson (then a graduate student!) and observed soon after that.
[48] For a more detailed discussion, including the derivation of Eq. (75), see e.g. EM Chapter 6.
[49] A detailed discussion of the Bose-Einstein condensation may be found, e.g., in SM Sec. 3.4.
[50] For more on this notion see, e.g. SM Sec. 6.3.

If the link between the superconductors is weak enough, the electric current I of the Cooper pairs (called the *supercurrent*) through the link may be approximately described by the following simple relation,

$$I = I_c \sin\varphi, \tag{1.74}$$

Josephson supercurrent

where I_c is some constant, dependent on the weak link's strength.[51] Now combining Eqs. (73) and (74), we see that if the applied voltage V is constant in time, the current oscillates sinusoidally, with the so-called *Josephson frequency*

$$\omega_J \equiv \frac{2e}{\hbar} V, \tag{1.75}$$

as high as ~484 MHz per microvolt of applied dc voltage. This effect may be readily observed experimentally: though its direct detection is a bit tricky, it is easy to observe the *phase locking* (synchronization)[52] of the Josephson oscillations by an external microwave signal of frequency ω. Such phase locking results in the relation $\omega_J = n\omega$ fulfilled within certain dc current intervals, and hence in the formation, on the weak link's dc I-V curve, of virtually vertical current steps at dc voltages

$$V_n = n\frac{\hbar\omega}{2e}, \tag{1.76}$$

where n is an integer.[53] Since frequencies may be stabilized and measured with very high precision, this effect is being used in highly accurate standards of dc voltage.

1.7. Spatial dependence

In contrast to the simple and universal time dependence (62) of the stationary states, their spatial wavefunctions $\psi_n(\mathbf{r})$ need to be calculated from the problem-specific stationary Schrödinger equation (65). The solution of this equation for various particular cases will be a major focus of the next two chapters. Here I will consider just one simple example, which nevertheless will be the basis for our discussion of more complex problems. Let a particle be confined inside a rectangular hard-wall box. Such confinement may be described by the following potential energy profile:[54]

$$U(\mathbf{r}) = \begin{cases} 0, & \text{for } 0 < x < a_x, \quad 0 < y < a_y, \quad \text{and } 0 < z < a_z, \\ +\infty, & \text{otherwise}. \end{cases} \tag{1.77}$$

Hard-wall box: potential

[51] In some cases, the function $I(\varphi)$ may somewhat deviate from Eq. (74), but these deviations do not affect its fundamental 2π-periodicity, and hence the fundamental relations (75)-(76). (To the best of the author's knowledge, no corrections to them have been found yet.)

[52] For the discussion of this very general effect, see, e.g., CM Sec. 5.4.

[53] The size of these dc current steps (frequently called the *Shapiro steps*) may be readily calculated from Eqs. (73) and (74). Let me leave this task for the reader's exercise.

[54] Another common name for such potential profiles, especially of lower dimensionality, is the *potential well* – in our current case (77), with a flat bottom and vertical, infinitely high walls. Note also that sometimes such potential profiles are called "quantum wells". The last term is very unfortunate because it seems to imply that particle confinement in potential wells is an effect specific to quantum mechanics. However, as we will repeatedly see in this course, the opposite is true: quantum effects do as much as they only can to *overcome* a particle's confinement in a well, sometimes letting it penetrate the "classically forbidden" regions beyond its walls.

The only way to keep the product $U(\mathbf{r})\psi_n$ in Eq. (65) finite outside the box, is to have $\psi = 0$ in these regions. Also, the function has to be continuous everywhere, to avoid the divergence of the kinetic-energy term $(-\hbar^2/2m)\nabla^2\psi_n$. Hence, in this case, we may solve the stationary Schrödinger equation (65) just inside the box, i.e. with $U = 0$, so it takes a simple form

$$-\frac{\hbar^2}{2m}\nabla^2\psi_n = E_n\psi_n, \tag{1.78a}$$

with zero boundary conditions on all the walls.[55] For our particular geometry, it is natural to express the Laplace operator in the Cartesian coordinates $\{x, y, z\}$ aligned with the box sides, with the origin at one of the corners of its rectangular $a_x \times a_y \times a_z$ volume, so our boundary problem becomes:

$$-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)\psi_n = E_n\psi_n, \quad \text{for } 0<x<a_x, \quad 0<y<a_y, \quad \text{and } 0<z<a_z, \tag{1.78b}$$
$$\text{with } \psi_n = 0 \text{ for } x = 0 \text{ and } x = a_x, \quad y = 0 \text{ and } y = a_y, \quad z = 0 \text{ and } z = a_z.$$

This problem may be readily solved using the same variable separation method as was used in Sec. 5 – now to separate the Cartesian spatial variables from each other, by looking for a partial solution of Eq. (78) in the form

$$\psi(\mathbf{r}) = X(x)Y(y)Z(z). \tag{1.79}$$

(Let us postpone assigning the function indices for a minute.) Plugging this expression into Eq. (78b) and dividing all terms by the product XYZ, we get

$$-\frac{\hbar^2}{2m}\frac{1}{X}\frac{d^2X}{dx^2}-\frac{\hbar^2}{2m}\frac{1}{Y}\frac{d^2Y}{dy^2}-\frac{\hbar^2}{2m}\frac{1}{Z}\frac{d^2Z}{dz^2} = E. \tag{1.80}$$

Now let us repeat the standard argumentation of the variable separation method: since each term on the left-hand side of this equation may be only a function of the corresponding argument, the equality is possible only if each of them is a constant – in our case, with the dimensionality of energy. Calling these constants E_x etc., we get three similar 1D equations

$$-\frac{\hbar^2}{2m}\frac{1}{X}\frac{d^2X}{dx^2} = E_x, \qquad -\frac{\hbar^2}{2m}\frac{1}{Y}\frac{d^2Y}{dy^2} = E_y, \qquad -\frac{\hbar^2}{2m}\frac{1}{Z}\frac{d^2Z}{dx^2} = E_z, \tag{1.81}$$

with Eq. (80) turning into the following energy-matching condition:

$$E_x + E_y + E_z = E. \tag{1.82}$$

All three ordinary differential equations (81), and hence their solutions, are similar. For example, for $X(x)$, we have the following 1D Helmholtz equation

$$\frac{d^2X}{dx^2} + k_x^2 X = 0, \qquad \text{with } k_x^2 \equiv \frac{2mE_x}{\hbar^2}, \tag{1.83}$$

[55] Rewritten as $\nabla^2 f + k^2 f = 0$, Eq. (78a) is just the *Helmholtz equation*, which describes waves of any nature (with the wave vector $\mathbf{k}$) in a uniform, isotropic, linear medium – see, e.g., EM Secs. 7.5-7.9 and 8.5.

with simple boundary conditions: $X(0) = X(a_x) = 0$.[56] Let me hope that the reader knows how to solve this well-known 1D boundary problem – describing, for example, the usual mechanical waves on a guitar string. The problem allows an infinite number of sinusoidal standing-wave eigenfunctions,[57]

$$X \propto \sin k_x x, \quad \text{with } k_x = \frac{\pi n_x}{a_x}, \quad \text{i.e. } X = \left(\frac{2}{a_x}\right)^{1/2} \sin\frac{\pi n_x x}{a_x}, \quad \text{with } n_x = 1, 2, \ldots, \tag{1.84}$$

Rectangular potential well: 1D eigen-functions

corresponding to the following eigenenergies:

$$E_x = \frac{\hbar^2}{2m} k_x^2 = \frac{\pi^2 \hbar^2}{2m a_x^2} n_x^2 \equiv E_{x1} n_x^2 . \tag{1.85}$$

Figure 8 shows these simple results, using a somewhat odd but very graphic and popular representation, in that the eigenenergy values (frequently called the *energy levels*) are used as horizontal axes for plotting the eigenfunctions – despite their different dimensionality.

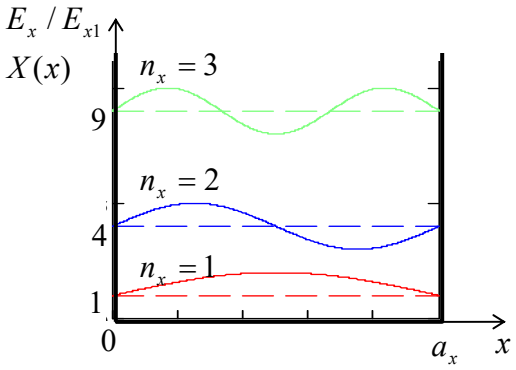


Fig. 1.8. The lowest eigenfunctions (solid lines) and eigenvalues (dashed lines) of Eq. (83) for a 1D potential well of length a_x. Solid black lines show the effective potential energy profile for this 1D eigenproblem.

Due to the similarity of all Eqs. (81), $Y(y)$ and $Z(z)$ are absolutely similar functions of their arguments, and may also be numbered by integers (say, n_y and n_z) independent of n_x, so the spectrum of values of the total energy (82) is

$$E_{n_x,n_y,n_z} = \frac{\pi^2 \hbar^2}{2m}\left(\frac{n_x^2}{a_x^2} + \frac{n_y^2}{a_y^2} + \frac{n_z^2}{a_z^2}\right). \tag{1.86}$$

Rectangular potential well: energy levels

Thus, in this 3D problem, the role of the index n in the general Eq. (69) is played by a set of three independent integers $\{n_x, n_y, n_z\}$. In quantum mechanics, such integers play a key role and thus have a special name, *quantum numbers*. Using them, for our current simple problem that general solution, may be represented as the following sum:

$$\Psi(\mathbf{r},t) = \sum_{n_x,n_y,n_z=1}^{\infty} c_{n_x,n_y,n_z} \sin\frac{\pi n_x x}{a_x} \sin\frac{\pi n_y y}{a_y} \sin\frac{\pi n_z z}{a_z} \exp\left\{-i\frac{E_{n_x,n_y,n_z}}{\hbar} t\right\}, \tag{1.87}$$

Rectangular potential well: general solution

[56] Please notice that we would also arrive at this 1D boundary problem if we considered a 1D analog of our 3D problem (77), i.e. a 1D particle placed in a hard-wall, flat-bottom potential well of length a_x. In quantum mechanics, such 1D problems play an important role, and will be the subject of extensive discussions in the next chapter.

[57] The front coefficient in the last expression for X enforces the (ortho)normality condition (66).

with the front coefficients that may be readily calculated from the initial wavefunction $\Psi(\mathbf{r}, 0)$, using Eq. (68) – again with the replacement $n \to \{n_x, n_y, n_z\}$.

This simple problem is a good illustration of typical results the wave mechanics gives for spatially-confined motion, including the discrete energy spectrum, and (in this case, evidently) orthogonal eigenfunctions. Perhaps most importantly, its solution shows that the lowest value of the particle's kinetic energy (86), reached in the so-called *ground state* (in our problem, the state with $n_x = n_y = n_z = 1$) is above zero for any finite size of the confining volume.

An example of the opposite case of a *continuous spectrum* for the *unconfined motion* of a free particle is given by the plane waves (29). With the account of relations $E = \hbar\omega$ and $\mathbf{p} = \hbar\mathbf{k}$, such wavefunction may be viewed as the product of the time-dependent factor (62) by the eigenfunction,

Free particle: eigen-functions

$$\psi_{\mathbf{k}} = a_{\mathbf{k}} \exp\{i\mathbf{k}\cdot\mathbf{r}\}, \tag{1.88}$$

which is the solution of the stationary Schrödinger equation (78a) if it is valid in the whole space.[58] The reader should not be worried too much by the fact that the fundamental solution (88) in free space is a *traveling* wave (having, in particular, a non-zero value of the probability current $\mathbf{j}$), while those inside a quantum box are *standing* waves with $\mathbf{j} = 0$, even though the free space may be legitimately considered as the ultimate limit of a quantum box with volume $V = a_x \times a_y \times a_z \to \infty$. Indeed, due to the linearity of wave mechanics, two traveling-wave solutions (88) with equal and opposite values of the momentum (and hence with the same energy) may be readily combined to give a standing-wave solution,[59] for example, $\exp\{i\mathbf{k}\cdot\mathbf{r}\} + \exp\{-i\mathbf{k}\cdot\mathbf{r}\} = 2\cos(\mathbf{k}\cdot\mathbf{r})$, with the net current $\mathbf{j} = 0$. Thus, depending on the convenience for a particular problem, we may represent its general solution as a sum of either traveling-wave or standing-wave eigenfunctions. Since in the unlimited free space, there are no boundary conditions to satisfy, the Cartesian components of the wave vector $\mathbf{k}$ in Eq. (88) can take any real values. (This is why it is more convenient to label these wavefunctions, and the corresponding eigenenergies

Free particle: eigen-energies

$$E_{\mathbf{k}} = \frac{\hbar^2 k^2}{2m} \geq 0, \tag{1.89}$$

with their wave vector $\mathbf{k}$ rather than an integer index.)

However, one aspect of continuous-spectrum systems requires a bit more caution with mathematics: the summation (69) should be replaced by the integration over a continuous index or indices – in our current case, the three Cartesian components of the vector $\mathbf{k}$. The main rule of such replacement may be readily extracted from Eq. (84): according to this relation, for standing-wave solutions, the eigenvalues of k_x are *equidistant*, i.e. separated by equal intervals $\Delta k_x = \pi/a_x$, with similar relations for the other two Cartesian components of $\mathbf{k}$. Hence the number of different eigenvalues of the standing-wave vector $\mathbf{k}$ (with $k_x, k_y, k_z \geq 0$), within a volume $d^3k >> 1/V$ of the $\mathbf{k}$ space is $dN = d^3k/(\Delta k_x \Delta k_x \Delta k_x) = (V/\pi^3)d^3k$. Frequently, it is more convenient to work with traveling waves (88); in this case, we should take into account that, as was just discussed, there are two different traveling wave

[58] In some systems (e.g., a particle interacting with a potential well of a finite depth), a discrete energy spectrum within a certain energy interval may coexist with a continuous spectrum in a complementary interval. However, the conceptual philosophy of eigenfunctions and eigenvalues remains the same even in this case.

[59] This is, of course, the general property of waves of any physical nature, propagating in a linear medium – see, e.g., CM Sec. 6.5 and/or EM Sec. 7.3.

numbers (say, $+k_x$ and $-k_x$) corresponding to each standing wave vector's $k_x > 0$. Hence the same number of physically different states corresponds to a $2^3 = 8$-fold larger **k**-space or, equivalently, to an 8-fold smaller number of states per unit volume d^3k:

$$dN = \frac{V}{(2\pi)^3} d^3k . \tag{1.90}$$

Number of 3D states

For $dN >> 1$, this expression is independent of the boundary conditions and is frequently represented as the following *summation rule*

$$\lim_{k^3V \to \infty} \sum_{\mathbf{k}} f(\mathbf{k}) = \int f(\mathbf{k}) dN = \frac{V}{(2\pi)^3} \int f(\mathbf{k}) d^3k , \tag{1.91}$$

Summation over 3D states

where $f(\mathbf{k})$ is any function of **k**. Note that if the same wave vector **k** corresponds to several internal quantum states (such as spin – see Chapter 4), the right-hand side of Eq. (91) requires its multiplication by the corresponding *degeneracy factor* of orbital states.[60]

Finally, note that in systems with reduced wavefunction dimensionality, Eq. (90) for the number of states at large **k** (i.e., for an essentially free particle motion) should be replaced accordingly: in a 2D system of area $A >> 1/k^2$,

$$dN = \frac{A}{(2\pi)^2} d^2k , \tag{1.92}$$

Number of 2D states

while in a 1D system of length $l >> 1/k$,

$$dN = \frac{l}{2\pi} dk , \tag{1.93}$$

Number of 1D states

with the corresponding changes in the summation rule (91). This change has important implications for the density of states on the energy scale, dN/dE: it is straightforward (and hence left for the reader :-) to use Eqs. (90), (99), and (100) to show that for free 3D particles, the density increases with E (proportionally to $E^{1/2}$), for free 2D particles, it does not depend on energy at all, while for free 1D particles, it scales as $E^{-1/2}$, i.e. decreases with energy.

1.8. Exercise problems

1.1. The actual postulate made by N. Bohr in his original 1913 paper was not directly Eq. (8), but rather the assumption that at quantum leaps between adjacent electron orbits with $n >> 1$, the hydrogen atom either emits or absorbs the energy $\Delta E = \hbar\omega$, where ω is its classical radiation frequency – according to classical electrodynamics, equal to the angular velocity of the electron's rotation.[61] Prove that this postulate, complemented with the natural requirement that $L = 0$ at $n = 0$, is equivalent to Eq. (8).

1.2. Generalize the Bohr theory for a hydrogen-like atom/ion with a nucleus with the electric charge $Q = Ze$, to the relativistic case.

[60] The front factor 2 in Eq. (1) for the number of electromagnetic wave modes is just one embodiment of the degeneracy factor, in that case describing two different polarizations of the waves with the same wave vector.
[61] See, e.g., EM Sec. 8.2.

1.3. A hydrogen atom, initially in the lowest excited state, returns to its ground state by emitting a photon propagating in a certain direction. Use the same approach as in Sec. 1(iv) to calculate the photon's frequency reduction due to atomic recoil.

1.4. Use Eq. (53) to prove that the linear operators of quantum mechanics are commutative: $\hat{A}_2 + \hat{A}_1 = \hat{A}_1 + \hat{A}_2$, and associative: $\left(\hat{A}_1 + \hat{A}_2\right) + \hat{A}_3 = \hat{A}_1 + \left(\hat{A}_2 + \hat{A}_3\right)$.

1.5. Prove that for any time-independent Hamiltonian operator $\hat{H}$ and two arbitrary complex functions $f(\mathbf{r})$ and $g(\mathbf{r})$,

$$\int f(\mathbf{r})\hat{H}g(\mathbf{r})d^3r = \int \hat{H}f(\mathbf{r})g(\mathbf{r})d^3r\,.$$

1.6. Prove that the Schrödinger equation (25) with the Hamiltonian operator given by Eq. (41) is Galilean form-invariant, provided that the wavefunction is transformed as

$$\Psi'(\mathbf{r}',t') = \Psi(\mathbf{r},t)\exp\left\{-i\frac{m\mathbf{v}\cdot\mathbf{r}}{\hbar} + i\frac{mv^2t}{2\hbar}\right\},$$

where the prime sign marks the variables observed in the reference frame 0′ that moves, without rotation and with a constant velocity **v**, relative to the "lab" frame 0. Give a physical interpretation of this transformation.

1.7.* Prove the so-called *Hellmann-Feynman theorem*:[62]

$$\frac{\partial E_n}{\partial \lambda} = \left\langle \frac{\partial H}{\partial \lambda} \right\rangle_n,$$

where λ is some *c*-number parameter, on which the time-independent Hamiltonian $\hat{H}$, and hence its eigenenergies E_n, depend.

1.8.* Use Eqs. (73) and (74) to analyze the effect of phase locking of Josephson oscillations on the dc current flowing through a weak link between two superconductors (frequently called the *Josephson junction*), assuming that an external source applies to the junction a sinusoidal ac voltage with frequency ω and amplitude A.

1.9. Calculate $\langle x\rangle$, $\langle p_x\rangle$, δx, and δp_x for the eigenstate $\{n_x, n_y, n_z\}$ of a particle in a rectangular hard-wall box described by Eq. (77) and compare the product $\delta x \delta p_x$ with the Heisenberg's uncertainty relation.

1.10. Looking at the lowest (red) line in Fig. 8, it seems plausible that the lowest-energy eigenfunction (84) of the 1D boundary problem (83) may be well approximated with an inverted quadratic parabola: $X(x) \approx Cx(a_x - x)$, where C is a normalization constant. Explore how good this approximation is.

[62] Despite this common name, H. Hellmann (in 1937) and R. Feynman (in 1939) were not the first ones in the long list of physicists who had (apparently, independently) discovered this equality. Indeed, it has been traced back to a 1922 paper by W. Pauli and was carefully proved by P. Güttinger in 1931.

1.11. A particle placed in a hard-wall rectangular box with sides $\{a_x, a_y, a_z\}$ is in its ground state. Calculate the average force it exerts on each face of the box. Can these forces be characterized by a certain pressure?

1.12. A 1D quantum particle was initially in the ground state of a very deep, flat-bottom potential well of width a:

$$U(x) = \begin{cases} 0, & \text{for } -a/2 < x < +a/2, \\ +\infty, & \text{otherwise.} \end{cases}$$

At some instant, the well's width is abruptly increased to a new value $a' > a$, leaving the potential symmetric with respect to the point $x = 0$, and then is kept constant. Calculate the probability that after the change, the particle is still in the ground state of the system.

1.13. At $t = 0$, a 1D particle of mass m is placed into a hard-wall, flat-bottom potential well

$$U(x) = \begin{cases} 0, & \text{for } 0 < x < a, \\ +\infty, & \text{otherwise,} \end{cases}$$

in a 50/50 linear superposition of the lowest-energy (ground) state and the first excited state. Calculate:

(i) the normalized wavefunction $\Psi(x, t)$ for arbitrary time $t \geq 0$, and
(ii) the time evolution of the expectation value $\langle x \rangle$ of the particle's coordinate.

1.14. Calculate the potential profiles $U(x)$ for which the following wavefunctions,

(i) $\Psi = c \exp\{-ax^2 - ibt\}$, and
(ii) $\Psi = c \exp\{-a|x| - ibt\}$

(with real coefficients $a > 0$ and b), satisfy the 1D Schrödinger equation for a particle with mass m. For each case, calculate $\langle x \rangle$, $\langle p_x \rangle$, δx, and δp_x, and compare the product $\delta x \delta p_x$ with Heisenberg's uncertainty relation.

1.15. The wavefunction of an excited stationary state of a 1D particle moving in a potential profile $U(x)$ is related to that of its ground state as $\psi_e(x) \propto x\psi_g(x)$. Calculate the function $U(x)$.

1.16. A 1D particle of mass m, moving in a potential well $U(x)$, has the following stationary eigenfunction: $\psi(x) = C/\cosh\kappa x$, where C is the normalization constant and κ is a given real constant. Calculate the function $U(x)$ and the state's eigenenergy E.

1.17. Calculate the density dN/dE of traveling-wave quantum states inside large hard-wall rectangular boxes of various dimensions: $d = 1$, 2, and 3.

1.18.* A 1D particle is confined in a potential well of width a, with a flat bottom and hard, infinitely high walls. Use the finite-difference method with steps $a/2$ and $a/3$ to find as many eigenenergies as possible. Compare the results with each other, and with the exact formula. [63]

[63] You may like to start by reading about the finite-difference method – see, e.g., CM Sec. 8.5 or EM Sec. 2.11.

Chapter 2. 1D Wave Mechanics

Even the simplest, 1D version of wave mechanics enables quantitative analysis of many important quantum-mechanical effects. The order of their discussion in this chapter is dictated mostly by mathematical convenience – going from the simplest potential profiles to more complex ones, so that we may build upon the previous results. However, the reader is advised to focus not on the math, but rather on the physics of the non-classical phenomena it describes, ranging from particle penetration into classically-forbidden regions, to quantum-mechanical tunneling, to the metastable state decay, to covalent bonding, to quantum oscillations, to energy bands and gaps.

2.1. Basic relations

In many important cases, the wavefunction may be represented in the form $\Psi(x, t)\chi(y, z)$, where $\Psi(x, t)$ satisfies the 1D version of the Schrödinger equation,[1]

Schrödinger equation

$$i\hbar\frac{\partial\Psi(x,t)}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\Psi(x,t)}{\partial x^2} + U(x,t)\Psi(x,t)\,. \tag{2.1}$$

If the transverse factor $\chi(y, z)$ is normalized:

$$\int |\chi(y,z)|^2\, dydz = 1\,, \tag{2.2}$$

then the similar integration of Eq. (1.22b) over the $[y, z]$ plane gives the following probability of finding the particle on a segment $[x_1, x_2]$:

Probability

$$W(t) \equiv \int_{x_1}^{x_2} \Psi(x,t)\Psi^*(x,t)dx\,. \tag{2.3a}$$

In particular, if the particle under analysis is definitely somewhere inside the system, the normalization of its 1D wavefunction $\Psi(x, t)$ may be provided by extending this integral to the whole axis x:

Normalization

$$\int_{-\infty}^{+\infty} w(x,t)dx = 1, \quad \text{where } w(x,t) \equiv \Psi(x,t)\Psi^*(x,t)\,. \tag{2.3b}$$

Similarly, the $[y, z]$-integration of Eq. (1.23) shows that in this case, the expectation value of any observable depending only on the coordinate x (and possibly time), may be expressed as

Expectation value

$$\langle A\rangle(t) = \int_{-\infty}^{+\infty} \Psi^*(x,t)\hat{A}\Psi(x,t)dx\,, \tag{2.4}$$

and that of Eq. (1.47) makes it valid for the whole *probability current* along the x-axis (a *scalar*):

[1] Note that for this reduction, it is *not* sufficient for the potential energy $U(\mathbf{r}, t)$ to depend on just one spatial coordinate (x). Actually, Eq. (1) is a more robust model for the description of the opposite situations when the potential energy changes within the $[y, z]$ much *faster* than in the x-direction, so that the transverse factor $\chi(y, z)$ is confined in space much more than Ψ, if the confining potential profile is independent of x and t. Let me leave a semi-quantitative analysis of this issue for the reader's exercise. (See also Sec. 3.1.)

© K. Likharev

$$I(x,t) \equiv \int j_x dydz = \frac{\hbar}{m}\mathrm{Im}\left(\Psi^* \frac{\partial}{\partial x}\Psi\right) = \frac{\hbar}{m}|\Psi(x,t)|^2 \frac{\partial \varphi}{\partial x}. \quad (2.5)$$

Probability current

Then the continuity equation (1.48) for any segment $[x_1, x_2]$ takes the form

$$\frac{dW}{dt} + I(x_2) - I(x_1) = 0. \quad (2.6)$$

Continuity equation

The above formulas are sufficient for the analysis of 1D problems of wave mechanics, but before proceeding to particular cases, let me deliver on my earlier promise to prove that Heisenberg's uncertainty relation (1.35) is indeed valid for any wavefunction $\Psi(x, t)$. For that, let us consider the following positive (or at least non-negative) integral:

$$J(\lambda) \equiv \int_{-\infty}^{+\infty} \left| x\Psi + \lambda \frac{\partial \Psi}{\partial x} \right|^2 dx \geq 0, \quad (2.7)$$

where λ is an arbitrary real constant, and assume that at $x \to \pm\infty$ the wavefunction vanishes, together with its first derivative – as we will see below, a very common case. Then the left-hand side of Eq. (7) may be recast as

$$\begin{aligned} J(\lambda) &\equiv \int_{-\infty}^{+\infty} \left| x\Psi + \lambda \frac{\partial \Psi}{\partial x} \right|^2 dx = \int_{-\infty}^{+\infty} \left(x\Psi + \lambda \frac{\partial \Psi}{\partial x} \right)\left(x\Psi + \lambda \frac{\partial \Psi}{\partial x} \right)^* dx \\ &= \int_{-\infty}^{+\infty} x^2 \Psi\Psi^* dx + \lambda \int_{-\infty}^{+\infty} x\left(\Psi \frac{\partial \Psi^*}{\partial x} + \frac{\partial \Psi}{\partial x}\Psi^* \right) dx + \lambda^2 \int_{-\infty}^{+\infty} \frac{\partial \Psi}{\partial x}\frac{\partial \Psi^*}{\partial x} dx. \end{aligned} \quad (2.8)$$

According to Eq. (4), the first term in the last form of Eq. (8) is just $\langle x^2 \rangle$, while the second and the third integrals may be worked out by parts:

$$\int_{-\infty}^{+\infty} x\left(\Psi \frac{\partial \Psi^*}{\partial x} + \frac{\partial \Psi}{\partial x}\Psi^* \right) dx \equiv \int_{-\infty}^{+\infty} x \frac{\partial}{\partial x}\left(\Psi\Psi^*\right) dx = \int_{x=-\infty}^{x=+\infty} x d\left(\Psi\Psi^*\right) = \Psi\Psi^* x \Big|_{x=-\infty}^{x=+\infty} - \int_{-\infty}^{+\infty} \Psi\Psi^* dx = -1, \quad (2.9)$$

$$\int_{-\infty}^{+\infty} \frac{\partial \Psi}{\partial x}\frac{\partial \Psi^*}{\partial x} dx = \int_{x=-\infty}^{x=+\infty} \frac{\partial \Psi}{\partial x} d\Psi^* = \frac{\partial \Psi}{\partial x}\Psi^* \Big|_{x=-\infty}^{x=+\infty} - \int_{-\infty}^{+\infty} \Psi^* \frac{\partial^2 \Psi}{\partial x^2} dx = \frac{1}{\hbar^2}\int_{-\infty}^{+\infty} \Psi^* \hat{p}_x^2 \Psi dx = \frac{\langle p_x^2 \rangle}{\hbar^2}. \quad (2.10)$$

As a result, Eq. (7) takes the following form:

$$J(\lambda) = \langle x^2 \rangle - \lambda + \lambda^2 \frac{\langle p_x^2 \rangle}{\hbar^2} \geq 0, \quad \text{i.e. } \lambda^2 + a\lambda + b \geq 0, \quad \text{with } a \equiv -\frac{\hbar^2}{\langle p_x^2 \rangle},\ b \equiv \frac{\hbar^2 \langle x^2 \rangle}{\langle p_x^2 \rangle}. \quad (2.11)$$

This inequality should be valid for any real λ, so the corresponding quadratic equation, $\lambda^2 + a\lambda + b = 0$, can have either one (degenerate) real root or no real roots at all. This is only possible if its discriminant $a^2 - 4b$, is non-positive, leading to the following requirement:

$$\langle x^2 \rangle \langle p_x^2 \rangle \geq \frac{\hbar^2}{4}. \quad (2.12)$$

In particular, if $\langle x \rangle = 0$ and $\langle p_x \rangle = 0$, then according to Eq. (1.33), Eq. (12) takes the form

$$\langle \tilde{x}^2 \rangle \langle \tilde{p}_x^2 \rangle \geq \frac{\hbar^2}{4}, \quad (2.13)$$

Heisenberg's uncertainty relation

which, according to the definition (1.34) of the r.m.s. uncertainties, is equivalent to Eq. (1.35).[2]

Now let us notice that Heisenberg's uncertainty relation looks very similar to the *commutation relation* between the corresponding operators:

$$[\hat{x}, \hat{p}_x]\Psi \equiv (\hat{x}\hat{p}_x - \hat{p}_x\hat{x})\Psi = x\left(-i\hbar\frac{\partial\Psi}{\partial x}\right) - \left(-i\hbar\frac{\partial}{\partial x}\right)(x\Psi) = i\hbar\Psi\,. \tag{2.14a}$$

Coordinate/ momentum operators' commutator

Since this relation is valid for any wavefunction $\Psi(x, t)$, it may be represented as operator equality:

$$\boxed{[\hat{x}, \hat{p}_x] = i\hbar \neq 0\,.} \tag{2.14b}$$

In Sec. 4.5 we will see that the relation between Eqs. (13) and (14) is just a particular case of a general relation between the expectation values of non-commuting operators and their commutators.

2.2. Free particle: Wave packets

Let us start our discussion of particular problems with the free 1D motion, i.e. with $U(x, t) = 0$. From Eq. (1.29), it is evident that in the 1D case, a similar "fundamental" (i.e. a particular but the most important) solution of the Schrödinger equation (1) is a sinusoidal ("monochromatic") wave

$$\Psi_0(x,t) = \text{const} \times \exp\{i(k_0 x - \omega_0 t)\}. \tag{2.15}$$

According to Eqs. (1.32), it describes a particle with a definite momentum[3] $p_0 = \hbar k_0$ and energy $E_0 = \hbar\omega_0 = \hbar^2 k_0^2/2m$. However, for this wavefunction, the product $\Psi^*\Psi$ does not depend on either x or t, so the particle is completely delocalized, i.e. the probability to find it the same along all axis x, at all times.

In order to describe a space-localized state, let us form, at the initial moment of time ($t = 0$), a wave packet of the type shown in Fig. 1.6, by multiplying the sinusoidal waveform (15) by some smooth *envelope function* $A(x)$. As the most important particular example, consider the *Gaussian wave packet*

Gaussian wave packet: $t = 0$

$$\boxed{\Psi(x,0) = A(x)e^{ik_0 x}, \qquad \text{with } A(x) = \frac{1}{(2\pi)^{1/4}(\delta x)^{1/2}}\exp\left\{-\frac{x^2}{(2\delta x)^2}\right\}.} \tag{2.16}$$

(By the way, Fig. 1.6a shows exactly such a packet.) The pre-exponential factor in this envelope function has been selected to have the initial probability density,

$$w(x,0) \equiv \Psi^*(x,0)\Psi(x,0) = A^*(x)A(x) = \frac{1}{(2\pi)^{1/2}\delta x}\exp\left\{-\frac{x^2}{2(\delta x)^2}\right\}, \tag{2.17}$$

normalized as in Eq. (3b), for any parameters δx and k_0.[4]

---

[2] Eq. (13) may be proved even if $\langle x\rangle$ and $\langle p_x\rangle$ are not equal to zero, by making the replacements $x \to x - \langle x\rangle$ and $\partial/\partial x \to \partial/\partial x + i\langle p\rangle/\hbar$ in Eq. (7), and then repeating all the calculations – which in this case become somewhat bulky. In Chapter 4, equipped with the bra-ket formalism, we will derive a more general uncertainty relation, which includes Heisenberg's relation (13) as a particular case, in a more efficient way.

[3] From this point on to the end of this chapter, I will drop index x in the x-components of the vectors **k** and **p**.

[4] This fact may be readily proved using the well-known integral of the Gaussian function (17), in infinite limits – see, e.g., MA Eq. (6.9b). It is also straightforward to use MA Eq. (6.9c) to prove that for the wave packet (16), the parameter δx is indeed the r.m.s. uncertainty (1.34) of the coordinate x, thus justifying its notation.

To explore the evolution of this wave packet in time, we could try to solve Eq. (1) with the initial condition (16) directly, but in the spirit of the discussion in Sec. 1.5, it is easier to proceed differently. Let us first represent the initial wavefunction (16) as a sum (1.67) of the eigenfunctions $\psi_k(x)$ of the corresponding stationary 1D Schrödinger equation (1.60), in our current case

$$-\frac{\hbar^2}{2m}\frac{d^2\psi_k}{dx^2} = E_k\psi_k, \qquad \text{with } E_k \equiv \frac{\hbar^2 k^2}{2m}, \tag{2.18}$$

which are simply monochromatic waves,

$$\psi_k = a_k e^{ikx}. \tag{2.19}$$

Since (as was discussed in Sec. 1.7) at the unconstrained motion the spectrum of possible wave numbers k is continuous, the sum (1.67) should be replaced with an integral:[5]

$$\Psi(x,0) = \int a_k e^{ikx}\,dk. \tag{2.20}$$

Now let us notice that from the point of view of mathematics, Eq. (20) is just the usual Fourier transform from the variable k to the "conjugate" variable x, and we can use the well-known formula of the reciprocal Fourier transform to write

$$a_k = \frac{1}{2\pi}\int \Psi(x,0)e^{-ikx}dx = \frac{1}{2\pi}\frac{1}{(2\pi)^{1/4}(\delta x)^{1/2}}\int \exp\left\{-\frac{x^2}{(2\delta x)^2} - i\tilde{k}x\right\}dx, \quad \text{where } \tilde{k} \equiv k - k_0. \tag{2.21}$$

This *Gaussian integral* may be worked out by the following standard method, which will be used many times in this course. Let us complement the exponent to the full square of a linear combination of x and k, adding a compensating term independent of x:

$$-\frac{x^2}{(2\delta x)^2} - i\tilde{k}x \equiv -\frac{1}{(2\delta x)^2}\left[x + 2i(\delta x)^2\tilde{k}\right]^2 - \tilde{k}^2(\delta x)^2. \tag{2.22}$$

Since the integration in the right-hand side of Eq. (21) should be performed at constant $\tilde{k}$, in the infinite limits of x, its result would not change if we replace dx with $d\boldsymbol{x} \equiv d[x + 2i(\delta x)^2\tilde{k}]$. As a result, we get:[6]

$$\begin{aligned} a_k &= \frac{1}{2\pi}\frac{1}{(2\pi)^{1/4}(\delta x)^{1/2}}\exp\left\{-\tilde{k}^2(\delta x)^2\right\}\int \exp\left\{-\frac{\boldsymbol{x}^2}{(2\delta x)^2}\right\}d\boldsymbol{x} \\ &= \left(\frac{1}{2\pi}\right)^{1/2}\frac{1}{(2\pi)^{1/4}(\delta k)^{1/2}}\exp\left\{-\frac{\tilde{k}^2}{(2\delta k)^2}\right\}, \end{aligned} \tag{2.23}$$

so a_k also has a Gaussian distribution, now along the k-axis, centered to the value k_0 (Fig. 1.6b), with the constant δk defined as

$$\delta k \equiv 1/2\delta x. \tag{2.24}$$

Thus we may represent the initial wave packet (16) as

[5] For the notation brevity, from this point on the infinite limit signs will be dropped in all 1D integrals.

[6] The fact that the argument's shift is imaginary is not important. (Let me leave proof of this fact for the reader's exercise.)

$$\Psi(x,0)=\left(\frac{1}{2\pi}\right)^{1/2}\frac{1}{(2\pi)^{1/4}(\delta k)^{1/2}}\int\exp\left\{-\frac{(k-k_0)^2}{(2\delta k)^2}\right\}e^{ikx}dk\,. \tag{2.25}$$

From the comparison of this formula with Eq. (16), it is evident that the r.m.s. uncertainty of the wave number k in this packet is indeed equal to δk defined by Eq. (24), thus justifying the notation. The comparison of the last relation with Eq. (1.35) shows that the Gaussian packet represents the ultimate case in which the product $\delta x\delta p = \delta x(\hbar\delta k)$ has the lowest possible value ($\hbar/2$); for any other envelope's shape, the uncertainty product may only be larger.

We could of course get the same result for δk from Eq. (16) using the definitions (1.23), (1.33), and (1.34); the real advantage of Eq. (25) is that it can be readily generalized to $t > 0$. Indeed, we already know that the time evolution of the wavefunction is always given by Eq. (1.69), for our current case[7]

Gaussian wave packet: arbitrary time

$$\Psi(x,t)=\left(\frac{1}{2\pi}\right)^{1/2}\frac{1}{(2\pi)^{1/4}(\delta k)^{1/2}}\int\exp\left\{-\frac{(k-k_0)^2}{(2\delta k)^2}\right\}e^{ikx}\exp\left\{-i\frac{\hbar k^2}{2m}t\right\}dk\,. \tag{2.26}$$

Fig. 1 shows several snapshots of the real part of the wavefunction (26), for a particular case $\delta k = 0.1\,k_0$.

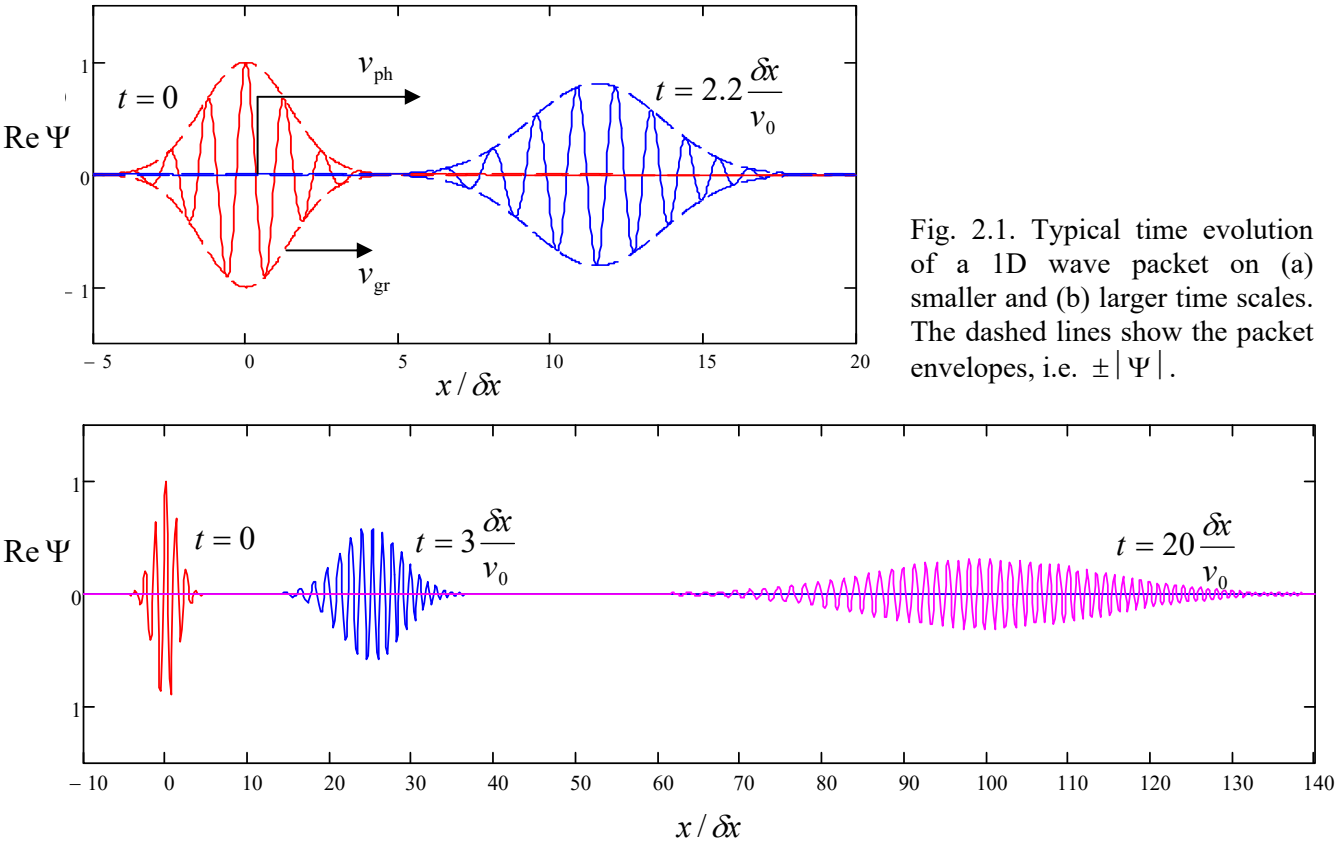


Fig. 2.1. Typical time evolution of a 1D wave packet on (a) smaller and (b) larger time scales. The dashed lines show the packet envelopes, i.e. $\pm|\Psi|$.

The plots clearly show the following effects:

[7] Note that Eq. (26) differs from Eq. (16) only by an exponent of a purely imaginary number, and hence this wavefunction is also properly normalized to 1 – see Eq. (3). Hence the wave packet introduction offers a natural solution to the problem of traveling de Broglie wave's normalization, which was mentioned in Sec. 1.2.

(i) the wave packet as a whole (as characterized by its envelope) moves along the x-axis with a certain *group velocity* v_{gr},

(ii) the "carrier" quasi-sinusoidal wave inside the packet moves with a different, *phase velocity* v_{ph} that may be defined as the velocity of the spatial points where the wave's phase $\varphi(x, t) \equiv \arg\Psi$ takes a certain fixed value (say, $\varphi = \pi/2$ where $\mathrm{Re}\Psi$ vanishes), and

(iii) the wave packet's spatial width gradually increases with time – the packet *spreads*.

All these effects are common for waves of any physical nature.[8] Indeed, let us consider a 1D wave packet of the type (26) but a more general one:

$$\Psi(x,t) = \int a_k e^{i(kx-\omega t)} dk, \tag{2.27}$$

Arbitrary 1D wave packet

propagating in a medium with an arbitrary (but smooth!) dispersion relation $\omega(k)$, and assume that the wave number distribution a_k is narrow: $\delta k << \langle k \rangle \equiv k_0$ – see Fig. 1.6b. Then we may expand the function $\omega(k)$ into the Taylor series near the central wave number k_0, and keep only three of its leading terms:

$$\omega(k) \approx \omega_0 + \frac{d\omega}{dk}\tilde{k} + \frac{1}{2}\frac{d^2\omega}{dk^2}\tilde{k}^2, \qquad \text{where } \tilde{k} \equiv k - k_0,\ \omega_0 \equiv \omega(k_0), \tag{2.28}$$

where the derivatives have to be evaluated at the point $k = k_0$. In this approximation,[9] the expression in the parentheses on the right-hand side of Eq. (27) may be rewritten as

$$\begin{aligned} kx - \omega(k)t &\approx k_0 x + \tilde{k}x - \left(\omega_0 + \frac{d\omega}{dk}\tilde{k} + \frac{1}{2}\frac{d^2\omega}{dk^2}\tilde{k}^2\right)t \\ &\equiv (k_0 x - \omega_0 t) + \tilde{k}\left(x - \frac{d\omega}{dk}t\right) - \frac{1}{2}\frac{d^2\omega}{dk^2}\tilde{k}^2 t, \end{aligned} \tag{2.29}$$

and Eq. (27) becomes

$$\Psi(x,t) \approx e^{i(k_0 x - \omega_0 t)} \int a_k \exp\left\{i\left[\tilde{k}\left(x - \frac{d\omega}{dk}t\right) - \frac{1}{2}\frac{d^2\omega}{dk^2}\tilde{k}^2 t\right]\right\} d\tilde{k}\,. \tag{2.30}$$

First, let us neglect the last term in the square brackets (which is much smaller than the first term if the dispersion relation is smooth enough), and compare the result with the initial form of the wave packet (27):

$$\Psi(x,0) = \int a_k e^{ikx} dk = A(x) e^{ik_0 x}, \quad \text{with } A(x) \equiv \int a_k e^{i\tilde{k}x} d\tilde{k}\,. \tag{2.31}$$

The comparison shows that in this approximation, Eq. (30) is reduced to

$$\Psi(x,t) = A(x - v_{gr}t) e^{ik_0(x - v_{ph}t)}, \tag{2.32}$$

where v_{gr} and v_{ph} are two constants with the dimension of velocity:

$$v_{gr} \equiv \left.\frac{d\omega}{dk}\right|_{k=k_0}, \qquad v_{ph} \equiv \left.\frac{\omega}{k}\right|_{k=k_0}. \tag{2.33a}$$

Group and phase velocities

[8] See, e.g., brief discussions in CM Sec. 6.3 and EM Sec. 7.2.

[9] By the way, in the particular case of de Broglie waves described by the dispersion relation (1.30), Eq. (28) is exact, because $\omega = E/\hbar$ is a quadratic function of $k = p/\hbar$, and all higher derivatives of ω over k vanish for any k_0.

Clearly, Eq. (32) describes the effects (i) and (ii) listed above. For the particular case of the de Broglie waves, whose dispersion law is given by Eq. (1.30),

$$v_{\rm gr} \equiv \frac{d\omega}{dk}\Big|_{k=k_0} = \frac{\hbar k_0}{m} \equiv v_0, \qquad v_{\rm ph} \equiv \frac{\omega}{k}\Big|_{k=k_0} = \frac{\hbar k_0}{2m} = \frac{v_{\rm gr}}{2}. \tag{2.33b}$$

We see that (very fortunately for the correspondence principle :-) the velocity of the wave packet's envelope is equal to v_0 – the classical velocity of the same particle.

Next, the last term in the square brackets of Eq. (30) describes the effect (iii), the wave packet's spread. It may be readily evaluated if the packet (27) is initially Gaussian, as in our example (25):

$$a_k = \text{const} \times \exp\left\{-\frac{\tilde{k}^2}{(2\delta k)^2}\right\}. \tag{2.34}$$

In this case, the integral (30) is Gaussian, and may be worked out exactly as the integral (21), i.e. by representing the merged exponents under the integral as a full square of a linear combination of x and k:

$$\begin{aligned} &-\frac{\tilde{k}^2}{(2\delta k)^2} + i\tilde{k}(x - v_{\rm gr}t) - \frac{i}{2}\frac{d^2\omega}{dk^2}\tilde{k}^2 t \\ &\equiv -\Delta(t)\left(\tilde{k} + i\frac{x - v_{\rm gr}t}{2\Delta(t)}\right)^2 - \frac{(x - v_{\rm gr}t)^2}{4\Delta(t)} + ik_0 x - \frac{i}{2}\frac{d^2\omega}{dk^2}k_0^2 t, \end{aligned} \tag{2.35}$$

where I have introduced the following complex function of time:

$$\Delta(t) \equiv \frac{1}{4(\delta k)^2} + \frac{i}{2}\frac{d^2\omega}{dk^2}t = (\delta x)^2 + \frac{i}{2}\frac{d^2\omega}{dk^2}t, \tag{2.36}$$

and used Eq. (24). Now integrating over $\tilde{k}$, we get

$$\Psi(x,t) \propto \exp\left\{-\frac{(x - v_{\rm gr}t)^2}{4\Delta(t)} + i\left(k_0 x - \frac{1}{2}\frac{d^2\omega}{dk^2}k_0^2 t\right)\right\}. \tag{2.37}$$

The imaginary part of the ratio $1/\Delta(t)$ in this exponent gives just an additional contribution to the wave's phase and does not affect the resulting probability distribution

$$w(x,t) = \Psi^*\Psi \propto \exp\left\{-\frac{(x - v_{\rm gr}t)^2}{2}\,\mathrm{Re}\frac{1}{\Delta(t)}\right\}. \tag{2.38}$$

This is again a Gaussian distribution over the x-axis, centered to point $\langle x\rangle = v_{\rm gr}t$, with the variance

$$(\delta x')^2 \equiv \left\{\mathrm{Re}\left[\frac{1}{\Delta(t)}\right]\right\}^{-1} = (\delta x)^2 + \left(\frac{1}{2}\frac{d^2\omega}{dk^2}t\right)^2 \frac{1}{(\delta x)^2}. \tag{2.39a}$$

In the particular case of de Broglie waves, $d^2\omega/dk^2 = \hbar/m$, so

Wave packet's spread

$$(\delta x')^2 = (\delta x)^2 + \left(\frac{\hbar t}{2m}\right)^2 \frac{1}{(\delta x)^2}. \tag{2.39b}$$

The physics of the packet spreading is very simple: if $d^2\omega/dk^2 \neq 0$, the group velocity $d\omega/dk$ of each small group dk of the monochromatic components of the wave is different, resulting in the gradual (eventually, linear) accumulation of the differences of the distances traveled by the groups. The most curious feature of Eq. (39) is that the packet width at $t > 0$ depends on its initial width $\delta x'(0) = \delta x$ in a non-monotonic way, tending to infinity at both $\delta x \to 0$ and $\delta x \to \infty$. Because of that, for a given time interval t, there is an optimal value of δx that minimizes $\delta x'$:

$$(\delta x')_{\min} = \sqrt{2}\,(\delta x)_{\text{opt}} = \left(\frac{\hbar t}{m}\right)^{1/2}. \tag{2.40}$$

This expression may be used to estimate the spreading effect's magnitude. Due to the smallness of the Planck constant $\hbar$ on the human scale of things, for macroscopic bodies the spreading is extremely small even for very long time intervals; however, for light particles, it may be very noticeable: for an electron ($m = m_e \approx 10^{-30}$ kg), and $t = 1$ s, Eq. (40) yields $(\delta x')_{\min} \sim 1$ cm.

Note also that for any $t \neq 0$, the wave packet retains its Gaussian envelope, but the ultimate relation (24) is *not* satisfied, $\delta x' \delta p > \hbar/2$, due to a gradually accumulated phase shift between the component monochromatic waves.

The last remark on this topic: in quantum mechanics, the wave packet spreading is *not* a ubiquitous effect! For example, in Chapter 5 we will see that in a quantum oscillator, the spatial width of a Gaussian packet (for that system, called the *Glauber state* of the oscillator) does not grow monotonically but rather either stays constant or oscillates in time.

Now let us briefly discuss the case when the initial wave packet is not Gaussian but is described by an arbitrary initial wavefunction. To make the forthcoming result more aesthetically pleasing, it is beneficial to generalize our calculations to an arbitrary initial time t_0; it is evident that if U does not depend on time explicitly, it is sufficient to replace t with $(t - t_0)$ in the above formulas. With this replacement, Eq. (27) becomes

$$\Psi(x,t) = \int a_k e^{i\left[kx - \omega(t - t_0)\right]} dk\,, \tag{2.41}$$

and the reciprocal transform (21) reads

$$a_k = \frac{1}{2\pi}\int \Psi(x,t_0)\, e^{-ikx} dx\,. \tag{2.42}$$

If we want to express these two formulas with one relation, i.e. plug Eq. (42) into Eq. (41), we should give the integration variable x some other name, e.g., x_0. (Such notation is appropriate because this variable describes the coordinate argument in the initial wave packet.) The result is

$$\Psi(x,t) = \frac{1}{2\pi}\int dk \int dx_0\, \Psi(x_0,t_0)\, e^{i\left[k(x - x_0) - \omega(t - t_0)\right]}\,. \tag{2.43}$$

Changing the order of integration, this expression may be represented in the following general form:

$$\boxed{\Psi(x,t) = \int G(x,t;x_0,t_0)\,\Psi(x_0,t_0)\,dx_0\,,} \tag{2.44}$$

1D propagator: definition

where the function G, usually called *kernel* in mathematics, in quantum mechanics is called the *propagator*.[10] Its physical sense may be understood by considering the following special initial condition:[11]

$$\Psi(x_0,t_0)=\delta(x_0-x'), \tag{2.45}$$

where x' is a certain point within the particle's motion domain. In this particular case, Eq. (44) gives

$$\Psi(x,t)=G(x,t;x',t_0). \tag{2.46}$$

Hence, the propagator, considered as a function of its arguments x and t only, is just the wavefunction of the particle, at the δ-functional initial conditions (45). Thus, just as Eq. (41) may be understood as a mathematical expression of the linear superposition principle in the momentum (i.e., reciprocal) space domain, Eq. (44) is an expression of this principle in the direct space domain: the system's "response" $\Psi(x,t)$ to an arbitrary initial condition $\Psi(x_0,t_0)$ is just a sum of its responses to elementary spatial "slices" of this initial function, with the propagator $G(x,\ t;\ x_0,\ t_0)$ representing the weight of each slice in the final sum.

According to Eqs. (43) and (44), in the case of a free particle, the propagator is equal to

$$G(x,t;x_0,t_0)=\frac{1}{2\pi}\int e^{i\left[k(x-x_0)-\omega(t-t_0)\right]}dk\,, \tag{2.47}$$

Calculating this integral, one should remember that here ω is not a constant but a function of k, given by the dispersion relation for the partial waves. In particular, for the de Broglie waves, with $\hbar\omega=\hbar^2k^2/2m$,

$$G(x,t;x_0,t_0)\equiv\frac{1}{2\pi}\int\exp\left\{i\left[k(x-x_0)-\frac{\hbar k^2}{2m}(t-t_0)\right]\right\}dk\,. \tag{2.48}$$

This is a Gaussian integral again, and it may be readily calculated just it was done (twice) above, by completing the exponent to the full square. The result is

Free particle's propagator

$$G(x,t;x_0,t_0)=\left[\frac{m}{2\pi i\hbar(t-t_0)}\right]^{1/2}\exp\left\{-\frac{m(x-x_0)^2}{2i\hbar(t-t_0)}\right\}. \tag{2.49}$$

Please note the following features of this complex function:

(i) It depends only on the differences $(x-x_0)$ and $(t-t_0)$. This is natural because the free-particle propagation problem is *translation-invariant* both in space and time.

(ii) The function's shape (Fig. 2) does not depend on its arguments – they just rescale the same function: as a function of x, it just becomes broader and lower with time. It is curious that the spatial broadening scales as $(t-t_0)^{1/2}$ – just as at the classical diffusion, indicating a deep mathematical analogy between quantum mechanics and classical statistics – to be discussed further in Chapter 7.

[10] Its standard notation by letter G stems from the fact that the propagator is essentially the spatial-temporal *Green's function* of the Schrödinger equation (1), defined very similarly to Green's functions of other ordinary and partial differential equations describing various physics systems – see, e.g., CM Sec. 5.1 and/or EM Sec. 2.7 and 7.3.

[11] Note that this initial condition is mathematically *not* equivalent to a δ-functional initial probability density (3).

(iii) In accordance with the uncertainty relation, the ultimately compressed wave packet (45) has an infinite width of momentum distribution, and the quasi-sinusoidal tails of the free-particle's propagator, clearly visible in Fig. 2, are the results of the free propagation of the fastest (highest-momentum) components of that distribution, in both directions from the packet's center.

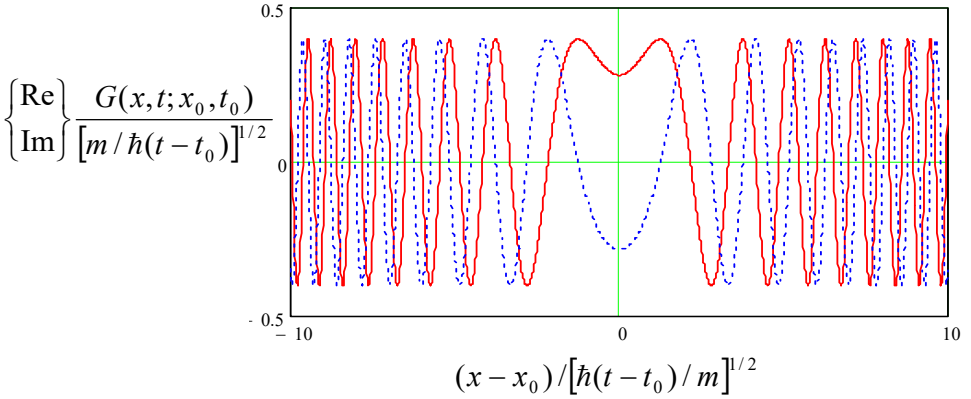


Fig. 2.2. The real (solid line) and imaginary (dotted line) parts of the 1D free particle's propagator (49).

In the following sections, I will mostly focus on monochromatic wavefunctions (which, for unconfined motion, may be interpreted as wave packets of a very large spatial width δx), and only rarely discuss wave packets. My best excuse is the linear superposition principle, i.e. our conceptual ability to restore the general solution from that of monochromatic waves of all possible energies. However, the reader should not forget that, as the above discussion has illustrated, mathematically such restoration is not always trivial.

2.3. Particle reflection and tunneling

Now, let us proceed to the cases when a 1D particle moves in various potential profiles $U(x)$ that are constant in time. Conceptually, the simplest of such profiles is a potential step – see Fig. 3.

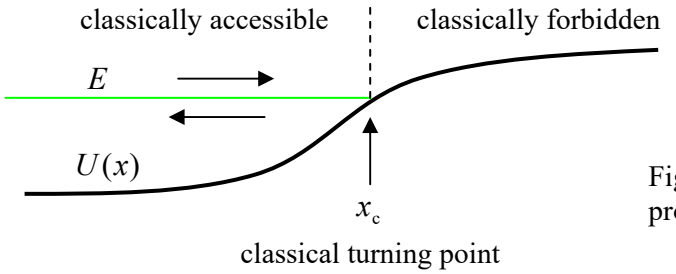


Fig. 2.3. Classical 1D motion in a potential profile $U(x)$.

As I am sure the reader knows, in classical mechanics the particle's kinetic energy $p^2/2m$ cannot be negative, so if the particle is incident on such a step (in Fig. 3, from the left), it can only move within the *classically accessible* region where its (conserved) full energy,

$$E = \frac{p^2}{2m} + U(x), \tag{2.50}$$

is larger than the local value $U(x)$. Let, for example, the initial velocity $v = p/m$ be positive, i.e. directed toward the step. Before it has reached the *classical turning point* x_c, defined by equality

$$U(x_c) = E\,, \tag{2.51}$$

the particle's kinetic energy $p^2/2m$ is positive, so it continues to move in the initial direction. On the other hand, a classical particle cannot penetrate that *classically forbidden region* $x > x_c$, because there, its kinetic energy would be negative. Hence when the particle reaches the point $x = x_c$, its velocity has to change its sign, i.e. the particle is reflected back from the classical turning point.

In order to see what the wave mechanics says about this situation, let us start from the simplest, sharp potential step shown with the bold black line in Fig. 4:

$$U(x) = U_0\theta(x) \equiv \begin{cases} 0, & \text{at } x < 0, \\ U_0, & \text{at } 0 < x. \end{cases} \tag{2.52}$$

For this choice, and any energy within the interval $0 < E < U_0$, the classical turning point is $x_c = 0$.

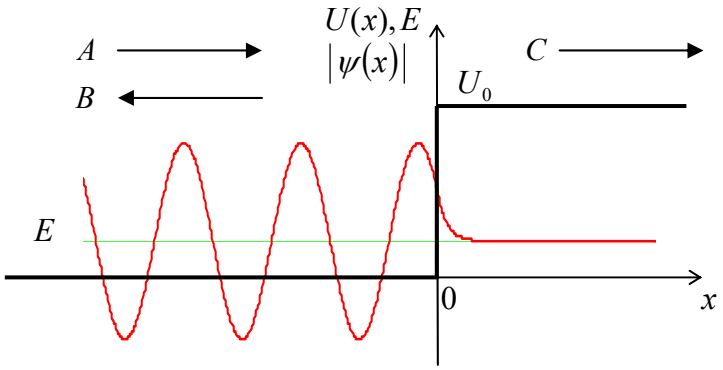


Fig. 2.4. Reflection of a monochromatic de Broglie wave from a potential step $U_0 > E$. (This particular wavefunction's shape is for $U_0 = 5E$.) The wavefunction is plotted with the same schematic vertical offset by E as those in Fig. 1.8.

Let us represent the incident particle with a wave packet so long that the spread $\delta k \sim 1/\delta x$ of its wave-number spectrum is sufficiently small to make the energy uncertainty $\delta E = \hbar\delta\omega = \hbar(d\omega/dk)\delta k$ negligible in comparison with its average value $E < U_0$, as well as with $(U_0 - E)$. In this case, E may be considered as a given constant, the time dependence of the wavefunction is given by Eq. (1.62), and we can calculate its spatial factor $\psi(x)$ from the 1D version of the stationary Schrödinger equation (1.65):[12]

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + U(x)\psi = E\psi\,. \tag{2.53}$$

At $x < 0$, i.e. at $U = 0$, the equation is reduced to the Helmholtz equation (1.78), and may be satisfied with either of two traveling waves, proportional to, respectively, $\exp\{+ikx\}$ and $\exp\{-ikx\}$, with k satisfying the dispersion equation (1.30):

$$k^2 \equiv \frac{2mE}{\hbar^2}. \tag{2.54}$$

Thus the general solution of Eq. (53) in this region may be represented as

Incident and reflected waves

$$\boxed{\psi_-(x) = Ae^{+ikx} + Be^{-ikx}, \qquad \text{for } x < 0\,.} \tag{2.55}$$

[12] Note that this is *not* an eigenproblem like the one we have solved in Sec. 1.4 for a potential well. Indeed, now the energy E is considered given – e.g., by the initial conditions that launch a long wave packet upon the potential step – in Fig. 4, from the left side.

The second term on the right-hand side of Eq. (55) evidently describes a (formally, infinitely long) wave packet traveling to the left, arising because of the particle's reflection from the potential step. If $B = -A$, Eq. (55) is reduced to Eq. (1.84) for a potential well with infinitely high walls, but for our current case of a finite step height U_0, the relation between the coefficients B and A may be different.

To show this, let us solve Eq. (53) for $x > 0$, where $U = U_0 > E$. In this region, the equation may be rewritten as

$$\frac{d^2\psi_+}{dx^2} = \kappa^2 \psi_+ , \tag{2.56}$$

where κ is a real and positive constant defined by a formula similar in structure to Eq. (54):

$$\kappa^2 \equiv \frac{2m(U_0 - E)}{\hbar^2} > 0. \tag{2.57}$$

Decay in classically forbidden region

The general solution of Eq. (56) is the sum of $\exp\{+\kappa x\}$ and $\exp\{-\kappa x\}$, with arbitrary pre-exponential coefficients. However, in our particular case the wavefunction should be finite at $x \to +\infty$, so only the latter exponent is acceptable:

$$\psi_+(x) = Ce^{-\kappa x}, \qquad \text{for } x > 0 . \tag{2.58}$$

Such penetration of the wavefunction into the classically forbidden region, and hence a non-zero probability to find the particle there, is one of the most fascinating predictions of quantum mechanics, which has been repeatedly observed in experiment – for example, via tunneling experiments, to be discussed in the next section.[13] From Eq. (58), it is evident that the constant κ, defined by Eqs. (57), may be interpreted as the reciprocal penetration depth. Even for the lightest particles, this depth is usually very small. Indeed, for any $E << U_0$ that relation yields

$$\delta \equiv \frac{1}{\kappa}\Big|_{E\to 0} = \frac{\hbar}{(2mU_0)^{1/2}} . \tag{2.59}$$

For example, let us consider a conduction electron in a typical metal, which runs, at the metal's surface, into a sharp potential step whose height is equal to the metal's workfunction $U_0 \approx 5$ eV – see the discussion of the photoelectric effect in Sec. 1.1. In this case, according to Eq. (59), δ is close to 0.1 nm, i.e. is close to a typical size of an atom. For heavier elementary particles (e.g., protons) the penetration depth is correspondingly lower, and for macroscopic bodies, it is hardly measurable.

Returning to Eqs. (55) and (58), we still should relate the coefficients B and C to the amplitude A of the incident wave, using the boundary conditions at $x = 0$. Since E is a finite constant, and $U(x)$ is a finite function, Eq. (53) says that $d^2\psi/dx^2$ should be finite as well. This means that the first derivative should be continuous:

$$\lim_{\varepsilon\to 0}\left(\frac{d\psi}{dx}\Big|_{x=+\varepsilon} - \frac{d\psi}{dx}\Big|_{x=-\varepsilon}\right) = \lim_{\varepsilon\to 0}\int_{-\varepsilon}^{+\varepsilon}\frac{d^2\psi}{dx^2}dx = \frac{2m}{\hbar^2}\lim_{\varepsilon\to 0}\int_{-\varepsilon}^{+\varepsilon}[U(x) - E]\psi\, dx = 0 . \tag{2.60}$$

Repeating such calculation for the wavefunction $\psi(x)$ itself, we see that it also should be continuous at all points, including the border point $x = 0$, so the boundary conditions in our problem are

[13] Note that this effect is also pertinent to *classical waves* of any type, including mechanical waves (see, e.g., CM Secs. 6.4 and 7.7) and electromagnetic waves (see, e.g., EM Secs. 7.3-7.7), but not *classical particles*.

$$\psi_-(0) = \psi_+(0), \qquad \frac{d\psi_-}{dx}(0) = \frac{d\psi_+}{dx}(0). \tag{2.61}$$

Plugging Eqs. (55) and (58) into Eqs. (61), we get a system of two linear equations

$$A + B = C, \qquad ikA - ikB = -\kappa C, \tag{2.62}$$

whose (easy :-) solution allows us to express B and C via A :

$$B = A\frac{k - i\kappa}{k + i\kappa}, \qquad C = A\frac{2k}{k + i\kappa}. \tag{2.63}$$

We immediately see that the numerator and denominator in the first of these fractions have equal moduli, so $|B| = |A|$. This means that, as we could expect, a particle with energy $E < U_0$ is totally reflected from the step – just as in classical mechanics. As a result, our solution (55) for $x < 0$ may be represented as a standing wave:

$$\psi_- = 2iAe^{i\theta}\sin(kx - \theta), \qquad \text{with } \theta \equiv \tan^{-1}\frac{k}{\kappa}. \tag{2.64}$$

Note that the shift $\Delta x \equiv \theta/k = (\tan^{-1}k/\kappa)/k$ of the standing wave to the right, due to the partial penetration of the wavefunction under the potential step, is commensurate with, but generally not equal to the penetration depth $\delta \equiv 1/\kappa$. The red line in Fig. 4 shows the exact behavior of the wavefunction, for a particular case $E = U_0/5$, at which $k/\kappa \equiv [E/(U_0 - E)]^{1/2} = 1/2$.

According to Eq. (59), as the particle's energy E is increased to approach U_0, the penetration depth $1/\kappa$ diverges. This raises an important question: what happens at $E > U_0$, i.e. if there is no classically forbidden region in the problem? In classical mechanics, the incident particle would continue to move to the right, though with a reduced velocity corresponding to the new kinetic energy $E - U_0$, so there would be no reflection. In quantum mechanics, however, the situation is different. To analyze it, it is not necessary to the whole problem again; it is sufficient to note that all our calculations, and hence Eqs. (63) are still valid if we take[14]

$$\kappa = -ik', \qquad \text{with} \qquad k'^2 \equiv \frac{2m(E - U_0)}{\hbar^2} > 0. \tag{2.65}$$

With this replacement, Eq. (63) becomes[15]

$$B = A\frac{k - k'}{k + k'}, \quad C = A\frac{2k}{k + k'}. \tag{2.66}$$

The most important result of this change is that now the particle's reflection is *not* total: $|B| < |A|$. To evaluate this effect quantitatively, it is fairer to use not the B/A or C/A ratios, but rather that of the probability currents (5) carried by the de Broglie waves traveling to the right, with amplitudes C and A, in the corresponding regions (respectively, for $x > 0$ and $x < 0$):

[14] Our earlier discarding of the particular solution $\exp\{\kappa x\}$, now becoming $\exp\{-ik'x\}$, is still valid, but now on different grounds: this term would describe a wave packet incident on the potential step from the right, and this is not the problem under our current consideration.

[15] These formulas are completely similar to those describing the partial reflection of classical waves from a sharp interface between two uniform media, at normal incidence (see, e.g., CM Sec. 6.4 and EM Sec. 7.4), with the effective impedance Z of de Broglie waves being proportional to their wave number k.

$$\mathcal{T} \equiv \frac{I_C}{I_A} = \frac{k'|C|^2}{k|A|^2} = \frac{4kk'}{(k+k')^2} \equiv \frac{4[E(E-U_0)]^{1/2}}{\left[E^{1/2}+(E-U_0)^{1/2}\right]^2}. \quad (2.67)$$

Potential step's transparency

(The parameter $\mathcal{T}$ so defined is called the *transparency* of the system, in our current case of the potential step of height U_0, at particle's energy E.) The result given by Eq. (67) is plotted in Fig. 5a as a function of the U_0/E ratio. Note its most important features:

(i) At $U_0 = 0$, the transparency is full, $\mathcal{T} = 1$ – naturally, because there is no step at all.

(ii) At $U_0 \to E$, the transparency drops to zero, giving a proper connection to the case $E < U_0$.

(iii) Nothing in our solution's procedure prevents us from using Eq. (67) even for $U_0 < 0$, i.e. for the *step-down* (or "cliff") potential profile – see Fig. 5b. Very counter-intuitively, the particle is (partly) reflected even from such a cliff, and the transmission diminishes (though rather slowly) at $U_0 \to -\infty$.

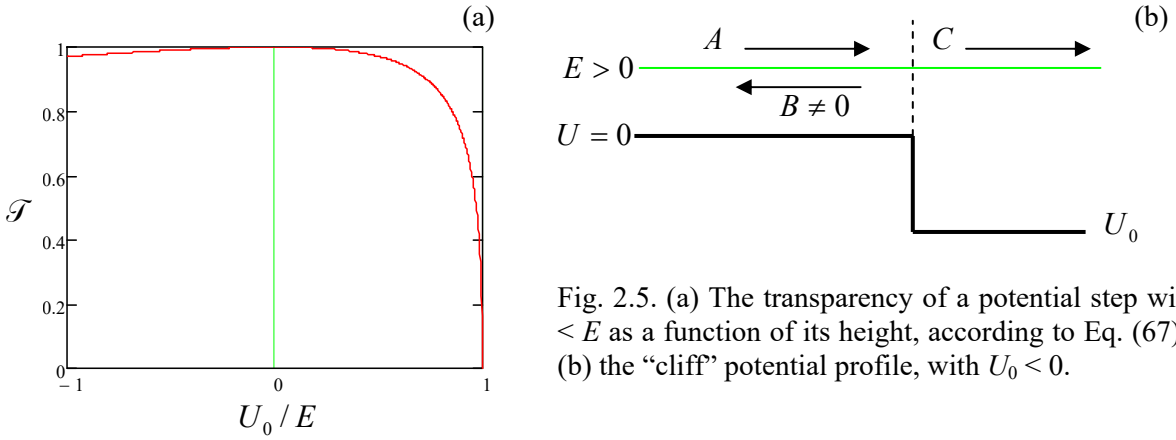


Fig. 2.5. (a) The transparency of a potential step with $U_0 < E$ as a function of its height, according to Eq. (67), and (b) the "cliff" potential profile, with $U_0 < 0$.

The most important conceptual conclusion of this analysis is that the quantum particle is *partly reflected* from a potential step with $U_0 < E$, in the sense that there is a non-zero probability $\mathcal{T} < 1$ to find it passed over the step, while there is also some probability, $(1 - \mathcal{T}) > 0$, to have it reflected.

The last property is exhibited, but for *any* relation between E and U_0, by another simple potential profile $U(x)$, the famous *potential* (or "tunnel") *barrier*. Fig. 6 shows its simple flat-top ("rectangular") version:

$$U(x) = \begin{cases} 0, & \text{for } x < -d/2, \\ U_0, & \text{for } -d/2 < x < +d/2, \\ 0, & \text{for } +d/2 < x. \end{cases} \quad (2.68)$$

To analyze this problem, it is sufficient to look for the solution to the Schrödinger equation in the form (55) at $x \le -d/2$. At $x > +d/2$, i.e., behind the barrier, we may use the arguments presented above (no wave source on the right!) to keep just one traveling wave, now with the same wave number:

$$\psi_+(x) = Fe^{ikx}. \quad (2.69)$$

However, under the barrier, i.e. at $-d/2 \le x \le +d/2$, we should generally keep both exponential terms,

$$\psi_b(x) = Ce^{-\kappa x} + De^{+\kappa x}, \quad (2.70)$$

because our previous argument used in the potential step problem's solution is no longer valid. (Here k and κ are still defined, respectively, by Eqs. (54) and (57).) In order to express the coefficients B, C, D, and F via the amplitude A of the incident wave, we need to plug these solutions into the boundary conditions similar to Eqs. (61), but now at two boundary points, $x = \pm d/2$.

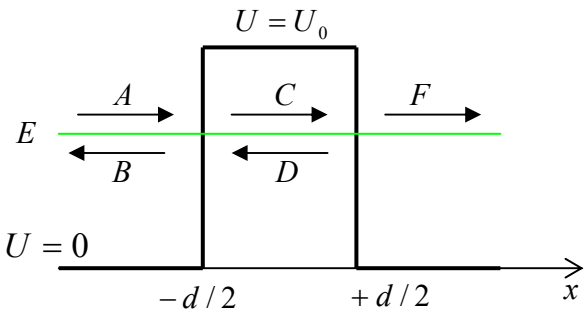


Fig. 2.6. A rectangular potential barrier, and the de Broglie waves taken into account in its analysis.

Solving the resulting system of 4 linear equations, we get 4 ratios B/A, C/A, etc.; in particular,

$$\frac{F}{A} = \left[\cosh \kappa d + \frac{i}{2}\left(\frac{\kappa}{k} - \frac{k}{\kappa}\right)\sinh \kappa d\right]^{-1} e^{-ikd}, \tag{2.71a}$$

and hence the barrier's transparency

Rectangular tunnel barrier's transparency

$$\mathcal{T} \equiv \left|\frac{F}{A}\right|^2 = \left[\cosh^2 \kappa d + \left(\frac{\kappa^2 - k^2}{2\kappa k}\right)^2 \sinh^2 \kappa d\right]^{-1} \equiv \left[1 + \left(\frac{\kappa^2 + k^2}{2\kappa k}\right)^2 \sinh^2 \kappa d\right]^{-1}. \tag{2.71b}$$

So, quantum mechanics indeed allows particles with energies $E < U_0$ to pass "through" the potential barrier – see Fig. 6 again. This is the famous effect of *quantum-mechanical tunneling*. Fig. 7a shows the barrier transparency as a function of the particle energy E, for several characteristic values of its thickness d, or rather of the ratio d/δ, with δ defined by Eq. (59).[16]

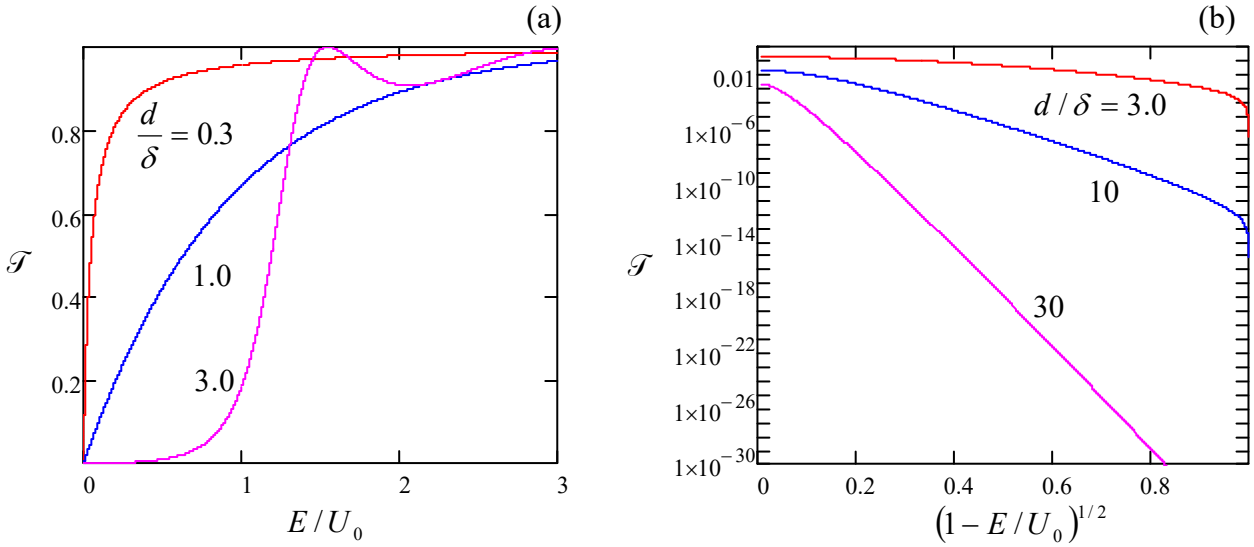


Fig. 2.7. The transparency of a rectangular potential barrier as a function of the particle's energy E.

[16] The branches with $E > U_0$ have been computed from Eq. (71b) by using the replacement (65).

The plots show that generally, the transparency grows with the particle's energy. This growth is natural because the penetration constant κ decreases with the growth of E, i.e., the wavefunction penetrates more and more into the barrier, so more and more of it is "picked up" at the second interface ($x = +d/2$) and transferred into the wave $F\exp\{ikx\}$ propagating behind the barrier.

Now let us consider the important limit of a very thin and high rectangular barrier with $d << \delta$, $E << U_0$, giving $k << \kappa << 1/d$. In this limit, the second form of Eq. (71b) yields

$$\mathcal{T} \to \frac{1}{1+\alpha^2}, \qquad \text{where } \alpha \equiv \left(\frac{\kappa^2+k^2}{2\kappa k}\right)\kappa d \approx \frac{\kappa^2 d}{2k} \approx \frac{m}{\hbar^2 k}U_0 d\,, \tag{2.72}$$

The last product, $U_0 d$, is just the *weight* (or the "energy area")

$$\mathcal{W} \equiv \int_{U(x)>E} U(x)dx \tag{2.73}$$

of the barrier for our particular case (68). This fact implies that the very simple result (72) may be correct for a barrier of any shape, provided that it is sufficiently thin and high.

To examine this guess, let us consider the tunneling problem for a very thin barrier with κd, $kd << 1$, by approximating it with the Dirac's δ-function (Fig. 8):

$$U(x) = \mathcal{W}\delta(x)\,, \tag{2.74}$$

so that the parameter $\mathcal{W}$ satisfies Eq. (73).

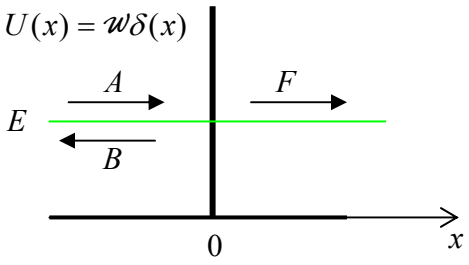


Fig. 2.8. A delta-functional potential barrier.

The solutions of the tunneling problem at all points but $x = 0$ still may be taken in the form of Eqs. (55) and (69), so we only need to analyze the boundary conditions at that point. However, due to the special character of the δ-function, we should be careful here. Indeed, instead of Eq. (60) we now get

$$\begin{aligned}\lim_{\varepsilon\to 0}\left(\frac{d\psi}{dx}\Big|_{x=+\varepsilon} - \frac{d\psi}{dx}\Big|_{x=-\varepsilon}\right) &= \lim_{\varepsilon\to 0}\int_{-\varepsilon}^{+\varepsilon}\frac{d^2\psi}{dx^2}dx = \lim_{\varepsilon\to 0}\frac{2m}{\hbar^2}\int_{-\varepsilon}^{+\varepsilon}[U(x)-E]\psi\,dx \\ &= \frac{2m}{\hbar^2}\mathcal{W}\psi(0).\end{aligned} \tag{2.75}$$

According to this relation, at a finite $\mathcal{W}$, the derivatives $d\psi/dx$ are also finite, so the wavefunction itself is still continuous:

$$\lim_{\varepsilon\to 0}\left(\psi\big|_{x=+\varepsilon} - \psi\big|_{x=+\varepsilon}\right) = \lim_{\varepsilon\to 0}\int_{-\varepsilon}^{+\varepsilon}\frac{d\psi}{dx}dx = 0. \tag{2.76}$$

Using these two boundary conditions, we readily get the following system of two linear equations,

$$A + B = F, \qquad ikF - (ikA - ikB) = \frac{2m\mathcal{W}}{\hbar^2} F, \tag{2.77}$$

whose solution yields

$$\frac{B}{A} = \frac{-i\alpha}{1+i\alpha}, \quad \frac{F}{A} = \frac{1}{1+i\alpha}, \qquad \text{where } \alpha \equiv \frac{m\mathcal{W}}{\hbar^2 k}. \tag{2.78}$$

(Taking Eq. (73) into account, this definition of α coincides with that in Eq. (72).) For the barrier transparency $\mathscr{T} \equiv |F/A|^2$, this result again gives the first of Eqs. (72), which is therefore general for such thin barriers. That formula may be recast to give the following simple expression (valid only for $E << U_{max}$):

Transparency: thin barrier

$$\mathscr{T} = \frac{1}{1+\alpha^2} \equiv \frac{E}{E+E_0}, \qquad \text{where } E_0 \equiv \frac{m\mathcal{W}^2}{2\hbar^2}, \tag{2.79}$$

which shows that as energy becomes larger than the constant E_0, the transparency approaches 1 – as it eventually does for *any* tunnel barrier. Another general behavior of Eq. (79) is that at $E \to 0$ (i.e. $k \to 0$), $\mathscr{T} \propto E$. Indeed, at $ka << 1$, *any* barrier of a finite width a may be well approximated with Eq. (74).

As Eq. (71) shows, in the opposite limit of relatively thick barriers ($d >> \delta$), the transparency is dominated by what is called the *tunnel exponent*:

Transparency: thick barrier

$$\mathscr{T} = \left(\frac{4k\kappa}{k^2+\kappa^2} \right)^2 e^{-2\kappa d} \tag{2.80}$$

– the behavior which may be clearly seen as the straight-line segments in semi-log plots (Fig. 7b) of $\mathscr{T}$ as a function of the combination $(1 - E/U_0)^{1/2}$ which is proportional to κ – see Eq. (57).

The exponential dependence of the barrier's transparency on its thickness is the most important factor for various applications of quantum-mechanical tunneling – from the field emission of electrons to vacuum[17] to the scanning tunneling microscopy.[18] Note also substantial negative implications of the effect on the electronic technology progress. Most importantly, it imposes limits on the so-called *Dennard scaling* of field-effect transistors in semiconductor integrated circuits (which was the technological basis of the well-known *Moore's law*), due to the increase of tunneling both through the gate oxide and along the channel of the transistors, from source to drain.[19]

Finally, one more feature visible in Fig. 7a (for case $d = 3\delta$) are the oscillations of the transparency as a function of energy, at $E > U_0$, with $\mathscr{T} = 1$, i.e. the reflection completely vanishing, at some points.[20] This is our first glimpse at one more interesting quantum effect: *resonant tunneling*. This

[17] See, e.g., G. Fursey, *Field Emission in Vacuum Microelectronics*, Kluwer, New York, 2005.

[18] See, e.g., G. Binning and H. Rohrer, *Helv. Phys. Acta* **55**, 726 (1982).

[19] See, e.g., V. Sverdlov *et al.*, *IEEE Trans. on Electron Devices* **50**, 1926 (2003), and references therein. (A brief discussion of the field-effect transistors, and literature for further reading, may be found in SM Sec. 6.4.)

[20] Let me mention in passing the curious case of the potential well $U(x) = -(\hbar^2/2m)\nu(\nu+1)/\cosh^2(x/a)$, with any positive integer ν and any real a, which is reflection-free ($\mathscr{T} = 1$) for the incident de Broglie wave of *any* energy E, and hence for any incident wave packet. (The well is called the *Pöschl-Teller potential*, though it was first described in a 1930 paper by P. Epstein, before the 1933 publication by G. Pöschl and E. Teller.)

effect will be discussed in more detail in Sec. 5 below, by using another potential profile where it is more clearly pronounced.

2.4. Motion in soft potentials[21]

Before moving on to explore other quantum-mechanical effects, let us examine how the results discussed in the previous section are modified in the opposite limit of the so-called *soft* (or "smooth") potential profiles, like the one sketched in Fig. 3.[22] The most efficient analytical tool to study this limit is the so-called *WKB* (or "JWKB", or "quasiclassical") *approximation* developed by H. Jeffrey, G. Wentzel, A. Kramers, and L. Brillouin in 1925-27. In order to derive its 1D version, let us rewrite the Schrödinger equation (53) in a simpler form

$$\frac{d^2\psi}{dx^2} + k^2(x)\psi = 0, \tag{2.81}$$

where the local wave number $k(x)$ is defined similarly to Eq. (65),

$$k^2(x) \equiv \frac{2m[E - U(x)]}{\hbar^2}, \tag{2.82}$$

Local wave number

besides that now it may be a function of x. We already know that for $k(x) = \text{const}$, the fundamental solutions of this equation are $A\exp\{+ikx\}$ and $B\exp\{-ikx\}$, which may be represented in a single form

$$\psi(x) = e^{i\Phi(x)}, \tag{2.83}$$

where $\Phi(x)$ is a complex function, in these two simplest cases being equal, respectively, to $(kx - i\ln A)$ and $(-kx - i\ln B)$. This is why we may try to use Eq. (83) to look for a solution of Eq. (81) even in the general case when $k(x) \neq \text{const}$. Differentiating Eq. (83) twice, we get

$$\frac{d\psi}{dx} = i\frac{d\Phi}{dx}e^{i\Phi}, \qquad \frac{d^2\psi}{dx^2} = \left[i\frac{d^2\Phi}{dx^2} - \left(\frac{d\Phi}{dx}\right)^2\right]e^{i\Phi}. \tag{2.84}$$

Plugging the last expression into Eq. (81) and canceling the common factor $\exp\{i\Phi(x)\} \neq 0$, we get

$$i\frac{d^2\Phi}{dx^2} - \left(\frac{d\Phi}{dx}\right)^2 + k^2(x) = 0. \tag{2.85}$$

This is still an exact, general equation. Superficially, it looks even harder to solve than the initial equation (81) because Eq. (85) is nonlinear. However, it is ready for simplification in the limit when the potential profile is soft, $dU/dx \to 0$. Indeed, for a uniform potential, $d^2\Phi/dx^2 = 0$. Hence, in the so-called *0^{th} approximation*, $\Phi(x) \to \Phi_0(x)$, we may try to keep that equality, so Eq. (85) is reduced to

$$\left(\frac{d\Phi_0}{dx}\right)^2 = k^2(x), \qquad \text{i.e. } \frac{d\Phi_0}{dx} = \pm k(x), \qquad \Phi_0(x) = \pm i\int^x k(x')dx', \tag{2.86}$$

so its general solution is a linear superposition of the two functions (83), with Φ replaced with Φ_0:

[21] Following tradition, I will frequently use this shorthand for "potential energy", returning to the full term in cases where there is any chance of confusion between this notion and another (say, electrostatic) potential.

[22] Quantitative conditions of the "softness" will be formulated later in this section.

$$\psi_0(x) = A\exp\left\{+i\int^x k(x')dx'\right\} + B\exp\left\{-i\int^x k(x')dx'\right\}, \tag{2.87}$$

where the choice of the lower limits of integration affects only the constants A and B. The physical sense of this result is simple: it is a sum of the forward- and back-propagating de Broglie waves, with the coordinate-dependent local wave number $k(x)$ adjusted to the potential profile.

Let me emphasize the non-trivial nature of this approximation.[23] First, any attempt to address the problem with the standard perturbation approach (say, $\psi = \psi_0 + \psi_1 + \ldots$, with ψ_n proportional to the n^{th} power of some small parameter) would fail for most potentials, because as Eq. (86) shows, even a slight but persisting deviation of $U(x)$ from a constant leads to a gradual accumulation of the phase Φ_0, impossible to describe by any small perturbation of ψ. Second, the dropping of the term $d^2\Phi/dx^2$ in Eq. (85) is not too easy to justify. Indeed, since we are committed to the "soft potential limit" $dU/dx \to 0$, we should be ready to assume the characteristic length a of the spatial variation of Φ to be large, and neglect the terms that are the smallest ones in the limit $a \to \infty$. However, both first terms in Eq. (85) are apparently of the same order in a, namely $O(a^{-2})$; why have we neglected just one of them?

The price we have paid for such a "sloppy" treatment is substantial: Eq. (87) does *not* satisfy the fundamental property of solutions of the stationary Schrödinger equation, namely the probability current's conservation: $I(x)$ = const. However, this is not true for any component of Eq. (87); for example for the first, forward-propagating component on its right-hand side, Eq. (5) yields

$$I_0(x) = \frac{\hbar}{m}\left|A\right|^2 k(x), \tag{2.88}$$

evidently not a constant if $k(x) \neq$ const. The brilliance of the WKB theory is that the problem may be fixed without a full revision of the 0th approximation, just by amending it. Indeed, let us explore the next, 1st approximation:

$$\Phi(x) \to \Phi_{\text{WKB}}(x) \equiv \Phi_0(x) + \Phi_1(x), \tag{2.89}$$

where Φ_0 still obeys Eq. (86), while Φ_1 describes a 0th approximation's correction that is small in the following sense:[24]

$$\left|\frac{d\Phi_1}{dx}\right| << \left|\frac{d\Phi_0}{dx}\right| = k(x). \tag{2.90}$$

Plugging Eq. (89) into Eq. (85), with the account of the definition (86), we get

$$i\left(\frac{d^2\Phi_0}{dx^2} + \frac{d^2\Phi_1}{dx^2}\right) - \frac{d\Phi_1}{dx}\left(2\frac{d\Phi_0}{dx} + \frac{d\Phi_1}{dx}\right) = 0. \tag{2.91}$$

Using the condition (90), we may neglect $d^2\Phi_1/dx^2$ in comparison with $d^2\Phi_0/dx^2$ inside the first parentheses, and $d\Phi_1/dx$ in comparison with $2d\Phi_0/dx$ inside the second parentheses. As a result, we get the following (still approximate!) result:

[23] Philosophically, this space-domain method is very close to the time-domain *van der Pol method* in classical mechanics, and the very similar *rotating wave approximation* (RWA) in quantum mechanics – see, e.g., CM Secs. 5.2-5.5, and also Secs. 6.5, 7.6, 9.2, and 9.4 of this course.

[24] For certainty, I will use the discretion given by Eq. (82) to define $k(x)$ as the *positive* root of its right-hand side.

$$\frac{d\Phi_1}{dx} = \frac{i}{2}\frac{d^2\Phi_0}{dx^2}\Big/\frac{d\Phi_0}{dx} \equiv \frac{i}{2}\frac{d}{dx}\left(\ln\frac{d\Phi_0}{dx}\right) = \frac{i}{2}\frac{d}{dx}\left[\ln k(x)\right] \equiv i\frac{d}{dx}\left[\ln k^{1/2}(x)\right], \tag{2.92}$$

$$i\Phi\big|_{\text{WKB}} \equiv i\Phi_0 + i\Phi_1 = \pm i\int^x k(x')dx' + \ln\frac{1}{k^{1/2}(x)}, \tag{2.93}$$

$$\psi_{\text{WKB}}(x) = \frac{a}{k^{1/2}(x)}\exp\left\{i\int^x k(x')dx'\right\} + \frac{b}{k^{1/2}(x)}\exp\left\{-i\int^x k(x')dx'\right\}. \qquad \text{for } k^2(x) > 0. \tag{2.94}$$

WKB wave-function

(Again, the lower integration limit is arbitrary, because its choice may be incorporated into the complex constants a and b.) This modified approximation overcomes the problem of current continuity; for example, for the forward-propagating wave, Eq. (5) gives

$$I_{\text{WKB}}(x) = \frac{\hbar}{m}|a|^2 = \text{const}. \tag{2.95}$$

WKB probability current

Physically, the factor $k^{1/2}$ in the denominator of the WKB wavefunction's pre-exponent may be easily understood. The smaller the local group velocity (32) of the wave packet, $v_{\text{gr}}(x) = \hbar k(x)/m$, the "easier" (more probable) it should be to find the particle within a certain interval dx. This is exactly the result that the WKB approximation gives: $w(x) = \psi\psi^* \propto 1/k(x) \propto 1/v_{\text{gr}}$. Another value of the 1st approximation is a clarification of the WKB theory's validity condition: it is given by Eq. (90). Plugging into this relation the first form of Eq. (92), and estimating $|d^2\Phi_0/dx^2|$ as $|d\Phi_0/dx|/a$, where a is the spatial scale of a substantial change of $|\,d\Phi_0/dx\,| = k(x)$, we may write the condition as

$$ka >> 1. \tag{2.96}$$

WKB: first condition of validity

In plain English, this means that the region where $U(x)$, and hence $k(x)$, change substantially should contain many de Broglie wavelengths $\lambda = 2\pi/k$.

So far I have implied that $k^2(x) \propto E - U(x)$ is positive, i.e. particle moves in the classically accessible region. Now let us extend the WKB approximation to situations where the difference $E - U(x)$ may change sign, for example to the reflection problem sketched in Fig. 3. Just as we did for the sharp potential step, we first need to find the appropriate solution in the classically forbidden region, in this case for $x > x_c$. For that, there is again no need to redo our calculations, because they are still valid if we, just as in the sharp-step problem, take $k(x) = i\kappa(x)$, where

$$\kappa^2(x) \equiv \frac{2m[U(x) - E]}{\hbar^2} > 0, \qquad \text{for } x > x_c, \tag{2.97}$$

and keep just one of two possible solutions (with $\kappa > 0$), in analogy with Eq. (58). The result is

$$\psi_{\text{WKB}}(x) = \frac{c}{\kappa^{1/2}(x)}\exp\left\{-\int^x \kappa(x')dx'\right\}, \qquad \text{for } k^2 < 0, \text{ i.e. } \kappa^2 > 0, \tag{2.98}$$

with the lower limit at some point with $\kappa^2 > 0$ as well. This is a really wonderful formula! It describes the quantum-mechanical penetration of the particle into the classically forbidden region and provides a natural generalization of Eq. (58) – leaving intact our estimates of the depth $\delta \sim 1/\kappa$ of such penetration.

Now we have to do what had been done for the sharp-step problem in Sec. 2: use the boundary conditions at the classical turning point $x = x_c$ to relate the constants a, b, and c. However, now this operation is a tad more complex, because both WKB functions (94) and (98) diverge, albeit weakly, at the point, because here both $k(x)$ and $\kappa(x)$ tend to zero. This *connection problem* may be solved in the following way. [25]

Let us use our commitment to the potential's "softness", assuming that it allows us to keep just two leading terms in the Taylor expansion of the function $U(x)$ at the point x_c:

$$U(x) \approx U(x_c) + \left.\frac{dU}{dx}\right|_{x=x_c}(x - x_c) \equiv E + \left.\frac{dU}{dx}\right|_{x=x_c}(x - x_c). \tag{2.99}$$

Using this truncated expansion, and introducing the following dimensionless variable for the coordinate's deviation from the classical turning point,

$$\zeta \equiv \frac{x - x_c}{x_0}, \qquad \text{with } x_0 \equiv \left[\frac{\hbar^2}{2m\left(dU/dx\right)_{x=x_c}}\right]^{1/3}, \tag{2.100}$$

we reduce the Schrödinger equation (81) to the so-called *Airy equation*

Airy equation

$$\boxed{\frac{d^2\psi}{d\zeta^2} - \zeta\psi = 0}. \tag{2.101}$$

This simple linear, ordinary, homogenous differential equation of the second order has been very well studied. Its general solution may be represented as a linear combination of two fundamental solutions, the *Airy functions*, $\mathrm{Ai}(\zeta)$ and $\mathrm{Bi}(\zeta)$, shown in Fig. 9a.[26]

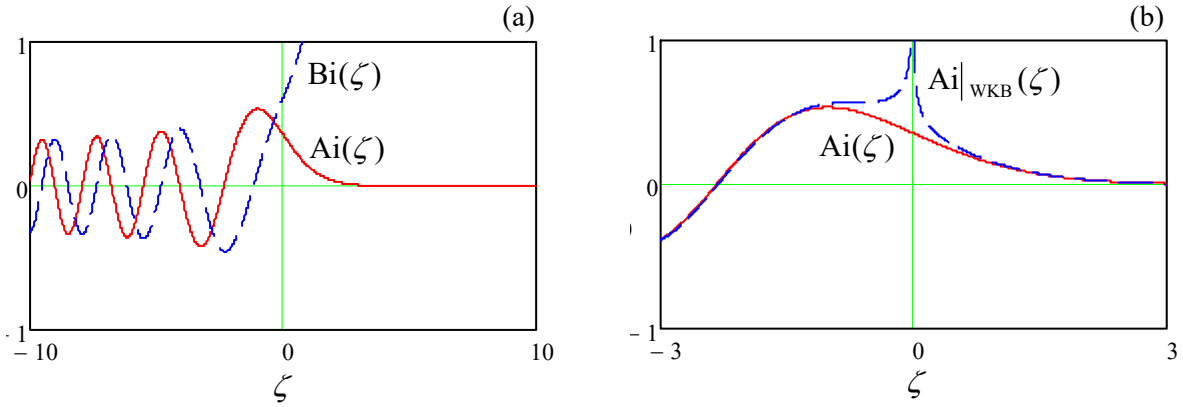


Fig. 2.9. (a) The Airy functions Ai and Bi, and (b) the WKB approximation for the function $\mathrm{Ai}(\zeta)$.

[25] An alternative way to solve the connection problem, without involving the Airy functions but using an analytical extension of WKB formulas to the complex-argument plane, may be found, e.g., in Sec. 47 of the textbook by L. Landau and E. Lifshitz, *Quantum Mechanics*, *Non-Relativistic Theory*, 3rd ed. Pergamon, 1977.

[26] Note the following (exact) integral formulas,

$$\mathrm{Ai}(\zeta) = \frac{1}{\pi}\int_0^\infty \cos\left(\frac{\xi^3}{3} + \zeta\xi\right)d\xi, \qquad \mathrm{Bi}(\zeta) = \frac{1}{\pi}\int_0^\infty \left[\exp\left\{-\frac{\xi^3}{3} + \zeta\xi\right\} + \sin\left(\frac{\xi^3}{3} + \zeta\xi\right)\right]d\xi,$$

frequently more convenient for practical calculations of the Airy functions than the differential equation (101).

The latter function diverges at $\zeta \to +\infty$, and thus is not suitable for our current problem (Fig. 3), while the former function has the following asymptotic behaviors at $|\zeta| >> 1$:

$$\mathrm{Ai}(\zeta) \to \frac{1}{\pi^{1/2}|\zeta|^{1/4}} \times \begin{cases} \frac{1}{2}\exp\left\{-\frac{2}{3}\zeta^{3/2}\right\}, & \text{for } \zeta \to +\infty, \\ \sin\left\{\frac{2}{3}(-\zeta)^{3/2} + \frac{\pi}{4}\right\}, & \text{for } \zeta \to -\infty. \end{cases} \tag{2.102}$$

Now let us apply the WKB approximation to the Airy equation (101). Taking the classical turning point ($\zeta = 0$) for the lower limit, for $\zeta > 0$ we get

$$\kappa^2(\zeta) = \zeta, \qquad \kappa(\zeta) = \zeta^{1/2}, \qquad \int_0^{\zeta} \kappa(\zeta')d\zeta' = \frac{2}{3}\zeta^{3/2}, \tag{2.103}$$

i.e. exactly the exponent in the top line of Eq. (102). Making a similar calculation for $\zeta < 0$, with the natural assumption $|b| = |a|$ (full reflection from the potential step), we arrive at the following result:

$$\mathrm{Ai}_{\mathrm{WKB}}(\zeta) = \frac{1}{|\zeta|^{1/4}} \times \begin{cases} c'\exp\left\{-\frac{2}{3}\zeta^{3/2}\right\}, & \text{for } \zeta > 0, \\ a'\sin\left\{\frac{2}{3}(-\zeta)^{3/2} + \varphi\right\}, & \text{for } \zeta < 0. \end{cases} \tag{2.104}$$

This approximation differs from the exact solution at small values of ζ, i.e. close to the classical turning point – see Fig. 9b. However, at $|\zeta| >> 1$, Eqs. (104) describe the Airy function exactly, provided that

$$\varphi = \frac{\pi}{4}, \qquad c' = \frac{a'}{2}. \tag{2.105}$$

WKB: connection formulas

These *connection formulas* may be used to rewrite Eq. (104) as

$$\mathrm{Ai}_{\mathrm{WKB}}(\zeta) = \frac{a'}{2|\zeta|^{1/4}} \times \begin{cases} \exp\left\{-\frac{2}{3}\zeta^{3/2}\right\}, & \text{for } \zeta > 0, \\ \frac{1}{i}\left[\exp\left\{+i\frac{2}{3}\zeta^{3/2} + i\frac{\pi}{4}\right\} - \exp\left\{-i\frac{2}{3}\zeta^{3/2} - i\frac{\pi}{4}\right\}\right], & \text{for } \zeta < 0, \end{cases} \tag{2.106}$$

and hence may be described by the following two simple mnemonic rules:

(i) If the classical turning point is taken for the lower limit in the WKB integrals in the classically allowed and the classically forbidden regions, then the moduli of the quasi-amplitudes of the exponents are equal.

(ii) Reflecting from a "soft" potential step, the wavefunction acquires an additional phase shift $\Delta\varphi = \pi/2$, if compared with its reflection from a "hard", infinitely high potential wall located at point x_c (for which, according to Eq. (63) with $\kappa = \infty$, we have $B = -A$).

In order for the connection formulas (105)-(106) to be valid, deviations from the linear approximation (99) of the potential profile should be relatively small within the region where the WKB approximation differs from the exact Airy function: $|\zeta| \sim 1$, i.e. $|x - x_c| \sim x_0$. These deviations may be estimated using the next term of the Taylor expansion, dropped in Eq. (99): $(d^2U/d^2x)(x - x_c)^2/2$. As a result, the condition of validity of the connection formulas (i.e. of the "softness" of the reflecting

potential profile) may be expressed as $\left|d^2U/d^2x\right|_{x_0} << \left|dU/dx\right|$ at $x \approx x_c$ – meaning the ~x_0–wide vicinity of the point x_c). With the account of Eq. (100) for x_0, this condition becomes

Connection formulas' validity

$$\left|\frac{d^2U}{dx^2}\right|^3_{x\approx x_c} << \frac{2m}{\hbar^2}\left(\frac{dU}{dx}\right)^4_{x\approx x_c}. \tag{2.107}$$

As an example of a very useful application of the WKB approximation, let us use the connection formulas to calculate the energy spectrum of a 1D particle in a soft 1D potential well (Fig. 10).

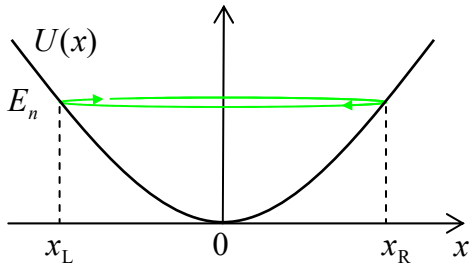


Fig. 2.10. The WKB treatment of an eigenstate of a particle in a soft 1D potential well.

As was discussed in Sec. 1.7, we may consider the standing wave describing an eigenfunction ψ_n (corresponding to an eigenenergy E_n) as a sum of two traveling de Broglie waves going back and forth between the walls, being sequentially reflected from each of them. Let us apply the WKB approximation to such traveling waves. First, according to Eq. (94), propagating from the left classical turning point x_L to the right such point x_R, it acquires the phase change

$$\Delta\varphi_{\rightarrow} = \int_{x_L}^{x_R} k(x)dx\,. \tag{2.108}$$

At the reflection from the soft wall at x_R, according to the mnemonic rule (ii), the wave acquires an additional shift $\pi/2$. Now, traveling back from x_R to x_L, the wave gets a shift similar to one given by Eq. (108): $\Delta\varphi_{\leftarrow} = \Delta\varphi_{\rightarrow}$. Finally, at the reflection from x_L, it gets one more $\pi/2$-shift. Summing up all these contributions at the wave's roundtrip, we may write the self-consistency condition (that the wavefunction "catches its own tail with its teeth") in the form

$$\Delta\varphi_{\text{total}} \equiv \Delta\varphi_{\rightarrow} + \frac{\pi}{2} + \Delta\varphi_{\leftarrow} + \frac{\pi}{2} \equiv 2\int_{x_L}^{x_R} k(x)dx + \pi = 2\pi m, \quad \text{with } n = 1, 2, ... \tag{2.109}$$

Rewriting this result in terms of the particle's momentum $p(x) = \hbar k(x)$, we arrive at the so-called *Wilson-Sommerfeld* (or, less fairly, "Bohr-Sommerfeld") *quantization rule*

Wilson-Sommerfeld quantization rule

$$\oint_C p(x)dx = 2\pi\hbar\left(n - \frac{1}{2}\right), \tag{2.110}$$

where the closed path C means the full period of classical motion.[27]

[27] Note that at the motion in more than one dimension, a closed classical trajectory may have no classical turning points. In this case, the constant ½, arising from the turns, should be dropped from Eqs. (110) written for the scalar product $\mathbf{p}(\mathbf{r})\cdot d\mathbf{r}$, giving the so-called *Bohr quantization* rule. It was suggested by N. Bohr as early as 1913 as an interpretation of Eq. (1.8) for the circular motion of the electron around the proton, while its 1D modification (110) is due to W. Wilson (1915) and A. Sommerfeld (1916).

Let us see what this quantization rule gives for the very important particular case of a quadratic potential profile of a harmonic oscillator of frequency ω_0. In this case,

$$U(x) = \frac{m}{2}\omega_0^2 x^2, \tag{2.111}$$

and the classical turning points (where $U(x) = E$) are the roots of a simple equation

$$\frac{m}{2}\omega_0^2 x_c^2 = E_n, \qquad \text{so that } x_R = \frac{1}{\omega_0}\left(\frac{2E_n}{m}\right)^{1/2} > 0, \quad x_L = -x_R < 0. \tag{2.112}$$

Due to the potential's symmetry, the integration required by Eq. (110) is also simple:

$$\begin{aligned} \int_{x_L}^{x_R} p(x)dx &= \int_{x_L}^{x_R} \{2m[E_n - U(x)]\}^{1/2} dx \equiv (2mE_n)^{1/2} 2\int_0^{x_R}\left(1 - \frac{x^2}{x_R^2}\right)^{1/2} dx \\ &\equiv (2mE_n)^{1/2} 2x_R \int_0^1 (1-\xi^2)^{1/2} d\xi = (2mE_n)^{1/2} 2x_R \frac{\pi}{4} \equiv \frac{\pi E_n}{\omega_0}, \end{aligned} \tag{2.113}$$

so Eq. (110) yields

$$E_n = \hbar\omega_0\left(n' + \frac{1}{2}\right), \qquad \text{with } n' \equiv n - 1 = 0, 1, 2, \ldots. \tag{2.114}$$

To estimate the validity of this result, we have to check the condition (96) at all points of the classically allowed region, and Eq. (107) at the turning points. The checkup shows that both conditions are valid only for $n >> 1$. However, we will see in Sec. 9 below that Eq. (114) is actually exactly correct for *all* energy levels – thanks to the special properties of the potential profile (111).

Now let us use the mnemonic rule (i) to examine the particle's penetration into the classically forbidden region of an abrupt potential step of a height $U_0 > E$. For this case, the rule, i.e. the second of Eqs. (105), yields the following relation of the quasi-amplitudes in Eqs. (94) and (98): $|c| = |a|/2$. If we now naively applied this relation to the sharp step sketched in Fig. 4, forgetting that it does not satisfy Eq. (107), we would get the following relation of the full amplitudes defined by Eqs. (55) and (58):

$$\left|\frac{C}{\kappa}\right| = \frac{1}{2}\left|\frac{A}{k}\right|. \qquad \text{(WRONG!)} \tag{2.115}$$

This result differs from the correct Eq. (63), and hence we may expect that the WKB approximation's prediction for more complex potentials, most importantly for tunneling through a soft potential barrier (Fig. 11) should be also different from the exact result (71) for the rectangular barrier shown in Fig. 6.

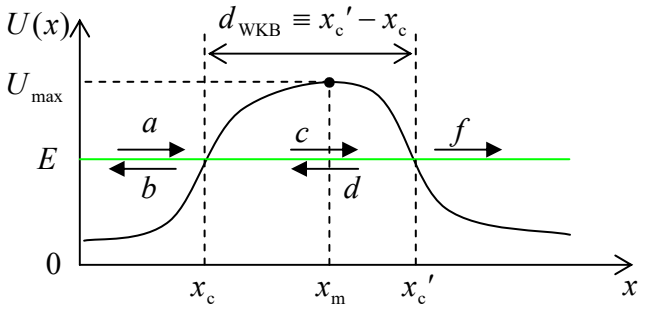


Fig. 2.11. Tunneling through a soft 1D potential barrier.

To analyze tunneling through such a soft barrier, we need (just as in the case of a rectangular barrier) to take into consideration five partial waves, but now they should be taken in the WKB form:

$$\psi_{\rm WKB} = \begin{cases} \dfrac{a}{k^{1/2}(x)}\exp\left\{i\int^{x} k(x')dx'\right\} + \dfrac{b}{k^{1/2}(x)}\exp\left\{-i\int^{x} k(x')dx'\right\}, & \text{for } x < x_{\rm c}, \\ \dfrac{c}{\kappa^{1/2}(x)}\exp\left\{-\int^{x} \kappa(x')dx'\right\} + \dfrac{d}{\kappa^{1/2}(x)}\exp\left\{\int^{x} \kappa(x')dx'\right\}, & \text{for } x_{\rm c} < x < x_{\rm c}', \\ \dfrac{f}{k^{1/2}(x)}\exp\left\{i\int^{x} k(x')dx'\right\}, & \text{for } x_{\rm c}' < x, \end{cases} \tag{2.116}$$

where the lower limits of integrals are arbitrary (each within the corresponding range of x). Since on the right of the left classical point, we have two exponents rather than one, and on the right of the second point, one traveling wave rather than two, the connection formulas (105) have to be generalized by using asymptotic formulas not only for Ai(ζ), but also for the second Airy function, Bi(ζ). The analysis, similar to that carried out above (though naturally a bit bulkier),[28] gives a remarkably simple result:

Soft potential barrier: transparency

$$\mathcal{T}_{\rm WKB} \equiv \left|\frac{f}{a}\right|^2 = \exp\left\{-2\int_{x_{\rm c}}^{x_{\rm c}'} \kappa(x)dx\right\} \equiv \exp\left\{-\frac{2}{\hbar}\int_{x_{\rm c}}^{x_{\rm c}'} \left(2m[U(x)-E]\right)^{1/2} dx\right\}, \tag{2.117}$$

with the pre-exponential factor equal to 1 – the fact that might be readily expected from the mnemonic rule (i) of the connection formulas.

This formula is broadly used in applied quantum mechanics, despite the approximate character of its pre-exponential coefficient for insufficiently soft barriers that do not satisfy Eq. (107). For example, Eq. (80) shows that for a rectangular barrier with thickness $d >> \delta$, the WKB approximation (117) with $d_{\rm WKB} = d$ underestimates $\mathcal{T}$ by a factor of $[4k\kappa/(k^2 + \kappa^2)]^2$ – equal, for example, 4, if $k = \kappa$, i.e. if $U_0 = 2E$. However, on the appropriate logarithmic scale (see Fig. 7b), such a factor, smaller than an order of magnitude, is just a small correction.

Note also that when E approaches the barrier's top $U_{\rm max}$ (Fig. 11), the points $x_{\rm c}$ and $x_{\rm c}'$ merge, so according to Eq. (117), $\mathcal{T}_{\rm WKB} \to 1$, i.e. the particle reflection vanishes at $E = U_{\rm max}$. So, the WKB approximation does not describe the effect of the over-barrier reflection at $E > U_{\rm max}$. (This fact could be noticed already from Eq. (95): in the absence of the classical turning points, the WKB probability current is constant for any barrier profile.) This conclusion is incorrect even for apparently smooth barriers where one could naively expect the WKB approximation to work perfectly. Indeed, near the point $x = x_{\rm m}$ where the potential reaches maximum (i.e. $U(x_{\rm m}) = U_{\rm max}$), we may approximate almost any smooth function $U(x)$ with the quadratic term of the Taylor expansion, i.e. with an inverted parabola:

$$U(x) \approx U_{\rm max} - \frac{m\omega_0^2 (x - x_{\rm m})^2}{2}. \tag{2.118}$$

Calculating the derivatives dU/dx and d^2U/dx^2 for this function and plugging the results into the condition (107), we may see that the WKB approximation is only valid if $|U_{\rm max} - E| >> \hbar\omega_0$. Just for the

[28] For the most important case $\mathcal{T}_{\rm WKB} << 1$, Eq. (117) may be simply derived from Eqs. (105)-(106) – the exercise left for the reader.

reader's reference, an exact analysis of tunneling through the barrier (118) gives the following *Kemble formula*:[29]

$$\mathcal{T} = \frac{1}{1+\exp\{-2\pi(E-U_{\max})/\hbar\omega_0\}}, \tag{2.119}$$

Kemble formula

valid for any sign of the difference $(E - U_{max})$. This formula describes a gradual approach of $\mathcal{T}$ to 1, i.e. a *gradual* reduction of reflection, at the particle energy's increase, with $\mathcal{T} = ½$ at $E = U_{max}$.

To conclude this section, one last remark: the WKB approximation opens a straight way toward an alternative formulation of quantum mechanics, based on the *Feynman path integral*. However, I will postpone its discussion until a more compact notation has been introduced in Chapter 4.

2.5. Resonant tunneling, and metastable states

Now let us move to other, conceptually different quantum effects taking place in more elaborate potential profiles. Neither piecewise-constant nor smooth-potential models of $U(x)$ are convenient for their quantitative description because they both require "stitching" partial de Broglie waves at each classical turning point, which may lead to cumbersome calculations. However, we may get a very good insight into the physics of quantum effects that may take place in such profiles, while avoiding overcomplicated math, by approximating them with sets of Dirac's delta functions.

Big help in this way is provided by the notions of *scattering* and *transfer matrices*, very useful for other purposes as well. Let us consider an arbitrary but finite-length potential "bump" (formally called a *scatterer*), localized somewhere between points x_1 and x_2, on a flat potential background, say $U = 0$ (Fig. 12).

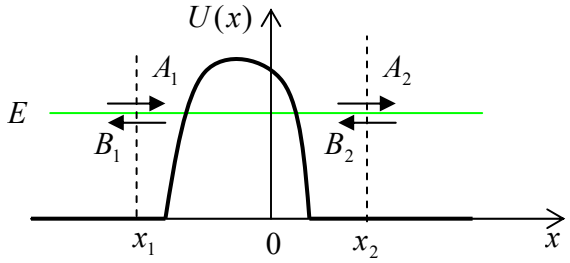


Fig. 2.12. De Broglie wave amplitudes near a single 1D scatterer.

From Sec. 2, we know that the general solutions of the stationary Schrödinger equation, with a certain energy E, in the regions outside the interval $[x_1, x_2]$, are sets of two sinusoidal waves, traveling in opposite directions. Let us represent them in the form

$$\psi_j = A_j e^{ik(x-x_j)} + B_j e^{-ik(x-x_j)}, \tag{2.120}$$

where the index j equals either 1 or 2 (for now), and $(\hbar k)^2/2m = E$. Note that each of the two wave pairs (129) has, in this notation, its own reference point x_j, because this will be very convenient for what

[29] This formula was derived (in a more general form, valid for an arbitrary soft potential barrier) by E. Kemble in 1935. In some communities, it is known as the "Hill-Wheeler formula", after D. Hill and J. Wheeler's 1953 paper in that the Kemble formula was spelled out for the quadratic profile (118). Note that mathematically Eq. (119) is similar to the Fermi distribution in statistical physics, with an effective temperature $T_{ef} = \hbar\omega_0/2\pi k_B$. This coincidence has some curious implications for the Fermi particle tunneling statistics.

follows. As we have already discussed, if the de Broglie wave/particle is incident from the left (i.e. $B_2 = 0$), the solution of the *linear* Schrödinger equation within the scatterer range ($x_1 < x < x_2$) can provide only *linear* expressions of the transmitted (A_2) and reflected (B_1) wave amplitudes via the incident wave amplitude A_1:

$$A_2 = S_{21}A_1, \quad B_1 = S_{11}A_1, \tag{2.121}$$

where S_{11} and S_{21} are certain (generally, complex) coefficients. Alternatively, if a wave with amplitude B_2 is incident on the scatterer from the right (i.e. if $A_1 = 0$), it can induce a transmitted wave (B_1) and a reflected wave (A_2), with amplitudes

$$B_1 = S_{12}B_2, \quad A_2 = S_{22}B_2, \tag{2.122}$$

where the coefficients S_{22} and S_{12} are generally different from S_{11} and S_{21}. Now we can use the linear superposition principle to argue that if the waves with amplitudes A_1 and B_2 are simultaneously incident on the scatterer (say, because the wave B_2 has been partly reflected back by some other scatterer located at $x > x_2$), the resulting *scattered wave* amplitudes A_2 and B_1 are just the sums of their values for separate incident waves:

$$\begin{aligned} B_1 &= S_{11}A_1 + S_{12}B_2, \\ A_2 &= S_{21}A_1 + S_{22}B_2. \end{aligned} \tag{2.123}$$

These two relations may be conveniently represented using the so-called *scattering matrix* S:

Scattering matrix: definition

$$\boxed{\begin{pmatrix} B_1 \\ A_2 \end{pmatrix} = \mathrm{S}\begin{pmatrix} A_1 \\ B_2 \end{pmatrix}, \qquad \text{with } \mathrm{S} \equiv \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}.} \tag{2.124}$$

Scattering matrices, duly generalized, are an important tool for the analysis of wave scattering in higher dimensions than one; for 1D problems, however, another matrix is often more convenient to represent the same linear relations (123). Indeed, let us solve this system for A_2 and B_2. The result is

Transfer matrix: definition

$$\boxed{\begin{aligned} A_2 &= T_{11}A_1 + T_{12}B_1, \\ B_2 &= T_{21}A_1 + T_{22}B_1, \end{aligned} \quad \text{i.e.} \quad \begin{pmatrix} A_2 \\ B_2 \end{pmatrix} = \mathrm{T}\begin{pmatrix} A_1 \\ B_1 \end{pmatrix},} \tag{2.125}$$

where T is the *transfer matrix*, with the following elements:

$$T_{11} = S_{21} - \frac{S_{11}S_{22}}{S_{12}}, \quad T_{12} = \frac{S_{22}}{S_{12}}, \quad T_{21} = -\frac{S_{11}}{S_{21}}, \quad T_{22} = \frac{1}{S_{12}}. \tag{2.126}$$

The matrices S and T have some universal properties, valid for an arbitrary (but time-independent) scatterer; they may be readily found from the probability current conservation and the time-reversal symmetry of the Schrödinger equation. Let me leave finding these relations for the reader's exercise. The results show, in particular, that the scattering matrix may be rewritten in the following form:

$$\mathrm{S} = e^{i\theta}\begin{pmatrix} re^{i\varphi} & t \\ t & -re^{-i\varphi} \end{pmatrix}, \tag{2.127a}$$

where the four real parameters r, t, θ, and φ satisfy the following universal relation:

$$r^2 + t^2 = 1, \tag{2.127b}$$

so only three of these parameters are independent. As a result of this symmetry, T_{11} may be also represented in a simpler form similar to T_{22}: $T_{11} = \exp\{i\theta\}/t = 1/S_{12}^{*} = 1/S_{21}^{*}$. The last form allows a ready expression of the scatterer's transparency via just one coefficient of the transfer matrix:

$$\mathscr{T} \equiv \left|\frac{A_2}{A_1}\right|^2_{B_2=0} = \left|S_{21}\right|^2 = \left|T_{11}\right|^{-2}. \tag{2.128}$$

In our current context, the most important property of the 1D transfer matrices is that to find the total transfer matrix T of a system consisting of several (say, N) sequential arbitrary scatterers (Fig. 13), it is sufficient to multiply their matrices.

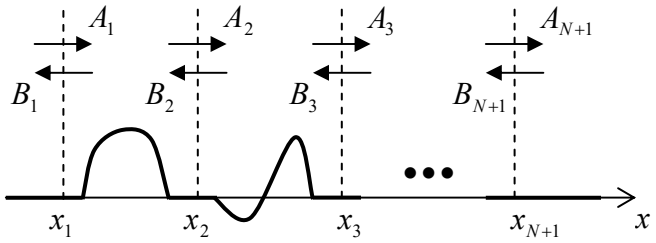


Fig. 2.13. A sequence of several 1D scatterers.

Indeed, extending the definition (125) to other points x_j ($j = 1, 2, \ldots, N + 1$), we can write

$$\begin{pmatrix} A_2 \\ B_2 \end{pmatrix} = \mathrm{T}_1 \begin{pmatrix} A_1 \\ B_1 \end{pmatrix}, \qquad \begin{pmatrix} A_3 \\ B_3 \end{pmatrix} = \mathrm{T}_2 \begin{pmatrix} A_2 \\ B_2 \end{pmatrix} = \mathrm{T}_2\mathrm{T}_1 \begin{pmatrix} A_1 \\ B_1 \end{pmatrix}, \quad \text{etc.} \tag{2.129}$$

(where the matrix indices correspond to the scatterers' order on the x-axis), so

$$\begin{pmatrix} A_{N+1} \\ B_{N+1} \end{pmatrix} = \mathrm{T}_N \mathrm{T}_{N-1} \ldots \mathrm{T}_1 \begin{pmatrix} A_1 \\ B_1 \end{pmatrix}. \tag{2.130}$$

But we can also define a *total transfer matrix* similarly to Eq. (125), i.e. as

$$\begin{pmatrix} A_{N+1} \\ B_{N+1} \end{pmatrix} \equiv \mathrm{T} \begin{pmatrix} A_1 \\ B_1 \end{pmatrix}, \tag{2.131}$$

so comparing Eqs. (130) and (131) we get

$$\boxed{\mathrm{T} = \mathrm{T}_N \mathrm{T}_{N-1} \ldots \mathrm{T}_1 .} \tag{2.132}$$

Transfer matrix: composite scatterer

This formula is valid even if the flat-potential gaps between component scatterers are shrunk to zero, so it may be applied to a scatterer with an arbitrary profile $U(x)$, by fragmenting its length into many small segments $\Delta x = x_{j+1} - x_j$, and treating each fragment as a rectangular barrier of the average height $(U_j)_{\text{ef}} = [U(x_{j+1}) - U(x_j)]/2$ – see Fig. 14. Since very efficient numerical algorithms are readily available for fast multiplication of matrices (especially as small as 2×2 in our case), this approach is broadly used in practice for the computation of transparency of 1D potential barriers with complicated profiles $U(x)$. (This procedure is much more efficient than the direct numerical solution of the stationary Schrödinger equation.)

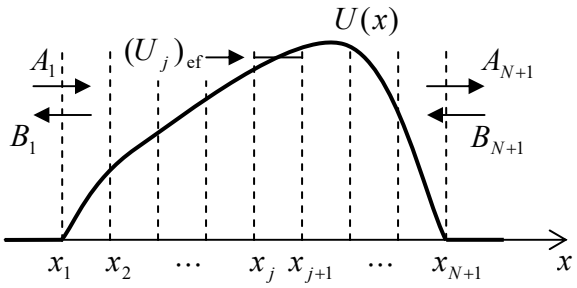


Fig. 2.14. The transfer matrix approach to a potential barrier with an arbitrary profile.

In order to apply this approach to several particular, conceptually important systems, let us calculate the transfer matrices for a few elementary scatterers, starting from the delta-functional barrier located at $x = 0$ – see Fig. 8. Taking $x_1, x_2 \to 0$, we can merely change the notation of the wave amplitudes in Eq. (78) to get

$$S_{11} = \frac{-i\alpha}{1+i\alpha}, \qquad S_{21} = \frac{1}{1+i\alpha}. \tag{2.133}$$

An absolutely similar analysis of the wave incidence from the left yields

$$S_{22} = \frac{-i\alpha}{1+i\alpha}, \qquad S_{12} = \frac{1}{1+i\alpha}, \tag{2.134}$$

and using Eqs. (126), we get

Transfer matrix: short scatterer

$$\boxed{\mathrm{T}_\alpha = \begin{pmatrix} 1-i\alpha & -i\alpha \\ i\alpha & 1+i\alpha \end{pmatrix}}. \tag{2.135}$$

As a sanity check, Eq. (128) applied to this result, immediately brings us back to Eq. (79).

The next example may seem strange at first glance: what if there is no scatterer at all between points x_1 and x_2? If the points coincide, the answer is indeed trivial and can be obtained, e.g., from Eq. (135) by taking $\alpha = 0$, i.e. $\alpha = 0$:

Identity transfer matrix

$$\boxed{\mathrm{T}_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \equiv \mathrm{I}} \tag{2.136}$$

– the so-called *identity matrix*. However, we are free to choose the reference points $x_{1,2}$ participating in Eq. (120) as we wish. For example, what if $x_2 - x_1 = a$? Let us first take the forward-propagating wave alone: $B_2 = 0$ (and hence $B_1 = 0$); then

$$\psi_2 = \psi_1 = A_1 e^{ik(x-x_1)} \equiv A_1 e^{ik(x_2-x_1)} e^{ik(x-x_2)}. \tag{2.137}$$

The comparison of this expression with the definition (120) for $j = 2$ shows that $A_2 = A_1 \exp\{ik(x_2 - x_1)\} = A_1 \exp\{ika\}$, i.e. $T_{11} = \exp\{ika\}$. Repeating the calculation for the back-propagating wave, we see that $T_{22} = \exp\{-ika\}$, and since the space interval provides no particle reflection, we finally get

Transfer matrix: spatial interval

$$\boxed{\mathrm{T}_a = \begin{pmatrix} e^{ika} & 0 \\ 0 & e^{-ika} \end{pmatrix}}, \tag{2.138}$$

independently of a common shift of points x_1 and x_2. At $a = 0$, we naturally recover the special case (136).

Now let us use these simple results to analyze the *double-barrier system* shown in Fig. 15. We could of course calculate its properties as before, writing down explicit expressions for all five traveling waves symbolized by arrows in Fig. 15, then using the boundary conditions (124) and (125) at each of the points $x_{1,2}$ to get a system of four linear equations, and finally, solving it for four amplitude ratios.

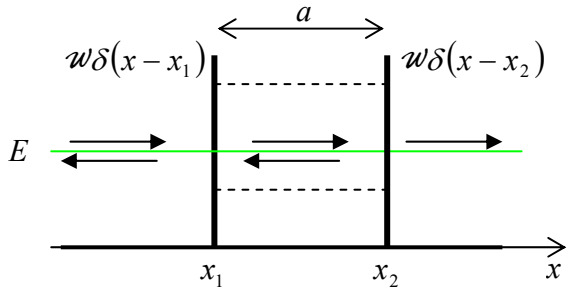


Fig. 2.15. The double-barrier system. The dashed lines show (schematically) the quasi-levels of the metastable-state energies.

However, the transfer matrix approach simplifies the calculations, because we may immediately use Eqs. (132), (135), and (138) to write

$$\mathrm{T} = \mathrm{T}_\alpha \mathrm{T}_a \mathrm{T}_\alpha = \begin{pmatrix} 1-i\alpha & -i\alpha \\ i\alpha & 1+i\alpha \end{pmatrix} \begin{pmatrix} e^{ika} & 0 \\ 0 & e^{-ika} \end{pmatrix} \begin{pmatrix} 1-i\alpha & -i\alpha \\ i\alpha & 1+i\alpha \end{pmatrix}. \tag{2.139}$$

Let me hope that the reader remembers the "row by column" rule of the multiplication of square matrices;[30] using it for the last two matrices, we may reduce Eq. (139) to

$$\mathrm{T} = \begin{pmatrix} 1-i\alpha & -i\alpha \\ i\alpha & 1+i\alpha \end{pmatrix} \begin{pmatrix} (1-i\alpha)e^{ika} & -i\alpha e^{ika} \\ i\alpha e^{-ika} & (1+i\alpha)e^{-ika} \end{pmatrix}. \tag{2.140}$$

Now there is no need to calculate all elements of the full product T, because, according to Eq. (128), for the calculation of the barrier's transparency T we need only one of its elements, T_{11}:

$$\mathscr{T} = \frac{1}{|T_{11}|^2} = \frac{1}{\left|\alpha^2 e^{-ika} + (1-i\alpha)^2 e^{ika}\right|^2}. \tag{2.141}$$

Double barrier: transparency

This result describes oscillations of the transparency: at a fixed parameter α, $\mathscr{T}$ is a π-periodic function of the product ka, reaching its maximum ($\mathscr{T} = 1$) at some point of each period – see Fig. 16a. Indeed, the denominator in Eq. (141) may be interpreted as the squared length of the difference between two 2D vectors, one of length α^2 and another of length $|(1 - i\alpha)^2| = 1 + \alpha^2$, with the angle $\theta = 2ka +$ const between them – see Fig. 16b. At the resonance, the vectors are aligned, and their difference is smallest (equal to 1) so $\mathscr{T}_{\max} = 1$. (This result is exact only if the two barriers are exactly equal.) This means that, rather counter-intuitively, the maximum transparency of the system is *perfect* even at $\alpha >> 1$, i.e. in the case of a very low transparency of each of the component barriers. This is the famous *resonant tunneling* effect.[31]

[30] In the analytical form: $(\mathrm{AB})_{jj'} = \sum_{j''=1}^{N} A_{jj''} B_{j''j'}$, where N is the square matrix rank (in our current case, $N = 2$).

[31] In older literature, it is sometimes called the *Townsend* (or "Ramsauer-Townsend") *effect*. However, nowadays it is more common to use the last term only for a similar effect at 3D scattering – to be discussed in Chapter 3.

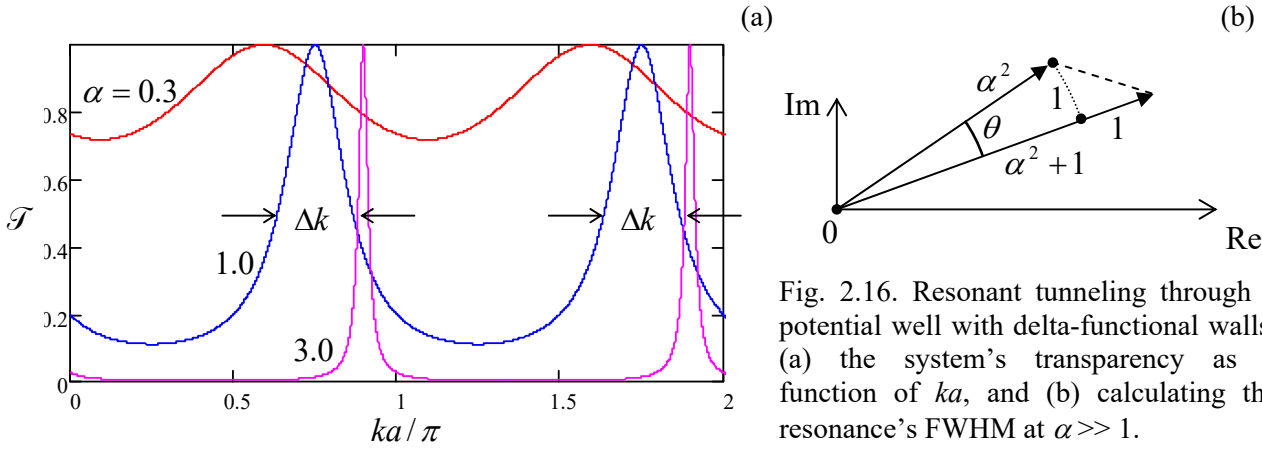


Fig. 2.16. Resonant tunneling through a potential well with delta-functional walls: (a) the system's transparency as a function of ka, and (b) calculating the resonance's FWHM at $\alpha >> 1$.

Its physics is the *constructive interference* of de Broglie waves, similar to that of electromagnetic waves (for example, light) in a *Fabry-Perot resonator* formed by two parallel semi-transparent mirrors.[32] Namely, the incident de Broglie wave may be thought to undertake, on its way through the system, several sequential reflections from these semi-transparent walls. At $k = k_n$, i.e. at $2ka = 2k_n a = 2\pi n$, the phase differences between all these partial waves are multiples of 2π, so they add up in phase – "constructively". (At large but finite α, the resonance condition slightly deviates from $ka = \pi n$.) Note that the same constructive interference of numerous reflections from the walls may be used to interpret the standing-wave eigenfunctions (1.84), so the resonant tunneling at $\alpha >> 1$ in our current system may be also considered a result of the resonant induction of such a standing wave, with a very large amplitude, in the space between the barriers, with the transmitted wave's amplitude proportionately increased.

The resonance peaks of the transparency may be very narrow. Their so-called FWHM (the common acronym for the *Full Width at Half-Maximum*), for the most interesting case for $\alpha >> 1$, may readily calculated by using the same vector diagram shown in Fig. 16b. By definition, FWHM is the difference $\Delta k = k_+ - k_-$ between such two values of k, on the opposite slopes of the same resonance curve, at that $\mathcal{T} = \mathcal{T}_{max}/2$ – see the arrows in Fig. 16a. Let the two vectors in Fig. 16b be misaligned by a small angle $\theta << 1$, so the length of the difference vector is much smaller than α^2. To double its length squared, and hence to reduce $\mathcal{T}$ by a factor of two in comparison with its maximum value of 1, the arc between the vectors, equal to $|\alpha^2\theta|$, should also become equal to 1, i.e. $\alpha^2(2k_{\pm}a + \text{const})$ to become equal to ± 1. Subtracting these two equalities from each other, we get

$$\Delta k \equiv k_+ - k_- = \frac{1}{a\alpha^2} << k_{\pm}. \qquad (2.142)$$

Now let us use the simple system shown in Fig. 15 to discuss an issue of large conceptual significance. For that, consider what would happen if at some initial moment (say, $t = 0$) we placed a 1D quantum particle inside the double-barrier well with $\alpha >> 1$, and left it there alone, without any incident wave. To simplify the analysis, let us assume that the initial state of the particle coincides with one of the stationary states of the infinite-wall well of the same size – see Eq. (1.84):

[32] See, e.g., EM Sec. 7.9. Note also that as Eqs. (2.71) and Fig. 7 show, similar resonant tunneling takes place on the top of the rectangular barrier of height $U_0 < E$, thanks to the step-down reflection from its borders.

$$\Psi(x,0)=\psi_n(x)=\left(\frac{2}{a}\right)^{1/2}\sin\left[k_n(x-x_1)\right],\qquad \text{where } k_n=\frac{\pi n}{a},\quad n=1,2,\ldots. \tag{2.143}$$

At $\alpha\to\infty$, this is just an eigenstate of the system, and from our analysis in Sec. 1.5 we know the time evolution of its wavefunction:

$$\Psi(x,t)=\psi_n(x)\exp\{-i\omega_n t\}=\left(\frac{2}{a}\right)^{1/2}\sin\left[k_n(x-x_1)\right]\exp\{-i\omega_n t\},\qquad \text{with } \omega_n=\frac{E_n}{\hbar}=\frac{\hbar k_n^2}{2m}, \tag{2.144}$$

telling us that the particle remains in the well at all times with a constant probability: $W(t) = W(0) = 1$.

However, if the parameter α is large but finite, the de Broglie wave would slowly "leak out" from the well, so $W(t)$ would slowly decrease. Such a state is called *metastable*. Let us derive the law of its time evolution, assuming that at the slow leakage, with a characteristic time $\tau >> 1/\omega_n$, does not affect the instant wave distribution inside the well, besides the gradual, slow reduction of W.[33] Then we can generalize Eq. (144) as

$$\Psi(x,t)=\left(\frac{2W}{a}\right)^{1/2}\sin\left[k_n(x-x_1)\right]\exp\{-i\omega_n t\}\equiv A\exp\{i(k_n x-\omega_n t)\}+B\exp\{-i(k_n x+\omega_n t)\}, \tag{2.145}$$

making the probability of finding the particle in the well equal to some $W \le 1$. As the last form of Eq. (145) shows, this is the sum of two traveling waves, with equal magnitudes of their amplitudes and hence equal but opposite probability currents (5):

$$|A|=|B|=\left(\frac{W}{2a}\right)^{1/2},\qquad I_A=\frac{\hbar}{m}|A|^2k_n=\frac{\hbar}{m}\frac{W}{2a}\frac{\pi n}{a},\qquad I_B=-I_A. \tag{2.146}$$

But we already know from Eq. (79) that at $\alpha >> 1$, the delta-functional wall's transparency $\mathscr{T}$ equals $1/\alpha^2$, so the wave carrying the current I_A, incident on the right wall from the inside, induces an outcoming wave outside of the well (Fig. 17) with the following probability current:

$$I_{\rm R}=\mathscr{T}I_A=\frac{1}{\alpha^2}I_A=\frac{1}{\alpha^2}\frac{\pi n\hbar W}{2ma^2}. \tag{2.147}$$

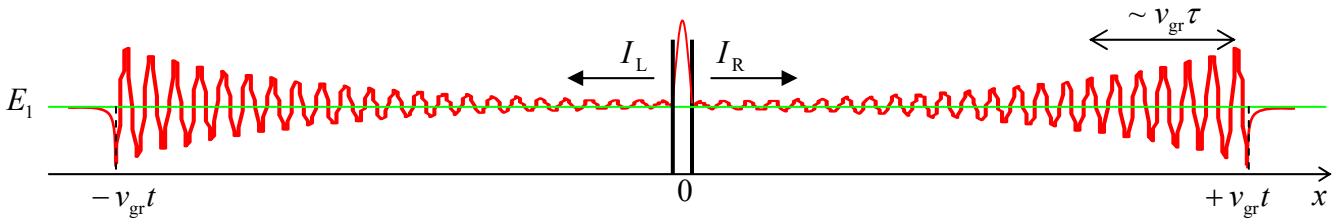


Fig. 2.17. A schematic snapshot of the wavefunction (say, Reψ) in the simple model shown in Fig. 15, at time $t > \tau >> 1/\omega_1$ after the particle's placement into the lowest metastable state with energy E_1.

Absolutely similarly,

$$I_{\rm L}=\frac{1}{\alpha^2}I_B=-I_{\rm R}. \tag{2.148}$$

[33] This (virtually evident) assumption finds its formal justification in the perturbation theory to be discussed in Chapter 6.

Now we may combine the 1D version (6) of the probability conservation law for the well's interior:

$$\frac{dW}{dt} + I_R - I_L = 0, \tag{2.149}$$

with Eqs. (147)-(148) to write

$$\frac{dW}{dt} = -\frac{1}{\alpha^2}\frac{\pi n\hbar}{ma^2}W. \tag{2.150}$$

This is just the standard differential equation,

Metastable state: decay law

$$\boxed{\frac{dW}{dt} = -\frac{1}{\tau}W}, \tag{2.151}$$

of an exponential decay, $W(t) = W(0)\exp\{-t/\tau\}$, where the constant τ is called the *metastable state's lifetime*. In our particular case,

$$\tau = \frac{ma^2}{\pi n\hbar}\alpha^2, \tag{2.152}$$

. Using Eq. (2.33b) for the de Broglie waves' group velocity, for our particular wave vector giving $v_{gr} = \hbar k_n/m = \pi n\hbar/ma$, Eq. (152) may be rewritten in a more general form,

Metastable state: lifetime

$$\boxed{\tau = \frac{t_a}{\mathscr{T}}}, \tag{2.153}$$

where the *attempt time* t_a is equal to a/v_{gr}, and (in our particular case) $\mathscr{T} = 1/\alpha^2$. Eq. (153) is valid for a broad class of similar metastable systems;[34] it may be interpreted in the following semi-classical way. The particle travels back and forth between the confining potential barriers, with the time interval t_a between the sequential moments of incidence, each time attempting to leak through the wall, with the success probability equal to $\mathscr{T}$, so the reduction of W per each incidence is $\Delta W = -W\mathscr{T}$. In the limit $\mathscr{T} << 1$, this equality immediately leads to the decay equation (151) with the lifetime (153).

Another useful look at Eq. (152) may be taken by returning to the resonant tunneling problem in the same system, and expressing the resonance width (142) in terms of the incident particle's energy:

$$\Delta E = \Delta\left(\frac{\hbar^2k^2}{2m}\right) \approx \frac{\hbar^2k_n}{m}\Delta k = \frac{\hbar^2k_n}{m}\frac{1}{a\alpha^2} = \frac{\pi n\hbar^2}{ma^2\alpha^2}. \tag{2.154}$$

Comparing Eqs. (152) and (154), we get a remarkably simple, parameter-independent formula[35]

Energy-time uncertainty relation

$$\boxed{\Delta E\cdot\tau = \hbar}. \tag{2.155}$$

[34] Essentially the only requirement for the attempt time t_a is to be much longer than the effective time (the so-called *instanton time*, see Sec. 5.3 below) of tunneling through the barrier. In the delta-functional approximation for the barrier, the latter time vanishes, so that this requirement is always fulfilled.

[35] Note that Eq. (2.151) may be formally obtained from the basic Schrödinger equation (1.61) by adding an imaginary part, equal to $(-\Delta E/2)$, to its eigenenergy E_n. Indeed, in this case Eq. (1.62) becomes $a_n(t) = \text{const}\times\exp\{-i(E_n - i\Delta E/2\}t/\hbar\} \equiv \text{const}\times\exp\{-iE_nt/\hbar\}\times\exp\{-\Delta Et/2\hbar\} = \text{const}\times\exp\{-iE_nt/\hbar\}\times\exp\{-t/2\tau\}$, so that $W(t) \propto |a_n(t)|^2 \propto \exp\{-t/\tau\}$. Such formalism, which hides the physical origin of the state's decay, may be convenient for some calculations, but misleading in other cases, and I will not use it in this course.

This *energy-time uncertainty relation* is certainly *more* general than our simple model; for example, it is valid for the lifetime and resonance tunneling width of any metastable state in the potential profile of any shape. This seems very natural since because of the energy identification with frequency, $E = \hbar\omega$, pertinent to quantum mechanics, Eq. (155) may be rewritten as $\Delta\omega\cdot\tau = 1$ and seems to follow directly from the Fourier transform in time, just as the Heisenberg's uncertainty relation (1.35) follows from the Fourier transform in space. In some cases, even those not involving any state decay, these two relations are indeed interchangeable. For example, Eq. (24) for the Gaussian wave packet width may be rewritten as $\delta E\cdot\Delta t = \hbar$, where $\delta E = \hbar(d\omega/dk)\delta k = \hbar v_{gr}\delta k$ is the r.m.s. spread of energies of monochromatic components of the packet, while $\Delta t \equiv \delta x/v_{gr}$ is the time scale of the packet's passage through a fixed observation point x.

However, Eq. (155) is much *less* general than Heisenberg's uncertainty relation (1.35). Indeed, the Cartesian coordinates of a particle, the Cartesian components of its momentum, and the energy E are regular observables, represented by operators. In contrast, in the non-relativistic quantum mechanics we are studying now, time is treated as a *c*-number argument, and is *not* represented by an operator, so Eq. (155) cannot be derived in such general assumptions as Eq. (1.35). Thus the time-energy uncertainty relation should be used with caution. Unfortunately, not everybody is so careful. One can find, for example, claims that due to this relation, the energy of a system cannot be measured, during a time interval Δt, with an accuracy better than $\hbar/\Delta t$. These claims are wrong.[36] Another incorrect statement is that the energy dissipated by any system performing an elementary (single-bit) calculation during a time interval Δt has to be larger than $\hbar/\Delta t$.[37]

Now that we have a quantitative mathematical description of the metastable state's decay (valid, again, only at $\alpha >> 1$, i.e. at $\tau >> t_a$), we may use it to discuss two important conceptual issues of quantum mechanics. First, the decay is one of the simplest examples of systems that may be considered, from two different points of view, as either Hamiltonian (and hence *time-reversible*), or open (and hence *irreversible*). Indeed, from the former point of view, our particular system is certainly described by a time-independent Hamiltonian (1.41), with the potential energy

$$U(x) = \mathcal{W}\left[\delta(x - x_1) + \delta(x - x_2)\right] \tag{2.156}$$

– see Fig. 15 again. In this point of view, the total probability of finding the particle *somewhere* on the x-axis remains equal to 1, and the full system's energy, calculated from Eq. (1.23),

$$\langle E \rangle = \int_{-\infty}^{+\infty} \Psi^*(x,t)\hat{H}\Psi(x,t)d^3x, \tag{2.157}$$

remains constant. On the other hand, since the "emitted" wave packets would never return to the potential well,[38] it makes much sense to look at the well's region alone. For such a truncated, *open*

[36] See, e.g., V. Braginsky and F. Khalili, *Quantum Measurement*, Cambridge U. Press, 1992.

[37] Here I dare to refer to my own old work K. Likharev, *Int. J. Theor. Phys*. **21**, 311 (1982), which provided a constructive proof (for a particular system) that at *reversible computations*, whose idea had been put forward in 1973 by C. Bennett (see, e.g., SM Sec. 2.3), energy dissipation may be lower than this apparent "quantum limit".

[38] For more realistic 2D and 3D systems, this statement is true even if the system as a whole is confined inside some closed volume, much larger than the potential well housing the metastable states. Indeed, if the walls providing such confinement are even slightly uneven, the emitted plane-wave packets will be reflected from them, but would never return to the well intact. (See SM Sec. 2.1 for a more detailed discussion of this issue.)

system (for which the space beyond the interval $[x_1, x_2]$ serves as its *environment*), the probability W of finding the particle inside this interval, and hence its energy $\langle E \rangle = WE_n$, decay exponentially per Eq. (151) – the decay equation typical for irreversible systems. We will return to the discussion of the dynamics of such open quantum systems in Chapter 7.

Second, the same model enables a preliminary discussion of one important aspect of quantum measurements. As Eq. (151) and Fig. 17 show, at $t >> \tau$, the well becomes virtually empty ($W \approx 0$), and the whole probability is localized in two clearly separated wave packets with equal amplitudes, moving from each other with speed v_{gr}, each "carrying the particle away" with a probability of 50%. Now assume that an experiment has detected the particle on the left side of the well. Though the formalisms suitable for quantitative analysis of the detection process will not be discussed until Chapter 7, due to the wide separation $\Delta x = 2v_{gr}t >> 2v_{gr}\tau$ of the packets, we may safely assume that such detection may be done without any actual physical effect on the counterpart wave packet.[39] But if we know that the particle has been found on the left side, there is no chance of finding it on the right side. If we attributed the full wavefunction to all stages of this particular experiment, this situation might be rather confusing. Indeed, that would mean that the wavefunction at the right packet's location should instantly turn into zero – the so-called *wave packet reduction* (or "collapse") – a hypothetical irreversible process that cannot be described by the Schrödinger equation for this system, even including the particle detectors.

However, if (as was already discussed in Sec. 1.3) we attribute the wavefunction to a certain statistical ensemble of similar experiments, there is no need to involve such artificial notions. The two-wave-packet picture we have calculated (Fig. 17) describes the full ensemble of experiments with all systems prepared in the initial state (143), i.e. does not depend on the particle detection results. On the other hand, the "reduced packet" picture (with no wave packet on the right of the well) describes only a sub-ensemble of such experiments, in which the particles have been detected on the left side. As was discussed in classical examples in Sec. 1.3, for such a redefined ensemble, the probability distribution may be rather different. So, the "wave packet reduction" is just a result of a purely accounting decision of the observer.[40] I will return to this important issue in Sec. 10.1 – on the basis of the forthcoming discussion of open systems in Chapters 7 and 8.

2.6. Localized state coupling, and quantum oscillations

Now let us discuss one more effect specific to quantum mechanics. Its mathematical description may be simplified using a model potential consisting of two very short and deep potential wells. For that, let us first analyze the properties of a single well of this type (Fig. 18), which may be modeled similarly to the short and high potential barrier – see Eq. (74), but with a negative "weight":

$$U(x) = -\mathcal{W}\delta(x), \qquad \text{with } \mathcal{W} > 0. \tag{2.158}$$

In contrast to its tunnel-barrier counterpart (74), such potential sustains a stationary state with a negative eigenenergy $E < 0$, and a *localized* eigenfunction ψ, with $|\psi| \to 0$ at $x \to \pm\infty$.

[39] This argument is especially convincing if the particle's detection time is much shorter than the time $t_c = 2v_{gr}t/c$, where c is the speed of light in vacuum, i.e. the maximum velocity of any information transfer.

[40] "The collapse of the wavefunction after measurement represents nothing more than the updating of that scientist's expectations." N. D. Mermin, *Phys. Today*, **72**, 53 (Jan. 2013).

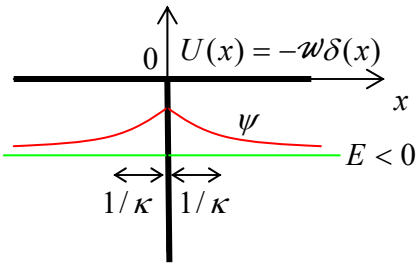


Fig. 2.18. Delta-functional potential well and its localized eigenstate (schematically).

Indeed, at $x \neq 0$, $U(x) = 0$, so the 1D Schrödinger equation is reduced to the Helmholtz equation (1.83), whose localized solutions with $E < 0$ are single exponents vanishing at large distances:[41]

$$\psi(x) = \psi_0(x) \equiv \begin{cases} Ae^{-\kappa x}, & \text{for } x > 0, \\ Ae^{+\kappa x}, & \text{for } x < 0, \end{cases} \qquad \text{with } \frac{\hbar^2\kappa^2}{2m} = -E, \quad \kappa > 0. \tag{2.159}$$

Here the pre-exponential coefficients are taken equal to satisfy the boundary condition (76) of the wavefunction's continuity at $x = 0$. Plugging Eq. (159) into the second boundary condition, given by Eq. (75) but now with the negative sign before $\mathcal{W}$, we get

$$(-\kappa A) - (+\kappa A) = -\frac{2m\mathcal{W}}{\hbar^2} A, \tag{2.160}$$

in which the common factor $A \neq 0$ may be canceled. This equation[42] has one solution for any $\mathcal{W} > 0$:

$$\kappa = \kappa_0 \equiv \frac{m\mathcal{W}}{\hbar^2}, \tag{2.161}$$

and hence the system has only one localized state, with the following eigenenergy:[43]

$$E = E_0 \equiv -\frac{\hbar^2\kappa_0^2}{2m} = -\frac{m\mathcal{W}^2}{2\hbar^2}. \tag{2.162}$$

Now we are ready to analyze the localized states of the two-well potential shown in Fig. 19:

$$U(x) = -\mathcal{W}\left[\delta\left(x - \frac{a}{2}\right) + \delta\left(x + \frac{a}{2}\right)\right], \qquad \text{with } \mathcal{W} > 0. \tag{2.163}$$

Here we may still use the single-exponent solutions similar to Eq. (159), for the wavefunction outside the interval $[-a/2, +a/2]$, but inside the interval, we need to take into account both possible exponents:

$$\psi = C_+ e^{\kappa x} + C_- e^{-\kappa x} \equiv C_A \sinh \kappa x + C_S \cosh \kappa x, \qquad \text{for } -\frac{a}{2} \le x \le +\frac{a}{2}, \tag{2.164}$$

with the parameter κ defined as in Eq. (159). The latter of these two equivalent expressions is more convenient because due to the symmetry of the potential (163) with respect to the central point $x = 0$, the system's eigenfunctions should be either symmetric (even) or antisymmetric (odd) functions of x (see

[41] See Eqs. (56)-(58), with $U_0 = 0$.

[42] Such algebraic equations are frequently called *characteristic*.

[43] Note that this E_0 is equal, by magnitude, to the constant E_0 that participates in Eq. (79). Note also that this result was actually already obtained, "backward", in the solution of Problem 1.12(ii), but that solution did not address the issue of whether the calculated potential (158) could sustain any other localized eigenstates.

Fig. 19), so they may be analyzed separately, only for one half of the system, say $x \geq 0$, and using just one of the hyperbolic functions (164) in each case.

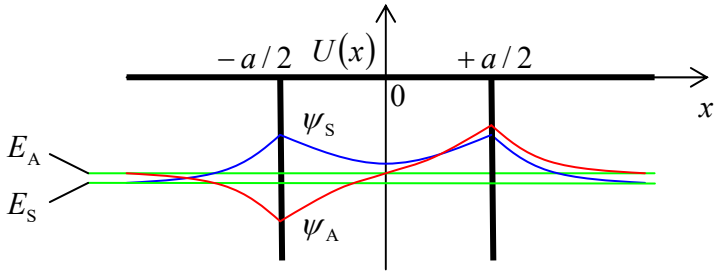


Fig. 2.19. A system of two coupled potential wells, and its localized eigenstates (schematically).

For the *antisymmetric* eigenfunction, Eqs. (159) and (164) yield

$$\psi_A \equiv C_A \times \begin{cases} \sinh \kappa x, & \text{for } 0 \leq x \leq \dfrac{a}{2}, \\ \sinh \dfrac{\kappa a}{2} \exp\left\{-\kappa\left(x - \dfrac{a}{2}\right)\right\}, & \text{for } \dfrac{a}{2} \leq x, \end{cases} \tag{2.165}$$

where the front coefficient in the lower line is adjusted to satisfy the condition (76) of the wavefunction's continuity at $x = +a/2$ – and hence at $x = -a/2$. What remains is to satisfy the condition (75), with a negative sign before $\mathcal{W}$, for the derivative's jump at that point. This condition yields the following characteristic equation:

$$\sinh \frac{\kappa a}{2} + \cosh \frac{\kappa a}{2} = \frac{2m\mathcal{W}}{\hbar^2 \kappa} \sinh \frac{\kappa a}{2}, \qquad \text{i.e. } 1 + \coth \frac{\kappa a}{2} = 2 \frac{(\kappa_0 a)}{(\kappa a)}, \tag{2.166}$$

where the κ_0, given by Eq. (161), is the value of κ for a single well, i.e. the reciprocal spatial width of its localized eigenfunction – see Fig. 18.

Figure 20a shows both sides of Eq. (166) as functions of the dimensionless product κa, for several values of the parameter $\kappa_0 a$, i.e. of the normalized distance between the two wells. The plots show, first of all, that as the parameter $\kappa_0 a$ is decreased, the left-hand side and right-hand side plots cross (i.e. Eq. (166) has a solution) at lower and lower values of κa. At $\kappa a << 1$, the left-hand side of the last form of this equation may be approximated as $2/\kappa a$. Comparing this expression with the right-hand side, we see that this transcendental characteristic equation has a solution (i.e. the system has an antisymmetric localized state) only if $\kappa_0 a > 1$, i.e. if the distance a between the two narrow potential wells is larger than the following value,

$$a_{\min} = \frac{1}{\kappa_0} \equiv \frac{\hbar^2}{m\mathcal{W}}, \tag{2.167}$$

which is equal to the characteristic spread of the wavefunction in a single well – see Fig. 18. (At $a \to a_{\min}$, $\kappa a \to 0$, meaning that the state's localization becomes weaker and weaker.)

In the opposite limit of large distances between the potential wells, i.e. $\kappa_0 a >> 1$, Eq. (166) shows that $\kappa a >> 1$ as well, so its left-hand side may be approximated as $2(1 + \exp\{-\kappa a\})$, and the equation yields

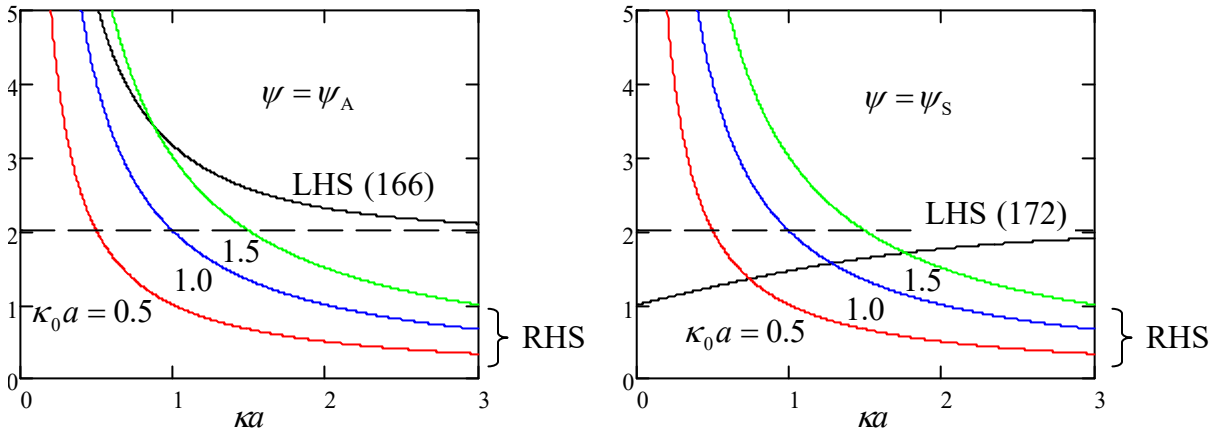


Fig. 2.20. Graphical solutions of the characteristic equations of the two-well system, for: (a) the antisymmetric eigenstate (165), and (b) the symmetric eigenstate (171).

$$\kappa \approx \kappa_0 \left(1 - \exp\{-\kappa_0 a\}\right) \approx \kappa_0 \ . \tag{2.168}$$

This result means that the eigenfunction is an antisymmetric superposition of two virtually unperturbed wavefunctions (159) of each partial potential well:

$$\psi_A(x) \approx \frac{1}{\sqrt{2}}\left[\psi_R(x) - \psi_L(x)\right], \qquad \text{with } \psi_R(x) \equiv \psi_0\left(x - \frac{a}{2}\right), \quad \psi_L(x) \equiv \psi_0\left(x + \frac{a}{2}\right), \tag{2.169}$$

where the front coefficient is selected in such a way that if the eigenfunction ψ_0 of each well is normalized, so is ψ_A. Plugging the middle (more exact) form of Eq. (168) into the last of Eqs. (159), we can see that in this limit the antisymmetric state's energy is slightly higher than the eigenenergy E_0 of a single well, given by Eq. (162):

$$E_A \approx E_0\left(1 - 2\exp\{-\kappa_0 a\}\right) \equiv E_0 + \delta, \qquad \text{where } \delta \equiv \frac{2m\mathcal{W}^2}{\hbar^2}\exp\{-\kappa_0 a\} > 0\,. \tag{2.170}$$

The s*ymmetric* eigenfunction has a form similar to Eq. (165), but is still different from it:

$$\psi = \psi_S \equiv C_S \times \begin{cases} \cosh \kappa x, & \text{for } 0 \le x \le \dfrac{a}{2}, \\ \cosh\dfrac{\kappa a}{2}\exp\left\{-\kappa\left(x - \dfrac{a}{2}\right)\right\}, & \text{for } \dfrac{a}{2} \le x, \end{cases} \tag{2.171}$$

giving a characteristic equation similar in structure to Eq. (166), but with a different left-hand side:

$$1 + \tanh\frac{\kappa a}{2} = 2\frac{(\kappa_0 a)}{(\kappa a)}\,. \tag{2.172}$$

Figure 20b shows both sides of this equation for several values of the parameter $\kappa_0 a$. It is evident that in contrast to Eq. (166), Eq. (172) has a unique solution (and hence the system has a localized symmetric eigenstate) for *any* value of the parameter $\kappa_0 a$, i.e. for any distance between the partial wells. In the limit of very close wells (i.e. their strong coupling), $\kappa_0 a << 1$, we get $\kappa a << 1$, $\tanh(\kappa a/2) \to 0$, and Eq. (172) yields $\kappa \to 2\kappa_0$, leading to a four-fold increase of the eigenenergy's magnitude in comparison with that of the single well:

$$E_S \approx 4E_0 \equiv -\frac{m(2\mathcal{W})^2}{2\hbar^2}, \qquad \text{for } \kappa_0 a << 1. \tag{2.173}$$

The physical meaning of this result is very simple: two very close potential wells act (on the symmetric eigenfunction only!) together, so their "weights" $\mathcal{W} \equiv \int U(x)dx$ just add up.

In the opposite, weak coupling limit, i.e. for $\kappa_0 a >> 1$, Eq. (172) shows that $\kappa a >> 1$ as well, so its left-hand side may be approximated as $2(1 - \exp\{-\kappa a\})$, and the equation yields

$$\kappa \approx \kappa_0 \left(1 + \exp\{-\kappa_0 a\}\right) \approx \kappa_0 . \tag{2.174}$$

In this limit, the eigenfunction is a symmetric superposition of two virtually unperturbed wavefunctions (159) of each partial potential well:

$$\psi_S(x) \approx \frac{1}{\sqrt{2}}\left[\psi_R(x) + \psi_L(x)\right], \tag{2.175}$$

and the eigenenergy is also close to the energy E_0 of each partial well, but is slightly lower than it:

$$E_S \approx E_0\left(1 + 2\exp\{-\kappa_0 a\}\right) \equiv E_0 - \delta, \qquad \text{so that } E_A - E_S = 2\delta , \tag{2.176}$$

where δ is again given by the last of Eqs. (170).

So, the eigenenergy of the symmetric state is always lower than that of the antisymmetric state. The physics of this effect (which remains qualitatively the same in more complex two-component systems, most importantly in diatomic molecules such as H_2) is evident from the sketch of the wavefunctions ψ_A and ψ_S, given by Eqs. (165) and (171), in Fig. 19. In the antisymmetric mode, the wavefunction has to vanish at the center of the system, so each its half is squeezed into one half of the system's spatial extension. Such a squeeze increases the function's gradient, and hence its kinetic energy (1.27), and hence its total energy. On the contrary, in the symmetric mode, the wavefunction effectively spreads into the counterpart well. As a result, it changes in space slower, and hence its kinetic energy is also lower.

Even more importantly, the symmetric state's energy level goes down as the distance a is decreased, corresponding to the effective attraction of the partial wells. This is a good toy model of the strongest (and most important) type of atomic cohesion – the *covalent* (or "chemical") *bonding*.[44] In the simplest case of the H_2 molecule, each of two electrons of the system, in its ground state,[45] reduces its kinetic energy by spreading its wavefunction around both hydrogen nuclei (protons), rather than being confined near one of them – as it had to be in a single atom. The resulting bonding is very strong: in chemical units, it is close to 430 kJ/mol, i.e. to 4.5 eV per molecule. Perhaps counter-intuitively, this quantum-mechanical bonding may be even stronger than the strongest classical (*ionic*) bonding due to electron transfer between atoms, leading to the Coulomb attraction of the resulting ions. (For example, the atomic cohesion in the NaCl molecule is 4.25 eV.)

[44] Historically, the development of the quantum theory of such bonding in the H_2 molecule (by Walter Heinrich Heitler and Fritz Wolfgang London in 1927) was the breakthrough decisive for the acceptance of the then-emerging quantum mechanics by the community of chemists.

[45] Due to the opposite spins of these electrons, the Pauli principle allows them to be in the same orbital ground state – see Chapter 8.

Now let us analyze the dynamic properties of our model system (Fig. 19) carefully because such a pair of weakly coupled potential wells is our first example of the very important class of *two-level systems*.[46] It is easiest to do in the weak-coupling limit $\kappa_0 a >> 1$, when the simple results (168)-(170) and (174)-(176) are quantitatively valid. In particular, Eqs. (169) and (175) enable us to represent the quasi-localized states of the particle in each partial well as linear combinations of its two eigenstates:

$$\psi_R(x) = \frac{1}{\sqrt{2}}[\psi_S(x) + \psi_A(x)], \qquad \psi_L(x) = \frac{1}{\sqrt{2}}[\psi_S(x) - \psi_A(x)]. \tag{2.177}$$

Let us perform the following thought ("gedanken") experiment: place a particle, at $t = 0$, into one of these quasi-localized states, say $\psi_R(x)$, and leave the system alone to evolve, so

$$\Psi(x,0) = \psi_R(x) = \frac{1}{\sqrt{2}}[\psi_S(x) + \psi_A(x)]. \tag{2.178}$$

According to the general solution (1.69) of the Schrödinger equation, the time dynamics of this wavefunction may be obtained simply by multiplying each eigenfunction by the corresponding complex-exponential time factor:

$$\Psi(x,t) = \frac{1}{\sqrt{2}}\left[\psi_S(x)\exp\left\{-i\frac{E_S}{\hbar}t\right\} + \psi_A(x)\exp\left\{-i\frac{E_A}{\hbar}t\right\}\right]. \tag{2.179}$$

From here, using Eqs. (170) and (176), and then Eqs. (169) and (175) again, we get

$$\begin{aligned}\Psi(x,t) &= \frac{1}{\sqrt{2}}\left(\psi_S(x)\exp\left\{\frac{i\delta t}{\hbar}\right\} + \psi_A(x)\exp\left\{-\frac{i\delta t}{\hbar}\right\}\right)\exp\left\{-\frac{iE_0 t}{\hbar}\right\} \\ &\equiv \left(\psi_R(x)\cos\frac{\delta t}{\hbar} + i\psi_L(x)\sin\frac{\delta t}{\hbar}\right)\exp\left\{-i\frac{E_0 t}{\hbar}\right\}.\end{aligned} \tag{2.180}$$

This result implies, in particular, that the probabilities W_R and W_L to find the particle, respectively, in the right and left wells change with time as

$$W_R = \cos^2\frac{\delta t}{\hbar}, \qquad W_L = \sin^2\frac{\delta t}{\hbar}, \tag{2.181}$$

Quantum oscillations

mercifully leaving the total probability constant: $W_R + W_L = 1$. (If our calculation had not passed this sanity check, we would be in big trouble.)

This is the famous effect of *quantum oscillations*[47] of the particle's wavefunction between two coupled localized states, with the frequency

$$\omega = \frac{2\delta}{\hbar} \equiv \frac{E_A - E_S}{\hbar}. \tag{2.182}$$

In its last form, this result does not depend on the assumption of weak coupling, though the simple form (181) of the oscillations, with its 100% probability variations, does. (Indeed, at a strong coupling of two

[46] As we will see later in Chapter 4, these properties are similar to those of spin-½ particles; hence two-level systems are sometimes called *spin-½-like* systems.

[47] Sometimes they are called the *Bloch oscillations*, but more commonly the last term is reserved for a related but different effect in spatially-periodic systems – to be discussed in Sec. 8 below.

subsystems, the very notion of the quasi-localized states ψ_R and ψ_L is ambiguous.) Qualitatively, this effect may be interpreted as follows: the particle placed into one of the potential wells tries to escape from it via tunneling through the potential barrier separating the wells. (In our particular system shown in Fig. 19, the barrier is formed by the spatial segment of length a, which has the potential energy, $U = 0$, higher than the eigenstate energy $-E_0$.) However, in the two-well system, the particle can only escape into the adjacent well. After the tunneling into that counterpart well, the particle tries to escape from it, and hence comes back, etc. – very much as a classical 1D oscillator, initially deflected from its equilibrium position, at negligible damping.

Some care is required in using such interpretation for quantitative conclusions. In particular, let us compare the period $\mathcal{T} \equiv 2\pi/\omega$ of the oscillations (181) with the metastable state's lifetime discussed in the previous section. For our particular model, we may use the second of Eqs. (170) to write

$$\omega = \frac{4|E_0|}{\hbar}\exp\{-\kappa_0 a\}, \quad \text{i.e. } \mathcal{T} = \frac{\pi\hbar}{\delta} = \frac{\pi\hbar}{2|E_0|}\exp\{\kappa_0 a\} \equiv \frac{t_a}{2}\exp\{\kappa_0 a\}, \qquad \text{for } \kappa_0 a >> 1, \quad (2.183)$$

where $t_a \equiv 2\pi/\omega_0 \equiv 2\pi\hbar/|E_0|$ is the effective attempt time. On the other hand, according to Eq. (80), the transparency $\mathcal{T}$ of our potential barrier, in this limit, scales as $\exp\{-2\kappa_0 a\}$,[48] so according to the general relation (153), the lifetime τ is of the order of $t_a\exp\{2\kappa_0 a\} >> \mathcal{T}$. This is a rather counter-intuitive result: the speed of particle tunneling into a similar adjacent well is much higher than that, through a similar barrier, to the free space!

In order to show that this important result is not an artifact of our simple delta-functional model of the potential wells, and also compare $\mathcal{T}$ and τ more directly, let us analyze the quantum oscillations between two weakly coupled wells, now assuming that the (symmetric) potential profile $U(x)$ is sufficiently soft (Fig. 21), so all its eigenfunctions ψ_S and ψ_A are at least differentiable at all points.[49]

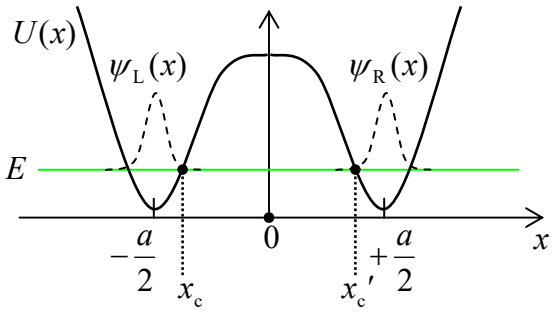


Fig. 2.21. Weak coupling between two similar soft potential wells.

If the barrier's transparency is low, the quasi-localized wavefunctions $\psi_R(x)$ and $\psi_L(x) = \psi_R(-x)$ and their eigenenergies may be found approximately by solving the Schrödinger equations in one of the wells, neglecting the tunneling through the barrier, but the calculation of δ requires a little bit more care. Let us write the stationary Schrödinger equations for the symmetric and antisymmetric solutions as

[48] It is hard to use Eq. (80) for a more exact evaluation of $\mathcal{T}$ in our current system, with its infinitely deep potential wells because the meaning of the wave number k is not quite clear. However, this is not too important, because, in the limit $\kappa_0 a >> 1$, the tunneling exponent makes the dominant contribution to the transparency – see, again, Fig. 2.7b.

[49] Such smooth wells may have more than one localized eigenstate, so the proper state (and energy) index n is implied in all remaining formulas of this section.

$$\left[E_{\rm A}-U(x)\right]\psi_{\rm A}=-\frac{\hbar^2}{2m}\frac{d^2\psi_{\rm A}}{dx^2},\qquad \left[E_{\rm S}-U(x)\right]\psi_{\rm S}=-\frac{\hbar^2}{2m}\frac{d^2\psi_{\rm S}}{dx^2},\tag{2.184}$$

multiply the former equation by $\psi_{\rm S}$ and the latter one by $\psi_{\rm A}$, subtract them from each other, and then integrate the result from 0 to ∞. The result is

$$(E_{\rm A}-E_{\rm S})\int_0^\infty \psi_{\rm S}\psi_{\rm A}\,dx=\frac{\hbar^2}{2m}\int_0^\infty\left(\frac{d^2\psi_{\rm S}}{dx^2}\psi_{\rm A}-\frac{d^2\psi_{\rm A}}{dx^2}\psi_{\rm S}\right)dx.\tag{2.185}$$

If $U(x)$, and hence $d^2\psi_{\rm A,S}/dx^2$, are finite for all x, we may integrate the right-hand side by parts to get

$$(E_{\rm A}-E_{\rm S})\int_0^\infty \psi_{\rm S}\psi_{\rm A}\,dx=\frac{\hbar^2}{2m}\left(\frac{d\psi_{\rm S}}{dx}\psi_{\rm A}-\frac{d\psi_{\rm A}}{dx}\psi_{\rm S}\right)_0^\infty.\tag{2.186}$$

So far, this result is exact (provided that the derivatives participating in it are finite at each point); for *weakly* coupled wells, it may be further simplified. Indeed, in this case, the left-hand side of Eq. (186) may be approximated as

$$(E_{\rm A}-E_{\rm S})\int_0^\infty \psi_{\rm S}\psi_{\rm A}\,dx\approx\frac{E_{\rm A}-E_{\rm S}}{2}\equiv\delta\,,\tag{2.187}$$

because this integral is dominated by the vicinity of point $x = a/2$, where the second terms in each of Eqs. (169) and (175) are negligible, so assuming the proper normalization of the function $\psi_{\rm R}(x)$, the integral is equal to ½. On the right-hand side of Eq. (186), the substitution at $x = \infty$ vanishes (due to the wavefunction's decay in the classically forbidden region), and so does the first term at $x = 0$, because for the antisymmetric solution, $\psi_{\rm A}(0) = 0$. As a result, the energy half-split δ may be expressed in any of the following (equivalent) forms:

$$\delta=\frac{\hbar^2}{2m}\psi_{\rm S}(0)\frac{d\psi_{\rm A}}{dx}(0)=\frac{\hbar^2}{m}\psi_{\rm R}(0)\frac{d\psi_{\rm R}}{dx}(0)=-\frac{\hbar^2}{m}\psi_{\rm L}(0)\frac{d\psi_{\rm L}}{dx}(0).\tag{2.188}$$

It is straightforward (and hence left for the reader's exercise) to show that within the limits of the WKB approximation's validity, Eq. (188) may be reduced to

$$\delta=\frac{\hbar}{t_{\rm a}}\exp\left\{-\int_{x_{\rm c}}^{x_{\rm c}'}\kappa(x')dx'\right\},\qquad \text{so that }\ \mathcal{T}\equiv\frac{\pi\hbar}{\delta}=\frac{t_{\rm a}}{2}\exp\left\{\int_{x_{\rm c}}^{x_{\rm c}'}\kappa(x')dx'\right\},\tag{2.189}$$

where $t_{\rm a}$ is the time period of the classical motion of the particle, with the energy $E \approx E_{\rm A} \approx E_{\rm S}$, inside each well, the function $\kappa(x)$ is defined by Eq. (82), and $x_{\rm c}$ and $x_{\rm c}'$ are the classical turning points limiting the potential barrier at the level E of the particle's eigenenergy – see Fig. 21. The result (189) is evidently a natural generalization of Eq. (183), so the strong relationship between the times of particle tunneling into the continuum of states and into a discrete eigenstate, is indeed not specific for the delta-functional model. We will return to this fact, in its more general form, at the end of Chapter 6.

2.7. Periodic systems: Energy bands and gaps

Let us now proceed to the discussion of one of the most consequential issues of wave mechanics: particle motion through a periodic system. As a precursor to this discussion, let us calculate the transparency of the potential profile shown in Fig. 22 (frequently called the *Dirac comb*): a sequence of N similar, equidistant delta-functional potential barriers separated by $(N – 1)$ potential-free intervals a.

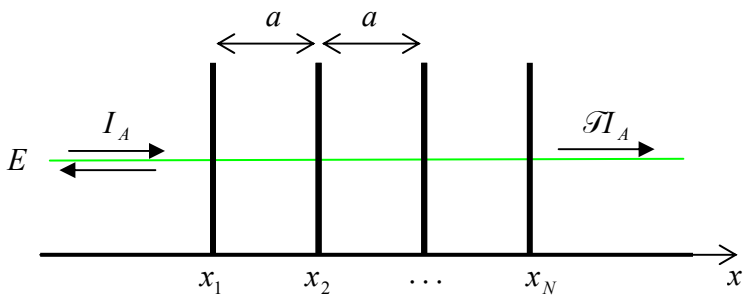


Fig. 2.22. Tunneling through a Dirac comb: a system of N similar, equidistant barriers, i.e. $(N-1)$ similar coupled potential wells.

According to Eq. (132), its transfer matrix is the following product

$$\mathrm{T} = \underbrace{\mathrm{T}_\alpha \mathrm{T}_a \mathrm{T}_\alpha \ldots \mathrm{T}_a \mathrm{T}_\alpha}_{N+(N-1)\ \text{operands}}, \tag{2.190}$$

with the component matrices given by Eqs. (135) and (138), and the barrier height parameter α defined by the last of Eqs. (78). Remarkably, this multiplication may be carried out analytically for arbitrary N,[50] giving

$$\mathcal{T} \equiv \left|T_{11}\right|^{-2} = \left[(\cos Nqa)^2 + \left(\frac{\sin ka - \alpha \cos ka}{\sin qa}\sin Nqa\right)^2\right]^{-1}, \tag{2.191a}$$

where q is a new parameter, with the wave number dimensionality, defined by the following relation:

$$\cos qa \equiv \cos ka + \alpha \sin ka. \tag{2.191b}$$

For $N = 1$, Eqs. (191) immediately yield our old result (79), while for $N = 2$ they may be readily reduced to Eq. (141) – see Fig. 16a. Fig. 23 shows their predictions for two larger numbers N, and several values of the dimensionless parameter α.

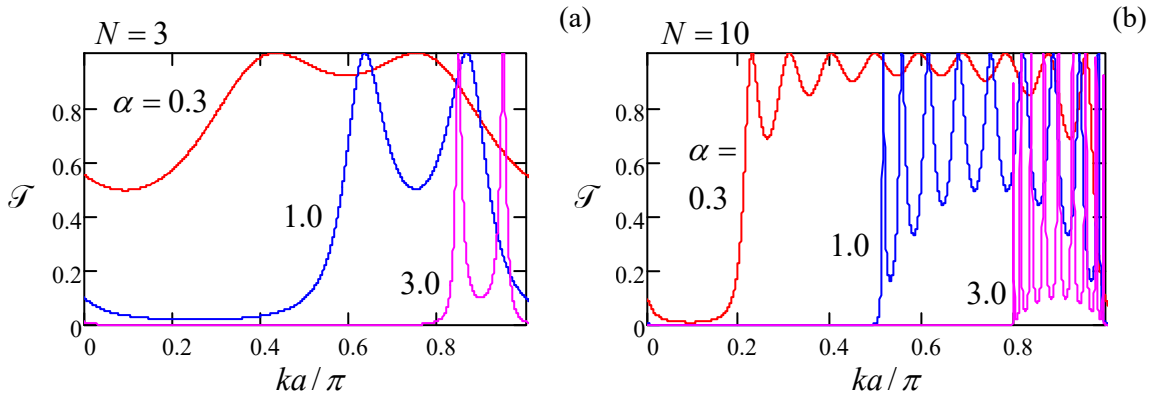


Fig. 2.23. The Dirac comb's transparency as a function of the product ka for three values of α. Since the function $\mathcal{T}(ka)$ is π-periodic (just like it is for $N = 2$, see Fig. 16a), only one period is shown.

Let us start the discussion of the plots from the case $N = 3$ when three barriers limit two coupled potential wells between them. Comparison of Fig. 23a and Fig. 16a shows that the transmission patterns,

[50] This formula will be easier to prove after we have discussed the properties of Pauli matrices in Chapter 4.

and their dependence on the parameter α, are very similar, besides that in the coupled-well system, each resonant tunneling peak splits into two, with the ka-difference between them scaling as $1/\alpha$. From the discussion in the last section, we may now readily interpret this result: each pair of resonance peaks of transparency corresponds to the alignment of the incident particle's energy E with the pair of energy levels of the symmetric (E_S) and antisymmetric (E_A) states of the system. However, in contrast to the system shown in Fig. 19, these states are metastable, because the particle may leak out from these states just as it could in the system studied in Sec. 5 – see Fig. 15 and its discussion. As a result, each of the resonant peaks has a non-zero energy width ΔE, obeying Eq. (155).

A further increase of N (see Fig. 23b) results in the increase of the number of resonant peaks per period to $(N - 1)$, and at $N \to \infty$ the peaks merge into the so-called *allowed energy bands* (frequently called just the "energy bands") with average transparency $\mathscr{T} \sim 1$, separated from similar bands in the adjacent periods of the function $\mathscr{T}(ka)$ by *energy gaps*[51] where $\mathscr{T} \to 0$. Notice the following important features of the pattern:

(i) at $N \to \infty$, the band/gap edges become sharp for any α, and tend to fixed positions (determined by α but independent of N);

(ii) the larger the well coupling (the smaller is α), the broader the allowed energy bands and the narrower the gaps between them.

Our previous discussion of the resonant tunneling gives us a clue for a semi-quantitative interpretation of these features: if $(N - 1)$ potential wells are weakly coupled by tunneling through the potential barriers separating them, the system's energy spectrum consists of groups of $(N - 1)$ metastable energy levels, each group being close to one of the unperturbed eigenenergies of the well. (According to Eq. (1.84), for our current example shown in Fig. 22, with its rectangular potential wells, these eigenenergies correspond to $k_n a = \pi n$.)

Now let us recall that in the case $N = 2$ analyzed in the previous section, the eigenfunctions (169) and (175) differed only by the phase shift $\Delta\varphi$ between their localized components $\psi_R(x)$ and $\psi_L(x)$, with $\Delta\varphi = 0$ for one of them (ψ_S) and $\Delta\varphi = \pi$ for its counterpart. Hence it is natural to expect that for other N as well, each metastable energy level corresponds to an eigenfunction that is a superposition of similar localized functions in each potential well, but with certain phase shifts $\Delta\varphi$ between them.

Moreover, we may expect that at $N \to \infty$, i.e. for *periodic structures*,[52] with

$$U(x + a) = U(x), \tag{2.192}$$

when the system does not have the ends that could affect its properties, the phase shifts $\Delta\varphi$ between the localized wavefunctions in all couples of adjacent potential wells should be equal, i.e.

$$\psi(x + a) = \psi(x) e^{i\Delta\varphi} \tag{2.193a}$$

for all x.[53] This equality is the much-celebrated *Bloch theorem*,[54] or rather its 1D version. Mathematical rigor aside,[55] it is a virtually evident fact because the particle's density $w(x) = \psi^*(x)\psi(x)$, which has to

[51] In solid-state (especially semiconductor) physics and electronics, the term *bandgaps* is more common.

[52] This is a reasonable 1D model, for example, for solid-state crystals, whose samples may feature up to ~10^9 similar atoms or molecules in each direction of the crystal lattice.

be periodic in this a-periodic system, may be so only $\Delta\varphi$ is constant. For what follows, it is more convenient to represent the real constant $\Delta\varphi$ in the form qa, so that the Bloch theorem takes the form

Bloch theorem: 1D version

$$\psi(x+a)=\psi(x)e^{iqa}. \tag{2.193b}$$

The physical sense of the parameter q will be discussed in detail below, but we may immediately notice that according to Eq. (193b), an addition of $(2\pi/a)$ to this parameter yields the same wavefunction; hence all observables have to be $(2\pi/a)$-periodic functions of q.[56]

Now let us use the Bloch theorem to calculate the eigenfunctions and eigenenergies for the infinite version of the system shown in Fig. 22, i.e. for an infinite set of delta-functional potential barriers – see Fig. 24.

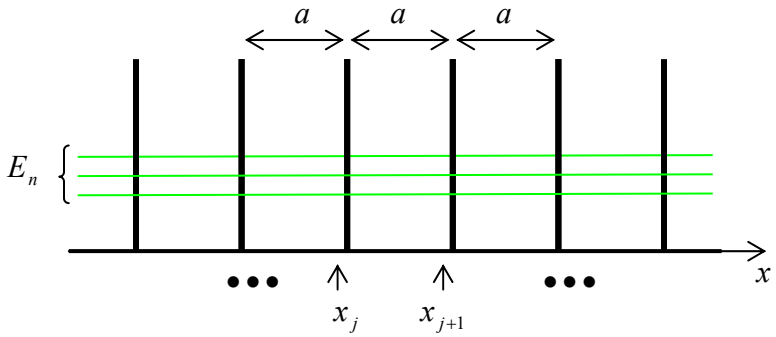


Fig. 2.24. The simplest periodic potential: an infinite Dirac comb.

To start, let us consider two points separated by one period a: one of them, x_j, just left of one of the barriers, and another one, x_{j+1}, just left of the following barrier – see Fig. 24 again. The eigenfunctions at each of the points may be represented as linear superpositions of two simple waves $\exp\{\pm ikx\}$, and the amplitudes of their components should be related by a 2×2 transfer matrix T of the potential fragment separating them. According to Eq. (132), this matrix may be found as the product of the matrix (135) of one delta-functional barrier by the matrix (138) of one zero-potential interval a:

$$\begin{pmatrix} A_{j+1} \\ B_{j+1} \end{pmatrix} = \mathrm{T}_a \mathrm{T}_\alpha \begin{pmatrix} A_j \\ B_j \end{pmatrix} = \begin{pmatrix} e^{ika} & 0 \\ 0 & e^{-ika} \end{pmatrix} \begin{pmatrix} 1-i\alpha & -i\alpha \\ i\alpha & 1+i\alpha \end{pmatrix} \begin{pmatrix} A_j \\ B_j \end{pmatrix}. \tag{2.194}$$

However, according to the Bloch theorem (193b), the component amplitudes should be also related as

---

[53] A reasonably fair classical image of $\Delta\varphi$ is the geometric angle between similar objects – e.g., similar paper clips – attached at equal distances to a long, uniform rubber band. If the band's ends are twisted, the twist is equally distributed between the structure's periods, representing the constancy of $\Delta\varphi$.

[54] Named after F. Bloch who applied this concept to wave mechanics in 1929, i.e. very soon after its formulation. Note, however, that an equivalent statement in mathematics, called the *Floquet theorem*, has been known since at least 1883.

[55] I will recover this rigor in two steps. Later in this section, we will see that the function obeying Eq. (193) is indeed *a* solution to the Schrödinger equation. However, to save time/space, it will be better for us to postpone until Chapter 4 the proof that *any* eigenfunction of the equation, with periodic boundary conditions, obeys the Bloch theorem. As a partial reward for this delay, that proof will be valid for an arbitrary spatial dimensionality.

[56] The product $\hbar q$, which has the linear momentum's dimensionality, is called either the *quasimomentum* or (especially in solid-state physics) the "crystal momentum" of the particle. Informally, it is very convenient (and common) to use the name "quasimomentum" for the bare q as well, despite its evidently different dimensionality.

$$\begin{pmatrix} A_{j+1} \\ B_{j+1} \end{pmatrix} = e^{iqa} \begin{pmatrix} A_j \\ B_j \end{pmatrix} \equiv \begin{pmatrix} e^{iqa} & 0 \\ 0 & e^{iqa} \end{pmatrix} \begin{pmatrix} A_j \\ B_j \end{pmatrix}. \tag{2.195}$$

The condition of self-consistency of these two equations gives the following characteristic equation:

$$\left| \begin{pmatrix} e^{ika} & 0 \\ 0 & e^{-ika} \end{pmatrix} \begin{pmatrix} 1-i\alpha & -i\alpha \\ i\alpha & 1+i\alpha \end{pmatrix} - \begin{pmatrix} e^{iqa} & 0 \\ 0 & e^{iqa} \end{pmatrix} \right| = 0. \tag{2.196}$$

In Sec. 5, we have already calculated the matrix product participating in this equation – see the second operand in Eq. (140). Using it, we see that Eq. (196) is reduced to the same simple Eq. (191b) that has jumped at us from the solution of the somewhat different (resonant tunneling) problem. Let us explore that simple result in detail. First of all, the left-hand side of Eq. (191b) is a sinusoidal function of the product qa with unit amplitude, while its right-hand side is a sinusoidal function of the product ka, with amplitude $(1 + \alpha^2)^{1/2} > 1$ – see Fig. 25.

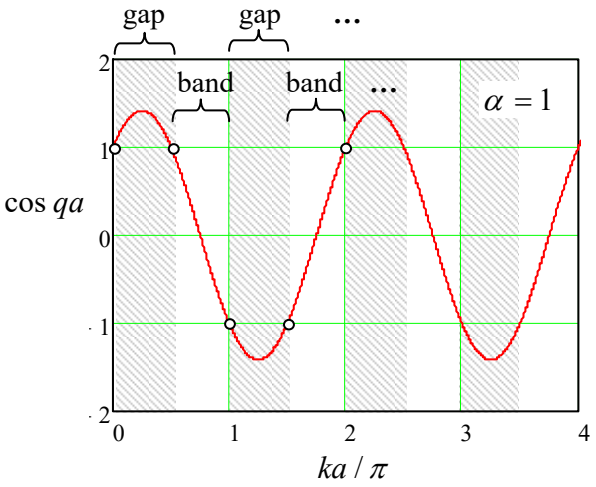


Fig. 2.25. The graphical representation of the characteristic equation (191b) for a fixed value of the parameter α. The ranges of ka that yield $|\cos qa| < 1$, correspond to allowed energy bands, while those with $|\cos qa| > 1$, correspond to energy gaps between them.

As a result, within each half-period $\Delta(ka) = \pi$ of the right-hand side, there is an interval where the magnitude of the right-hand side is larger than 1, so the characteristic equation does not have a real solution for q. These intervals correspond to the energy gaps (see Fig. 23 again) while the complementary intervals of ka, where a real solution for q exists, correspond to the allowed energy bands. In contrast, the parameter q can take *any* real values, so it is more convenient to plot the eigenenergy $E = \hbar^2k^2/2m$ as the function of the quasimomentum $\hbar q$ (or, even more conveniently, of the dimensionless parameter qa) rather than ka.[57] Before doing that, we need to recall that the parameter α, defined by the last of Eqs. (78), depends on the wave vector k as well, so if we vary q (and hence k), it is better to characterize the structure by another, k-independent dimensionless parameter, for example

$$\beta \equiv (ka)\alpha \equiv \frac{\mathcal{W}}{\hbar^2/ma}, \tag{2.197}$$

so our characteristic equation (191b) becomes

[57] A more important reason for taking q as the argument is that for a general periodic potential $U(x)$, the particle's momentum $\hbar k$ is not uniquely related to E, while (according to the Bloch theorem) the quasimomentum $\hbar q$ is.

Dirac comb: q vs k

$$\cos qa \equiv \cos ka + \beta \frac{\sin ka}{ka}. \tag{2.198}$$

Fig. 26 shows the plots of k and E, following from Eq. (198), as functions of qa, for a particular, moderate value of the parameter β. The first evident feature of the pattern is its 2π-periodicity in argument qa, which we have already predicted from the general Bloch theorem arguments. Due to this periodicity, the complete band/gap pattern may be studied, for example, on just one interval $-\pi \le qa \le +\pi$, called the *1st Brillouin zone* – the so-called *reduced zone picture*. For some applications, however, it is more convenient to use the *extended zone picture* with $-\infty \le qa \le +\infty$ – see, e.g., the next section.

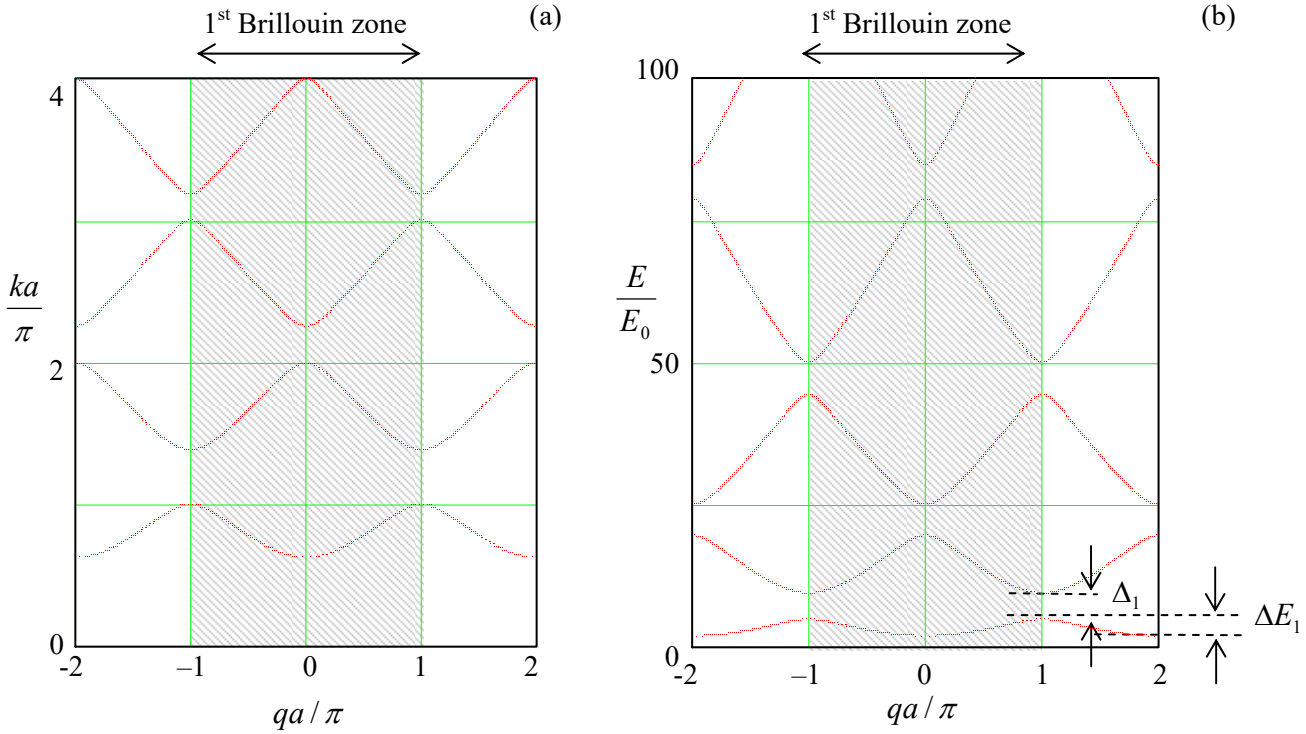


Fig. 2.26. (a) The "genuine" momentum k of a particle in an infinite Dirac comb (Fig. 24), and (b) its energy $E = \hbar^2k^2/2m$ (in the units of $E_0 \equiv \hbar^2/2ma^2$), as functions of the normalized quasimomentum, for a particular value ($\beta = 3$) of the dimensionless parameter defined by Eq. (197). Arrows in the lower right corner of panel b illustrate the definitions of the energy band (ΔE_n) and energy gap (Δ_n) widths.

However, maybe the most important fact, clearly visible in Fig. 26, is that there is an infinite number of energy bands, with different energies $E_n(q)$ for the same value of q. Mathematically, it is evident from Eq. (198) – or alternatively, from Fig. 25. Indeed, for each value of qa, there is a solution ka of this equation on each half-period $\Delta(ka) = \pi$. Each of such solutions (see Fig. 26a) gives a specific value of the particle's energy $E = \hbar^2k^2/2m$. A continuous set of similar solutions for various qa forms a particular energy band.

Since the energy band picture is one of the most practically important results of quantum mechanics, it is imperative to understand its physics. It is natural to describe this physics, in two opposite potential strength limits, in different ways. In parallel, we will use this discussion to obtain simpler expressions for the energy band/gap structure in each limit. An important advantage of this

approach is that both analyses may be carried out for an arbitrary periodic potential $U(x)$ rather than for the particular Dirac comb model shown in Fig. 24.

(i) *Tight-binding approximation*. This approximation works well when the eigenenergy E_n of the states quasi-localized at the energy profile minima is much lower than the height of the potential barriers separating them – see Fig. 27. As should be clear from our discussion in Sec. 6, essentially the only role of coupling between these states (via tunneling through the potential barriers separating the minima) is to establish a certain phase shift $\Delta\varphi \equiv qa$ between the adjacent quasi-localized wavefunctions $u_n(x - x_j)$ and $u_n(x - x_{j+1})$.

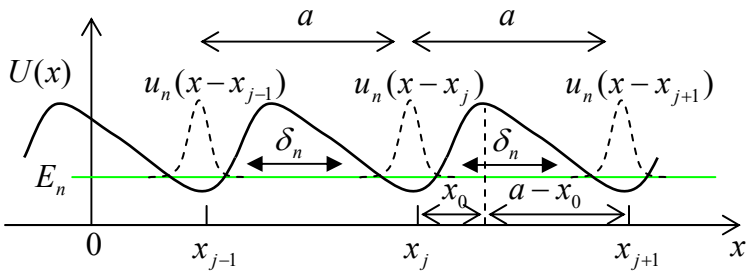


Fig. 2. 27. The tight-binding approximation (schematically).

To describe this effect quantitatively, let us first return to the problem of two coupled wells considered in Sec. 6, and recast the result (180), with the restored eigenstate index n, as

$$\Psi_n(x,t) = \left[a_{\mathrm{R}}(t)\psi_{\mathrm{R}}(x) + a_{\mathrm{L}}(t)\psi_{\mathrm{L}}(x)\right]\exp\left\{-i\frac{E_n}{\hbar}t\right\}, \tag{2.199}$$

where the probability amplitudes a_{R} and a_{L} oscillate sinusoidally in time:

$$a_{\mathrm{R}}(t) = \cos\frac{\delta_n}{\hbar}t, \qquad a_{\mathrm{L}}(t) = i\sin\frac{\delta_n}{\hbar}t. \tag{2.200}$$

This evolution satisfies the following system of two equations whose structure is similar to Eq. (1.61a):

$$i\hbar\dot{a}_{\mathrm{R}} = -\delta_n a_{\mathrm{L}}, \qquad i\hbar\dot{a}_{\mathrm{L}} = -\delta_n a_{\mathrm{R}}. \tag{2.201}$$

Eq. (199) may be readily generalized to the case of many similar coupled wells:

$$\Psi_n(x,t) = \left[\sum_j a_j(t)u_n(x-x_j)\right]\exp\left\{-i\frac{E_n}{\hbar}t\right\}, \tag{2.202}$$

where E_n are the eigenenergies and u_n the eigenfunctions of each well. In the tight-binding limit, only the adjacent wells are coupled, so instead of Eq. (201) we should write an infinite system of similar equations

$$i\hbar\dot{a}_j = -\delta_n a_{j-1} - \delta_n a_{j+1}, \tag{2.203}$$

for each well number j, where parameters δ_n describe the coupling between two adjacent potential wells. Repeating the calculation outlined at the end of the last section for our new situation, for a smooth potential we may get an expression essentially similar to the last form of Eq. (188):

$$\boxed{\delta_n = \frac{\hbar^2}{m}u_n(x_0)\frac{du_n}{dx}(a - x_0),} \tag{2.204}$$

Tight-binding limit: coupling energy

where x_0 is the distance between the well bottom and the middle of the potential barrier on the right of it – see Fig. 27. The only substantial new feature of this expression in comparison with Eq. (188) is that the sign of δ_n alternates with the level number n: $\delta_1 > 0$, $\delta_2 < 0$, $\delta_3 > 0$, etc. Indeed, the number of zeros (and hence, "wiggles") of the eigenfunctions $u_n(x)$ of any potential well increases as n – see, e.g., Fig. 1.8,[58] so the difference of the exponential tails of the functions, sneaking under the left and right barriers limiting the well also alternates with n.

The infinite system of ordinary differential equations (203) enables solutions of many important problems (such as the spread of the wavefunction that was initially localized in one well, etc.), but our task right now is just to find its stationary states, i.e. the solutions proportional to $\exp\{-i(\varepsilon_n/\hbar)t\}$, where ε_n is a still unknown, q-dependent addition to the background energy E_n of the n^{th} energy level. To satisfy the Bloch theorem (193) as well, such a solution should have the following form:

$$a_j(t) = a\exp\left\{iqx_j - i\frac{\varepsilon_n}{\hbar}t + \text{const}\right\}. \tag{2.205}$$

Plugging this solution into Eq. (203) and canceling the common exponent, we get

Tight-binding limit: energy bands

$$E = E_n + \varepsilon_n = E_n - \delta_n\left(e^{-iqa} + e^{iqa}\right) \equiv E_n - 2\delta_n\cos qa\,, \tag{2.206}$$

so in this approximation, the energy band width ΔE_n (see Fig. 26b) equals $4|\delta_n|$.

The relation (206), whose validity is restricted to $|\delta_n| << E_n$, describes the lowest energy bands plotted in Fig. 26b reasonably well. (For larger β, the agreement would be even better.) So, this calculation explains what the energy bands really are: in the tight-binding limit, they are best interpreted as isolated well's energy levels E_n broadened into bands by the interwell interaction. Also, this result gives clear proof that the energy band extremes correspond to $qa = 2\pi l$ and $qa = 2\pi(l + ½)$, with integer l. Finally, the sign alteration of the coupling coefficient δ_n (204) explains why the energy maxima of one band are aligned, on the qa axis, with energy minima of the adjacent bands – see Fig. 26.

(ii) *Weak-potential limit*. Amazingly, the energy-band structure is also compatible with a completely different physical picture that may be developed in the opposite limit. Let the particle's energy E be so high that the periodic potential $U(x)$ may be treated as a small perturbation. Naively, in this limit, we could expect a slightly and smoothly deformed parabolic dispersion relation $E = \hbar^2k^2/2m$. However, if we are plotting the stationary-state energy as a function of q rather than k, we need to add $2\pi l/a$, with an arbitrary integer l, to the argument. Let us show this by expanding all variables into the 1D-spatial Fourier series. For the potential energy $U(x)$ that obeys Eq. (192), such an expansion is straightforward:[59]

$$U(x) = \sum_{l''} U_{l''}\exp\left\{-i\frac{2\pi x}{a}l''\right\}, \tag{2.207}$$

where the summation is over all integers l'', from $-\infty$ to $+\infty$. However, for the wavefunction we should show due respect to the Bloch theorem (193), which shows that strictly speaking, $\psi(x)$ is *not* periodic.

[58] Below, we will see several other examples of this behavior. This alternation rule is also described by the Wilson-Sommerfeld quantization condition (110).

[59] The benefits of such an unusual notation of the summation index (l'' instead of, say, l) will be clear in a few lines.

To overcome this difficulty, let us define another function:

$$u(x) \equiv \psi(x) e^{-iqx}, \tag{2.208}$$

and study its periodicity:

$$u(x+a) = \psi(x+a) e^{-iq(x+a)} = \psi(x) e^{-iqx} = u(x). \tag{2.209}$$

We see that the new function is a-periodic, and hence we can use Eqs. (208)-(209) to rewrite the Bloch theorem in a different form:

1D Bloch theorem: alternative form

$$\boxed{\psi(x) = u(x) e^{iqx}, \quad \text{with} \quad u(x+a) = u(x).} \tag{2.210}$$

Now it is safe to expand the periodic function $u(x)$ exactly as $U(x)$:

$$u(x) = \sum_{l'} u_{l'} \exp\left\{-i \frac{2\pi x}{a} l'\right\}, \tag{2.211}$$

so, according to Eq. (210),

$$\psi(x) = e^{iqx} \sum_{l'} u_{l'} \exp\left\{-i \frac{2\pi x}{a} l'\right\} = \sum_{l'} u_{l'} \exp\left\{i\left(q - \frac{2\pi}{a} l'\right)x\right\}. \tag{2.212}$$

The only nontrivial part of using Eqs. (207) and (212) in the stationary Schrödinger equation (53) is how to handle the product term,

$$U(x)\psi = \sum_{l',l''} U_{l''} u_{l'} \exp\left\{i\left[q - \frac{2\pi}{a}(l' + l'')\right]x\right\}. \tag{2.213}$$

At fixed l', we may change the summation over l'' to that over $l \equiv l' + l''$ (so that $l'' \equiv l - l'$), and write:

$$U(x)\psi = \sum_{l} \exp\left\{i\left(q - \frac{2\pi}{a} l\right)x\right\} \sum_{l'} u_{l'} U_{l-l'}. \tag{2.214}$$

Now plugging Eq. (212) (with the summation index l' replaced with l) and Eq. (214) into the stationary Schrödinger equation (53), and requiring the coefficients of each spatial exponent to match, we get an infinite system of linear equations for u_l:

$$\sum_{l'} U_{l-l'} u_{l'} = \left[E - \frac{\hbar^2}{2m}\left(q - \frac{2\pi}{a} l\right)^2\right] u_l. \tag{2.215}$$

(Note that by this calculation we have essentially proved that the Bloch wavefunction (210) is indeed a solution of the Schrödinger equation, provided that the quasimomentum q is selected in a way to make the system of linear equation (215) compatible, i.e. is a solution of its characteristic equation.)

So far, the system of equations (215) is an equivalent alternative to the initial Schrödinger equation, for any potential's strength.[60] In the weak-potential limit, i.e. if all Fourier coefficients U_n are

[60] By the way, the system is very efficient for a fast numerical solution of the stationary Schrödinger equation for any periodic profile $U(x)$, even though to describe potentials with large U_n, this approach may require taking into account a correspondingly large number of Fourier amplitudes u_l.

small,[61] we can complete all the calculations analytically.[62] Indeed, in the so-called 0th approximation we can ignore *all* U_n, so in order to have at least one u_l different from 0, Eq. (215) requires that

$$E \to E_l \equiv \frac{\hbar^2}{2m}\left(q - \frac{2\pi l}{a}\right)^2 . \tag{2.216}$$

(u_l itself should be obtained from the normalization condition). This result means that in this approximation, the dispersion relation $E(q)$ has an infinite number of similar quadratic branches numbered by integer l – see Fig. 28.

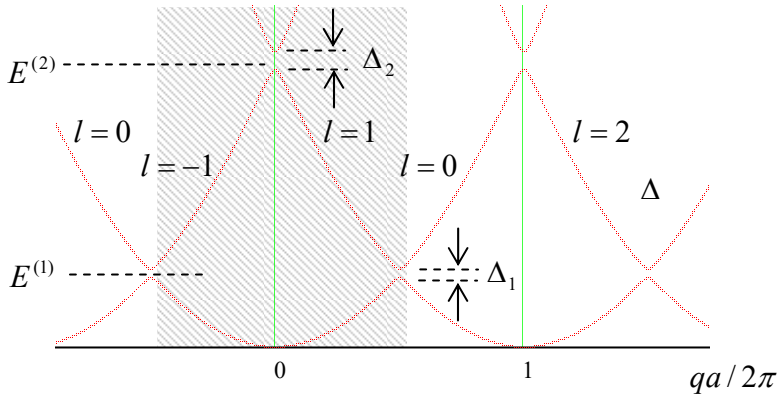


Fig. 2.28. A typical energy band/gap pattern in the weak-potential case, with the shading showing the 1st Brillouin zone.

On every branch, such eigenfunction has just one Fourier coefficient, i.e. is a monochromatic traveling wave

$$\psi_l \to u_l e^{ikx} = u_l \exp\left\{i\left(q - \frac{2\pi l}{a}\right)x\right\}. \tag{2.217}$$

Next, the above definition of E_l allows us to rewrite Eq. (215) in a more transparent form

$$\sum_{l' \neq l} U_{l-l'} u_{l'} = \left(E - E_l\right) u_l , \tag{2.218}$$

which may be formally solved for u_l:

$$u_l = \frac{1}{E - E_l} \sum_{l' \neq l} U_{l-l'} u_{l'} . \tag{2.219}$$

This formula shows that if the Fourier coefficients U_n are non-zero but small, the wavefunctions do acquire other Fourier components (besides the main one, with the index corresponding to the branch number), but these additions are all small, besides narrow regions near the points $E_l = E_{l'}$ where two branches (216) of the dispersion relation $E(q)$, with some specific numbers l and l', cross. According to Eq. (216), this happens when

[61] Besides, possibly, the average potential U_0, which, as was discussed in Chapter 1, may be always taken for the energy reference. In the following calculations, I will take $U_0 = 0$ to simplify the formulas.

[62] This method is so powerful that its multi-dimensional version is not much more complex than the 1D version described here – see, e.g., Sec. 3.2 in the classical textbook by J. Ziman, *Principles of the Theory of Solids*, 2nd ed., Cambridge U. Press, 1979.

$$\left(q-\frac{2\pi}{a}l\right)\approx-\left(q-\frac{2\pi}{a}l'\right), \tag{2.220}$$

i.e. at $q \approx q_m \equiv \pi m/a$ (with the integer $m \equiv l + l'$) [63] corresponding to

$$E_l \approx E_{l'} \approx \frac{\hbar^2}{2ma^2}\left[\pi(l+l')-2\pi l\right]^2 = \frac{\pi^2\hbar^2}{2ma^2}n^2 \equiv E^{(n)}, \tag{2.221}$$

Weak-potential limit: energy gap positions

with integer $n \equiv l - l'$. (According to their definitions, the index n is just the number of the branch crossing on the energy scale, while the index m numbers the position of the crossing points on the q-axis – see Fig. 28.) In such a region, E has to be close to both E_l and $E_{l'}$, so the denominator in just one of the infinite number of terms in Eq. (219) is very small, making the term substantial despite the smallness of U_n. Hence we can take into account only one term in each of the sums (written for l and l'):

$$\begin{aligned} U_n u_{l'} &= (E-E_l)u_l, \\ U_{-n}u_l &= (E-E_{l'})u_{l'}. \end{aligned} \tag{2.222}$$

Taking into account that for any real function $U(x)$, the Fourier coefficients in its Fourier expansion (207) have to be related as $U_{-n} = U_n^*$, Eq. (222) yields the following simple characteristic equation

$$\begin{vmatrix} E-E_l & -U_n \\ -U_n^* & E-E_{l'} \end{vmatrix} = 0, \tag{2.223}$$

with the following solution:

$$E_\pm = E_{\text{ave}} \pm \left[\left(\frac{E_l-E_{l'}}{2}\right)^2 + U_n U_n^*\right]^{1/2}, \qquad \text{with } E_{\text{ave}} \equiv \frac{E_l+E_{l'}}{2} = E^{(n)}. \tag{2.224}$$

Weak-potential limit: level anticrossing

According to Eq. (216), close to the branch crossing point $q_m = \pi(l + l')/a$, the fraction participating in this result may be approximated as[64]

$$\frac{E_l-E_{l'}}{2} \approx \gamma\tilde{q}, \qquad \text{with } \gamma \equiv \left.\frac{dE_l}{dq}\right|_{q=q_m} = \frac{\pi\hbar^2 n}{ma} = \frac{2aE^{(n)}}{\pi n}, \quad \text{and } \tilde{q} \equiv q-q_m, \tag{2.225}$$

while the parameters $E_{\text{ave}} = E^{(n)}$ and $U_n U_n^* = |U_n|^2$ do not depend on $\tilde{q}$, i.e. on the distance from the central point q_m. This is why Eq. (224) may be plotted as the famous *level anticrossing* (also called "avoided crossing", or "intended crossing", or "non-crossing") *diagram* (Fig. 29), with the energy gap width Δ_n equal to $2|U_n|$, i.e. just twice the magnitude of the n-th Fourier harmonic of the periodic potential $U(x)$. Such anticrossings are also clearly visible in Fig. 28, which shows the result of the exact solution of Eq. (198) for the particular case $\beta = 0.5$.[65]

[63] Let me hope that the difference between this new integer and the particle's mass, both called m, is absolutely clear from the context.

[64] Physically, $\gamma/\hbar \equiv \hbar(\pi n/a)/m = \hbar k^{(n)}/m$ is just the velocity of a free classical particle with energy $E^{(n)}$.

[65] From that figure, it is also clear that in the weak potential limit, the width ΔE_n of the n^{th} energy band is just $E^{(n)} - E^{(n-1)}$ – see Eq. (221). Note that this is exactly the distance between the adjacent energy levels of the simplest 1D potential well of infinite depth – cf. Eq. (1.85).

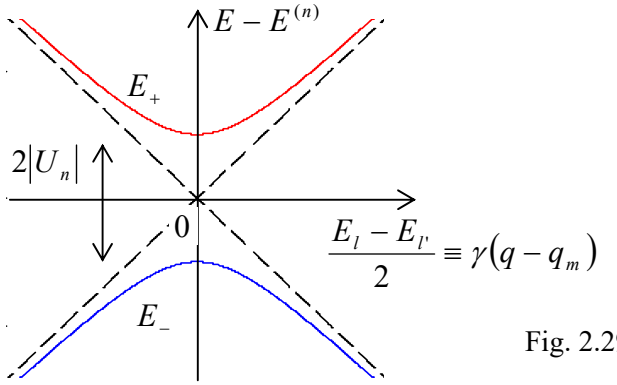


Fig. 2.29. The level anticrossing diagram.

We will run into the anticrossing diagram again and again in the course, notably at the discussion of spin-½ and other two-level systems. It is also repeatedly met in classical mechanics, for example at the calculation of frequencies of coupled oscillators.[66,67] In our current case of the weak potential limit of the band theory, the diagram describes the interaction of two traveling de Broglie waves (217), with oppositely directed wave vectors, l and $-l'$, via the $(l - l')^{th}$ (i.e. the n^{th}) Fourier harmonic of the potential profile $U(x)$.[68] This effect exists also in the classical wave theory and is known as the *Bragg reflection*, describing, for example, a 1D model of the X-wave reflection by a crystal lattice (see, e.g. Fig. 1.5) in the limit of weak interaction between the incident wave and each atom.

The anticrossing diagram shows that rather counter-intuitively, even a weak periodic potential changes the topology of the initially parabolic dispersion relation radically, connecting its different branches, and thus creating the energy gaps. Let me hope that the reader has enjoyed the elegant description of this effect, discussed above, as well as one more illustration of the wonderful ability of physics to give completely different interpretations (and different approximate approaches) to the same effect in opposite limits.

So, we have explained analytically (though only in two limits) the particular band structure shown in Fig. 26. Now one may wonder how general this structure is, i.e. how much of it is independent of the Dirac comb model (Fig. 24). For that, let us represent the band pattern, such as that shown in Fig. 26b (plotted for a particular value of the parameter β, characterizing the potential barrier strength) in a more condensed form, which would allow us to place the results for a range of β values on a single comprehensible plot. The way to do this should be clear from Fig. 26b: since the dependence of energy on the quasimomentum in each energy band is not too eventful, we may plot just the highest and the smallest values of the particle's energy $E = \hbar^2k^2/2m$ as functions of $\beta \equiv ma\mathcal{W}/\hbar^2$ – see Fig. 30, which may be obtained from Eq. (198) with $qa = 0$ and $qa = \pi$.

[66] See, e.g., CM Sec. 6.1 and in particular Fig. 6.2.

[67] Actually, we could readily obtain this diagram in the previous section, for the system of two weakly coupled potential wells (Fig. 21), if we assumed the wells to be slightly dissimilar.

[68] In the language of the de Broglie wave scattering, to be discussed in Sec. 3.3, Eq. (220) may be interpreted as the condition that each of these waves, scattered on the n^{th} Fourier harmonic of the potential profile, constructively interferes with its counterpart, leading to a strong enhancement of their interaction.

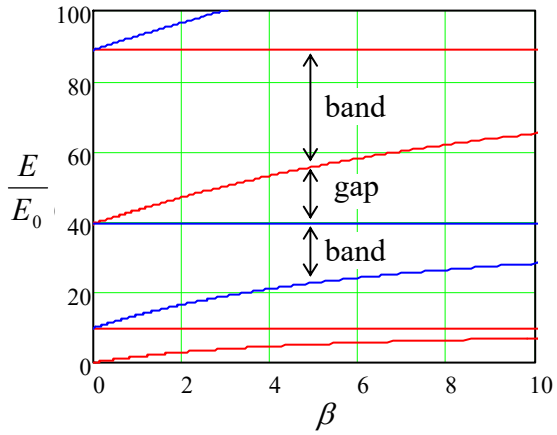


Fig. 2.30. Characteristic curves of the Schrödinger equation for the infinite Dirac comb (Fig. 24).

These plots (in mathematics, commonly called *characteristic curves*, while in applied physics and electronic engineering, *band-edge diagrams*) show, first of all, that at small β, all energy gap widths are equal and proportional to this parameter, and hence to $\mathcal{W}$. This feature is in a full agreement with the main conclusion (224) of our general analysis of the weak-potential limit, because for the Dirac comb potential (Fig. 24),

$$U(x) = \mathcal{W} \sum_{j=-\infty}^{+\infty} \delta(x - ja + \text{const}), \tag{2.226}$$

all Fourier harmonic amplitudes defined by Eq. (207), are equal by magnitude: $|U_l| = \mathcal{W}/a$. As β is further increased, the gaps grow and the allowed energy bands shrink, but rather slowly. This is also natural, because, as Eq. (79) shows, the transparency $\mathcal{T}$ of the delta-functional barriers separating the quasi-localized states (and hence the coupling parameters $\delta_n \propto \mathcal{T}^{1/2}$ participating in the general tight-binding limit's theory) decrease with $\mathcal{W} \propto \beta$ very gradually.

These features may be compared with those for more realistic and relatively simple periodic functions $U(x)$, for example, the sinusoidal potential $U(x) = A\cos(2\pi x/a)$ – see Fig. 31a.

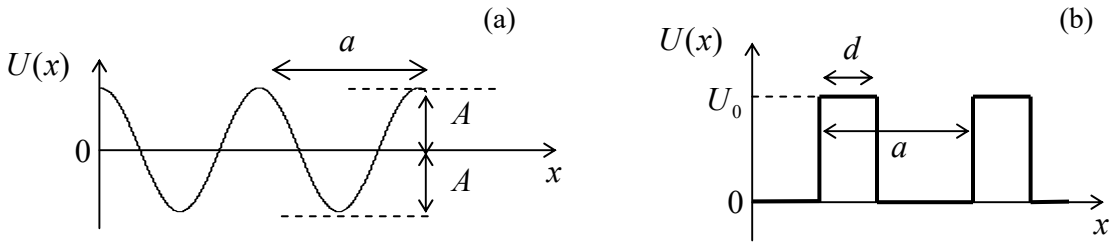


Fig. 2.31. Two other simple periodic potential profiles: (a) the sinusoidal ("Mathieu") potential and (b) the Kronig-Penney potential.

For this potential, the stationary Schrödinger equation (53) takes the following form:

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + A\cos\frac{2\pi x}{a}\psi = E\psi\,. \tag{2.227}$$

By the introduction of dimensionless variables

$$\xi \equiv \frac{\pi x}{a}, \quad \alpha \equiv \frac{E}{E^{(1)}}, \quad 2\beta \equiv \frac{A}{E^{(1)}}, \tag{2.228}$$

where $E^{(1)}$ is defined by Eq. (221) with $n = 1$,[69] Eq. (227) is reduced to the canonical form of the well-studied *Mathieu equation*[70]

Mathieu equation

$$\boxed{\frac{d^2\psi}{d\xi^2} + (\alpha - 2\beta\cos 2\xi)\psi = 0.} \tag{2.229}$$

Figure 32 shows the characteristic curves of this equation. We see that now at small β the first energy gap grows much faster than the higher ones: $\Delta_n \propto \beta^n$. This feature is in accord with the weak-coupling result $\Delta_1 = 2|U_1|$, which is valid only in the linear approximation in U_n, because for the Mathieu potential, $U_l = A(\delta_{l,+1} + \delta_{l,-1})/2$. Another clearly visible feature is the exponentially fast shrinkage of the allowed energy bands at $2\beta > \alpha$ (in Fig. 32, on the right from the dashed line), i.e. at $E < A$. It may be readily explained by our tight-binding approximation result (206): as soon as the eigenenergy drops significantly below the potential maximum $U_{max} = A$ (see Fig. 31a), the quantum states in the adjacent potential wells are connected only by tunneling through relatively high potential barriers separating these wells, so the coupling amplitudes δ_n become exponentially small – see, e.g., Eq. (189).

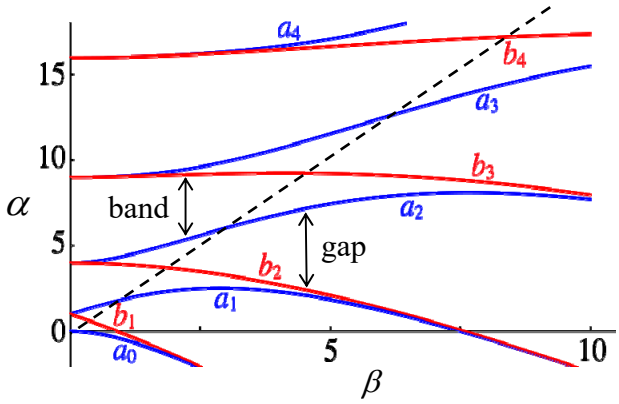


Fig. 2.32. Characteristic curves of the Mathieu equation. The dashed line corresponds to the equality $\alpha = 2\beta$, i.e. $E = A \equiv U_{max}$, separating the regions of under-barrier tunneling and over-barrier motion. Adapted from Fig. 28.2.1 at http://dlmf.nist.gov as a contribution by the US Government (not subject to copyright).

Another simple periodic profile is the *Kronig-Penney potential* shown in Fig. 31b, which gives relatively simple analytical expressions for the band/gap patterns (though transcendent equations for the characteristic curves). Its advantage over the Dirac comb (226) is a more realistic law of the decrease of the Fourier harmonics U_l at $l >> 1$, and hence of the energy gaps in the weak-potential limit:

$$\Delta_n \approx 2|U_n| \propto \frac{U_0}{n}, \quad \text{at } E \sim E^{(n)} >> U_0. \tag{2.230}$$

Leaving a detailed analysis of the Kronig-Penney potential for the reader's exercise, let me conclude this section by addressing the effect of potential modulation on the number of eigenstates in 1D systems of a large but finite length $l >> a, k^{-1}$. Perhaps surprisingly, the Bloch theorem makes the

[69] Note that this definition of β is quantitatively different from that for the Dirac comb (226), but in both cases, this parameter is proportional to the amplitude of the potential's periodic modulation.

[70] This equation, first studied in the 1860s by É. Mathieu in the context of a rather practical problem of vibrating elliptical drumheads (!), has many other important applications in physics and engineering, notably including the parametric excitation of oscillations – see, e.g., CM Sec. 5.5.

analysis of this problem elementary, for arbitrary $U(x)$. Indeed, let us assume that l is comprised of an integer number of periods a, and its ends are described by similar boundary conditions – both assumptions evidently inconsequential for $l >> a$. Then, according to Eq. (210), the boundary conditions impose, on the quasimomentum q, exactly the same quantization condition as we had for k for a free 1D motion. Hence, instead of Eq. (1.93), we can write

$$dN = \frac{l}{2\pi} dq \,, \tag{2.231}$$

1D number of states

with the corresponding change of the summation rule:

$$\sum_q f(q) \to \frac{l}{2\pi} \int f(q) dk \,. \tag{2.232}$$

As a result, the density of states in the 1D *q-space,* $dN/dq = l/2\pi$, does not depend on the potential profile at all! Note, however, that the profile does affect the density of states on the *energy* scale, dN/dE. As an extreme example, on the bottom and at the top of each energy band we have $dE/dq \to 0$, and hence

$$\frac{dN}{dE} = \frac{dN}{dq} \Big/ \frac{dE}{dq} = \frac{l}{2\pi} \Big/ \frac{dE}{dq} \to \infty \,. \tag{2.233}$$

This effect of state concentration at the band/gap edges (which survives in higher spatial dimensionalities as well) has important implications for the operation of several important electronic and optical devices, in particular semiconductor lasers and light-emitting diodes.

2.8. Periodic systems: Particle dynamics

The band structure of the energy spectrum of a particle moving in a periodic potential has profound implications not only for its density of states but also for its dynamics. Indeed, let us consider the simplest case of a wave packet composed of the Bloch functions (210), all belonging to the same (say, n^{th}) energy band. Similarly to Eq. (27) for a free particle, we can describe such a packet as

$$\Psi(x,t) = \int a_q u_q(x) e^{i[qx - \omega(q)t]} dq \,, \tag{2.234}$$

where the a-periodic functions $u(x)$, defined by Eq. (208), are now indexed to emphasize their dependence on the quasimomentum, and $\omega(q) \equiv E_n(q)/\hbar$ is the function of q describing the shape of the corresponding energy band – see, e.g., Fig. 26b or Fig. 28. If the packet is narrow in the q-space, i.e. if the width δq of the distribution a_q is much smaller than all the characteristic q-scales of the dispersion relation $\omega(q)$, in particular than π/a, we may simplify Eq. (234) exactly as it was done in Sec. 2 for a free particle, despite the presence of the periodic factors $u_q(x)$ under the integral. In the linear approximation of the Taylor expansion, we get a full analog of Eq. (32), but now with q rather than k, and

$$v_{gr} = \left. \frac{d\omega}{dq} \right|_{q=q_0} , \quad \text{and} \quad v_{ph} = \left. \frac{\omega}{q} \right|_{q=q_0} , \tag{2.235}$$

where q_0 is the central point of the quasimomentum's distribution. Despite the formal similarity with Eqs. (33) for the free particle, this result is much more eventful. For example, as evident from the dispersion relation's topology (see Figs. 26b, 28), the group velocity vanishes not only at $q = 0$, but at all

values of q that are multiples of (π/a), i.e. at the bottom and on the top of each energy band. Even more intriguing, the group velocity's sign changes periodically with q.

This group velocity alternation leads to fascinating, counter-intuitive phenomena if a particle placed in a periodic potential is the subject of an additional external force $F(t)$. (For an electron, this may be, for example, the force exerted by the applied electric field.) Let the force be relatively weak so that the product Fa (i.e. the scale of the energy increment from the additional force per one lattice period) is much smaller than both relevant energy scales of the dispersion relation $E(q)$ – see Fig. 26b:

$$Fa << \Delta E_n, \Delta_n . \tag{2.236}$$

This strong relation enables us to neglect the force-induced interband transitions, so the wave packet (234) includes the Bloch eigenfunctions belonging to only one (initial) energy band at all times. The time evolution of its center q_0 obeys an extremely simple equation of motion:[71]

Time evolution of quasi-momentum

$$\boxed{\dot{q}_0 = \frac{1}{\hbar} F(t) .} \tag{2.237}$$

This equation is physically very transparent: it is essentially the 2nd Newton law for the time evolution of the quasimomentum $\hbar q$ under the effect of the additional force $F(t)$ only, excluding the periodic force $-\partial U(x)/\partial x$ of the background potential $U(x)$. This is very natural, because as Eq. (210) implies, $\hbar q$ is essentially the particle's momentum $\hbar k$ averaged over the potential's period, and the periodic force effect drops out at such an averaging.

Despite the simplicity of Eq. (237), the results of its solution may be highly nontrivial. First, let us use Eqs. (235) and (237) to find the instant *group acceleration* of the particle (i.e. the acceleration of its wave packet's envelope):

$$a_{\rm gr} \equiv \frac{dv_{\rm gr}}{dt} \equiv \frac{d}{dt}\frac{d\omega(q_0)}{dq_0} \equiv \frac{d}{dq_0}\frac{d\omega(q_0)}{dq_0}\frac{dq_0}{dt} = \frac{d^2\omega(q_0)}{dq_0^2}\frac{dq_0}{dt} = \frac{1}{\hbar}\frac{d^2\omega}{dq^2}\Big|_{q=q_0} F(t) . \tag{2.238}$$

This means that the second derivative of the dispersion relation $\omega(q)$ (specific for each energy band) plays the role of the effective reciprocal mass of the particle at this particular value of q_0:

Effective mass

$$\boxed{m_{\rm ef} = \frac{\hbar}{d^2\omega/dq^2} \equiv \frac{\hbar^2}{d^2E_n/dq^2} .} \tag{2.239}$$

For the particular case of a free particle, for which Eq. (216) is exact, this expression is reduced to the original (and constant) mass m, but generally, the effective mass depends on the wave packet's momentum. According to Eq. (239), at the bottom of any energy band, $m_{\rm ef}$ is always positive but depends on the strength of the particle's interaction with the periodic potential. In particular, according to Eq. (206), in the tight-binding limit, the effective mass is very large:

$$\left|m_{\rm ef}\right|_{q=(\pi/a)n} = \frac{\hbar^2}{2\delta_n a^2} \equiv m\frac{E^{(1)}}{\pi^2\delta_n} >> m . \tag{2.240}$$

[71] The proof of Eq. (237) is not difficult but is more compact in the bra-ket formalism to be discussed in Chapter 4. This is why I recommend to the reader its proof as an exercise after reading that chapter.

On the contrary, in the weak-potential limit, the effective mass is close to m at most points of each energy band, but at the edges of the (narrow) bandgaps, it is much smaller. Indeed, expanding Eq. (224) in the Taylor series near point $q = q_m$, we get

$$E_{\pm}\Big|_{E\approx E^{(n)}} - E_{\text{ave}} \approx \pm|U_n| \pm \frac{1}{2|U_n|}\left(\frac{dE_1}{dq}\right)^2_{q=q_m} \tilde{q}^2 = \pm|U_n| \pm \frac{\gamma^2}{2|U_n|}\tilde{q}^2, \tag{2.241}$$

where γ and $\tilde{q}$ are defined by Eq. (225), and hence

$$|m_{\text{ef}}|_{q=q_m} = |U_n|\frac{\hbar^2}{\gamma^2} \equiv m\frac{|U_n|}{2E^{(n)}} << m\,. \tag{2.242}$$

The effective mass effects in real atomic crystals may be very significant. For example, the charge carriers in silicon have $m_{\text{ef}} \approx 0.19\ m_{\text{e}}$ in the lowest, normally-empty energy band (traditionally called the *conduction band*), and $m_{\text{ef}} \approx 0.98\ m_{\text{e}}$ in the adjacent lower, normally-filled *valence band*. In some semiconducting compounds, the conduction-band mass may be even smaller – down to $0.0145\ m_{\text{e}}$ in InSb!

However, the effective mass magnitude is not the most surprising effect. A more fascinating corollary of Eq. (239) is that on the top of each energy band, the effective mass is *negative* – please revisit Figs. 26b, 28, and 29 again. This means that the particle (or more strictly, its wave packet's envelope) is accelerated in the direction *opposite* to the applied force. This is exactly what electronic engineers, working with electrons in semiconductors, call *holes*, characterizing them by a *positive mass* $|m_{\text{ef}}|$, but compensating this sign change by taking their charge e positive. If the particle stays in close vicinity of the energy band's top (say, due to frequent scattering effects, typical for the semiconductors used in engineering practice), such double sign flip does not lead to an error in calculations of hole's dynamics, because the electric field's force is proportional to the particle's charge, so the particle's acceleration a_{gr} is proportional to the charge-to-mass ratio.[72]

However, in some phenomena such simple representation is unacceptable.[73] For example, let us form a narrow wave packet at the bottom of the lowest energy band,[74] and then exert on it a constant force $F > 0$ – say, due to a constant external electric field directed along the x-axis. According to Eq. (237), this force would lead to linear growth of q_0 in time, so in the quasimomentum space, the packet's center would slide, with a constant speed, along the q axis – see Fig. 33a. Close to the energy band's bottom, this motion would correspond to a positive effective mass (possibly, somewhat different than the genuine particle's mass m), and hence be close to the free particle's acceleration. However, as soon as q_0 has reached the inflection point where $d^2E_1/dq^2 = 0$, the effective mass, and hence its acceleration (238) change signs to negative, i.e. the packet starts to slow down (in the direct space), while still moving ahead with the same velocity in the quasimomentum space. Finally, at the energy band's top, the particle stops at a certain x_{max}, while continuing to move forward in the q-space.

[72] More discussion of this issue may be found in SM Sec. 6.4.

[73] The balance of this section describes effects that are not discussed in most quantum mechanics textbooks. Though, in my opinion, every educated physicist should be aware of them, some readers may skip them at the first reading, jumping directly to the next Sec. 9.

[74] Physical intuition tells us (and the theory of open systems, to be discussed in Chapter 7, confirms) that this may be readily done, for example, by weakly coupling the system to a relatively low-temperature environment, and letting it relax to the lowest possible energy.

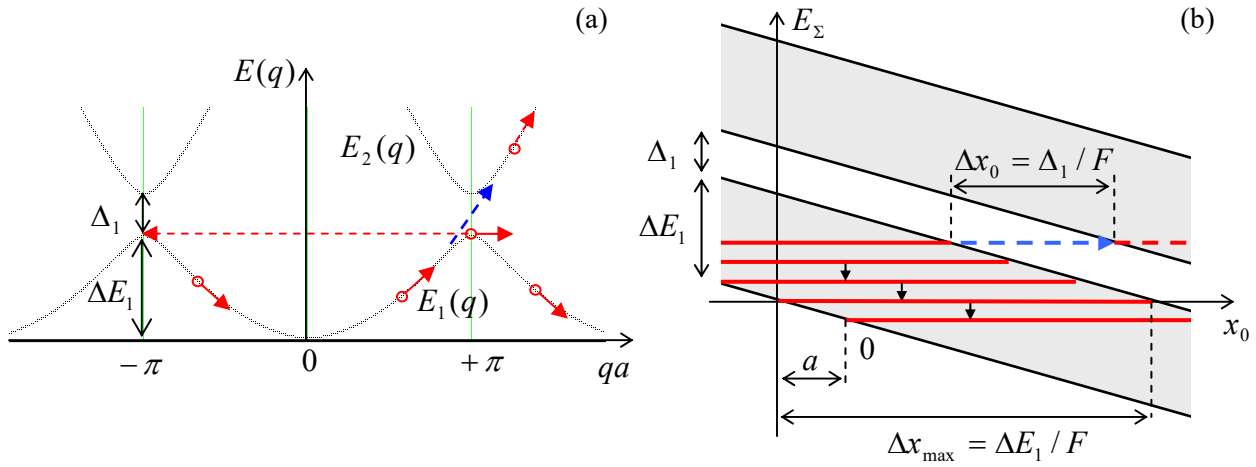


Fig. 2.33. The Bloch oscillations (red lines) and the Landau-Zener tunneling (blue arrows) represented in: (a) the reciprocal space of q, and (b) the direct space. On panel (b), the tilted gray strips show the allowed energy bands, while the bold red lines, the Wannier-Stark ladder's steps.

Now we have two alternative ways to look at the further time evolution of the wave packet along the quasimomentum's axis. From the extended zone picture (which is the simplest for this analysis, see Fig. 33a),[75] we may say that the particle crosses the 1st Brillouin zone's boundary and continues to go forward in q-space, i.e. down the lowest energy band. According to Eq. (235), this region (up to the next energy minimum at $qa = 2\pi$) corresponds to a negative group velocity. After q_0 has reached that minimum, the whole process repeats again – and again, and again.

These are the famous *Bloch oscillations* – the effect which had been predicted, by the same F. Bloch, as early as 1929 but evaded experimental observation until the 1980s (see below) due to the strong scattering effects in real solid-state crystals. The time period of the oscillations may be readily found from Eq. (237):

$$\Delta t_B = \frac{\Delta q}{dq/dt} = \frac{2\pi/a}{F/\hbar} = \frac{2\pi\hbar}{Fa}, \tag{2.243}$$

so their frequency may be expressed by a very simple formula

Bloch oscillations: frequency

$$\boxed{\omega_B \equiv \frac{2\pi}{\Delta t_B} = \frac{Fa}{\hbar}}, \tag{2.244}$$

and hence is independent of any peculiarities of the energy band/gap structure.

The direct-space motion of the wave packet's center $x_0(t)$ during the Bloch oscillation process may be analyzed by integrating the first of Eqs. (235) over some time interval Δt, and using Eq. (237):

[75] This phenomenon may be also discussed from the point of view of the reduced zone picture, but then it requires the introduction of instant jumps between the Brillouin zone boundary points (see the dashed red line in Fig. 33) that correspond to physically equivalent states of the particle. Evidently, for the description of this particular phenomenon, this language is more artificial.

$$\Delta x_0(t) \equiv \int_0^{\Delta t} v_{gr}\,dt = \int_0^{\Delta t} \frac{d\omega(q_0)}{dq_0}\,dt \equiv \int_0^{\Delta t} \frac{d\omega(q_0)}{dq_0/dt} = \frac{\hbar}{F}\int_{t=0}^{t=\Delta t} d\omega = \frac{\hbar}{F}\Delta\omega(q_0). \tag{2.245}$$

If the interval Δt is equal to the Bloch oscillation period Δt_B (243), the initial and final values of $E(q_0) = \hbar\omega(q_0)$ are equal, giving $\Delta x_0 = 0$: in the end of the period, the wave packet returns to its initial position in space. However, if we carry out this integration only from the smallest to the largest values of $\omega(q_0)$, i.e. the adjacent points where the group velocity vanishes, we get the following Bloch oscillation swing:

Bloch oscillations: spatial swing

$$\boxed{\Delta x_{max} = \frac{\hbar}{F}(\omega_{max} - \omega_{min}) \equiv \frac{\Delta E_1}{F}.} \tag{2.246}$$

This simple result may be interpreted using an alternative energy diagram (Fig. 33b), which results from the following arguments. The additional force F may be described not only via the 2nd Newton law's version (237), but, alternatively, by its contribution $-Fx$ to the Gibbs potential energy[76]

$$U_\Sigma(x) = U(x) - Fx \tag{2.247}$$

The exact solution of the Schrödinger equation (61) with such a potential may be hard to find directly, but if the force F is sufficiently weak, as we are assuming throughout this discussion, the second term in Eq. (247) may be considered as a constant on the scale of $a << \Delta x_{max}$. In this case, our quantum-mechanical treatment of the periodic potential $U(x)$ is still virtually correct, but with an energy shift depending on the "global" position x_0 of the packet's center. In this approximation, the total energy of the wave packet is

$$E_\Sigma = E(q_0) - Fx_0. \tag{2.248}$$

In a plot of such energy as a function of x_0 (Fig. 33b), the energy dependence on q_0 is hidden, but as was discussed above, it is rather uneventful and may be well characterized by the position of band-gap edges on the energy axis.[77] In this representation, the Bloch oscillations keep the full energy E_Σ of the particle constant, i.e. follow a horizontal line in Fig. 33b, limited by the classical turning points corresponding to the bottom and the top of the allowed energy band. The distance Δx_{max} between these points is evidently given by Eq. (246).

Besides this alternative look at the Bloch oscillation swing, the total energy diagram shown in Fig. 33b enables one more remarkable result. Let a wave packet be so narrow in the momentum space that $\delta x \sim 1/\delta q >> \Delta x_{max}$; then it may be well represented by a definite energy, i.e. by a horizontal line in Fig. 33b. But Eq. (247) is exactly invariant with respect to the following simultaneous translation of the coordinate and the energy:

$$x \to x + a, \quad E \to E - Fa. \tag{2.249}$$

[76] Physically, this is just the relevant part of the potential energy of the total system comprised of our particle (in the periodic potential) and the source of the force F – see, e.g., CM Sec. 1.4.

[77] In semiconductor physics and engineering, such spatial *band-edge diagrams* are virtually unavoidable components of almost every discussion/publication. In this series, a few more examples of such diagrams may be found in SM Sec. 6.4.

This means that it is satisfied by an infinite set of similar solutions, each corresponding to one of the horizontal red lines shown in Fig. 33b. This is the famous *Wannier-Stark ladder*,[78] with the step height

Wannier-Stark ladder

$$\Delta E_{\rm WS} = Fa\,. \tag{2.250}$$

The importance of this alternative representation of the Bloch oscillations is due to the following fact. In most experimental realizations, the power of electromagnetic radiation with frequency (244), which may be extracted from the oscillations of a charged particle, is very low, so their direct detection represents a hard problem.[79] However, let us apply to a Bloch oscillator an additional ac field at frequency $\omega \approx \omega_{\rm B}$. As these frequencies are brought close together, the external signal should synchronize ("phase-lock") the Bloch oscillations,[80] resulting in certain changes of time-independent observables – for example, a resonant change of absorption of the external radiation. Now let us notice that the combination of Eqs. (244) and (250) yield the following simple relation:

$$\Delta E_{\rm WS} = \hbar\omega_{\rm B}\,. \tag{2.251}$$

This means that the phase-locking at $\omega \approx \omega_{\rm B}$ allows for an alternative (but equivalent) interpretation – as the result of ac-field-induced quantum transitions[81] between the steps of the Wannier-Stark ladder. (Again, such occasions when two very different languages may be used for alternative interpretations of the same effect is one of the most beautiful features of physics.)

This phase-locking effect has been used for the first experimental confirmations of the Bloch oscillation theory.[82] For this purpose, the natural periodic structures, solid-state crystals, are inconvenient due to their very small period $a \sim 10^{-10}$ m. Indeed, according to Eq. (244), such structures require very high forces F (and hence very high electric fields $\mathscr{E} = F/e$) to bring $\omega_{\rm B}$ to an experimentally convenient range. This problem has been overcome using artificial periodic structures (*superlattices*) of certain semiconductor compounds, such as $Ga_{1-x}Al_xAs$ with various degrees x of the gallium-to-aluminum atom replacement, whose layers may be grown over each other epitaxially, i.e., with very few crystal structure violations. Such superlattices, with periods $a \sim 10$ nm, have enabled a clear observation of the resonance at $\omega \approx \omega_{\rm B}$, and hence a measurement of the Bloch oscillation frequency, in particular its proportionality to the applied dc electric field, predicted by Eq. (244).

Very soon after this discovery, the Bloch oscillations were observed[83] in small Josephson junctions, where they result from the quantum dynamics of the Josephson phase difference φ in a 2π-periodic potential profile, created by the junction. A straightforward translation of Eq. (244) to this case (left for the reader's exercise) shows that the frequency of such Bloch oscillations is

---

[78] This effect was first discussed in detail by G. Wannier in his 1959 monograph on solid-state physics, while the name of J. Stark is traditionally associated with virtually any electric field effect on atomic systems after he had discovered the first of such effects in 1913 – see its discussion in Sec. 6.2 below.

[79] In systems with many independent particles (such as electrons in semiconductors), the detection problem is exacerbated by the phase incoherence of the Bloch oscillations performed by each particle. This drawback is absent in atomic Bose-Einstein condensates whose Bloch oscillations (in a periodic potential created by standing optical waves) were eventually observed by M. Ben Dahan *et al.*, *Phys. Rev. Lett.* **76**, 4508 (1996).

[80] A simple analysis of the phase locking of a classical oscillator may be found, e.g., in CM Sec. 5.4. (See also the brief discussion of the phase locking of the Josephson oscillations at the end of Sec. 1.6 of this course.)

[81] A quantitative theory of such transitions will be discussed in Sec. 6.6 and then in Chapter 7.

[82] E. Mendez *et al.*, *Phys. Lev. Lett.* **60**, 2426 (1988).

[83] D. Haviland *et al.*, *Z. Phys. B* **85**, 339 (1991).

$$\omega_{\rm B} = \frac{\pi \bar{I}}{2e}, \qquad \text{i.e. } f_{\rm B} \equiv \frac{\omega_{\rm B}}{2\pi} = \frac{\bar{I}}{2e}, \tag{2.252}$$

where $\bar{I}$ is the dc current passed through the junction – the effect not to be confused with the "classical" Josephson oscillations with frequency (1.75). It is curious that Eq. (252) may be legitimately interpreted as a result of a periodic transfer, through the Josephson junction, of discrete Cooper pairs (with electric charge $-2e$ each), between two coherent Bose-Einstein condensates in the superconducting electrodes of the junction.[84]

So far, our discussion of the Bloch oscillations was based on the premise that the wave packet of the particle stays within one (say, the lowest) energy band. However, just one look at Fig. 28 shows that this assumption becomes unrealistic if the energy gap separating this band from the next one becomes very small, $\Delta_1 \to 0$. Indeed, in the weak-potential approximation, which is adequate in this limit, $|U_1| \to 0$, the two dispersion curve branches (216) cross without any interaction, so if our particle (meaning its the wave packet) is driven to approach that point, it should continue to move up in energy – see the dashed blue arrow in Fig. 33a. Similarly, in the real-space representation shown in Fig. 33b, it is intuitively clear that at $\Delta_1 \to 0$, the particle residing at one of the steps of the Wannier-Stark ladder should be able to somehow overcome the vanishing spatial gap $\Delta x_0 = \Delta_1/F$ and to "leak" into the next band – see the horizontal dashed blue arrow on that panel.

This process, called the *Landau-Zener* (or "interband", or "band-to-band") *tunneling*,[85] is indeed possible. To analyze it, let us first take $F = 0$, and consider what happens if a quantum particle, described by an x-long (and hence E-narrow) wave packet, is incident from free space upon a periodic structure of a large but finite length $l = Na >> a$ – see, e.g., Fig. 22. If the packet's energy E is within one of the energy bands, it may evidently propagate through the structure (though may be partly reflected from its ends). The corresponding quasimomentum may be found by solving the dispersion relation for q; for example, in the weak-potential limit, Eq. (224) (which is valid near the gap) yields

$$q = q_m + \tilde{q}, \quad \text{with} \quad \tilde{q} = \pm \frac{1}{\gamma}\left[\tilde{E}^2 - |U_n|^2\right]^{1/2}, \quad \text{for } |U_n|^2 \le \tilde{E}^2, \tag{2.253}$$

where $\tilde{E} \equiv E_{\pm} - E^{(n)}$ and $\gamma = 2aE^{(n)}/\pi n$ – see the second of Eqs. (225).

Now, if the energy E is inside one of the energy gaps Δ_n, the wave packet's propagation in an infinite periodic lattice is impossible, so it is completely reflected from it. However, our analysis of the potential step problem in Sec. 3 implies that the packet's wavefunction should still have an exponential tail protruding into the structure and decaying on some length δ – see Eq. (58) and Fig. 2.4. Indeed, a straightforward review of the calculations leading to Eq. (253) shows that it remains valid for energies within the gap as well, if the quasimomentum is understood as a purely imaginary number:

$$\tilde{q} \to \pm i\kappa, \quad \text{where } \kappa \equiv \frac{1}{\gamma}\left[|U_n|^2 - \tilde{E}^2\right]^{1/2}, \qquad \text{for } \tilde{E}^2 \le |U_n|^2. \tag{2.254}$$

[84] See, e.g., D. Averin *et al.*, *Sov. Phys. – JETP* **61**, 407 (1985). This effect is qualitatively similar to the transfer of single electrons, with a similar frequency $f = \bar{I}/e$, in tunnel junctions between "normal" (non-superconducting) metals – see, e.g., EM Sec. 2.9 and references therein.

[85] It was predicted, apparently independently, by L. Landau, C. Zener, E. Stueckelberg, and E. Majorana in 1932.

With this replacement, the Bloch solution (193b) indeed describes an exponential decay of the wavefunction at length $\delta \sim 1/\kappa$.

Returning to the effects of weak force F, in the real-space approach described by Eq. (248) and illustrated in Fig. 33b, we may recast Eq. (254) as

$$\kappa \to \kappa(x) = \frac{1}{\gamma}\left[|U_n|^2 - (F\widetilde{x})^2\right]^{1/2}, \tag{2.255}$$

where $\widetilde{x}$ is the particle's (i.e. its wave packet center's) deviation from the midgap point. Thus the gap creates a potential barrier of a finite width $\Delta x_0 = 2|U_n|/F$, through which the wave packet may tunnel with a non-zero probability. As we already know, in the WKB approximation (in our case requiring $\kappa\Delta x_0 >> 1$) this probability is just the potential barrier's transparency $\mathscr{T}$, which may be calculated from Eq. (117):

$$-\ln \mathscr{T} = 2\int_{\kappa(x)^2>0} \kappa(x)dx = \frac{2}{\gamma}\int_{-x_c}^{x_c}\left[|U_n|^2 - (F\widetilde{x})^2\right]^{1/2} d\widetilde{x} = \frac{2|U_n|}{\gamma} 2x_c \int_0^1 \left(1-\xi^2\right)^{1/2} d\xi\,. \tag{2.256}$$

where $\pm x_c \equiv \pm\Delta x_0/2 = \pm|U_n|/F$ are the classical turning points. Working out this simple integral (or just noticing that it is a quarter of the unit circle's area, and hence is equal to $\pi/4$), we get

Landau-Zener tunneling probability

$$\boxed{\mathscr{T} = \exp\left\{-\frac{\pi|U_n|^2}{\gamma F}\right\}}. \tag{2.257}$$

This famous result may be also obtained in a more complex way, whose advantage is a constructive proof that Eq. (257) is valid for an arbitrary relation between γF and $|U_n|^2$, i.e. arbitrary $\mathscr{T}$, while our simple derivation was limited to the WKB approximation, valid only at $\mathscr{T} << 1$.[86] Using Eq. (225), we may rewrite the product γF participating in Eq. (257), as

$$\gamma F = \frac{1}{2}\left|\frac{d(E_l - E_{l'})}{dq_0}\right|_{E_l = E_{l'} = E^{(n)}} \hbar\frac{dq_0}{dt} = \frac{\hbar}{2}\left|\frac{d(E_l - E_{l'})}{dt}\right|_{E_l = E_{l'} = E^{(n)}} \equiv \frac{\hbar u}{2}, \tag{2.258}$$

where u has the meaning of the "speed" of the energy level crossing in the absence of the gap. Hence, Eq. (257) may be rewritten in the form

$$\mathscr{T} = \exp\left\{-\frac{2\pi|U_n|^2}{\hbar u}\right\}, \tag{2.259}$$

which is more transparent physically. Indeed, the fraction $2|U_n|/u = \Delta_n/u$ gives the time scale Δt of the energy's crossing the gap region, and according to the Fourier transform, its reciprocal, $\omega_{max} \sim 1/\Delta t$ gives the upper cutoff of the frequencies essentially involved in the Bloch oscillation process. Hence Eq. (259) means that

$$-\ln \mathscr{T} \approx \frac{\Delta_n}{\hbar\omega_{max}}. \tag{2.260}$$

[86] In Chapter 6 below, Eq. (257) will be derived using a different method, based on the so-called *Golden Rule* of quantum mechanics, but also in the weak-potential limit, i.e. for hyperbolic dispersion law (253).

This formula allows us to interpret the Landau-Zener tunneling as the system's excitation across the energy gap Δ_n by the highest-energy quantum $\hbar\omega_{max}$ available from the Bloch oscillation process. This interpretation remains valid even in the opposite, tight-binding limit, in which, according to Eqs. (206) and (237), the Bloch oscillations are purely sinusoidal, so the Landau-Zener tunneling is completely suppressed at $\hbar\omega_B < \Delta_1$.

Such interband tunneling is an important ingredient of several physical phenomena and even some practical electron devices, for example, the *tunneling* (or "Esaki") *diodes*. This simple device is just a junction of two semiconductor electrodes, one of them so strongly *n*-doped by electron donors that some electrons form a degenerate Fermi gas at the bottom of the conduction band. [87] Similarly, the counterpart semiconductor electrode is *p*-doped so strongly that the Fermi level in the valence band is shifted below the band edge – see Fig. 34.

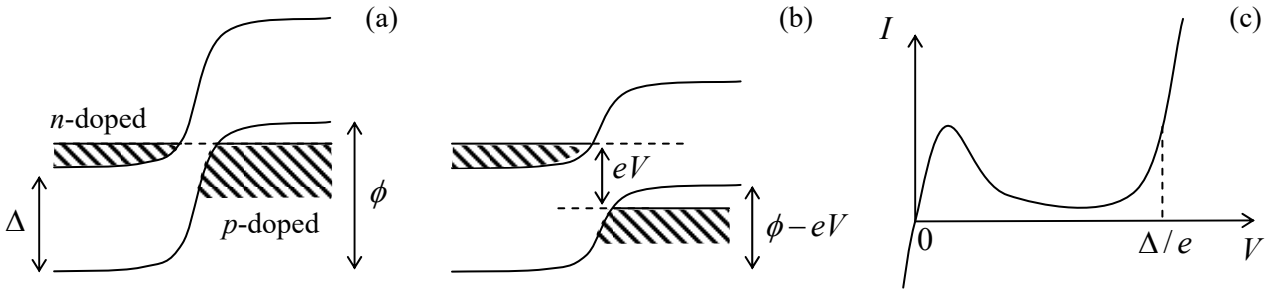


Fig. 2.34. The tunneling ("Esaki") diode: (a) the band-edge diagram of the device at zero bias; (b) the same diagram at a modest positive bias $eV \sim \Delta/2$, and (c) the *I-V* curve of the device (schematically). Dashed lines on panels (a) and (b) show the Fermi-level positions.

In thermal equilibrium, and in the absence of external voltage bias, the Fermi levels of the two electrodes self-align, leading to the build-up of the *contact potential difference* ϕ/e, with ϕ a bit larger than the energy bandgap Δ – see Fig. 34a. This potential difference creates an internal electric field that tilts the energy bands (just as the external field did in Fig. 33b), and leads to the formation of the so-called *depletion layer*, in which the Fermi level is located within the energy gap and hence there are no charge carriers ready to move. In the usual *p-n* junctions, this layer is broad and prevents any current at applied voltages *V* lower than $\sim\Delta/e$. In contrast, in a tunneling diode the depletion layer is so thin (below ~10 nm) that the interband tunneling is possible and provides a substantial Ohmic current at small applied voltages – see Fig. 34c. However, at larger positive biases, with $eV \sim \Delta/2$, the conduction band is aligned with the middle of the energy gap in the *p*-doped electrode, and electrons cannot tunnel there. Similarly, there are no electrons in the *n*-doped semiconductor to tunnel into the available states just above the Fermi level in the *p*-doped electrode – see Fig. 34b. As a result, at such voltages the current drops significantly, to grow again only when *eV* exceeds $\sim\Delta$, enabling electron motion within each energy band. Thus the junction's *I-V* curve has a part with *negative differential resistance* ($dV/dI < 0$) – see Fig. 34c. This phenomenon, equivalent in its effect to negative kinematic friction in mechanics, may

[87] Here I have to rely on the reader's background knowledge of basic semiconductor physics; this picture will be discussed in more detail in SM Sec. 6.4.

be used for amplification of weak analog signals, for self-excitation of electronic oscillators[88] (i.e. an ac signal generation), and for signal swing restoration in digital electronics.

2.9. Harmonic oscillator: Brute force approach

To complete our review of the basic 1D wave mechanics, we have to consider the famous harmonic oscillator, i.e. a 1D particle moving in the quadratic-parabolic potential (111). For it, the stationary Schrödinger equation (53) reads

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2}+\frac{m\omega_0^2x^2}{2}\psi=E\psi\,. \tag{2.261}$$

Conceptually, on the background of the fascinating quantum effects discussed in the previous sections, this is not a very interesting system: Eq. (261) is just a standard 1D eigenproblem, resulting in a discrete energy spectrum E_n, with smooth eigenfunctions $\psi_n(x)$ vanishing at $x \to \pm\infty$ (because the potential energy tends to infinity there).[89] However, as we will repeatedly see later in the course, this problem's solutions have an enormous range of applications, so we have to know their basic properties.

The direct analytical solution of the problem is not very simple (see below), so let us start by trying some indirect approaches to it. First, as was discussed in Sec. 4, the WKB-approximation-based Wilson-Sommerfeld quantization rule (110), applied to this potential, yields the eigenenergy spectrum (114). With the common quantum number convention, this result is

Harmonic oscillator: energy levels

$$E_n=\hbar\omega_0\left(n+\frac{1}{2}\right),\quad \text{with } n=0,1,2,\ldots, \tag{2.262}$$

so (in contrast to the 1D rectangular potential well) the ground-state energy corresponds to $n = 0$. However, as was discussed in the end of Sec. 4, for the quadratic potential (111) the WKB approximation's conditions are strictly satisfied only at $E_n >> \hbar\omega_0$, so at this point, we can only trust Eq. (262) for high levels, with $n >> 1$, rather than for the (most important) ground state.

This is why let me use Eq. (261) to demonstrate another approximate approach, called the *variational method*, whose simplest form is aimed at finding ground states. The method is based on the following observation. (Here I am presenting its 1D wave mechanics form, though the method is much more general.) Let ψ_n be the exact, full, and orthonormal set of stationary wavefunctions of the system under study, and E_n the set of the corresponding energy levels, satisfying Eq. (1.60):

$$\hat{H}\psi_n=E_n\psi_n\,. \tag{2.263}$$

Then we may use this set for the unique expansion of an arbitrary *trial wavefunction*:

$$\psi_{\text{trial}}=\sum_n\alpha_n\psi_n,\qquad \text{so that } \psi^*_{\text{trial}}=\sum_n\alpha_n^*\psi_n^*, \tag{2.264}$$

[88] See, e.g., CM Sec. 5.4.

[89] The stationary state of the harmonic oscillator (which, as will be discussed in Secs. 5.4 and 7.1, may be considered as the state with a definite number of identical bosonic excitations) is sometimes called its *Fock state* – after V. A. Fock. (This term is also used in a more general sense, for definite-particle-number states of systems with indistinguishable bosons of any kind – see Sec. 8.3.)

where α_n are some (generally, complex) coefficients. Let us require the trial function to be normalized, using the condition (1.66) of orthonormality of the eigenfunctions ψ_n:

$$\int \psi_{\text{trial}}^* \psi_{\text{trial}} d^3x \equiv \sum_{n,n'} \int \alpha_n^* \psi_n^* \alpha_{n'} \psi_{n'} d^3x \equiv \sum_{n,n'} \alpha_{n'} \alpha_n^* \int \psi_n^* \psi_{n'} d^3x \equiv \sum_{n,n'} \alpha_{n'} \alpha_n^* \delta_{n,n'} \equiv \sum_n W_n = 1, \quad (2.265)$$

where each of the coefficients W_n, defined as

$$W_n \equiv \alpha_n^* \alpha_n \equiv |\alpha_n|^2 \geq 0, \quad (2.266)$$

may be interpreted as the probability for the particle, in the trial state, to be found in the n^{th} genuine stationary state. Now let us use Eq. (1.23) for a similar calculation of the expectation value of the system's Hamiltonian in the trial state:

$$\begin{aligned} \langle H \rangle_{\text{trial}} &= \int \psi_{\text{trial}}^* \hat{H} \psi_{\text{trial}} d^3x \equiv \sum_{n,n'} \int \alpha_n^* \psi_n^* \hat{H} \alpha_{n'} \psi_{n'} d^3x \equiv \sum_{n,n'} \alpha_{n'} \alpha_n^* E_{n'} \int \psi_n^* \psi_{n'} d^3x \\ &\equiv \sum_{n,n'} \alpha_{n'} \alpha_n^* E_{n'} \delta_{n,n'} \equiv \sum_n W_n E_n . \end{aligned} \quad (2.267)$$

Since the exact ground state energy E_g is, by definition, the lowest one of the set E_n, i.e. $E_n \geq E_g$, Eqs. (265) and (267) yield the following inequality:

$$\langle H \rangle_{\text{trial}} \geq \sum_n W_n E_g \equiv E_g \sum_n W_n = E_g . \quad (2.268)$$

Variational method's justification

Thus, the genuine ground state energy of the system is always lower than (or equal to) its energy in any trial state. Hence, if we make several attempts with reasonably selected trial wavefunctions, we may expect the lowest of the results to approximate the genuine ground state energy reasonably well. Even more conveniently, if we select some reasonable class of trial wavefunctions dependent on a free parameter λ, then we may use the necessary condition of the minimum of $\langle H \rangle_{\text{trial}}$,

$$\frac{\partial \langle H \rangle_{\text{trial}}}{\partial \lambda} = 0, \quad (2.269)$$

to find the closest of them to the genuine ground state. Sometimes, even better results may be obtained using trial wavefunctions dependent on several parameters. Note, however, that the variational method does not tell us how exactly the trial function should be selected, or how close its final result is to the genuine ground-state function. In this sense, this method has "uncontrollable accuracy", and differs from both the WKB approximation and the perturbation methods (to be discussed in Chapter 6), for which we have certain accuracy criteria. Because of this drawback, the variational method is typically used as the last resort – though sometimes (as in the example that follows) it works remarkably well.[90]

Let us apply this method to the harmonic oscillator. Since the potential (111) is symmetric with respect to point $x = 0$, and continuous at all points (so, according to Eq. (261), $d^2\psi/dx^2$ has to be continuous as well), the most natural selection of the ground-state trial function is the Gaussian function

[90] The variational method may be used also to estimate the first excited state (or even a few lowest excited states) of the system, by requiring the new trial function to be orthogonal to the previously calculated eigenfunctions of the lower-energy states. However, the method's error typically grows with the state number.

$$\psi_{\text{trial}}(x) = C\exp\{-\lambda x^2\}, \tag{2.270}$$

with some real $\lambda > 0$. The normalization coefficient C may be immediately found either from the standard Gaussian integration of $|\psi_{\text{trial}}|^2$, or just from the comparison of this expression with Eq. (16), in which $\lambda = 1/(2\delta x)^2$, i.e. $\delta x = 1/2\lambda^{1/2}$, giving $|C|^2 = (2\lambda/\pi)^{1/2}$. Now the expectation value of the particle's Hamiltonian,

$$\hat{H} = \frac{\hat{p}^2}{2m} + U(x) = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{m\omega_0^2 x^2}{2}, \tag{2.271}$$

in the trial state, may be calculated as

$$\begin{aligned}\langle H\rangle_{\text{trial}} &\equiv \int_{-\infty}^{+\infty}\psi_{\text{trial}}^*\left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{m\omega_0^2 x^2}{2}\right)\psi_{\text{trial}}\,dx \\ &= \left(\frac{2\lambda}{\pi}\right)^{1/2}\left[\frac{\hbar^2\lambda}{m}\int_0^{\infty}\exp\{-2\lambda x^2\}dx + \left(\frac{m\omega_0^2}{2} - \frac{2\hbar^2\lambda^2}{m}\right)\int_0^{\infty} x^2\exp\{-2\lambda x^2\}dx\right].\end{aligned} \tag{2.272}$$

Both involved integrals are of the same well-known Gaussian type,[91] giving

$$\langle H\rangle_{\text{trial}} = \frac{\hbar^2}{2m}\lambda + \frac{m\omega_0^2}{8\lambda}. \tag{2.273}$$

As a function of λ, this expression has a single minimum at the value λ_{opt} that may be found from the requirement (269), giving $\lambda_{\text{opt}} = m\omega_0/2\hbar$. The resulting minimum of $\langle H\rangle_{\text{trial}}$ is *exactly* equal to ground-state energy following from Eq. (262),

Harmonic oscillator: ground state energy

$$\boxed{E_0 = \frac{\hbar\omega_0}{2}}. \tag{2.274}$$

Such a coincidence of results of the WKB approximation and of the variational method is rather unusual and implies (though does not prove) that Eq. (274) is exact. As a minimum, this coincidence gives a strong motivation to verify the trial wavefunction (270), with $\lambda = \lambda_{\text{opt}}$, i.e.

Harmonic oscillator: ground state wavefunction

$$\boxed{\psi_0 = \left(\frac{m\omega_0}{\pi\hbar}\right)^{1/4}\exp\left\{-\frac{m\omega_0 x^2}{2\hbar}\right\}}, \tag{2.275}$$

and its energy (274), by plugging them into the Schrödinger equation (261). Such substitution[92] shows that the equation is indeed exactly satisfied.

According to Eq. (275), the characteristic scale of the wavefunction's spatial spread[93] is

Harmonic oscillator: spatial scale

$$\boxed{x_0 \equiv \left(\frac{\hbar}{m\omega_0}\right)^{1/2}}. \tag{2.276}$$

Due to the importance of this scale, let us give its crude estimates for several representative systems:[94]

[91] See, e.g., MA Eqs. (6.9b) and (6.9c).

[92] Actually, this is a twist on one of the tasks of Problem 1.13.

[93] Quantitatively, as was already mentioned in Sec. 2.1, $x_0 = \sqrt{2}\delta x = \langle 2x^2\rangle^{1/2}$.

(i) For atom-bound electrons in solids and fluids, $m \sim 10^{-30}$ kg, and $\omega_0 \sim 10^{15}$ s^{-1}, giving $x_0 \sim 0.3$ nm, of the order of the typical inter-atomic distances in condensed matter. As a result, classical mechanics is not valid at all for the analysis of their motion.

(ii) For atoms in solids, $m \approx 10^{-24}$-10^{-26} kg, and $\omega_0 \sim 10^{13}$ s^{-1}, giving $x_0 \sim 0.01 - 0.1$ nm, i.e. somewhat smaller than inter-atomic distances. Because of that, the methods based on classical mechanics (e.g., molecular dynamics) are approximately valid for the analysis of atomic motion, though they may miss some effects exhibited by lighter atoms – e.g., the so-called *quantum diffusion* of hydrogen atoms, due to their tunneling through the energy barriers of the potential profiles created by other atoms.

(iii) Recently, the progress of patterning technologies has enabled the fabrication of high-quality *micromechanical* oscillators, still consisting of zillions of atoms. For example, the oscillator used in one of the pioneering experiments in this field[95] was a ~1-μm thick membrane with a 60-μm diameter, and had $m \sim 2\times10^{-14}$ kg and $\omega_0 \sim 3\times10^{10}$ s^{-1}, so $x_0 \sim 4\times10^{-16}$ m. It is remarkable that despite such extreme smallness of x_0 (much smaller than not only any atom but even any atomic nucleus!), quantum states of such oscillators may be manipulated and measured, using their coupling to electromagnetic (in particular, optical) resonant cavities.[96]

Returning to the Schrödinger equation (261), in order to analyze its higher eigenstates, we will need more help from mathematics. Let us recast this equation into a dimensionless form by introducing the natural dimensionless variable $\xi \equiv x/x_0$. This gives

$$-\frac{d^2\psi}{d\xi^2} + \xi^2\psi = \varepsilon\psi\,, \tag{2.277}$$

where $\varepsilon \equiv 2E/\hbar\omega_0 = E/E_0$. In this notation, the ground state's wavefunction (275) is proportional to $\exp\{-\xi^2/2\}$. Using this clue, let us look for solutions of Eq. (277) in the form

$$\psi = C\exp\left\{-\frac{\xi^2}{2}\right\}H(\xi)\,, \tag{2.278}$$

where $H(\xi)$ is a new function, and C is the normalization constant. With this substitution, Eq. (277) yields

$$\frac{d^2H}{d\xi^2} - 2\xi\frac{dH}{d\xi} + (\varepsilon - 1)H = 0\,. \tag{2.279}$$

It is evident that H = const and $\varepsilon = 1$ is one of its solutions, describing the ground-state eigenfunction (275) and energy (274), but what are the other eigenstates and eigenvalues? Fortunately, the linear differential equation (279) was studied in detail in the mid-1800s by C. Hermite who has shown that all its eigenvalues are given by the set

---

[94] By order of magnitude, such estimates are also valid for the systems whose dynamics is substantially different from that of harmonic oscillators, if a typical frequency of their quantum transitions is taken for ω_0.

[95] A. O'Connell *et al*., *Nature* **464**, 697 (2010).

[96] See a review of such experiments by M. Aspelmeyer *et al*., *Rev. Mod. Phys*. **86**, 1391 (2014), and also more recent experiments with nanoparticles placed in much "softer" potential wells – e.g., by U. Delić *et al*., *Science* **367**, 892 (2020).

$$\varepsilon_n - 1 = 2n, \quad \text{with} \quad n = 0, 1, 2, \ldots, \tag{2.280}$$

so Eq. (262) is indeed exact for any n. The eigenfunction of Eq. (279), corresponding to the eigenvalue ε_n, is a polynomial (called the *Hermite polynomial*) of degree n, which may be most conveniently calculated using the following explicit formula:

Hermite polynomials

$$H_n = (-1)^n \exp\{\xi^2\} \frac{d^n}{d\xi^n} \exp\{-\xi^2\}. \tag{2.281}$$

It is easy to use this formula to spell out several lowest-degree polynomials – see Fig. 35a:

$$H_0 = 1, \quad H_1 = 2\xi, \quad H_2 = 4\xi^2 - 2, \quad H_3 = 8\xi^3 - 12\xi, \quad H_4 = 16\xi^4 - 48\xi^2 + 12, \ldots \tag{2.282}$$

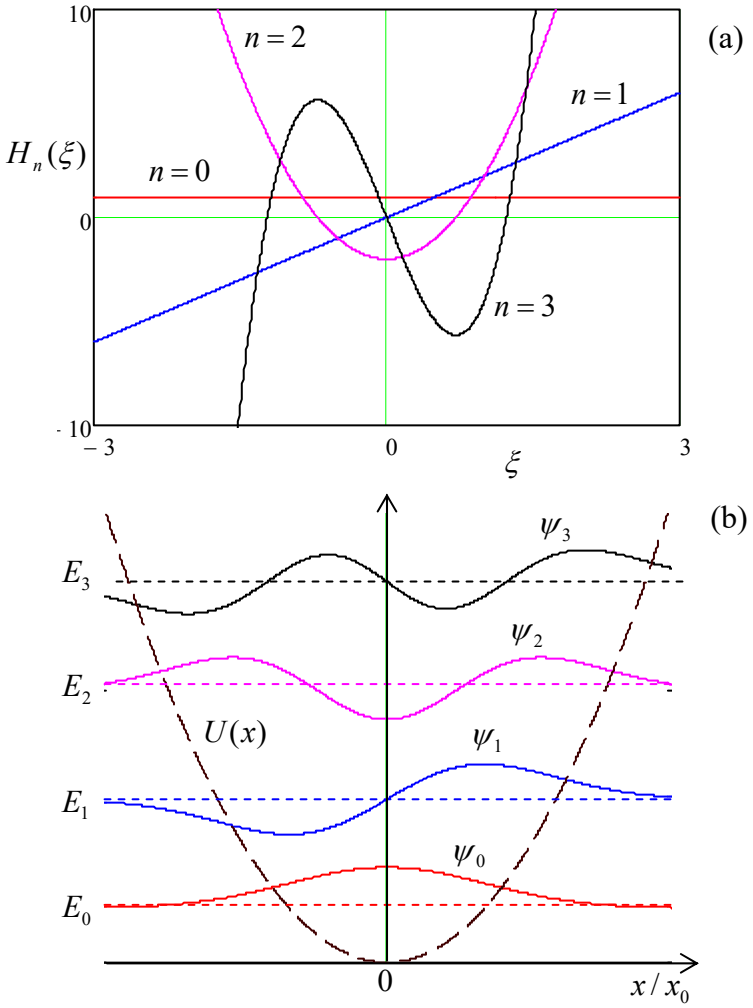


Fig. 2.35. (a) A few lowest Hermite polynomials and (b) the corresponding eigenenergies (horizontal dashed lines) and eigenfunctions (solid lines) of the harmonic oscillator. The dashed black curve shows the potential profile $U(x)$ drawn on the same scale as the energies E_n, so its crossings with the energy levels correspond to classical turning points.

The properties that are most important for applications are as follows:

(i) the function $H_n(\xi)$ has exactly n zeros (i.e. its plot crosses the ξ-axis exactly n times); as a result, the "parity" (odd-even) of these functions alternates with n, and

(ii) the polynomials are mutually orthonormal in the following sense:

$$\int_{-\infty}^{+\infty} H_n(\xi) H_{n'}(\xi) \exp\{-\xi^2\} d\xi = \pi^{1/2} 2^n n! \delta_{n,n'}. \quad (2.283)$$

Using the last property, we may readily calculate, from Eq. (278), the normalized eigenfunctions $\psi_n(x)$ of the harmonic oscillator – see Fig.35b:

$$\psi_n(x) = \frac{1}{\left(2^n n!\right)^{1/2} \pi^{1/4} x_0^{1/2}} \exp\left\{-\frac{x^2}{2x_0^2}\right\} H_n\left(\frac{x}{x_0}\right). \quad (2.284)$$

Harmonic oscillator: eigen-functions

At this point, it is instructive to compare these eigenfunctions with those of a 1D rectangular potential well, with its ultimately hard walls – see Fig. 1.8. Let us list their *common* features:

(i) The wavefunctions oscillate in the classically allowed regions with $E_n > U(x)$, while dropping exponentially beyond the boundaries of that region. (For the rectangular well with infinite walls, the latter regions are infinitesimally narrow.)

(ii) Each step up the energy level ladder increases the number of the oscillation half-waves (and hence the number of its zeros), by one.[97]

And here are the major features *specific* for a soft (e.g., quadratic-parabolic) confinement:

(i) The spatial spread of the wavefunction grows with n, following the gradual widening of the classically allowed region.

(ii) Correspondingly, E_n exhibits a slower growth than the $E_n \propto n^2$ law given by Eq. (1.85), because the gradual reduction of spatial confinement moderates the kinetic energy's growth.

Unfortunately, the "brute-force" approach to the harmonic oscillator problem, discussed above, is not too appealing. First, the proof of Eq. (281) is rather longish – so I do not have time/space for it. More importantly, it is hard to use Eq. (284) for the calculation of the expectation values of observables including the so-called *matrix elements* of the system – as we will see in Chapter 4, virtually the only numbers important for most applications. Finally, it is also almost evident that there has to be some straightforward math leading to any formula as simple as Eq. (262) for E_n. Indeed, there is a much more efficient, operator-based approach to this problem; it will be described in Sec. 5.4.

2.10. Exercise problems

2.1. As was mentioned in Sec. 2.1 of the lecture notes, Eq. (2.1) may be incorrect if the particle's potential energy depends on just one spatial coordinate: $U = U(x, t)$, and is much more reliable for particles strongly but uniformly confined in the transverse directions y, z. Explain why.

2.2. Prove that the final form of Eq. (2.23) of the lecture notes is correct even though x' has an (x-independent) imaginary part.

Hint: This is a good exercise in using the Cauchy theorem.[98]

[97] In mathematics, a slightly more general statement, valid for a broader class of ordinary linear differential equations, is frequently called the *Sturm oscillation theorem* and is a part of the *Sturm-Liouville theory* of such equations – see, e.g., Chapter 10 in the handbook by G. Arfken *et al*., cited in MA Sec. 16.

[98] See, e.g., MA Eq. (15.1).

2.3. The initial wave packet of a free 1D particle is described by Eq. (20): $\Psi(x,0) = \int a_k e^{ikx} dk$.

(i) Obtain a compact expression for the expectation value $\langle p \rangle$ of the particle's momentum at an arbitrary moment $t > 0$.

(ii) Calculate $\langle p \rangle$ for the case when the function $|a_k|^2$ is symmetric with respect to some value k_0.

2.4. Calculate the function a_k defined by Eq. (20), for the wave packet with a rectangular spatial envelope:

$$\Psi(x,0) = \begin{cases} C \exp\{ik_0 x\}, & \text{for } -a/2 \le x \le +a/2, \\ 0, & \text{otherwise.} \end{cases}$$

Analyze the result in the limit $k_0 a \to \infty$.

2.5. Prove Eq. (49) for the 1D propagator of a free quantum particle, by starting from Eq. (48).

2.6. Express the 1D propagator defined by Eq. (44) via the eigenfunctions and eigenenergies of a particle moving in an arbitrary stationary potential $U(x)$.

2.7. Calculate the change of a 1D particle's wavefunction, resulting from a short pulse of an external classical force that may be well approximated by a delta function: $F(t) = P\delta(t)$.

2.8. Calculate the transparency $\mathscr{T}$ of the rectangular potential barrier (68),

$$U(x) = \begin{cases} 0, & \text{for } x < -d/2, \\ U_0, & \text{for } -d/2 < x < +d/2, \\ 0, & \text{for } d/2 < x, \end{cases}$$

for a 1D particle of energy $E > U_0$. Analyze and interpret the result, taking into account that U_0 may be either positive or negative. (In the latter case, we are speaking about the particle's passage over a rectangular potential well of a finite depth $|U_0|$.)

2.9. Prove Eq. (117) for the case $\mathscr{T}_{\text{WKB}} << 1$, by using the connection formulas (105).

2.10. Spell out the stationary wavefunctions of a harmonic oscillator in the WKB approximation, and use them to calculate the expectation values $\langle x^2 \rangle$ and $\langle x^4 \rangle$ for an eigenstate number $n >> 1$.

2.11. Use the WKB approximation to express the expectation value of the kinetic energy of a 1D particle confined in a soft potential well, in its n^{th} stationary state, via the derivative dE_n/dn, for $n >> 1$.

2.12. Use the WKB approximation to calculate the transparency $\mathscr{T}$ of the following triangular potential barrier:

$$U(x) = \begin{cases} 0, & \text{for } x < 0, \\ U_0 - Fx, & \text{for } x > 0, \end{cases}$$

with F, $U_0 > 0$, as a function of the incident particle's energy E.

Hint: Be careful treating the sharp potential step at $x = 0$.

2.13. Prove that Eq. (2.67) of the lecture notes is valid even if the potential $U(x)$ changes, sufficiently slowly, on both sides of the potential step, provided that $U(x) < E$ everywhere.

2.14.* Prove that the symmetry of the 1D scattering matrix S describing an arbitrary time-independent scatterer allows its representation in the form (127).

2.15. Prove the universal relations between elements of the 1D transfer matrix T of a stationary (but otherwise arbitrary) scatterer, mentioned in Sec. 5.

2.16.* A k-narrow wave packet is incident on a finite-length 1D scatterer. Obtain a general expression for the time of its delay caused by the scatterer, and evaluate the time for the case of a very short but high potential barrier.

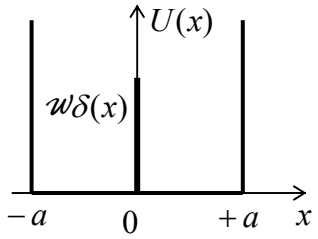
2.17. A 1D particle had been localized in a very narrow and deep potential well, with the "weight" $\int U(x)dx$ equal to $-\mathcal{W}$, where $\mathcal{W} > 0$. Then (say, at $t = 0$) the well's bottom is suddenly lifted up, so that the particle becomes completely free. Calculate the probability density to find the particle in a state with a certain wave number k at $t > 0$ and the total final energy of the system.

2.18. Calculate the lifetime of the metastable localized state of a 1D particle in the potential

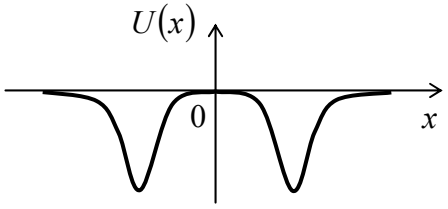
$$U(x) = -\mathcal{W}\delta(x) - Fx, \qquad \text{with } \mathcal{W} > 0,$$

in the WKB approximation. Formulate the condition of validity of the result.

2.19. Calculate the energy levels and the corresponding eigenfunctions of a 1D particle placed into a flat-bottom potential well of width $2a$, with infinitely high hard walls and a narrow potential barrier in the middle – see the figure on the right. Discuss the particle's dynamics in the limit when $\mathcal{W}$ is very large but still finite.



2.20.* Consider a symmetric system of two potential wells of the type shown in Fig. 21, but with $U(0) = U(\pm\infty) = 0$ – see the figure on the right. Derive a general expression for the well interaction force due to their sharing a quantum particle of mass m, and determine its sign for the cases when the particle is in:

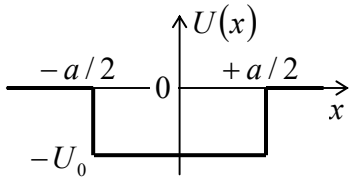


(i) a symmetric localized eigenstate: $\psi_S(-x) = \psi_S(x)$, and

(ii) an antisymmetric localized eigenstate: $\psi_A(-x) = -\psi_A(x)$.

Use a different approach to verify your conclusions for the particular case of delta-functional wells.

2.21. Derive and analyze the characteristic equation for localized eigenstates of a 1D particle in a rectangular potential well of a finite depth (see the figure on the right):



$$U(x) = \begin{cases} -U_0, & \text{for } |x| \le a/2, \\ 0, & \text{otherwise,} \end{cases} \qquad \text{with } U_0 > 0.$$

In particular, calculate the number of localized states as a function of the well's width a, and explore the limit $U_0 << \hbar^2/2ma^2$.

2.22. Calculate the energy of a 1D particle localized in a potential well of an arbitrary shape $U(x)$, provided that its width a is finite, and the average depth is very small:

$$\left|\overline{U}\right| << \frac{\hbar^2}{2ma^2}, \qquad \text{where } \overline{U} \equiv \frac{1}{a}\int_{\text{well}} U(x)dx .$$

2.23. A particle of mass m is moving in a field with the following potential:

$$U(x) = U_0(x) + \mathcal{W}\delta(x),$$

where $U_0(x)$ is a smooth symmetric function with $U_0(0) = 0$, growing monotonically at $x \to \pm\infty$. Use the WKB approximation to:

(i) derive the characteristic equation for the particle's energy spectrum, and

(ii) semi-quantitatively describe the spectrum's evolution at the increase of $|\mathcal{W}|$, for both signs of this parameter.

Spell out both results for the quadratic-parabolic potential (111): $U_0(x) = m\omega_0^2x^2/2$.

2.24. Prove Eq. (189).

2.25. For the problem discussed at the beginning of Sec. 7, i.e. the 1D particle's motion in an infinite Dirac comb potential (Fig. 24), write explicit expressions for the eigenfunctions at the very bottom and at the very top of the lowest energy band. Sketch both functions.

2.26. A 1D particle of mass m moves in an infinite periodic system of very narrow and deep potential wells that may be described by delta functions:

$$U(x) = \mathcal{W}\sum_{j=-\infty}^{+\infty}\delta(x - ja), \qquad \text{with } \mathcal{W} < 0 .$$

(i) Sketch the energy band structure of the system for very small and very large values of the potential well's "weight" $|\mathcal{W}|$, and

(ii) calculate explicitly the ground-state energy of the system in these two limits.

2.27. For the system discussed in the previous problem, write explicit expressions for the eigenfunctions of the system, corresponding to:

(i) the bottom of the lowest energy band,

(ii) the top of that band, and

(iii) the bottom of each higher energy band.

Sketch these functions.

2.28.* The 1D "crystal" analyzed in the last two problems, now extends only to $x > 0$, with a sharp step to a flat potential plateau at $x < 0$:

$$U(x) = \begin{cases} \mathcal{W}\sum_{j=1}^{+\infty} \delta(x - ja), & \text{with } \mathcal{W} < 0, \quad \text{for } x > 0, \\ U_0 > 0, & \text{for } x < 0. \end{cases}$$

Prove that the system can have a set of the so-called *Tamm states* localized near the "surface" $x = 0$, and calculate their energies in the limit when U_0 is very large but finite. (Quantify this condition.) [99]

2.29. Calculate the transfer matrix of the rectangular potential barrier specified by Eq. (68), for particle energies both below and above U_0.

2.30. Use the results of the previous problem to calculate the transfer matrix of one period of the periodic Kronig-Penney potential shown in Fig. 31b.

2.31. Using the results of the previous problem, derive the characteristic equations for a particle's motion in the periodic Kronig-Penney potential, for both $E < U_0$ and $E > U_0$. Try to bring the equations to a form similar to that obtained in Sec. 7 for the delta-functional barriers – see Eq. (198). Use the equations to formulate the conditions of applicability of the tight-binding and weak-potential approximations, in terms of the system's parameters and the particle's energy E.

2.32. For the Kronig-Penney potential, use the tight-binding approximation to calculate the widths of the allowed energy bands. Compare the results with those of the previous problem (in the corresponding limit).

2.33. For the same Kronig-Penney potential, use the weak-potential limit formulas to calculate the energy gap widths. Again, compare the results with those of Problem 31, in the corresponding limit.

2.34. 1D periodic chains of atoms may exhibit what is called the *Peierls instability*, leading to the *Peierls transition* to a phase in which atoms are slightly displaced, from the exact periodicity, by equal but sign-alternating shifts $\Delta x_j = (-1)^j \Delta x$, with $\Delta x << a$, where j is the atom's number in the chain, and a is its initial period. These displacements lead to an alternation of the coupling amplitudes δ_n (see Eq. (204)) between close values δ_n^+ and δ_n^-. Use the tight-binding approximation to calculate the resulting change of the n^{th} energy band, and discuss the result.

2.35.* Use Eqs. (1.73)-(1.74) to derive Eq. (252), and discuss the relation between these Bloch oscillations and the Josephson oscillations of frequency (1.75).

2.36. A 1D particle of mass m is placed into the following triangular potential well:

$$U(x) = \begin{cases} +\infty, & \text{for } x < 0, \\ Fx, & \text{for } x > 0, \end{cases} \quad \text{with } F > 0.$$

(i) Calculate its energy spectrum using the WKB approximation.

[99] In applications to electrons in solid-state crystals, the delta-functional potential wells model the attractive potentials of atomic nuclei, while U_0 represents the workfunction, i.e. the energy necessary for the extraction of an electron from the crystal to the free space – see, e.g., Sec. 1.1(ii), and also EM Sec. 2.6 and SM Sec. 6.3.

(ii) Estimate the ground state energy using the variational method, with two different trial functions.

(iii) Calculate the three lowest energy levels, and also the 10^{th} level, with an accuracy better than 0.1%, from the exact solution of the problem.

(iv) Compare and discuss the results.

Hint: The values of the first few zeros of the Airy function, necessary for Task (iii), may be found in many math handbooks, for example, in Table 9.9.1 of the open-access online version of the collection edited by Abramowitz and Stegun.[100]

2.37. Use the variational method to estimate the ground state energy E_g of a particle in the potential well

$$U(x) = -U_0 \exp\{-\alpha x^2\}, \qquad \text{with } \alpha > 0, \text{ and } U_0 > 0.$$

Spell out the results in the limits of small and large U_0, and give their interpretation.

2.38. For a 1D particle of mass m, in a potential well with the following profile,

$$U(x) = ax^{2s}, \quad \text{with } a > 0 \text{ and } s > 0,$$

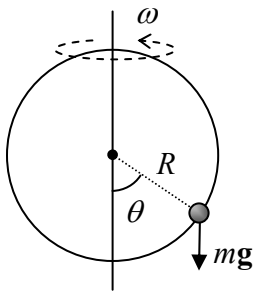
(i) calculate its energy spectrum using the WKB approximation, and

(ii) estimate the ground state energy using the variational method.

Compare the ground-state energy results.

2.39. Use the variational method to estimate the lowest excited energy level of a 1D harmonic oscillator.

2.40. Assuming the quantum effects to be small, calculate the lower part of the energy spectrum of the following system: a small bead of mass m, free to move without friction along a ring of radius R, which is rotated about its vertical diameter with a constant angular velocity ω – see the figure on the right. Formulate a quantitative condition of validity of your results.



Hint: This system was used as the "testbed problem" in the CM part of this series, and the reader is welcome to use any relations derived there.

2.41. A 1D harmonic oscillator with mass m and frequency ω_0 was in its ground state. At $t = 0$, an additional force F is suddenly exerted on it and then is kept constant. Calculate the probability of the oscillator staying in its ground state.

2.42. A 1D particle of mass m was placed into a quadratic potential well (111),

$$U(x) = \frac{m\omega_0^2 x^2}{2},$$

[100] See https://dlmf.nist.gov/9.9.

and allowed to relax into the ground state. At $t = 0$, the well is fast accelerated to move with velocity v, without changing its profile, so that at $t \geq 0$ the above formula for U is valid with the replacement $x \to x' \equiv x - vt$. Calculate the probability for the system to still be in the ground state at $t > 0$.

2.43. Initially, a 1D harmonic oscillator was in its ground state. At a certain moment of time, its spring constant κ is abruptly increased so that its frequency $\omega_0 = (\kappa/m)^{1/2}$ is increased by a factor of α, and then is kept constant at the new value. Calculate the probability that after the change, the oscillator is still in its ground state.

2.44. A 1D particle is in the following potential well:

$$U(x) = \begin{cases} +\infty, & \text{for } x < 0, \\ m\omega_0^2 x^2/2, & \text{for } x \geq 0. \end{cases}$$

(i) Find its eigenfunctions and eigenenergies.

(ii) The particle was let to relax into its ground state, and then the potential wall at $x < 0$ is rapidly removed so that the system is instantly turned into the usual harmonic oscillator (with the same m and ω_0). Find the probability for the particle to remain in the ground state.

2.45. Prove the following formula for the propagator of the 1D harmonic oscillator:

$$G(x,t;x_0,t_0) = \left\{\frac{m\omega_0}{2\pi i\hbar \sin[\omega_0(t-t_0)]}\right\}^{1/2} \exp\left\{\frac{im\omega_0}{2\hbar \sin[\omega_0(t-t_0)]}\left[\left(x^2 + x_0^2\right)\cos[\omega_0(t-t_0)] - 2xx_0\right]\right\}.$$

Discuss the relation between this formula and the propagator of a free 1D particle.

2.46. In the context of the Sturm oscillation theorem mentioned in Sec. 9, prove that the number of eigenfunction's zeros of a particle confined in an arbitrary but finite potential well always increases with the corresponding eigenenergy.

Hint: You may like to use the suitably modified Eq. (186).

2.47.* Use the WKB approximation to calculate the lifetime of the metastable ground state of a 1D particle of mass m in the "pocket" of the potential profile

$$U(x) = \frac{m\omega_0^2}{2}x^2 - \alpha x^3.$$

Contemplate the significance of this problem.

Chapter 3. Higher Dimensionality Effects

The descriptions of the basic quantum-mechanical effects, given in the previous chapter, may be extended to higher dimensions in an obvious way. This is why this chapter is focused on the phenomena (*such as the AB effect and the Landau levels*) *that cannot take place in one dimension due to topological reasons, and also on the key 3D problems* (*such as the Born approximation in the scattering theory, and the axially and spherically symmetric systems*) *that are important for numerous applications.*

3.1. Quantum interference and the AB effect

In the past two chapters, we have already discussed some effects of the de Broglie wave interference. For example, standing waves inside a potential well, or even on the top of a potential barrier, may be considered as a result of interference of incident and reflected waves. However, there are some remarkable new effects made possible by the spatial separation of such waves, and such separation requires a higher (either 2D or 3D) dimensionality. A good example of wave separation is provided by the Young-type experiment (Fig. 1) in which particles, emitted by the same source, are passed through two small holes (or narrow slits) in an otherwise opaque partition.

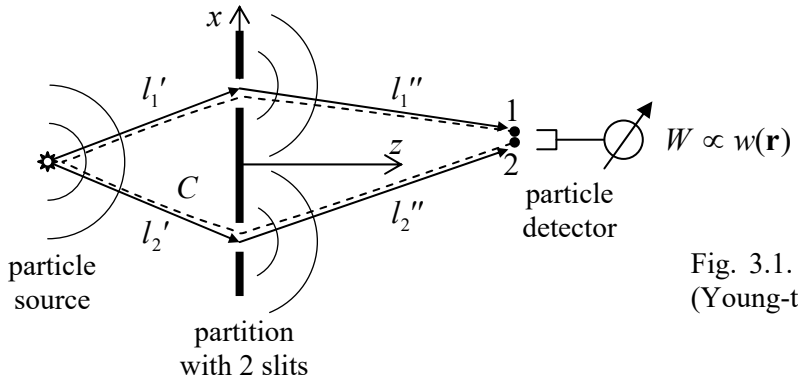


Fig. 3.1. The scheme of the "two-slit" (Young-type) interference experiment.

According to Eq. (1.22), if particle interactions are negligible (which is always true if the emission rate is sufficiently low), the average rate of particle counting by the detector is proportional to the probability density $w(\mathbf{r}, t) = \Psi(\mathbf{r}, t)\,\Psi^*(\mathbf{r}, t)$ to find a single particle at the detector's location $\mathbf{r}$, where $\Psi(\mathbf{r}, t)$ is the solution of the single-particle Schrödinger equation (1.25) for the system. Let us calculate this rate for the case when the incident particles may be represented by virtually monochromatic waves of energy E (e.g., very long wave packets), so their wavefunction may be taken in the form given by Eqs. (1.57) and (1.62): $\Psi(\mathbf{r}, t) = \psi(\mathbf{r})\exp\{-iEt/\hbar\}$. In this case, in the free-space parts of the system, where $U(\mathbf{r}) = 0$, $\psi(\mathbf{r})$ satisfies the stationary Schrödinger equation (1.78a):

$$-\frac{\hbar^2}{2m}\nabla^2\psi = E\psi\,. \tag{3.1a}$$

With the standard definition $k \equiv (2mE)^{1/2}/\hbar$, it may be rewritten as the 3D Helmholtz equation:

3D Helmholtz equation

$$\boxed{\nabla^2\psi + k^2\psi = 0\,.} \tag{3.1b}$$

The opaque parts of the partition may be well described as classically forbidden regions, so if their size scale a is much larger than the wavefunction penetration depth δ described by Eq. (2.59), we may use on their surface S the same boundary conditions as for the well's walls of infinite height:

$$\psi\big|_S = 0 . \tag{3.2}$$

Eqs. (1) and (2) describe the standard boundary problem of the theory of propagation of scalar waves of any nature. For an arbitrary geometry, this problem does not have a simple analytical solution. However, for a conceptual discussion of wave interference, we may use certain natural assumptions that will allow us to find its particular, approximate solution.

First, let us discuss the wave emission, into free space, by a small-size, isotropic source located at the origin of our reference frame. Naturally, the emitted wave should be spherically symmetric: $\psi(\mathbf{r}) = \psi(r)$. The well-known expression for the Laplace operator in spherical coordinates[1] shows that in this case, Eq. (1) is reduced to the following ordinary differential equation:

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d\psi}{dr}\right) + k^2\psi = 0 . \tag{3.3}$$

Let us introduce a new function, $f(r) \equiv r\psi(r)$. Plugging the reciprocal relation $\psi = f/r$ into Eq. (3), we see that for the function f, it gives the standard 1D wave equation:

$$\frac{d^2 f}{dr^2} + k^2 f = 0 . \tag{3.4}$$

As was discussed in Sec. 2.2, for a fixed k, the general solution of Eq. (4) may be represented in the form of two traveling waves:

$$f = f_+ e^{ikr} + f_- e^{-ikr} \tag{3.5}$$

so the full solution of Eq. (3) is

$$\psi(\mathbf{r}) = \frac{f_+}{r}e^{ikr} + \frac{f_-}{r}e^{-ikr}, \quad \text{i.e. } \Psi(\mathbf{r},t) = \frac{f_+}{r}e^{i(kr-\omega t)} + \frac{f_-}{r}e^{-i(kr+\omega t)}, \quad \text{with } \omega \equiv \frac{E}{\hbar} = \frac{\hbar k^2}{2m}. \tag{3.6}$$

If the source is located at point $\mathbf{r}' \neq 0$, the obvious generalization of Eq. (6) is

$$\Psi(\mathbf{r},t) = \frac{f_+}{R}e^{i(kR-\omega t)} + \frac{f_-}{R}e^{-i(kR+\omega t)}, \quad \text{with } R \equiv |\mathbf{R}|, \text{ where } \mathbf{R} \equiv \mathbf{r} - \mathbf{r}'. \tag{3.7}$$

The first term of this solution describes a spherically symmetric wave propagating from the source outward, while the second one represents a wave converging onto the source point $\mathbf{r}'$ from large distances. Though the latter solution is possible in some very special circumstances (say, when the outgoing wave is reflected back from a perfectly spherical shell), for our current problem, only the outgoing waves are relevant, so we may keep only the first term (proportional to f_+) in Eq. (7). Note that the factor R is the denominator (that was absent in the 1D geometry) has a simple physical sense: it provides the independence of the full probability current $I = 4\pi R^2 j(R)$, with $j(R) \propto k\Psi\Psi^* \propto 1/R^2$, of the distance R between the observation point and the source.

[1] See, e.g., MA Eq. (10.9) with $\partial/\partial\theta = \partial/\partial\varphi = 0$.

Now let us assume that the partition's geometry is not too complicated – for example, it is either planar as shown in Fig. 1, or nearly-planar, and consider the region of the particle detector location far behind the partition (at $z >> 1/k$), and at a relatively small angle to the normal: $|x| << z$. Then it should be physically clear that the spherical waves (7) emitted by each point inside the slit cannot be perturbed too much by the opaque parts of the partition, and their only role is the restriction of the set of such emitting points to the area of the slits. Hence, an approximate solution of the boundary problem is given by the following *Huygens principle*: the wave behind the partition looks as if it was the sum of the contributions (7) from two point sources located in the slits, with each source's strength f_+ proportional to the amplitude of the wave arriving at this pseudo-source from the real source – see Fig. 1. This principle finds its confirmation in the strict wave theory, which shows that with our assumptions, the solution of the boundary problem (1)-(2) may be represented as the following *Kirchhoff integral*:[2]

Kirchhoff integral

$$\psi(\mathbf{r}) = c\int_{\text{slits}} \frac{\psi(\mathbf{r}')}{R} e^{ikR} d^2r', \qquad \text{with } c = \frac{k}{2\pi i}. \tag{3.8}$$

If the source is also far from the partition, its wave's front is almost parallel to the slit plane, and if the slits are not too broad, we can take $\psi(\mathbf{r}')$ constant ($\psi_{1,2}$) at each slit, so Eq. (8) is reduced to

$$\psi(\mathbf{r}) = a''_1 \exp\{ikl''_1\} + a''_2 \exp\{ikl''_2\}, \qquad \text{with } a''_{1,2} = \frac{cA_{1,2}}{l''_{1,2}}\psi_{1,2}, \tag{3.9}$$

where $A_{1,2}$ are the slit areas, and $l''_{1,2}$ are the distances from the slits to the detector. The wavefunctions on the slits may be calculated approximately[3] by applying the same Eq. (7) to the region *before* the slits: $\psi_{1,2} \approx (f_+/l'_{1,2})\exp\{ikl'_{1,2}\}$, where $l'_{1,2}$ are the distances from the source to the slits – see Fig. 1. As a result, Eq. (9) may be rewritten as

Wave-function superposition

$$\psi(\mathbf{r}) = a_1 \exp\{ikl_1\} + a_2 \exp\{ikl_2\}, \qquad \text{with } l_{1,2} \equiv l'_{1,2} + l''_{1,2}; \quad a_{1,2} \equiv \frac{c f_+ A_{1,2}}{l'_{1,2} l''_{1,2}}. \tag{3.10}$$

(As Fig. 1 shows, each of $l_{1,2}$ is the full length of the classical path of the particle from the source, through the corresponding slit, and further to the observation point $\mathbf{r}$.)

According to Eq. (10), the resulting rate of particle counting at point $\mathbf{r}$ is proportional to

Quantum interference

$$w(\mathbf{r}) = \psi(\mathbf{r})\psi^*(\mathbf{r}) = |a_1|^2 + |a_2|^2 + 2|a_1 a_2|\cos\varphi_{12}, \tag{3.11}$$

where

$$\varphi_{12} \equiv k(l_2 - l_1) \tag{3.12}$$

is the difference between the total wave phase accumulations along each of the two alternative paths. The last expression may be evidently generalized as

[2] For the proof and a detailed discussion of Eq. (8), see, e.g., EM Sec. 8.5.

[3] A possible (and reasonable) concern about the application of Eq. (7) to the field in the slits is that it ignores the effect of opaque parts of the partition. However, as we know from Chapter 2, the main role of the classically forbidden region is reflecting the incident wave toward its source (i.e. to the left in Fig. 1). As a result, the contribution of this reflection to the field inside the slits is insignificant if $A_{1,2} >> \lambda^2$, and even in the opposite case provides just some rescaling of the amplitudes $a_{1,2}$, which is not important for our conceptual discussion.

$$\varphi_{12} = \oint_C \mathbf{k} \cdot d\mathbf{r}, \tag{3.13}$$

Quantum interference: phase difference

with integration along the virtually closed contour C (see the dashed line in Fig. 1), i.e. from point 1, in the positive (i.e. counterclockwise) direction all the way to point 2. (From our discussion of the 1D WKB approximation in Sec. 2.4, we may expect such generalization to be valid even if k changes, sufficiently slowly, along the paths.)

Our result (11)-(12) shows that the particle counting rate oscillates as a function of the difference $(l_2 - l_1)$, which in turn changes with the detector's position, giving the famous interference pattern, with its amplitude proportional to the product $|a_1 a_2|$, and hence vanishing if any of the slits is closed. For the wave theory, this is a well-known result,[4] but for particle physics, it was (and still is :-) rather shocking. Indeed, our analysis is valid for a very low particle emission rate, so there is no other way to interpret the pattern other than resulting from a particle's interference *with itself* – or rather the interference of its de Broglie waves passing through each of two slits.[5] As was already noted in Sec. 1.1(v), nowadays such interference is reliably observed not only for electrons but also for much heavier particles: atoms and molecules including very complex organic ones.

Let us now discuss a very interesting effect of magnetic field on quantum interference. To simplify our discussion, let us consider a slightly different version of the two-slit experiment, in which each of the two alternative paths is constricted to a narrow channel using lateral confinement – see Fig. 2. (In this arrangement, moving the particle detector without changing the channels' geometry and hence local values of k may be more problematic experimentally, so let us think about its position $\mathbf{r}$ as fixed.) In this case, because of the effect of the walls providing the path confinement, we cannot use Eqs. (10) for the amplitudes $a_{1,2}$. However, from the discussions in Sec. 1.6 and Sec. 2.2, it should be clear that the first of the expressions (10) remains valid, though maybe with a value of k specific for each channel.

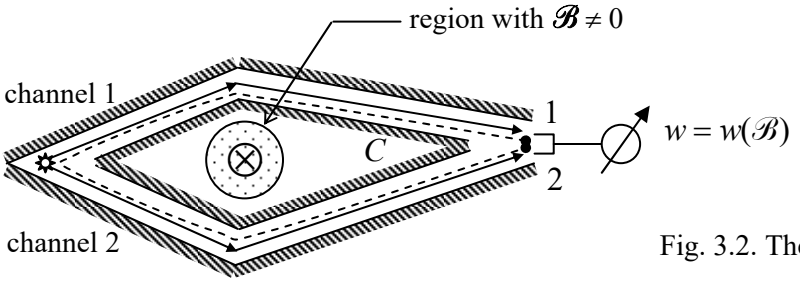


Fig. 3.2. The AB effect.

In this geometry, we can apply some local magnetic field $\mathscr{B}$, say normal to the plane of particle motion, whose lines would pierce but not touch the contour C drawn along the particle propagation channels – see the dashed line in Fig. 2. In classical electrodynamics,[6] the external magnetic field's effect on a particle with electric charge q is described by the *Lorentz force*

[4] See, e.g., a detailed discussion in EM Sec. 8.4.

[5] Here I have to mention the fascinating experiments (first performed in 1987 by C. Hong *et al.* with photons, and recently, in 2015, by R. Lopes *et al*., with non-relativistic particles – helium atoms) on the interference of de Broglie waves of *independent* but *identical* particles, in the same internal quantum state and virtually the same values of E and k. These experiments raise the important issue of particle *indistinguishability*, which will be discussed in Sec. 8.1.

[6] See, e.g., EM Sec. 5.1. Note that Eq. (14), as well as all other formulas of this course, are in the SI units.

$$\mathbf{F}_{\mathscr{B}} = q\mathbf{v}\times\mathscr{B}\,, \tag{3.14}$$

where $\mathscr{B}$ is the field value at the point of its particle's location, so for the experiment shown in Fig. 2, $\mathbf{F}_{\mathscr{B}} = 0$, and the field would not affect the particle motion at all. In quantum mechanics, this is not so, and the field does affect the probability density w, even if $\mathscr{B} = 0$ at all points where the wavefunction $\psi(\mathbf{r})$ is not equal to zero.

In order to describe this surprising effect, let us first develop a general framework for an account of electromagnetic field effects on a charged quantum particle, which will also give us some by-product results important for forthcoming discussions. To do that, we need to calculate the Hamiltonian of such a particle in electric and magnetic fields. For an electrostatic field, this is easy. Indeed, from classical electrodynamics[7] we know that this field may be represented as a gradient of its electrostatic potential ϕ,

$$\mathscr{E} = -\nabla\phi(\mathbf{r}), \tag{3.15}$$

so the force exerted by the field on a particle with electric charge q,

$$\mathbf{F}_{\mathscr{E}} = q\mathscr{E}, \tag{3.16}$$

may be described by adding the field-induced potential energy,

$$U(\mathbf{r}) = q\phi(\mathbf{r}), \tag{3.17}$$

to other possible components of the full potential energy. As was already discussed in Sec. 1.4, such potential energy may be included in the particle's Hamiltonian operator just by adding it to the kinetic energy operator – see Eq. (1.41).

However, the magnetic field's effect is peculiar: since its Lorentz force (14) is perpendicular to the classical particle's velocity, it cannot do any work on it:

$$d\mathscr{W}_{\mathscr{B}} \equiv \mathbf{F}_{\mathscr{B}}\cdot d\mathbf{r} = \mathbf{F}_{\mathscr{B}}\cdot\mathbf{v}dt = q(\mathbf{v}\times\mathscr{B})\cdot\mathbf{v}\,dt = 0, \tag{3.18}$$

and hence the field cannot be represented by any potential energy, so it may not be immediately clear how to account for it in the Hamiltonian. The crucial help comes from the analytical-mechanics approach to classical electrodynamics:[8] in the non-relativistic limit, the Hamiltonian function of a particle in an electromagnetic field looks like that in the electric field only:

$$H = \frac{mv^2}{2} + U = \frac{p^2}{2m} + q\phi\,; \tag{3.19}$$

however, the momentum $\mathbf{p} \equiv m\mathbf{v}$ that participates in this expression is now the difference

$$\mathbf{p} = \mathbf{P} - q\mathbf{A}\,. \tag{3.20}$$

Here $\mathbf{A}$ is the vector potential of the field, defined by the well-known relations:[9]

$$\mathscr{E} = -\nabla\phi - \frac{\partial\mathbf{A}}{\partial t}, \qquad \mathscr{B} = \nabla\times\mathbf{A}\,, \tag{3.21}$$

[7] Note that here (until Chapter 9) we are describing the *particle* quantum-mechanically but the *fields*, classically.
[8] See, e.g., EM Sec. 9.7, in particular, Eq. (9.196).
[9] See, e.g., EM Sec. 6.1, in particular Eqs. (6.7).

while **P** is the *canonical momentum*, whose Cartesian components may be calculated (in classics) from the Lagrangian function L by using the standard formula of analytical mechanics,

$$P_j \equiv \frac{\partial L}{\partial v_j}. \tag{3.22}$$

To emphasize the difference between the two momenta, $\mathbf{p} = m\mathbf{v}$ is frequently called the *kinematic momentum* (or "m**v**-momentum"). The distinction between **p** and $\mathbf{P} = \mathbf{p} + q\mathbf{A}$ becomes more clear if we notice that the vector potential is not *gauge-invariant*: according to the second of Eqs. (21), at the so-called *gauge transformation*

$$\mathbf{A} \to \mathbf{A} + \nabla\chi, \tag{3.23}$$

with an arbitrary single-valued scalar *gauge function* $\chi = \chi(\mathbf{r}, t)$, the magnetic field does not change. Moreover, according to the first of Eqs. (21), if we make the simultaneous replacement

$$\phi \to \phi - \frac{\partial \chi}{\partial t}, \tag{3.24}$$

the gauge transformation does not affect the electric field either. With that, the gauge function's choice does not affect the classical particle's equation of motion, and hence the velocity **v** and momentum **p**. Hence, the kinematic momentum is gauge-invariant, while **P** is not, because according to Eqs. (20) and (23), the introduction of χ changes it by $q\nabla\chi$.

Now the standard way of transfer to the wave mechanics is to use Eq. (1.26) for the operator of the canonical rather than the kinematic momentum:[10]

$$\boxed{\hat{\mathbf{P}} = -i\hbar\nabla}. \tag{3.25}$$

Canonical momentum: operator

Hence the Hamiltonian operator corresponding to the classical Hamiltonian function (19) is

$$\boxed{\hat{H} = \frac{1}{2m}\left(-i\hbar\nabla - q\mathbf{A}\right)^2 + q\phi \equiv -\frac{\hbar^2}{2m}\left(\nabla - i\frac{q}{\hbar}\mathbf{A}\right)^2 + q\phi}, \tag{3.26}$$

so the stationary Schrödinger equation (1.60) of a particle moving in an electromagnetic field (but otherwise free) is

Charged particle in EM field

$$\boxed{-\frac{\hbar^2}{2m}\left(\nabla - i\frac{q}{\hbar}\mathbf{A}\right)^2\psi + q\phi\psi = E\psi}, \tag{3.27}$$

We may now repeat all the calculations of Sec. 1.4 for the case $\mathbf{A} \neq 0$, and get the following generalized expression for the probability current density:

$$\mathbf{j} = \frac{\hbar}{2im}\left[\psi^*\left(\nabla - i\frac{q}{\hbar}\mathbf{A}\right)\psi - \text{c.c}\right] \equiv \frac{1}{2m}\left(\psi^*\hat{\mathbf{p}}\,\psi - \text{c.c.}\right) \equiv \frac{\hbar}{m}|\psi|^2\left(\nabla\varphi - \frac{q}{\hbar}\mathbf{A}\right). \tag{3.28}$$

We see that the current density is gauge-invariant (as required for any observable) only if at the transformation (23), the wavefunction's phase φ changes as

[10] The validity of this choice is clear from the fact that if the *kinetic* momentum was described by this differential operator, the Hamiltonian operator corresponding to the classical Hamiltonian function (19), and the corresponding Schrödinger equation would not describe the magnetic field effects at all.

$$\varphi \to \varphi + \frac{q}{\hbar}\chi \,. \tag{3.29}$$

This may be a point of conceptual concern: since quantum interference is described by the spatial dependence of the phase φ, can the observed interference pattern depend on the gauge function's choice? (That would not make any sense, because we may change the gauge in our mind.) Fortunately, this is not true, because according to Eq. (29), the spatial phase *difference* between two interfering paths, participating in Eq. (12), is gauge-transformed as

$$\varphi_{12} \to \varphi_{12} + \frac{q}{\hbar}(\chi_2 - \chi_1). \tag{3.30}$$

But χ has to be a single-valued function of coordinates; hence in the limit when the points 1 and 2 coincide, $\chi_1 = \chi_2$, so φ_{12} is gauge-invariant, and so is the interference pattern.

However, the difference φ_{12} *may be* affected by the magnetic field, even if it is localized outside the channels in which the particle propagates. Indeed, in this case, the field cannot affect the particle distribution across the channels, so, for an externally-fixed total flow of particles in each of them, we can take

$$\mathbf{j}(\mathbf{r})\Big|_{\mathscr{B}\neq 0} = \mathbf{j}(\mathbf{r})\Big|_{\mathscr{B}=0}\,, \tag{3.31}$$

and the last form of Eq. (28) yields

$$\nabla\varphi(\mathbf{r})\Big|_{\mathscr{B}\neq 0} = \nabla\varphi(\mathbf{r})\Big|_{\mathscr{B}=0} + \frac{q}{\hbar}\mathbf{A}\,. \tag{3.32}$$

Integrating this equation along the contour C (Fig. 2), for the phase difference between points 1 and 2 we get

$$\varphi_{12}\Big|_{\mathscr{B}\neq 0} = \varphi_{12}\Big|_{\mathscr{B}=0} + \frac{q}{\hbar}\oint_C \mathbf{A}\cdot d\mathbf{r}\,, \tag{3.33}$$

where the integral should be taken along the same contour C as before (in Fig. 2, from point 1, counterclockwise along the dashed line to point 2). But from classical electrodynamics, we know[11] that as points 1 and 2 tend to each other, i.e. the contour C becomes closed, the last integral is just the magnetic flux $\Phi \equiv \int \mathscr{B}_n d^2r$ through any smooth surface limited by this contour, so Eq. (33) may be rewritten as

AB effect

$$\boxed{\varphi_{12}\Big|_{\mathscr{B}\neq 0} = \varphi_{12}\Big|_{\mathscr{B}=0} + \frac{q}{\hbar}\Phi\,.} \tag{3.34a}$$

In terms of the interference pattern, this means a shift of interference fringes, proportional to the magnetic flux.

This phenomenon is usually called the "Aharonov-Bohm" (or just the *AB*) *effect*.[12] For particles with a single elementary charge, $q = \pm e$, this result is frequently represented as

[11] See, e.g., EM Sec. 5.3.

[12] I prefer the latter, less personable name because the effect had been actually predicted by Werner Ehrenberg and Raymond Siday in 1949, i.e. well before it was rediscovered (also theoretically) by Y Aharonov and D. Bohm in 1959. To be fair to Aharonov and Bohm, it was their work that triggered a wave of interest in the phenomenon, leading to its first experimental observation by Robert G. Chambers in 1960 and several other groups soon after that. Later, the experiments were improved using ferromagnetic cores and/or superconducting shielding to provide a better separation between the electrons and the applied field – as in the work whose result is shown in Fig. 3.

$$\varphi_{12}\Big|_{\mathcal{B}\neq 0} = \varphi_{12}\Big|_{\mathcal{B}=0} \pm 2\pi\frac{\Phi}{\Phi_0'}, \tag{3.34b}$$

where the fundamental constant $\Phi_0' \equiv 2\pi\hbar/e \approx 4.14\times10^{-15}$ Wb has the meaning of the magnetic flux necessary to change φ_{12} by 2π, i.e. to shift the interference pattern (11) by one period, and is called the *normal magnetic flux quantum* – "normal" because of the reasons we will soon discuss.

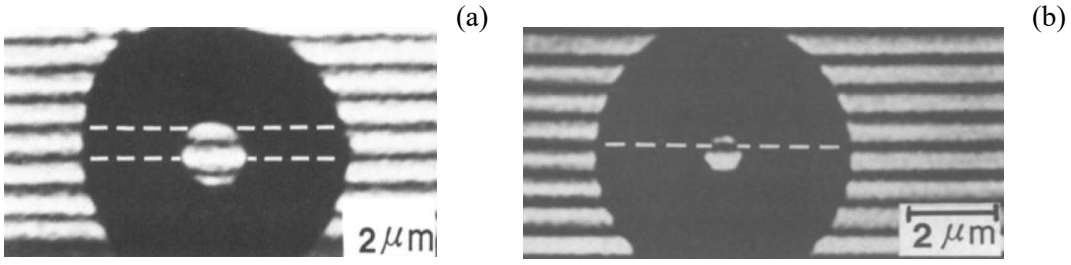


Fig. 3.3. Typical results of a two-paths interference experiment by A. Tonomura *et al*., *Phys. Rev. Lett*. **56**, 792 (1986), showing the AB effect for electrons shielded from the applied magnetic field. In this particular experimental geometry, the AB effect produces a relative shift of the interference patterns inside and outside the dark ring. (a) $\Phi = \Phi_0'/2$, (b) $\Phi = \Phi_0'$. © 1986 APS.

The AB effect may be "almost explained" classically, in terms of Faraday's electromagnetic induction. Indeed, a change $\Delta\Phi$ of magnetic flux in time induces a vortex-like electric field $\Delta\mathscr{E}$ around it. That field is *not* restricted to the magnetic field's location, i.e. may reach the particle's trajectories. The field's magnitude (or rather of its integral along the contour C) may be readily calculated by integration of the first of Eqs. (21):

$$\Delta V \equiv \oint_C \Delta\mathscr{E}\cdot d\mathbf{r} = -\frac{d\Phi}{dt}. \tag{3.35}$$

I hope that in this expression the reader readily recognizes the integral ("undergraduate") form of Faraday's induction law.[13] To calculate the effect of this electric field on the particles, let us assume that the variable separation described by Eq. (1.57) may be applied to the endpoints 1 and 2 of the particle's alternative trajectories as two independent systems,[14] and that the magnetic flux' change by a certain amount $\Delta\Phi$ does not change the spatial factors $\psi_{1,2}$, provided that the phases $\varphi_{1,2}$ are included into the time-dependent factors $a_{1,2}$. Then we may repeat the arguments that were used in Sec. 1.6 at the discussion of the Josephson effect, and since the change (35) leads to the change of the potential energy difference $\Delta U = q\Delta V$ between the two points, we may rewrite Eq. (1.72) as

$$\frac{d\varphi_{12}}{dt} = -\frac{\Delta U}{\hbar} = -\frac{q}{\hbar}\Delta V = \frac{q}{\hbar}\frac{d\Phi}{dt}. \tag{3.36}$$

Integrating this relation over the time of the magnetic field's change, we get

[13] See, e.g., EM Sec. 6.1.

[14] This assumption may seem a little bit of a stretch, but the resulting relation (37) may be indeed proven for a rather realistic model, though that would take more time/space than I can afford.

$$\Delta\varphi_{12} = \frac{q}{\hbar}\Delta\Phi\,, \tag{3.37}$$

– superficially, the same result as given by Eq. (34).

However, this interpretation of the AB effect is limited. Indeed, it requires the particle to be in the system (on the way from the source to the detector) during the flux change, i.e. when the induced electric field $\mathscr{E}$ may affect its dynamics. On the contrary, Eq. (34) predicts that the interference pattern would shift even if the field change has been made when there was no particle in the system, and hence the field $\mathscr{E}$ could not be felt by it. Experiment confirms the latter conclusion. Hence, there is *something* in the space where a particle propagates (i.e., outside of the magnetic field region), that transfers the information about even the static magnetic field to the particle. The standard interpretation of this surprising fact is as follows: the vector potential **A** is not just a convenient mathematical tool, but a physical reality (just as its scalar counterpart ϕ is), despite the large freedom of choice we have in prescribing specific spatial and temporal dependences of these potentials without affecting any observable – see Eqs. (23)-(24).

To conclude this section, let me briefly discuss the very interesting form taken by the AB effect in superconductivity. To be applied to this case, our results require two changes. The first one is simple: since superconductivity may be interpreted as a result of the Bose-Einstein condensate of Cooper pairs with electric charge $q = -2e$, Φ_0' has to be replaced by the so-called *superconducting flux quantum*[15]

Super-conducting flux quantum

$$\boxed{\Phi_0 \equiv \frac{\pi\hbar}{e} \approx 2.07\times10^{-15}\ \text{Wb} = 2.07\times10^{-7}\ \text{Gs}\cdot\text{cm}^2.} \tag{3.38}$$

Second, since the pairs are Bose particles and are all condensed in the same (ground) quantum state described by the same wavefunction, the total electric current density, proportional to the probability current density j, may be extremely large – in practical superconducting materials, up to ~10^{12} A/m^2. In these conditions, one cannot neglect the contribution of that current into the magnetic field and hence into its flux Φ, which (according to the Lenz rule of the Faraday induction law) tries to compensate for changes in external flux. To see possible results of this contribution, let us consider a closed superconducting loop (Fig. 4). Due to the Meissner effect (which is just another version of the flux self-compensation), the current and magnetic field penetrate into a superconductor by only a small distance (called the *London penetration depth*) $\delta_L \sim 10^{-7}$ m.[16] If the loop is made of a superconducting "wire" that is considerably thicker than δ_L, we may draw a contour deep inside the wire, at which the current density is negligible. According to the last form of Eq. (28), everywhere at the contour,

$$\nabla\varphi - \frac{q}{\hbar}\mathbf{A} = 0\,. \tag{3.39}$$

Integrating this equation along the contour as before (in Fig. 4, from some point 1, all the way around the ring to the virtually coinciding point 2), we need to have the phase difference φ_{12} equal to $2\pi n$, because the wavefunction $\psi \propto \exp\{i\varphi\}$ at the initial and final points 1 and 2 should be "essentially" the same, i.e. produce the same observables. As a result, we get

[15] One more bad, though common term: a wire may (super)conduct, but a quantum hardly can!
[16] For more detail, see EM Sec. 6.4.

$$\Phi \equiv \oint_C \mathbf{A}\cdot d\mathbf{r} = \frac{\hbar}{q} 2\pi n \equiv n\Phi_0 . \tag{3.40}$$

Flux quantization

This is the famous *flux quantization effect*,[17] which justifies the term "magnetic flux quantum" for the constant Φ_0 given by Eq. (38).

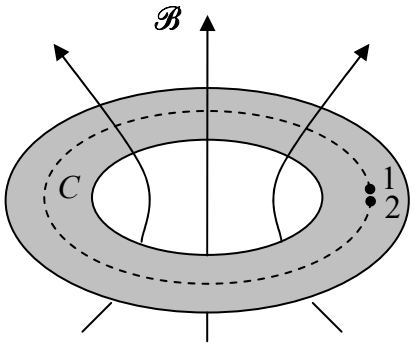


Fig. 3.4. The magnetic flux quantization in a superconducting loop (schematically).

Here I have to mention the very interesting effects of "flux semi-quantization" that arise when superconductor loops are closed with Josephson junctions, forming the so-called *Superconducting QUantum Interference Devices* ("SQUIDs"). Such devices may be used, in particular, for supersensitive magnetometry and ultrafast low-power computing,[18] and are currently explored as a possible basis for quantum computation and cryptography – see Sec. 8.5 below.

3.2. Landau levels and quantum Hall effect

In the last section, we have used the Schrödinger equation (27) for an analysis of static magnetic field effects in "almost-1D", circular geometries shown in Figs. 1, 2, and 4. However, this equation describes very interesting effects in fully higher dimensions as well, especially in the 2D case. Let us consider a quantum particle free to move within the [*x, y*] plane only (say, due to its strong confinement in the perpendicular direction *z* – see the discussion at the beginning of Sec. 2.1). In this case, Eq. (27) reduces to a similar equation but with the Laplace operator acting only in the directions *x* and *y*:

$$-\frac{\hbar^2}{2m}\left(\mathbf{n}_x\frac{\partial}{\partial x}+\mathbf{n}_y\frac{\partial}{\partial y}-i\frac{q}{\hbar}\mathbf{A}\right)^2\psi = E\psi . \tag{3.41}$$

Let us find its solutions for the simplest case when the applied static magnetic field is uniform and normal to the motion plane:

$$\mathscr{B} = \mathscr{B}\mathbf{n}_z . \tag{3.42}$$

According to the second of Eqs. (21), this relation imposes the following restriction on the choice of the vector potential:

$$\frac{\partial A_y}{\partial x}-\frac{\partial A_x}{\partial y} = \mathscr{B} , \tag{3.43}$$

[17] It was predicted in 1949 by Fritz London and experimentally discovered (independently and virtually simultaneously) in 1961 by two experimental groups: B. Deaver and W. Fairbank, and R. Doll and M. Näbauer.
[18] A brief review of these applications, and recommendations for further reading may be found in EM Sec. 6.5.

but the gauge transformations still give us a lot of freedom in its choice. The natural axially symmetric form, $\mathbf{A} = \mathbf{n}_\varphi \rho \mathscr{B}/2$, where $\rho = (x^2 + y^2)^{1/2}$ is the distance from some z-axis, leads to cumbersome math. In 1930, L. Landau realized that the energy spectrum of Eq. (41) may be obtained by making a much simpler, though very counter-intuitive gauge choice:

$$A_x = 0, \quad A_y = \mathscr{B}(x - x_0), \tag{3.44}$$

(with arbitrary x_0), which evidently satisfies Eq. (43), though ignores the physical symmetry of the x and y directions for the field (42).

Now, expanding the eigenfunction into the Fourier integral in the y-direction:

$$\psi(x,y) = \int X_k(x) \exp\{ik(y - y_0)\} dk, \tag{3.45}$$

we see that for each component of this expansion, Eq. (41) yields a specific equation

$$-\frac{\hbar^2}{2m}\left\{\mathbf{n}_x \frac{d}{dx} + i\mathbf{n}_y\left[k - \frac{q}{\hbar}\mathscr{B}(x - x_0)\right]\right\}^2 X_k = EX_k. \tag{3.46}$$

Since the two vectors inside the curly brackets are mutually perpendicular, its square has no cross-terms, so Eq. (46) reduces to

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}X_k + \frac{q^2}{2m}\mathscr{B}^2(x - x_0')^2 X_k = EX_k, \qquad \text{where } x_0' \equiv x_0 + \frac{\hbar k}{q\mathscr{B}}. \tag{3.47}$$

But this 1D Schrödinger equation is identical to Eq. (2.261) for a 1D harmonic oscillator,[19] with the center at point x_0', and the frequency ω_0 equal to

$$\omega_c = \frac{|q\mathscr{B}|}{m}. \tag{3.48}$$

In the last expression, it is easy to recognize the *cyclotron frequency* of the classical particle's rotation in the magnetic field. (It may be readily obtained using the 2nd Newton law for a circular orbit of radius r,

$$m\frac{v^2}{r} = F_{\mathscr{B}} \equiv qv\mathscr{B}, \tag{3.49}$$

and noting that the resulting ratio $v/r = |q\mathscr{B}|/m$ is just the radius-independent angular velocity ω_c of the particle's rotation.) Hence, the energy spectrum for each Fourier component of the expansion (45) is the same:

Landau levels

$$E_n = \hbar\omega_c\left(n + \frac{1}{2}\right), \tag{3.50}$$

independent of x_0, y_0, and k.

[19] This result may become a bit less puzzling if we recall that at the classical circular cyclotron motion of a particle, each of its Cartesian coordinates, including x, performs sinusoidal oscillations with frequency (48), just as a 1D harmonic oscillator with this frequency.

Hence, this is a good example of a highly *degenerate* system: for each eigenvalue E_n, there are many similarly-structured eigenfunctions that differ by the positions $\{x_0, y_0\}$ of their centers and the rate k of their phase change along the y-axis. They may be used to assemble a large variety of linear combinations, including 2D wave packets whose centers move along classical circular orbits. Note, however, that the radius of such rotation cannot be smaller than the so-called *Landau radius*,

$$r_L \equiv \left(\frac{\hbar}{m\omega_c}\right)^{1/2} \equiv \left(\frac{\hbar}{|q\mathscr{B}|}\right)^{1/2}, \quad (3.51)$$

Landau radius

which characterizes the minimum size of the wave packet, and follows from Eq. (2.276) after the replacement $\omega_0 \to \omega_c$. This radius is remarkably independent of the particle mass, and may be interpreted in the following way: the scale $\mathscr{B}A_{min}$ of the applied magnetic field's flux through the effective area $A_{min} = 2\pi r_L^2$ of the smallest wave packet is just one normal flux quantum $\Phi_0' \equiv 2\pi\hbar/|q|$.

A detailed analysis of such wave packets (for which we would not have time in this course) proves, in particular, the virtually evident fact: the applied magnetic field does not change the *average* density dN_2/dE of different 2D states on the energy scale, given by Eq. (1.92), but just "assembles" the states them on the Landau levels (see Fig. 5a), so the number of different orbital states on each Landau level (per unit area) is

$$n_L \equiv \frac{N_2}{A} = \frac{1}{A}\frac{dN_2}{dE}\Big|_{\mathscr{B}=0}\Delta E \equiv \frac{1}{A}\frac{dN_2}{d^2k}\Big|_{\mathscr{B}=0}\frac{d^2k}{dk}\frac{1}{dE/dk}\Delta E = \frac{1}{A}\frac{A}{(2\pi)^2}2\pi k\frac{1}{\hbar^2 k/m}\hbar\omega_c \equiv \frac{|q\mathscr{B}|}{2\pi\hbar}. \quad (3.52)$$

This expression may again be interpreted in terms of magnetic flux quanta: $n_L\Phi_0' = \mathscr{B}$, i.e. there is one particular state on each Landau level per each normal flux quantum.

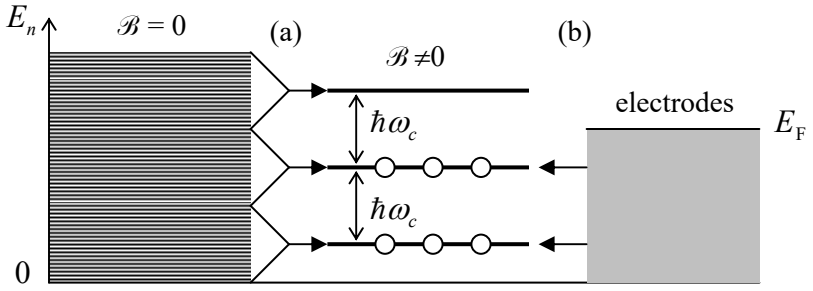


Fig. 3.5. (a) The "assembly" of 2D states on Landau levels, and (b) filling the levels with electrons at the quantum Hall effect.

The most famous application of the Landau levels picture is the explanation of the *quantum Hall effect*[20]. It is usually observed in the "Hall bar" geometry sketched in Fig. 6, where electric current I is passed through a rectangular conducting sample placed into magnetic field $\mathscr{B}$ perpendicular to the sample's plane. The classical analysis of the effect may be based on the notion of the Lorentz force (14). As the magnetic field is turned on, this force starts to deviate the effective charge carriers (electrons or holes) from their straight motion between the electrodes, bending them toward the insulated sides of the

[20] It was first observed in 1980 by a group led by Klaus von Klitzing, while the classical version (54) of the effect was first observed by Edwin Hall a century earlier – in 1879.

bar (in Fig. 6, parallel to the x-axis). Here the carriers accumulate, generating a gradually increasing electric field $\mathscr{E}$ until its force (16) exactly balances the Lorentz force (14):

$$q\mathscr{E}_y = qv_x\mathscr{B}, \tag{3.53}$$

where v_x is the drift velocity of the carriers along the bar (Fig. 6), providing the sustained balance condition $\mathscr{E}_y/v_x = \mathscr{B}$ at each point of the sample.

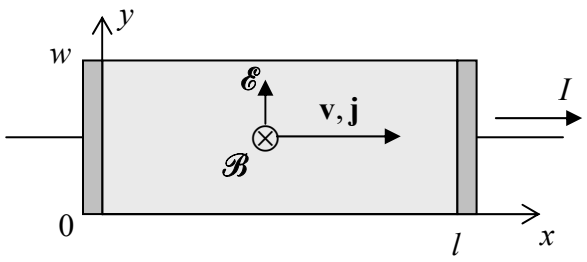


Fig. 3.6. The Hall bar geometry. Darker rectangles show external (3D) electrodes.

With n_2 carriers per unit area, in a sample of width w, this balance condition yields the following classical expression for the so-called *Hall resistance* R_H, remarkably independent of w and l:

Classical Hall effect

$$R_H \equiv \frac{V_y}{I_x} = \frac{\mathscr{E}_y w}{qn_2 v_x w} = \frac{\mathscr{B}}{qn_2}. \tag{3.54}$$

This formula is broadly used in practice for the measurement of the 2D density n_2 of the charge carriers and of the carrier type: electrons with $q = -e < 0$, or holes with the effective charge $q = +e > 0$.

However, in experiments with high-quality (low-defect) 2D samples, at sub-kelvin temperatures[21] and high magnetic fields, the linear growth of R_H with $\mathscr{B}$, described by Eq. (54), is interrupted by virtually horizontal plateaus (Fig. 7).

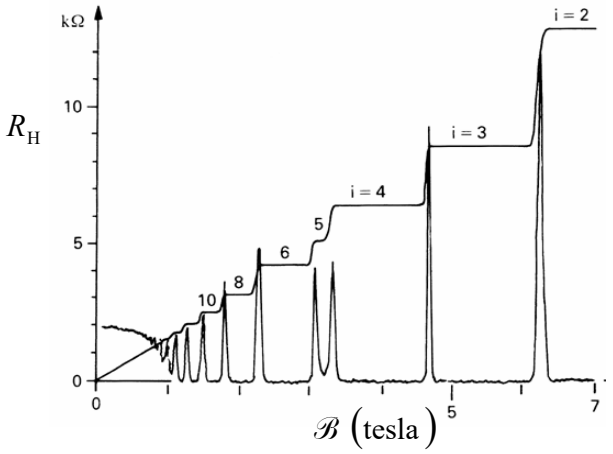


Fig. 3.7. A typical record of the integer quantum Hall effect. The lower trace (with sharp peaks) shows the diagonal element, V_x/I_x, of the resistance tensor. (Adapted from https://www.nobelprize.org/nobel_prizes/physics/laureates/1998/press.html).

[21] In some systems, such as the *graphene* (virtually perfect 2D sheets of carbon atoms – see Sec. 4 below), the effect may be more stable to thermal fluctuations, due to their topological properties, so it may be observed even at room temperature – see, e.g., K. Novoselov *et al., Science* **315,** 1379 (2007). Note also that in some spontaneously-magnetized ferromagnetic layers, the quantum Hall effect may be observed in the absence of an *external* magnetic field – see, e.g., M. Götz *et al.*, *Appl. Phys. Lett.* **112**, 072102 (2018) and references therein.

Most remarkably, the experimental values of R_H on these plateaus are reproduced with extremely high accuracy (up to ~10^{-9}) from sample to sample.[22] They are described by the following formula:

$$R_H = \frac{1}{i} R_K, \qquad \text{where } R_K \equiv \frac{2\pi\hbar}{e^2}, \tag{3.55}$$

Quantum Hall effect

with the following value:

$$R_K \approx 25.812\,807\,459\,304\,\text{k}\Omega, \tag{3.56}$$

and i is (only until the end of this section, following tradition!) the plateau number, i.e. a real integer.

This effect may be explained using the Landau-level picture. The 2D sample is typically in weak contact with 3D electrodes whose conductivity electrons, at low temperatures, fill all states with energies below a certain Fermi energy E_F – see Fig. 5b. According to Eqs. (48) and (50), as $\mathscr{B}$ is increased, the spacing $\hbar\omega_c$ between the Landau levels increases proportionately, so fewer and fewer of these levels are below E_F (and hence in equilibrium, all their states are filled), and within certain ranges of field variations, the number i of the filled levels is constant. (In the schematic Fig. 5b, $i = 2$.) So, plugging $n_2 = in_L$ and $q = -e$ into Eq. (54), and using Eq. (52) for n_L, we get

$$R_H = \frac{1}{i}\frac{\mathscr{B}}{qn_L} = \frac{1}{i}\frac{2\pi\hbar}{e^2}, \tag{3.57}$$

i.e. exactly the experimental result (55).

This admittedly oversimplified explanation of the quantum Hall effect does not take into account at least two important factors:

(i) the nonuniformity of the background potential $U(x, y)$ in realistic Hall bar samples, and the role of the quasi-1D *edge channels* this nonuniformity produces;[23] and

(ii) the Coulomb interaction of the electrons, in high-quality samples leading to the formation of R_H plateaus with not only integer but also fractional values of i (1/3, 2/5, 3/7, etc.).[24]

Unfortunately, a thorough discussion of these very interesting features is well beyond the framework of this course.[25,26]

---

[22] Due to this high accuracy (which is a rare exception in solid-state physics!), the *von Klitzing constant* R_K was used in metrology for the "legal" ohm's definition. Since 2018, the values of $\hbar$ and e, and hence of R_K, are considered exactly known and fixed – see *Appendix UCA*: *Selected Units and Constants.*

[23] Such quasi-1D regions, with the width of the order of r_L, form along the lines where the Landau levels cross the Fermi surface and are actually responsible for all the electron transfer at the quantum Hall effect (giving the pioneering example of what is nowadays called the *topological insulators*). The particle motion along these channels is effectively one-dimensional; because of this, it is unaffected by modest unintentional nonuniformities of the potential $U(x, y)$. This fact is responsible for the extraordinary accuracy of Eq. (55).

[24] This *fractional quantum Hall effect* was discovered in 1982 by D. Tsui, H. Stormer, and A. Gossard. In contrast, the effect described by Eq. (55) with integer i (Fig. 7) is now called the *integer quantum Hall effect*.

[25] For a comprehensive discussion of these effects, I can recommend either the monograph by D. Yoshioka, *The Quantum Hall Effect*, Springer, 1998, or the review by D. Yennie, *Rev. Mod. Phys.* **59,** 781 (1987). (See also the later publications cited above.)

[26] Note also that the quantum Hall effect is sometimes discussed in terms of the so-called *Berry phase*, one of the *geometric phases* – the notion apparently pioneered by S. Pancharatnam in 1956. However, in the "usual" quantum Hall effect the Berry phase equals zero, and I believe that this concept should be saved for the discussion of more topologically involved systems. Unfortunately, I will have no time/space for a discussion of such systems

3.3. Scattering and diffraction

The second class of quantum effects that become richer in multi-dimensional spaces, is typically referred to as either *diffraction* or *scattering* – depending on the context. In classical physics, these two terms are used to describe very different effects. The term "diffraction" is used for the interference of the *waves* re-emitted by elementary components of extended objects, under the effect of a single incident *wave*.[27] On the other hand, the term "scattering" is used in classical mechanics to describe the result of the interaction of a beam of *particles*[28] incident upon an object called the *scatterer* – see Fig. 8.

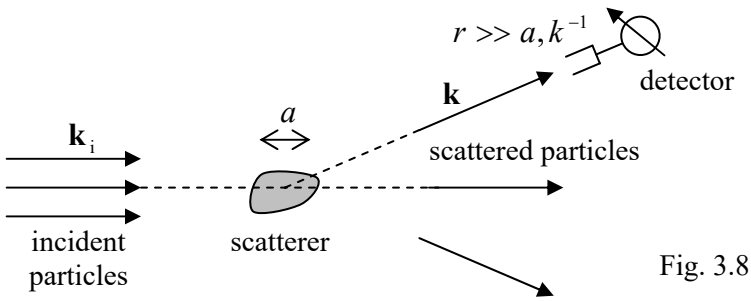


Fig. 3.8. Scattering (schematically).

Most commonly, the detector of the scattered particles is located at a large distance $r >> a$ from the scatterer. In this case, the main observable independent of r is the *flux* (the number per unit time) of particles scattered in a certain direction, i.e. their flux per unit solid angle Ω. Since it is proportional to the incident flux of particles per unit area, the efficiency of scattering in a particular direction may be characterized by the ratio of these two fluxes. This ratio is called the *differential cross-section* of the scatterer:

Differential cross-section

$$\frac{d\sigma}{d\Omega} \equiv \frac{\text{flux of scatterd particles per unit solid angle}}{\text{flux of incident particles per unit area}}. \tag{3.58}$$

Such terminology and notation stem from the fact that the integral of $d\sigma/d\Omega$ over all scattering angles,

Total cross-section

$$\sigma \equiv \oint \frac{d\sigma}{d\Omega} d\Omega = \frac{\text{total flux of scattered particles}}{\text{incident flux per per unit area}}, \tag{3.59}$$

evidently having the dimensionality of area, has a simple interpretation as the *total cross-section* of scattering. For the simplest case when a solid object scatters all classical particles hitting its surface but does not affect the particles flying by it, σ is just the geometric area of the scatterer, as observed from the direction of the incident particles. In classical mechanics, we first calculate the particle's scattering angle as a function of its *impact parameter b* and then average the result over all values of b, considered random. [29]

---

in this course, and have to refer the interested reader to special literature – see, e.g., either the key original papers collected by A. Shapere and F. Wilczek, *Geometric Phases in Physics*, *World Scientific*, 1992, or the monograph by A. Bohm *et al*., *The Geometric Phase in Quantum Systems*, Springer, 2003.

[27] The notion of interference is very close to diffraction, but the former term is typically reserved for the wave re-emission by just a few components, such as two slits in the Young experiment – see Figs. 1 and 2. A detailed discussion of diffraction and interference of electromagnetic waves may be found in EM Secs. 8.3-8.8.

[28] In the classical wave theory, the term "scattering" is typically reserved for wave interaction with disordered sets of small objects – see, e.g., EM Sec. 8.3.

[29] See, e.g., CM Sec. 3.5.

In quantum mechanics, due to the particle/wave duality, a relatively broad, parallel beam of incident particles of the same energy E may be fairly represented with a plane de Broglie wave (1.88):

$$\psi_{\rm i} = \left|\psi_{\rm i}\right| \exp\{i\mathbf{k}_{\rm i}\cdot\mathbf{r}\}, \tag{3.60}$$

with the free-space wave number $k_{\rm i} = k = (2mE)^{1/2}/\hbar$. As a result, the particle scattering becomes a synonym of the de Broglie wave diffraction, and (somewhat counter-intuitively) the description of the effect becomes simpler, excluding the notion of the impact parameter. Indeed, the wave (60) corresponds to a constant probability current density (1.49):

$$\mathbf{j}_{\rm i} = \left|\psi_{\rm i}\right|^2 \frac{\hbar}{m}\mathbf{k}_{\rm i}, \tag{3.61}$$

which is exactly the flux of incident particles per unit area that is used in the denominator of Eq. (58), while the numerator of that fraction may be simply expressed via the probability current density $\mathbf{j}_{\rm s}$ of the scattered de Broglie waves:

$$\frac{d\sigma}{d\Omega} = \frac{j_{\rm s} r^2}{j_{\rm i}}, \qquad \text{at } r >> a. \tag{3.62}$$

Hence the task of finding $d\sigma/d\Omega$ is reduced to the calculation of $j_{\rm s}$ at sufficiently large distances r from the scatterer. For the *elastic* scattering (when the energy E of the scattered particles is the same as that of the incident particles) this may be done by solving the stationary Schrödinger equation (1.65). Let us rewrite it in the form

$$\left(E - \hat{H}_0\right)\psi = U(\mathbf{r})\psi, \qquad \text{with } \hat{H}_0 \equiv -\frac{\hbar^2}{2m}\nabla^2, \text{ and } E = \frac{\hbar^2 k^2}{2m}, \tag{3.63}$$

where the potential energy $U(\mathbf{r})$ describes the scatterer's effect. Looking for the solution of Eq. (62) in the natural form

$$\psi = \psi_{\rm i} + \psi_{\rm s}, \tag{3.64}$$

where $\psi_{\rm i}$ is the incident wave (60) and $\psi_{\rm s}$ has the sense of the scattered wave, and taking into account that the former wave satisfies the free-space Schrödinger equation

$$\hat{H}_0\psi_{\rm i} = E\psi_{\rm i}, \tag{3.65}$$

we may reduce Eq. (63) to either of the following equivalent forms:

$$\left(E - \hat{H}_0\right)\psi_{\rm s} = U(\mathbf{r})(\psi_i + \psi_{\rm s}), \qquad \left(\nabla^2 + k^2\right)\psi_{\rm s} = \frac{2m}{\hbar^2}U(\mathbf{r})\psi. \tag{3.66}$$

For applications, an integral version of this equation is frequently more convenient. To derive it, we may look at the second of Eqs. (66) as a linear inhomogeneous differential equation for the function $\psi_{\rm s}$, thinking of its right-hand side as a known "source". The solution of such an equation obeys the linear superposition principle, i.e. we may represent it as the sum of the waves outcoming from all elementary volumes d^3r' of the scatterer. Mathematically, this sum may be expressed as either

$$\psi_{\rm s}(\mathbf{r}) = \frac{2m}{\hbar^2}\int U(\mathbf{r}')\psi(\mathbf{r}')G(\mathbf{r},\mathbf{r}')d^3r', \tag{3.67a}$$

or, equivalently, as[30]

$$\psi(\mathbf{r}) = \psi_i(\mathbf{r}) + \frac{2m}{\hbar^2}\int U(\mathbf{r}')\psi(\mathbf{r}')G(\mathbf{r},\mathbf{r}')d^3r', \tag{3.67b}$$

where $G(\mathbf{r}, \mathbf{r}')$ is the *spatial Green's function*, defined as such an elementary, spherically symmetric response of the 3D Helmholtz equation to a point source, i.e. the outward-propagating solution of the following equation[31]

$$\left(\nabla^2 + k^2\right)G = \delta(\mathbf{r} - \mathbf{r}'). \tag{3.68}$$

But we already know such a solution of this equation – see Eq. (7) and its discussion:

$$G(\mathbf{r},\mathbf{r}') = \frac{f_+}{R}e^{ikR}, \qquad \text{where } \mathbf{R} \equiv \mathbf{r} - \mathbf{r}', \tag{3.69}$$

so we need just to calculate the coefficient f_+ for Eq. (68). This can be done in several ways, for example by noticing that at $R << k^{-1}$, the second term on the left-hand side of Eq. (68) is negligible, so it is reduced to the well-known Poisson equation with a delta-functional right-hand side, which describes, for example, the electrostatic potential induced by a point electric charge. Either recalling the Coulomb law or applying the Gauss theorem,[32] we readily get the asymptote

$$G \to -\frac{1}{4\pi R}, \qquad \text{at } kR << 1, \tag{3.70}$$

which is compatible with Eq. (69) only if $f_+ = -1/4\pi$, i.e. if

$$G(\mathbf{r},\mathbf{r}') = -\frac{1}{4\pi R}e^{ikR}. \tag{3.71}$$

Plugging this result into Eq. (67a), we get the following formal solution of Eq. (66):

Scattering problem: formal solution

$$\boxed{\psi_s(\mathbf{r}) = -\frac{m}{2\pi\hbar^2}\int U(\mathbf{r}')\frac{\psi(\mathbf{r}')}{R}e^{ikR}d^3r'.} \tag{3.72}$$

Note that if the function $U(\mathbf{r})$ is smooth, the singularity in the denominator is integrable (i.e. not dangerous); indeed, the contribution of a sphere with some radius $\mathcal{R} \to 0$, with the center at point $\mathbf{r}'$, into this integral scales as

$$\int_{R<\mathcal{R}} \frac{d^3R}{R} \equiv 4\pi\int_0^{\mathcal{R}} \frac{R^2dR}{R} \equiv 4\pi\int_0^{\mathcal{R}} RdR = 2\pi\mathcal{R}^2 \to 0. \tag{3.73}$$

So far, our result (72) is exact, but its apparent simplicity is deceiving because the wavefunction ψ on its right-hand side generally includes not only the incident wave ψ_I but also the scattered wave ψ_s

[30] This formula is sometimes called the *Lipmann-Schwinger equation,* though more frequently this term is reserved for either its operator form or the resulting equation for the spatial Fourier components of ψ and ψ_i.

[31] Please notice both the similarity and difference between this Green's function and the propagator discussed in Sec. 2.1. In both cases, we use the linear superposition principle to solve wave equations, but while Eq. (67) gives the solution of the *inhomogeneous* equation (66), Eq. (2.44) does that for a *homogeneous* Schrödinger equation. In the latter case, the elementary wave sources are the elementary parts of the initial wavefunction, rather than of the equation's right-hand side as in our current problem.

[32] See, e.g., EM Sec. 1.2.

– see Eq. (64). The most straightforward, and most common simplification of this problem, called the *Born approximation*,[33] is possible if the scattering potential $U(\mathbf{r})$ is in some sense small. (We will derive the quantitative condition of this smallness in a minute.) Since at $U(\mathbf{r}) = 0$, the scattering wave ψ_s has to disappear, at small but non-zero $U(\mathbf{r})$, $|\psi_s|$ has to be much smaller than $|\psi_i|$. In this case, on the right-hand side of Eq. (73), we may ignore ψ_s in comparison with ψ_i, getting

$$\psi_s(\mathbf{r}) = -\frac{m}{2\pi\hbar^2}|\psi_i|\int U(\mathbf{r}')\frac{\exp\{i\mathbf{k}_i\cdot\mathbf{r}'\}}{R}e^{ikR}d^3r'. \qquad (3.74)$$

Born approximation

Actually, Eq. (74) gives us even *more* than we wanted: it evaluates the scattered wave at *any* point, including those within of the scattering object, while to spell out Eq. (62), we only need to find the wave *far* from the scatterer, at $r \to \infty$. However, before going to that limit, we can use Eq. (74) to find a quantitative criterion of the Born approximation's validity. For that, let us estimate the magnitude of the right-hand side of this equation for a scatterer of a linear size $\sim a$, and the potential magnitude's scale U_0. The results are different in the following two limits:

(i) If $ka << 1$, then inside the scatterer (i.e., at distances $\Delta r' \sim a$), both $\exp\{i\mathbf{k}\cdot\mathbf{r}'\}$ and the second exponent under the integral in Eq. (74) change little, and a crude but fair estimate of the solution's magnitude is

$$|\psi_s| \sim \frac{m}{2\pi\hbar^2}|\psi_i|U_0 a^2. \qquad (3.75)$$

(ii) In the opposite limit $ka >> 1$, the function under the integral is nearly periodic in one of the spatial directions (that of the incident wave's propagation), so the net integral accumulates only on distances of the order of the de Broglie wavelength, $\sim k^{-1}$, and the integral is correspondingly smaller:

$$|\psi_s| \sim \frac{m}{2\pi\hbar^2}|\psi_i|U_0\frac{a^2}{ka}. \qquad (3.76)$$

These relations allow us to spell out the Born approximation's condition, $|\psi_s| << |\psi_I|$, as

$$U_0 << \frac{\hbar^2}{ma^2}\max[ka, 1]. \qquad (3.77)$$

In the fraction on the right-hand side of this relation, we may readily recognize the scale of the kinetic (quantum-confinement) energy E_a of the particle inside a potential well of a size of the order of a, so the Born approximation is valid essentially if the potential energy of particle's interaction with the scatterer is smaller than E_a. Note, however, that the estimate (76) is not valid in some special situations when the effects of scattering *accumulate* in some direction. This is frequently the case for small angles θ of scattering by extended objects, when $ka >> 1$, but $ka\theta \lesssim 1$.

[33] Named after M. Born, who was the first to apply this approximation in quantum mechanics. However, the basic idea of this approach had been developed much earlier (in 1881) by Lord Rayleigh in the context of electromagnetic wave scattering – see, e.g., EM Sec. 8.3. Note also that the contents of that section repeat some aspects of our current discussion – perhaps regrettably but unavoidably so, because the Born approximation is a centerpiece of the theory of scattering/diffraction for both the electromagnetic waves and the de Broglie waves. Hence I felt I had to cover it in this course for the benefit of the readers who skipped the EM part of my series.

Now let us proceed to large distances $r >> r' \sim a$, and simplify Eq. (74) using an approximation similar to the dipole expansions in electrodynamics.[34] Namely, in the denominator's R, we may ignore r' in comparison with the much larger r, but the exponents require more care, because even if $r' \sim a << r$, the product $kr' \sim ka$ may still be of the order of 1. As Fig. 9a shows, in the first approximation in r', we may take

$$R \equiv |\mathbf{r} - \mathbf{r}'| \approx r - \mathbf{n}_r \cdot \mathbf{r}', \tag{3.78}$$

and since the directions of the vectors $\mathbf{k}$ and $\mathbf{r}$ coincide, i.e. $\mathbf{k} = k\mathbf{n}_r$, we get

$$kR \approx kr - \mathbf{k} \cdot \mathbf{r}', \quad \text{i.e. } e^{ikR} \approx e^{ikr} e^{-i\mathbf{k}\cdot\mathbf{r}'}. \tag{3.79}$$

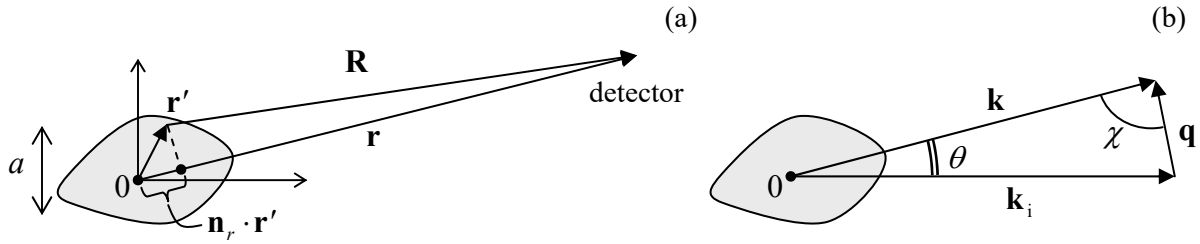


Fig. 3.9. (a) The long-range expansion of R, and (b) the definitions of $\mathbf{q}$, χ, and θ.

With this replacement, Eq. (74) yields

$$\psi_s(\mathbf{r}) = -\frac{m}{2\pi\hbar^2}\frac{|\psi_i|}{r} e^{ikr} \int U(\mathbf{r}') \exp\{-i(\mathbf{k} - \mathbf{k}_i)\cdot\mathbf{r}'\} d^3r'. \tag{3.80}$$

This relation is a particular case of a more general formula[35]

Scattering function: definition

$$\psi_s = |\psi_i| \frac{f(\mathbf{k}, \mathbf{k}_i)}{r} e^{ikr}, \tag{3.81}$$

where $f(\mathbf{k}, \mathbf{k}_i)$ is called the *scattering function*.[36] The physical sense of this function becomes clear from the calculation of the corresponding probability current density $\mathbf{j}_s$. For that, generally, we need to use Eq. (1.47) with the gradient operator having all spherical-coordinate components.[37] However, at $kr >> 1$, the main contribution to $\nabla\psi_s$, proportional to $k >> 1/r$, is provided by differentiating the factor e^{ikr}, which changes in the common direction of vectors $\mathbf{r}$ and $\mathbf{k}$, so

$$\nabla\psi_s \approx \mathbf{n}_r \frac{\partial}{\partial r}\psi_s \approx \mathbf{k}\psi_s, \qquad \text{at } kr >> 1, \tag{3.82}$$

and Eq. (1.47) yields

[34] See, e.g., EM Sec. 8.2.

[35] It is easy to prove that this form is an asymptotic form of *any* solution ψ_s of the scattering problem (even that beyond the Born approximation) at sufficiently large distances $r >> a, k^{-1}$.

[36] Note that the function f has the dimension of length, and does not account for the incident wave. This is why sometimes a dimensionless function, $S = 1 + 2ikf$, is used instead. This function S is called the *scattering matrix*, because it may be considered a natural generalization of the 1D matrix S defined by Eq. (2.124), to higher dimensionality.

[37] See, e.g., MA Eq. (10.8).

$$\mathbf{j}_s(\theta) \approx \frac{\hbar}{m}|\psi_i|^2 \frac{|f(\mathbf{k},\mathbf{k}_i)|^2}{r^2}\mathbf{k}\,. \tag{3.83}$$

Plugging this expression and also Eq. (61) into Eq. (62), for the differential cross-section we get simply

$$\frac{d\sigma}{d\Omega} = |f(\mathbf{k},\mathbf{k}_i)|^2, \tag{3.84}$$

while the total cross-section is

$$\sigma = \oint |f(\mathbf{k},\mathbf{k}_i)|^2 d\Omega, \tag{3.85}$$

so the scattering function $f(\mathbf{k}, \mathbf{k}_i)$ gives us everything we need – and in fact more because the function also contains information about the phase of the scattered wave.

Comparing Eqs. (80) and (81), we see that in the Born approximation, the scattering function is given by the so-called *Born integral*

$$f(\mathbf{k},\mathbf{k}_i) = -\frac{m}{2\pi\hbar^2}\int U(\mathbf{r}) e^{-i\mathbf{q}\cdot\mathbf{r}} d^3r, \tag{3.86}$$

Born integral

where, just for the notation simplicity, $\mathbf{r}'$ was replaced with $\mathbf{r}$, and $\mathbf{q}$ is the following *scattering vector*:

$$\mathbf{q} \equiv \mathbf{k} - \mathbf{k}_i, \tag{3.87}$$

with the length $q = 2k\sin(\theta/2)$, where θ is the *scattering angle* between the vectors $\mathbf{k}$ and $\mathbf{k}_i$ – see Fig. 9b. For the differential cross-section, Eqs. (84) and (86) yield[38]

$$\frac{d\sigma}{d\Omega} = \left(\frac{m}{2\pi\hbar^2}\right)^2 \left|\int U(\mathbf{r}) e^{-i\mathbf{q}\cdot\mathbf{r}} d^3r\right|^2. \tag{3.88}$$

$d\sigma/d\Omega$: Born approximation

This is the basic result of this section; it may be further simplified for spherically symmetric scatterers, with

$$U(\mathbf{r}) = U(r). \tag{3.89}$$

In this case, it is convenient to represent the exponent in the Born integral as $\exp\{-iqr\cos\chi\}$, where χ is the angle between the vectors $\mathbf{k}$ (i.e. the direction $\mathbf{n}_r$ toward the detector) and $\mathbf{q}$ (rather than the incident wave vector $\mathbf{k}_i$!) – see Fig. 9b. Now, for a fixed $\mathbf{q}$, we can take this vector's direction for the polar axis of a spherical coordinate system, and reduce Eq. (86) to a 1D integral:

$$\begin{aligned} f(\mathbf{k},\mathbf{k}_i) &= -\frac{m}{2\pi\hbar^2}\int_0^\infty r^2 dr U(r) \int_0^{2\pi} d\varphi \int_0^{\pi} \sin\chi d\chi \exp\{-iqr'\cos\chi\} \\ &\equiv -\frac{m}{2\pi\hbar^2}\int_0^\infty r^2 dr U(r)\; 2\pi \frac{2\sin qr}{qr} \equiv -\frac{2m}{\hbar^2 q}\int_0^\infty U(r)\sin(qr)\, r dr. \end{aligned} \tag{3.90}$$

[38] Note that according to Eq. (88), in the Born approximation, the scattering intensity does not depend on the sign of the potential U, and also that scattering in a certain direction is completely determined by a specific Fourier component of the function $U(\mathbf{r})$, namely by its harmonic with the wave vector equal to the scattering vector $\mathbf{q}$.

As a simple example, let us use the Born approximation to analyze scattering on the following spherically symmetric potential:

$$U(\mathbf{r}) = U_0 \exp\left\{-\frac{r^2}{2a^2}\right\}. \tag{3.91}$$

In this particular case, it is better to avoid the temptation to exploit the spherical symmetry by using Eq. (90), and instead, use the general Eq. (88) because it may be represented as a product of three similar Cartesian factors:

$$f(\mathbf{k},\mathbf{k}_i) = -\frac{mU_0}{2\pi\hbar^2} I_x I_y I_z, \qquad \text{with } I_x \equiv \int_{-\infty}^{+\infty} \exp\left\{-\left(\frac{x^2}{2a^2} + iq_x x\right)\right\} dx, \tag{3.92}$$

and similar integrals for I_y and I_z. From Chapter 2, we already know that the Gaussian integrals like I_x may be readily worked out by complementing the exponent to the full square, in our current case giving

$$\begin{aligned} I_x &= (2\pi)^{1/2} a \exp\left\{-\frac{q_x^2 a^2}{2}\right\}, \text{ etc.; so:} \\ \frac{d\sigma}{d\Omega} &= \left(\frac{mU_0}{2\pi\hbar^2} I_x I_y I_z\right)^2 = 2\pi a^2 \left(\frac{mU_0 a^2}{\hbar^2}\right)^2 e^{-q^2 a^2}. \end{aligned} \tag{3.93}$$

Now, the total cross-section σ is an integral of $d\sigma/d\Omega$ over all directions of the vector $\mathbf{k}$. Since in our case the scattering intensity does not depend on the azimuthal angle φ, the only nontrivial integration is over the scattering angle θ (see Fig. 9b again):

$$\begin{aligned} \sigma &\equiv \oint \frac{d\sigma}{d\Omega} d\Omega = 2\pi \int_0^{\pi} \frac{d\sigma}{d\Omega} \sin\theta d\theta = 4\pi^2 a^2 \left(\frac{mU_0 a^2}{\hbar^2}\right)^2 \int_0^{\pi} \exp\left\{-\left(2k\sin\frac{\theta}{2}\right)^2 a^2\right\} \sin\theta d\theta \\ &\equiv 4\pi^2 a^2 \left(\frac{mU_0 a^2}{\hbar^2}\right)^2 \int_{\theta=0}^{\theta=\pi} \exp\left\{-2k^2 a^2 (1-\cos\theta)\right\} d(1-\cos\theta) = \frac{2\pi^2}{k^2}\left(\frac{mU_0 a^2}{\hbar^2}\right)^2 \left(1 - e^{-4k^2 a^2}\right). \end{aligned} \tag{3.94}$$

Let us analyze these results. In the low-energy limit, $ka << 1$ (and hence $qa << 1$ for any scattering angle), the scattered wave is virtually isotropic: $d\sigma/d\Omega \approx$ const – a very typical feature of a scalar-wave scattering[39] by small objects, in any approximation. Note that according to Eq. (77), the Born expression for σ, following from Eq. (94) in this limit,

$$\sigma = 8\pi^2 a^2 \left(\frac{mU_0 a^2}{\hbar^2}\right)^2, \tag{3.95}$$

is only valid if σ is much smaller than the scale a^2 of the physical cross-section of the scatterer. In the opposite, high-energy limit $ka >> 1$, the scattering is dominated by small angles $\theta \approx q/k \sim 1/ka \sim \lambda/a$:

[39] Note that this is only true for scalar (e.g., the de Broglie) waves but not the vector ones, in particular, the electromagnetic waves where the intensity of the dipole radiation, and hence the scattering by small objects vanishes in the direction of the incident field's polarization – see, e.g., EM Eqs. (8.26) and (8.139).

$$\frac{d\sigma}{d\Omega} \approx 2\pi a^2 \left(\frac{mU_0 a^2}{\hbar^2} \right)^2 \exp\left\{ -k^2 a^2 \theta^2 \right\}. \tag{3.96}$$

This is, again, very typical for diffraction. Note, however, that due to the smooth character of the Gaussian potential (91), the diffraction pattern (98) exhibits no oscillations of $d\sigma/d\Omega$ as a function of the diffraction angle θ.

Such oscillations naturally appear for scatterers with sharp borders. Indeed, let us consider a uniform spherical scatterer described by the potential

$$U(\mathbf{r}) = \begin{cases} U_0, & \text{for } r < R, \\ 0, & \text{otherwise.} \end{cases} \tag{3.97}$$

In this case, integration by parts of Eq. (90) readily yields

$$f(\mathbf{k}, \mathbf{k}_i) = \frac{2mU_0}{\hbar^2 q^3} (qR \cos qR - \sin qR), \qquad \text{so } \frac{d\sigma}{d\Omega} = \left(\frac{2mU_0}{\hbar^2 q^3} \right)^2 (qR \cos qR - \sin qR)^2. \tag{3.98}$$

According to this result, the scattered wave's intensity drops very fast with q, so one needs a semi-log plot (such as that used in Fig. 10) to reveal small diffraction fringes,[40] with the n^{th} destructive interference (zero-intensity) point tending to $qR = \pi(n + ½)$ at $n \to \infty$. Since, as Fig. 9b shows, q may only change from 0 to $2k$, these intensity minima are only observable at sufficiently large values of the parameter kR when they correspond to real values of the scattering angle θ. (At $kR >> 1$, approximately kR/π of these minima, i.e. "dark rings" of low scattering probability, are observable.) On the contrary, at $kR << 1$ all allowed values of qR are much smaller than 1, so in this limit, the differential cross-section does not depend on qR, i.e. the scattering by the sphere (as by *any* object in this limit) is isotropic.

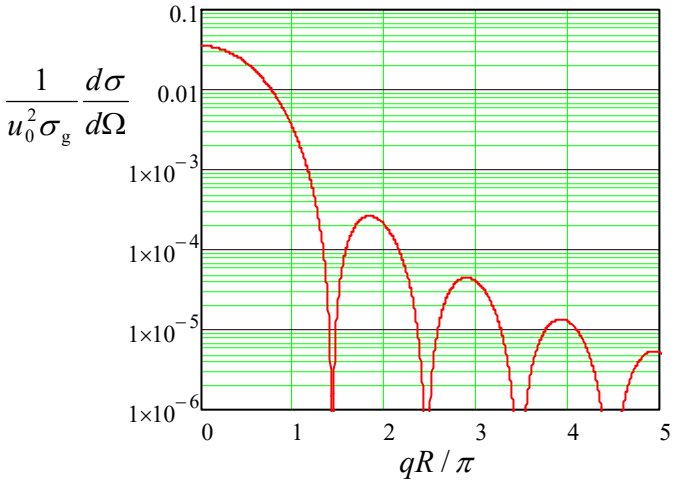


Fig. 3.10. The differential cross-section of the Born scattering of a particle by a "hard" (sharp-border) sphere (97), normalized to its geometric cross-section $\sigma_g \equiv \pi R^2$ and the square of the potential's magnitude parameter $u_0 \equiv U_0/(\hbar^2/2mR^2)$, as a function of the normalized magnitude of the scattering vector $\mathbf{q}$.

This example shows that in quantum mechanics, the notions of particle scattering and diffraction are essentially inseparable.

[40] Their physics is very similar to that of the Fraunhofer diffraction on a 1D scatterer – see, e.g., EM Sec. 8.4.

The Born approximation, while being very simple and used more than any other scattering theory, is not without shortcomings. For example, it is not too difficult (and hence is left for the reader's exercise) to prove the so-called *optical theorem*, valid for an arbitrary scatterer:

Optical theorem

$$\operatorname{Im} f(\mathbf{k}_i, \mathbf{k}_i) = \frac{k}{4\pi}\sigma . \tag{3.99}$$

However, Eq. (86) shows that in the Born approximation, the function f is purely real at $q = 0$ (i.e. for $\mathbf{k} = \mathbf{k}_i$), and hence cannot satisfy the optical theorem. Even more evidently, it cannot describe such a simple effect as a dark shadow ($\psi \approx 0$) cast by a virtually opaque object (say, with $U >> E$). There are several ways to improve the Born approximation, while still keeping its general idea of an approximate treatment of U.

(i) Instead of the main assumption $\psi_s \propto U_0$, we may use a complete perturbation series:

$$\psi_s = \psi_1 + \psi_2 + ... \tag{3.100}$$

with $\psi_n \propto U_0^n$, and find successive approximations ψ_n one by one. In the 1st approximation, we return to the Born formula, but already the 2nd approximation yields

$$\operatorname{Im} f_2(\mathbf{k}_i, \mathbf{k}_i) = \frac{k}{4\pi}\sigma_1 , \tag{3.101}$$

where σ_1 is the total cross-section calculated in the 1st approximation, so the optical theorem (99) is "almost satisfied".[41]

(ii) As was mentioned above, the Born approximation does not work very well for the objects stretched along the direction (say, x) of the initial wave vector $\mathbf{k}_i$. This deficiency may be corrected by the so-called *eikonal*[42] *approximation*, which replaces the plane-wave representation (60) of the incident wave with a WKB-like exponent, though still in the 1st approximation in $U \to 0$:

$$\begin{aligned} \exp\{i\mathbf{k}_i \cdot \mathbf{r}\} \equiv \exp\{ikx\} \to \exp\left\{i\int_0^x k(\mathbf{r}')dx'\right\} \equiv \exp\left\{i\int_0^x \frac{\{2m[E - U(\mathbf{r}')]\}^{1/2}}{\hbar}dx'\right\} \\ \approx \exp\left\{i\left[kx - \frac{m}{\hbar^2 k}\int_0^x U(\mathbf{r}')dx'\right]\right\}. \end{aligned} \tag{3.102}$$

The results of this approach satisfy the optical theorem (99) already in the 1st approximation.

Another way toward quantitative results in the theory of scattering, beyond the Born approximation, may be pursued for spherically symmetric potentials (89); I will discuss it in Sec. 8, after a general discussion of particle motion in such potentials in Sec. 7.

3.4. Energy bands in higher dimensions

In Sec. 2.7, we have discussed the 1D band theory for potential profiles $U(x)$ that obey the periodicity condition (2.192). For what follows, let us notice that the condition may be rewritten as

[41] An even simpler way to satisfy the theorem (even in the Born approximation) is to change the definition of the function $f(\mathbf{k}, \mathbf{k}_i)$ so that for forward scattering ($\mathbf{k} = \mathbf{k}_i$), it includes the incident de Broglie wave as well.

[42] From the Greek word εικον, meaning "image". In our current context, this term is purely historic.

$$U(x+X)=U(x), \tag{3.103}$$

where $X = \tau a$, with τ being an arbitrary integer. One may say that the set of points X forms a periodic *1D lattice* in the direct (**r**-) space. We have also seen that each Bloch state (i.e., each eigenstate of the Schrödinger equation for such periodic potential) is characterized by the quasimomentum $\hbar q$, and its energy does not change if q is changed by a multiple of $2\pi/a$. Hence if we form, in the reciprocal (**q**-) space, a 1D lattice of points $Q = lb$, with $b \equiv 2\pi/a$ and integer l, any two points from these two mutually reciprocal lattices satisfy the following rule:

$$\exp\{iQX\} = \exp\left\{ il\frac{2\pi}{a}\tau a \right\} \equiv e^{2\pi i l\tau} = 1, \tag{3.104}$$

because the product of any two integers l and τ is also an integer.

In this form, the results of Sec. 2.7 may be readily generalized to d-dimensional periodic potentials whose translational symmetry obeys the following natural generalization of Eq. (103):

$$U(\mathbf{r}+\mathbf{R})=U(\mathbf{r}), \tag{3.105}$$

where the points **R**, which may be numbered by d integers τ_j, form the so-called *Bravais lattice*:[43]

$$\mathbf{R}=\sum_{j=1}^{d}\tau_j\mathbf{a}_j, \tag{3.106}$$

Bravais lattice

with d *primitive vectors* $\mathbf{a}_j$. The simplest example of a 3D Bravais lattice is given by the *simple cubic lattice* (Fig. 11a), which may be described by a system of mutually perpendicular primitive vectors $\mathbf{a}_j$ of equal length. However, not in any lattice these vectors are perpendicular; for example, Figs. 11b and 11c show possible sets of the primitive vectors describing, respectively, the *face-centered cubic* (fcc) lattice and the *body-centered cubic* (bcc) lattice. In 3D, the science of crystallography based on group theory distinguishes, by their symmetry properties, 14 different Bravais lattices, which may be grouped into 7 distinct *lattice systems*.[44]

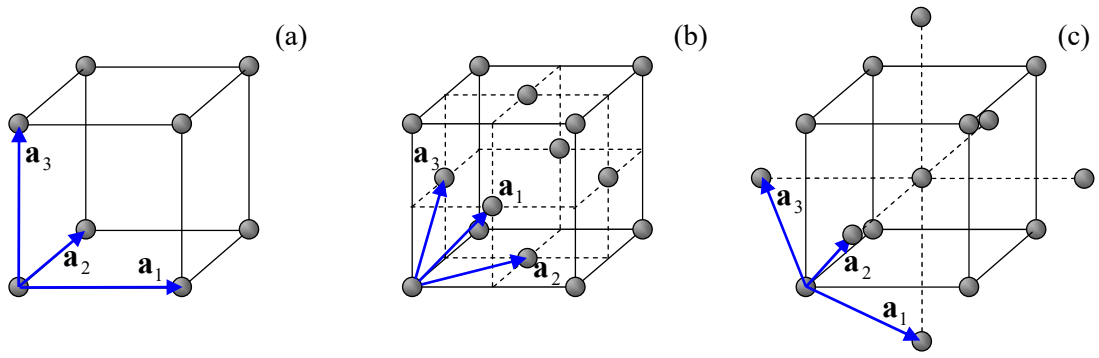


Fig. 3.11. The simplest (and most common) 3D Bravais lattices: (a) simple cubic, (b) face-centered cubic (fcc), and (c) body-centered cubic (bcc), and possible choices of their primitive vector sets (blue arrows).

[43] Named after A. Bravais, the crystallographer who introduced this notion in 1850.

[44] An exceptionally clear and well-illustrated introduction to the Bravais lattices is given in Chapters 4 and 7 of the famous textbook by N. Ashcroft and N. Mermin, *Solid State Physics*, Saunders College, 1976.

Note, however, not all highly symmetric sets of single points form Bravais lattices. As probably the most striking example, the nodes of the very simple 2D *honeycomb lattice* (Fig. 12a)[45] cannot be described by a Bravais lattice – while those of the 2D *hexagonal lattice* shown in Fig. 12b, can. The most prominent 3D case of such a lattice is the diamond structure (Fig. 12c), which describes, in particular, silicon crystals.[46] In cases like these, the band theory is much facilitated by the fact that the Bravais lattices using some point groups called *unit cells* (or "bases", or "cells with basis", or "motifs") may describe these systems.[47] For example, Fig. 12a shows a possible choice of the primitive vectors for the honeycomb lattice, with the unit cell formed by any two adjacent points of the original lattice (say, within the dashed ovals on that panel). Similarly, the diamond lattice may be described as an fcc Bravais lattice with a two-point unit cell – see Fig. 12c.

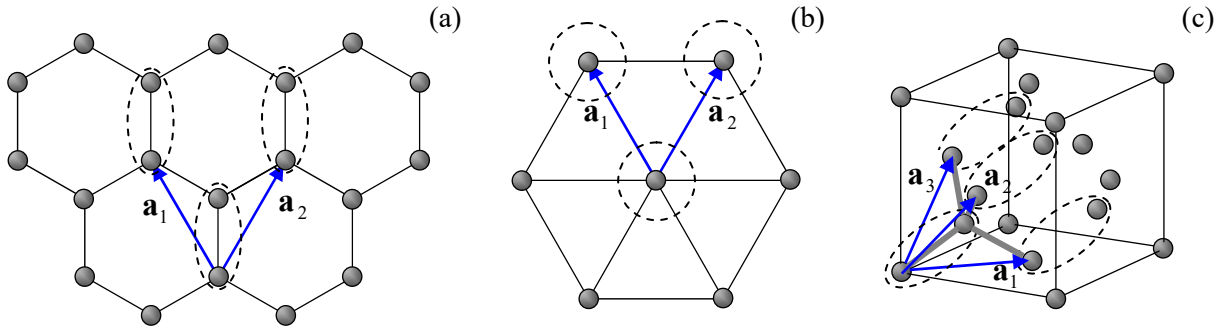


Fig. 3.12. Two important periodic structures that require two-point unit cells for their Bravais lattice representation: (a) 2D honeycomb lattice and (c) 3D diamond lattice, and their primitive vectors. For contrast, panel (b) shows the 2D hexagonal lattice that forms a Bravais lattice with a single-point unit cell.

Now we are ready for the following generalization of the 1D Bloch theorem, given by Eqs. (2.193) and (2.210), to higher dimensions: any eigenfunction of the Schrödinger equation describing a particle's motion in the spatially-unlimited periodic potential (105) may be represented either as

3D Bloch theorem

$$\psi(\mathbf{r}+\mathbf{R}) = \psi(\mathbf{r}) e^{i\mathbf{q}\cdot\mathbf{R}}, \tag{3.107}$$

or as

$$\psi(\mathbf{r}) = u(\mathbf{r}) e^{i\mathbf{q}\cdot\mathbf{r}}, \quad \text{with } u(\mathbf{r}+\mathbf{R}) = u(\mathbf{r}), \tag{3.108}$$

where the quasimomentum $\hbar\mathbf{q}$ is again a constant of motion, but now it is a vector. The key notion of the band theory in d dimensions is the *reciprocal lattice* in the wave vector (**q**-) space, formed as

Reciprocal lattice in **q**-space

$$\mathbf{Q} = \sum_{j=1}^{d} l_j \mathbf{b}_j, \tag{3.109}$$

[45] This structure describes, for example, the now-famous *graphene* – isolated monolayer sheets of carbon atoms arranged in a honeycomb lattice with an interatomic distance of 0.142 nm.

[46] This diamond structure may be best understood as an overlap of two fcc lattices of side a, mutually shifted by the vector $\{1, 1, 1\}\times a/4$, so the distances between each point of the combined lattice and its 4 nearest neighbors (see the solid gray lines in Fig. 12c) are all equal.

[47] A harder case is presented by so-called *quasicrystals* (whose idea may be traced down to medieval Islamic tilings, but was discovered in natural crystals, by D. Shechtman *et al.*, only in 1984), which obey a high (say, the 5-fold) rotational symmetry, but cannot be described by a Bravais lattice with *any* finite unit cell. For a popular review of quasicrystals, see, for example, P. Stephens and A. Goldman, *Sci. Amer.* **264**, #4, 24 (1991).

with integer l_j, and vectors $\mathbf{b}_j$ selected in such a way that the following natural generalization of Eq. (104) is valid for any two points of the direct and reciprocal lattices:

$$e^{i\mathbf{Q}\cdot\mathbf{R}} = 1. \tag{3.110}$$

One way to describe the physical sense of the lattice $\mathbf{Q}$ is to say that according to Eqs. (80) and/or (86), it gives the set of the vectors $\mathbf{q} \equiv \mathbf{k} - \mathbf{k}_i$ for that the interference of the waves scattered by all Bravais lattice points is constructive, and hence strongly enhanced.[48] Another way to look at the reciprocal lattice follows from the first formulation of the Bloch theorem, given by Eq. (107): if we add to the quasimomentum $\mathbf{q}$ of a particle any vector $\mathbf{Q}$ of the reciprocal lattice, the wavefunction does not change. This means, in particular, that all information about the system's eigenfunctions is contained inside just one elementary cell of the reciprocal space $\mathbf{q}$. Its most frequent choice, called the *1st Brillouin zone*, is the set of all points $\mathbf{q}$ that are closer to the origin than to any other point of the lattice $\mathbf{Q}$. (Evidently, the 1st Brillouin zone in one dimension, discussed in Sec. 2.7, falls under this definition – see, e.g., Figs. 2.26 and 2.28.)

It is easy to see that the primitive vectors $\mathbf{b}_j$ of the reciprocal lattice may be constructed as

Reciprocal lattice: primitive vectors

$$\mathbf{b}_1 = 2\pi\frac{\mathbf{a}_2\times\mathbf{a}_3}{\mathbf{a}_1\cdot(\mathbf{a}_2\times\mathbf{a}_3)}, \quad \mathbf{b}_2 = 2\pi\frac{\mathbf{a}_3\times\mathbf{a}_1}{\mathbf{a}_1\cdot(\mathbf{a}_2\times\mathbf{a}_3)}, \quad \mathbf{b}_3 = 2\pi\frac{\mathbf{a}_1\times\mathbf{a}_2}{\mathbf{a}_1\cdot(\mathbf{a}_2\times\mathbf{a}_3)}. \tag{3.111}$$

Indeed, from the "operand rotation rule" of the vector algebra[49] it is evident that $\mathbf{a}_j\cdot\mathbf{b}_{j'} = 2\pi\delta_{jj'}$. Hence, with the account of Eq. (109), the exponent on the left-hand side of Eq. (110) is reduced to

$$e^{i\mathbf{Q}\cdot\mathbf{R}} = \exp\{2\pi i(l_1\tau_1 + l_2\tau_2 + l_3\tau_3)\}. \tag{3.112}$$

Since all l_j and all τ_j are integers, the expression in the parentheses is also an integer, so the exponent indeed equals 1, thus satisfying the definition of the reciprocal lattice given by Eq. (110).

As the simplest example, let us return to the simple cubic lattice of a period a (Fig. 11a), oriented in space so that

$$\mathbf{a}_1 = a\mathbf{n}_x, \quad \mathbf{a}_2 = a\mathbf{n}_y, \quad \mathbf{a}_3 = a\mathbf{n}_z. \tag{3.113}$$

According to Eq. (111), its reciprocal lattice is also simple cubic:

$$\mathbf{Q} = \frac{2\pi}{a}(l_x\mathbf{n}_x + l_y\mathbf{n}_y + l_z\mathbf{n}_z), \tag{3.114}$$

so the 1st Brillouin zone is a cube with the side $b = 2\pi/a$.

Almost equally simple calculations show that the reciprocal lattice of fcc is bcc, and vice versa. Figure 13 shows the resulting 1st Brillouin zone of the fcc lattice.

The notion of the reciprocal lattice makes the multi-dimensional band theory not much more complex than that in 1D, especially for numerical calculations, at least for the single-point Bravais

[48] This is why the notion of the **Q**-lattice is also the main starting point of X-ray diffraction studies of crystals. Indeed, it allows rewriting the well-known Bragg condition for diffraction peaks in an extremely simple form: $\mathbf{k} = \mathbf{k}_i + \mathbf{Q}$, where $\mathbf{k}_i$ and $\mathbf{k}$ are the wave vectors of the, respectively, incident and diffracted waves – see, e.g., EM Sec. 8.4 (where it was more convenient for me to use the notation $\mathbf{k}_0$ for $\mathbf{k}_i$).

[49] See, e.g., MA Eq. (7.6).

lattices. Indeed, repeating all the steps that have led us to Eq. (2.218), but now with a d-dimensional Fourier expansion of the functions $U(\mathbf{r})$ and $u(\mathbf{r})$, we readily get its generalization:

$$\sum_{\mathbf{l}'\neq\mathbf{l}} U_{\mathbf{l}'-\mathbf{l}} u_{\mathbf{l}'} = (E - E_{\mathbf{l}}) u_{\mathbf{l}}, \tag{3.115}$$

where $\mathbf{l}$ is now a d-dimensional vector of integer indices l_j. The summation in Eq. (115) should be carried over all essential components of this vector (i.e. over all relevant nodes of the reciprocal lattice), so writing a corresponding computer code requires a bit more care than in 1D. However, this is just a homogeneous system of linear equations for coefficients $u_{\mathbf{l}}$, and numerous routines of finding its eigenvalues E are readily available from both public sources and commercial software packages.

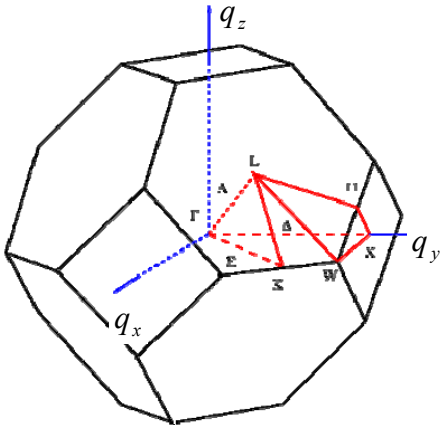


Fig. 3.13. The 1st Brillouin zone of the fcc lattice, and the traditional notation of its main directions. Adapted from http://en.wikipedia.org/wiki/Band_structure, as a public domain material.

What is indeed more complex than in 1D is the representation (and hence comprehension :-) of the calculated results and experimental data. Typically, the representation is limited to plotting the Bloch state eigenenergy as a function of the vector q's components along certain special directions in the reciprocal space of quasimomentum (see, e.g., the red lines in Fig. 13), typically on a single panel. Fig. 14 shows perhaps the most famous (and certainly the most practically important) of such plots, the band structure of electrons in crystalline silicon. The dashed horizontal lines mark the so-called *indirect gap* of width ~1.12 eV between the lower "valence" (nominally occupied) and the next "conduction" (nominally unoccupied) energy bands.

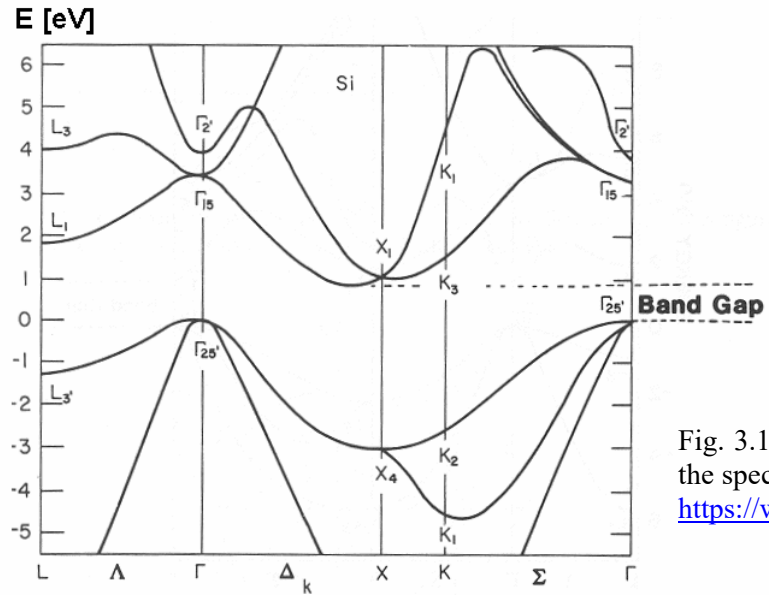


Fig. 3.14. The band structure of silicon, plotted along the special directions shown in Fig. 13. (Adapted from https://www.tf.uni-kiel.de/matwis/amat/semi_en/.)

In order to understand the reason for such complexity, let us see how would we start to calculate such a picture in the weak-potential approximation, for the simplest case of a 2D square lattice – which is a subset of the cubic lattice (106), with $\tau_3 = 0$. Its 1st Brillouin zone is of course also a square, of the area $(2\pi/a)^2$ – see the dashed lines in Fig. 15. Let us draw the lines of the constant energy of a free particle ($U = 0$) in this zone. Repeating the arguments of Sec. 2.7 (see especially Fig. 2.28 and its discussion), we may conclude that Eq. (2.216) should be now generalized as follows,

$$E = \frac{\hbar^2 k^2}{2m} = \frac{\hbar^2}{2m}\left[\left(q_x - \frac{2\pi l_x}{a}\right)^2 + \left(q_y - \frac{2\pi l_y}{a}\right)^2\right], \tag{3.116}$$

with all possible integers l_x and l_y. Considering this result only within the 1st Brillouin zone, we see that as the particle's energy E grows, the lines of equal energy, for the lowest energy band, evolve as shown in Fig. 15. Just like in 1D, the weak-potential effects are only important at the Brillouin zone boundaries and may be crudely represented as the appearance of narrow energy gaps. However, one can see that the band structure in the **q**-space is complex enough even without these effects – and becomes even more involved at higher E.

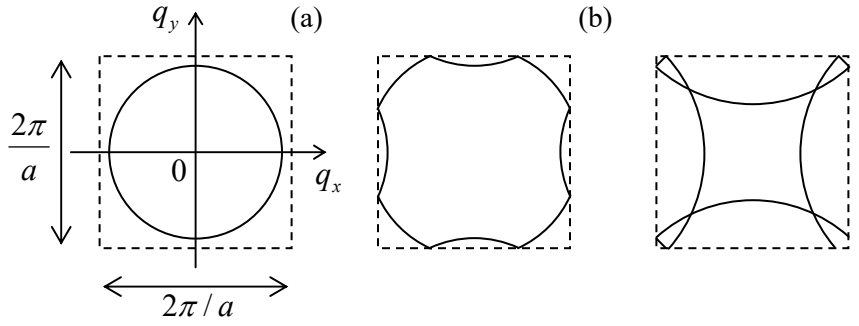


Fig. 3.15. The lines of constant energy E of a free particle, within the 1st Brillouin zone of a square Bravais lattice, for: (a) $E/E_1 \approx 0.95$, (b) $E/E_1 \approx 1.05$; and (c) $E/E_1 \approx 2.05$, where $E_1 \equiv \pi^2\hbar^2/2ma^2$.

The tight-binding approximation is usually easier to follow. For example, for the same square 2D lattice, we may repeat the arguments that have led us to Eq. (2.203), to write [50]

$$i\hbar\dot{a}_{0,0} = -\delta_n\left(a_{-1,0} + a_{+1,0} + a_{0,+1} + a_{0,-1}\right), \tag{3.117}$$

where the indices correspond to the deviations of the integers τ_x and τ_y from an arbitrarily selected minimum of the potential energy – and hence of the wavefunction's "hump" that is quasi-localized at this minimum. Now, looking for the stationary solution of these equations, that would obey the Bloch theorem (107), instead of Eq. (2.206) we get

$$E = E_n + \varepsilon_n = E_n - \delta_n\left(e^{iq_x a} + e^{-iq_x a} + e^{iq_y a} + e^{-iq_y a}\right) \equiv E_n - 2\delta_n\left(\cos q_x a + \cos q_y a\right). \tag{3.118}$$

Figure 16 shows this result, within the 1st Brillouin zone, in two forms: as color-coded lines of equal energy, and as a 3D plot. It is evident that the plots of this function along different lines on the **q**-plane, for example along one of the axes (say, q_x) and along a diagonal of the 1st Brillouin zone (say, with $q_x = q_y$) give different curves $E(q)$, qualitatively similar to those of silicon (Fig. 14).

[50] Actually, using the same values of δ_n in both directions (x and y) implies some sort of symmetry of the quasi-localized states. For example, the s-states of axially-symmetric potentials (see the next section) always have such symmetry.

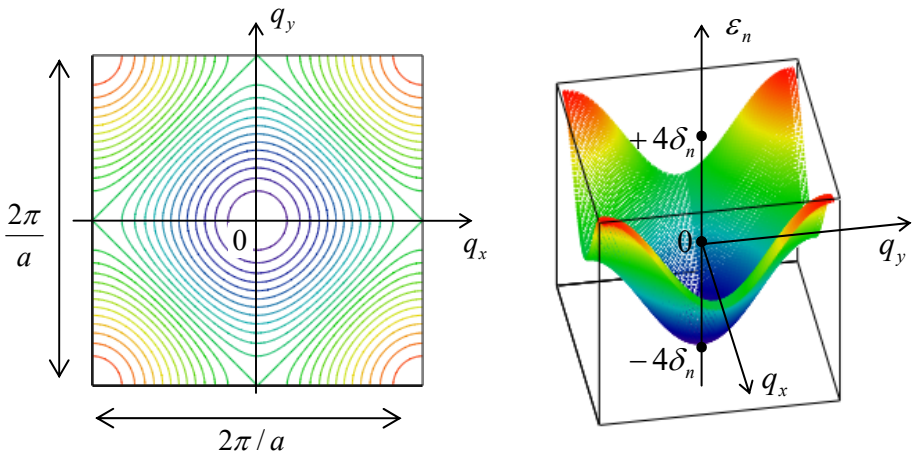


Fig. 3.16. The allowed band energy $\varepsilon_n \equiv E - E_n$ for a square 2D lattice, in the tight-binding approximation.

However, the latter structure is further complicated by the fact that the unit cell of its Bravais lattice contains two atoms – see Fig. 12c and its discussion. In this case, even the tight-binding picture becomes more complex. Indeed, even if the atoms at different positions of the unit cell are similar (as they are, for example, in both graphene and silicon), and hence the potential wells near those points and the corresponding local wavefunctions $u(\mathbf{r})$ are similar as well, the Bloch theorem (which only pertains to Bravais lattices!) does not forbid them from having different complex probability amplitudes $a(t)$ whose time evolution should be described by a specific differential equation.

As the simplest example, to describe the honeycomb lattice shown in Fig. 12a, we have to prescribe different probability amplitudes to the "top" and "bottom" points of its unit cell – say, α and β, correspondingly. Since each of these points is surrounded (and hence weakly interacts) with three neighbors of the opposite type, instead of Eq. (117) we have to write two equations:

$$i\hbar\dot{\alpha} = -\delta_n \sum_{j=1}^{3} \beta_j, \qquad i\hbar\dot{\beta} = -\delta_n \sum_{j'=1}^{3} \alpha_{j'}, \tag{3.119}$$

where each summation is over three nearest-neighbor points. (In these two sums, I am using different summation indices just to emphasize that these directions are different for the "top" and "bottom" points of the unit cell – see Fig. 12a.) Now using the Bloch theorem (107) in the form similar to Eq. (2.205), we get two coupled systems of linear algebraic equations:

$$(E - E_n)\alpha = -\delta_n \beta \sum_{j=1}^{3} e^{i\mathbf{q}\cdot\mathbf{r}_j}, \qquad (E - E_n)\beta = -\delta_n \alpha \sum_{j'=1}^{3} e^{i\mathbf{q}\cdot\mathbf{r}'_{j'}}, \tag{3.120}$$

where $\mathbf{r}_j$ and $\mathbf{r}'_{j'}$ are the nearest-neighbor positions, as seen from the top and bottom points, respectively. Writing the condition of consistency of this system of homogeneous linear equations, we get two equal and opposite values for energy correction for each value of $\mathbf{q}$:

$$E_{\pm} = E_n \pm \delta_n \Sigma^{1/2}, \qquad \text{where } \Sigma \equiv \sum_{j,j'=1}^{3} e^{i\mathbf{q}\cdot\left(\mathbf{r}_j + \mathbf{r}'_{j'}\right)}. \tag{3.121}$$

According to Eq. (120), these two energy bands correspond to the phase shifts (on the top of the regular Bloch shift $\mathbf{q}\cdot\Delta\mathbf{r}$) of either 0 or π between the adjacent quasi-localized wavefunctions $u(\mathbf{r})$.

The most interesting corollary of such energy symmetry, augmented by the honeycomb lattice's symmetry, is that for certain values $\mathbf{q}_D$ of the vector $\mathbf{q}$ (that turn out to be in each of six corners of the honeycomb-shaped 1st Brillouin zone), the double sum Σ vanishes, i.e. the two band surfaces $E_\pm(\mathbf{q})$ touch each other. As a result, in the vicinities of these so-called *Dirac points*,[51] the dispersion relation is linear:

$$E_\pm\Big|_{\mathbf{q}\approx\mathbf{q}_D} \approx E_n \pm \hbar v_n \left|\tilde{\mathbf{q}}\right|, \quad \text{where } \tilde{\mathbf{q}} \equiv \mathbf{q} - \mathbf{q}_D, \tag{3.122}$$

with $v_n \propto \delta_n$ being a constant with the dimension of velocity – for graphene, close to 10^6 m/s. Such a linear dispersion relation ensures several interesting transport properties of graphene, in particular of the quantum Hall effect in it – as was already mentioned in Sec. 2. For their more detailed discussion, I have to refer the reader to special literature.[52]

3.5. Axially symmetric systems

I cannot conclude this chapter (and hence our review of wave mechanics) without addressing the exact solutions of the stationary Schrödinger equation[53] possible in the cases of highly symmetric functions $U(\mathbf{r})$. Such solutions are very important, in particular, for atomic and nuclear physics, and will be used in the later chapters of this course.

In some rare cases, such symmetries may be exploited by the separation of variables in Cartesian coordinates. The most famous (and rather important) example is the d-dimensional isotropic harmonic oscillator – a particle moving inside the potential well

$$U = \frac{m\omega_0^2}{2}\sum_{j=1}^{d} r_j^2 \ . \tag{3.123}$$

Separating the variables exactly as we did in Sec. 1.7 for the rectangular hard-wall box (1.77), for each degree of freedom we get the Schrödinger equation (2.261) of a 1D oscillator, whose eigenfunctions are

[51] This term is based on a (rather indirect) analogy with the Dirac theory of relativistic quantum mechanics, to be discussed in Chapter 9 below.

[52] See, e.g., the reviews by A. Castro Neto *et al*., *Rev. Mod. Phys*. **81**, 109 (2009) and by X. Lu *et al*., *Appl. Phys. Rev*. **4**, 021306 (2017). Note that the transport properties of graphene are determined by coupling of $2p_z$-states of its carbon atom electrons (see Secs. 6 and 7 below), whose wavefunctions are proportional to $\exp\{\pm i\varphi\}$ rather than are axially symmetric as implied by Eqs. (120). However, due to the lattice symmetry, this fact does not affect the above dispersion relation $E(\mathbf{q})$.

[53] This is my best chance to mention, in passing, that the eigenfunctions $\psi_n(\mathbf{r})$ of any such problem do not feature the instabilities typical for the *deterministic chaos* effects of classical mechanics – see, e.g., CM Chapter 9. (This is why the term *quantum mechanics of classically chaotic systems* is preferable to the occasionally used term "quantum chaos".) It is curious that at the initial stages of the time evolution of the wavefunctions of such systems, their certain correlation functions still grow exponentially, reminding the *Lyapunov exponents* λ of their classical chaotic dynamics. This growth stops at the so-called *Ehrenfest times* $t_E \sim \lambda^{-1}\ln(S/\hbar)$, where S is the action scale of the problem – see, e.g., I. Aleiner and A. Larkin, *Phys. Rev. E* **55**, R1243 (1997). In a stationary quantum state, the most essential trace of the classical chaos in a system is the unusual statistics of its eigenvalues, in particular of the energy spectra. We will have a chance for a brief look at such statistics in Chapter 5, but unfortunately, I will not have time/space to discuss this field in much detail. Perhaps the best available book for further reading is the monograph by M. Gutzwiller, *Chaos in Classical and Quantum Mechanics*, Springer, 1991.

given by Eq. (2.284), and the energy spectrum is described by Eq. (2.162). As a result, the total energy spectrum may be indexed by a vector $\mathbf{n} = \{n_1, n_2, \ldots, n_d\}$ of d independent integer quantum numbers n_j:

$$E_{\mathbf{n}} = \hbar\omega_0\left(\sum_{j=1}^{d} n_j + \frac{d}{2}\right), \tag{3.124}$$

each ranging from 0 to ∞. Note that every energy level of this system, with the only exception of its ground state,

$$\psi_g = \prod_{j=1}^{d}\psi_0(r_j) = \frac{1}{\pi^{d/4}x_0^{d/2}}\exp\left\{-\frac{1}{2x_0^2}\sum_{j=1}^{d} r_j^2\right\}, \tag{3.125}$$

is *degenerate*: several different wavefunctions, each with its own different set of the quantum numbers n_j, but the same value of their sum, have the same energy.

However, the harmonic oscillator problem is an exception: for other axially and spherically symmetric problems, the solution is made much easier by using the appropriate curvilinear coordinates. Let us start with the simplest axially symmetric problem: the so-called *planar rotor* (or "rotator"), i.e. a particle of mass $\mathscr{m}$,[54] constrained to move along a plane circle of radius R (Fig. 17).[55]

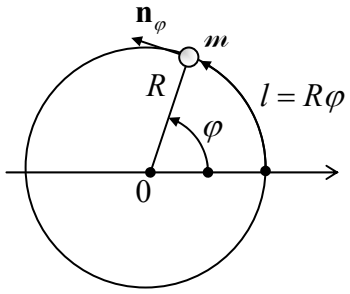


Fig. 3.17. A planar rigid rotor.

The classical state of such planar rotor may be described by just *one* coordinate, say, the angular displacement φ (or equivalently, the arc displacement $l \equiv R\varphi$) from some reference direction, with the energy (and the Hamiltonian function) $H = p^2/2\mathscr{m}$, where $\mathbf{p} \equiv \mathscr{m}\mathbf{v} = \mathscr{m}\mathbf{n}_\varphi(dl/dt)$, $\mathbf{n}_\varphi$ being the unit vector in the azimuthal direction – see Fig. 17. This function is similar to that of a free 1D particle (with the replacement $x \to l \equiv R\varphi$), and hence the rotor's quantum properties may be described by a similar Hamiltonian operator:

$$\hat{H} = \frac{\hat{p}^2}{2\mathscr{m}}, \qquad \text{with } \hat{\mathbf{p}} = -i\hbar\mathbf{n}_\varphi\frac{\partial}{\partial l} \equiv -i\frac{\hbar}{R}\mathbf{n}_\varphi\frac{\partial}{\partial\varphi}, \tag{3.126}$$

whose eigenfunctions have a similar structure:

$$\psi = Ce^{ikl} \equiv Ce^{ikR\varphi}. \tag{3.127}$$

[54] From this point on (until the chapter's end), I will use this exotic font for the particle's mass, to avoid any chance of its confusion with the impending "magnetic" quantum number m, traditionally used in axially-symmetric problems.

[55] This is a reasonable model for the confinement of light atoms, notably hydrogen, in some organic compounds, but I am addressing this system mostly as the basis for the forthcoming, more complex problems.

The "only" new feature is that in the rotor, all observables should be 2π-periodic functions of the angle φ. Hence, as we have already discussed in the context of the magnetic flux quantization (see Fig. 4 and its discussion), as the particle makes one turn around the central point 0, its wavefunction's phase $kR\varphi$ may only change by $2\pi m$, with an arbitrary integer m (ranging from $-\infty$ to $+\infty$):

$$\psi_m(\varphi + 2\pi) = \psi_m(\varphi) e^{2\pi i m}. \tag{3.128}$$

With the eigenfunctions (127), this periodicity condition immediately gives $2\pi kR = 2\pi m$. Thus, the wave number k can take only quantized values $k_m = m/R$, so the eigenfunctions should be indexed by this *magnetic* quantum number m:

$$\psi_m = C_m \exp\left\{ im \frac{l}{R} \right\} \equiv C_m \exp\{ im\varphi \}, \tag{3.129}$$

Planar rotor: eigenfunctions

and the energy spectrum is discrete[56]:

$$E_m = \frac{p_m^2}{2m} = \frac{\hbar^2 k_m^2}{2m} = \frac{\hbar^2 m^2}{2mR^2}. \tag{3.130}$$

Planar rotor: eigenenergies

This simple model allows exact analysis of an external magnetic field's effect on a confined motion of an electrically charged particle. Indeed, in the simplest case when this field is axially symmetric (or just uniform) and directed normally to the rotor's plane, it does not violate the axial symmetry of the system. According to Eq. (26), in this case, we have to generalize Eq. (126) as

$$\hat{H} = \frac{1}{2m}\left(-i\hbar \mathbf{n}_\varphi \frac{\partial}{\partial l} - q\mathbf{A} \right)^2 \equiv \frac{1}{2m}\left(-i\frac{\hbar}{R} \mathbf{n}_\varphi \frac{\partial}{\partial \varphi} - q\mathbf{A} \right)^2. \tag{3.131}$$

Here, in contrast to the Cartesian gauge choice (44), which was so instrumental for the solution of the Landau level problem, it is beneficial to take the vector potential in the axially symmetric form $\mathbf{A} = A(\rho)\mathbf{n}_\varphi$, where $\boldsymbol{\rho} \equiv \{x, y\}$ is the 2D radius vector, with the magnitude $\rho = (x^2 + y^2)^{1/2}$. Using the well-known expression for the curl operator in the cylindrical coordinates,[57] we can readily check that the requirement $\nabla \times \mathbf{A} = \mathcal{B}\mathbf{n}_z$, with $\mathcal{B}$ = const, is satisfied by the following function (which was already mentioned in Sec. 2):

$$\mathbf{A} = \mathbf{n}_\varphi \frac{\mathcal{B}\rho}{2}. \tag{3.132}$$

For the planar rotor, $\rho = R$ = const, so the stationary Schrödinger equation becomes

$$\frac{1}{2m}\left(-i\frac{\hbar}{R}\frac{\partial}{\partial \varphi} - q\frac{\mathcal{B}R}{2} \right)^2 \psi_m = E_n \psi_m. \tag{3.133}$$

A little bit surprisingly, this equation is still satisfied with the eigenfunctions (127)! Moreover, since the periodicity condition (128) is also unaffected by the applied magnetic field, we return to the periodic eigenfunctions (129), independent of $\mathcal{B}$. However, the field does affect the system's eigenenergies:

[56] Note that E_m does not include the radial confinement energy. (See Sec. 2.1 and the solution of Problem 2.1.)
[57] See, e.g., MA Eq. (10.5).

Planar rotor: magnetic field's effect

$$E_m = \frac{1}{\psi_m}\frac{1}{2m}\left(-i\frac{\hbar}{R}\frac{\partial}{\partial\varphi} - q\frac{\mathcal{B}R}{2}\right)^2\psi_m = \frac{1}{2m}\left(\frac{\hbar m}{R} - q\frac{\mathcal{B}R}{2}\right)^2 \equiv \frac{\hbar^2}{2mR^2}\left(m - \frac{\Phi}{\Phi_0'}\right)^2, \qquad (3.134)$$

where $\Phi \equiv \pi R^2\mathcal{B}$ is the magnetic flux through the area limited by the particle's trajectory, and $\Phi_0' \equiv 2\pi\hbar/q$ is the "normal" magnetic flux quantum we have already met in the AB effect's context – see Eq. (34) and its discussion. The field also changes the circular electric current of the particle for each m:

$$I_m = q\frac{\hbar}{2imR}\left[\psi_m^*\left(\frac{\partial}{\partial\varphi} - \frac{iqR\mathcal{B}}{2\hbar}\right)\psi_m - \text{c.c.}\right] = q\frac{\hbar}{mR}|C_m|^2\left(m - \frac{\Phi}{\Phi_0'}\right). \qquad (3.135)$$

Normalizing the wavefunction (129) to have $W_m = 1$, we get $|C_m|^2 = 1/2\pi R$, so Eq. (135) becomes

$$I_m = \left(m - \frac{\Phi}{\Phi_0'}\right)I_0, \qquad \text{with } I_0 \equiv \frac{\hbar q}{2\pi m R^2}. \qquad (3.136)$$

The functions $E_m(\Phi)$ and $I_m(\Phi)$ are shown in Fig. 18. Note that since $\Phi_0' \propto 1/q$, for any sign of the particle's charge q, $dI_m/d\Phi < 0$. It is easy to verify that this means that the current is *diamagnetic* for any sign of q:[58] the field-induced current flows in such a direction that its own magnetic field tries to compensate for the external magnetic flux applied to the loop. This result may be interpreted as a different manifestation of the AB effect.[59] In contrast to the interference experiment that was discussed in Sec. 1, in the situation shown in Fig. 17 the particle is not absorbed by the detector but travels around the ring continuously. As a result, its wavefunction is "rigid": due to the periodicity condition (128), the quantum number m is discrete, and the applied magnetic field cannot change the wavefunction gradually. In this sense, the system is similar to a superconducting loop – see Fig. 4 and its discussion. The difference between these systems is two-fold:

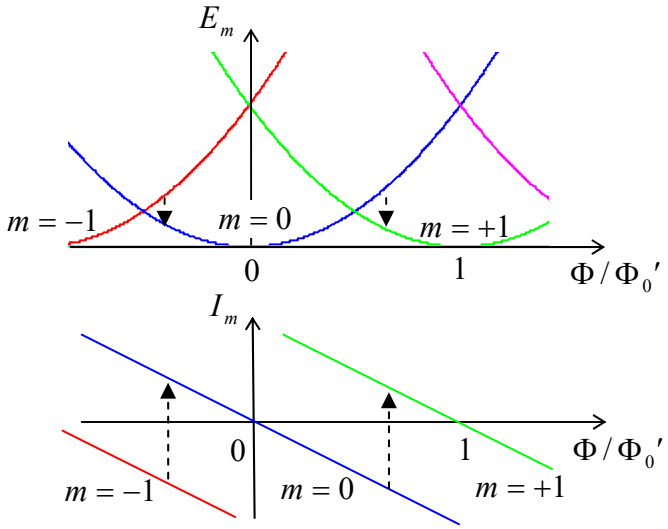


Fig. 3.18. The magnetic field's effect on a charged planar rotor. Dashed arrows show possible inelastic transitions between metastable and ground states, due to weak interaction with the environment, as the external magnetic field is slowly increased.

[58] This effect, whose qualitative features remain the same for all 2D or 3D localized states (see Chapter 6 below), is frequently referred to as *orbital diamagnetism*. In magnetic materials consisting of particles with uncompensated spins, this effect competes with an opposite effect, *spin paramagnetism* – see, e.g., EM Sec. 5.5.

[59] It is straightforward to check that the final forms of Eqs. (134)-(136) remain valid even if the magnetic field is localized well inside the rotor's circumference so its lines do not touch the particle's trajectory.

(i) For a single charged particle, in macroscopic systems with practicable values of q, R, and m, the scale I_0 of the induced current is very small. For example, for $m = m_e$, $q = -e$, and $R = 1$ μm, Eq. (136) yields $I_0 \approx 3$ pA.[60] With the ring's inductance $\mathscr{L}$ of the order of $\mu_0 R$,[61] the contribution $\Phi_I = \mathscr{L} I \sim \mu_0 R I_0 \sim 10^{-24}$ Wb of such a small current to the net magnetic flux Φ is negligible in comparison with $\Phi_0' \sim 10^{-15}$ Wb, so the wavefunction quantization does not lead to the constancy of the total magnetic flux.

(ii) As soon as the magnetic field raises the eigenstate energy E_m above that of another eigenstate $E_{m'}$, the former state becomes metastable, and a weak interaction of the system with its environment (which is neglected in our simple model, but will be discussed in Chapter 7) may induce a quantum transition of the system to the lower-energy state, thus reducing the diamagnetic current's magnitude – see the dashed lines in Fig. 18. The flux quantization in superconductors is much more robust to such perturbations.[62]

Now let us return, once again, to the key Eq. (129), and see what it gives for one more important observable, the particle's angular momentum

$$\mathbf{L} \equiv \mathbf{r} \times \mathbf{p}\,, \tag{3.137}$$

In this particular geometry, the vector $\mathbf{L}$ has just one component, normal to the rotor plane:

$$L_z = Rp\,. \tag{3.138}$$

In classical mechanics, the rotor's L_z should be conserved (due to the absence of an external torque), but it may take arbitrary values. In quantum mechanics, the situation changes: with $p = \hbar k$, our result $k_m = m/R$ for the m^{th} eigenstate may be rewritten as

$$\boxed{(L_z)_m = R\hbar k_m = \hbar m\,.} \tag{3.139}$$

Angular momentum quantization

Thus, the angular momentum is quantized: it may be only a multiple of the Planck constant $\hbar$ – confirming the N. Bohr's guess – see Eq. (1.8). As we will see in Chapter 5, this result is very general (though it may be modified by spin effects), and the wavefunctions (129) may be interpreted as eigenfunctions of the angular momentum operator.

Let us see whether this quantization persists in more general but still axial-symmetric systems. To implement the planar rotor in our 3D world, we needed to provide rigid confinement of the particle both in the motion plane and along the 2D radius ρ. Let us consider a more general situation when only the former confinement is strict, i.e. the case when a 2D particle moves in an arbitrary axially symmetric potential

$$U(\boldsymbol{\rho}) = U(\rho)\,. \tag{3.140}$$

[60] Such weak persistent, macroscopic diamagnetic currents in non-superconducting systems have been experimentally observed by measuring the weak magnetic field induced by the currents, in systems of a large number (~10^7) of similar conducting rings – see L. Lévy *et al.*, *Phys. Rev. Lett.* **64**, 2074 (1990). Due to the dephasing effects of electron scattering by phonons and other electrons (unaccounted for in our simple theory), the effect's observation required submicron rings and millikelvin temperatures.

[61] See, e.g., EM Sec. 5.3.

[62] Interrupting a superconducting ring with a weak link (Josephson junction), i.e. forming a SQUID, we may get a switching behavior similar to that shown with dashed arrows in Fig. 18 – see, e.g., EM Sec. 6.5.

Using the well-known expression for the 2D Laplace operator in polar coordinates,[63] we may represent the 2D stationary Schrödinger equation in the form

$$-\frac{\hbar^2}{2m}\left[\frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial}{\partial\rho}\right)+\frac{1}{\rho^2}\frac{\partial^2}{\partial\varphi^2}\right]\psi+U(\rho)\psi=E\psi\,. \tag{3.141}$$

Separating the radial and angular variables as[64]

$$\psi=\mathcal{R}(\rho)\mathcal{F}(\varphi)\,, \tag{3.142}$$

we get, after the division of all terms by ψ and their multiplication by ρ^2, the following equation:

$$-\frac{\hbar^2}{2m}\left[\frac{\rho}{\mathcal{R}}\frac{d}{d\rho}\left(\rho\frac{d\mathcal{R}}{d\rho}\right)+\frac{1}{\mathcal{F}}\frac{d^2\mathcal{F}}{d\varphi^2}\right]+\rho^2U(\rho)=\rho^2E\,. \tag{3.143}$$

The fraction $(d^2\mathcal{F}/d\varphi^2)/\mathcal{F}$ should be a constant (because all other terms of the equation may be functions only of ρ), so for the function $\mathcal{F}(\varphi)$ we get an ordinary differential equation,

$$\frac{d^2\mathcal{F}}{d\varphi^2}+\nu^2\mathcal{F}=0\,, \tag{3.144}$$

where ν^2 is the variable separation constant. The fundamental solutions of Eq. (144) are evidently $\mathcal{F}\propto\exp\{\pm i\nu\varphi\}$. Now requiring, as we did for the planar rotor, the 2π periodicity of any observable, i.e.

$$\mathcal{F}(\varphi+2\pi)=\mathcal{F}(\varphi)e^{2\pi i m}\,, \tag{3.145}$$

where m is an integer, we see that the constant ν has to be equal to m. Thus we get, for the angular factor, the same result as for the full wavefunction of the planar rotor – cf. Eq. (129):

$$\mathcal{F}_m=C_m e^{im\varphi}\,,\qquad \text{with } m=0,\pm1,\pm2,... \tag{3.146}$$

Plugging the resulting relation $(d^2\mathcal{F}/d\varphi^2)/\mathcal{F}=-m^2$ back into Eq. (143), we may rewrite it as

$$-\frac{\hbar^2}{2m}\left[\frac{1}{\rho\mathcal{R}}\frac{d}{d\rho}\left(\rho\frac{d\mathcal{R}}{d\rho}\right)-\frac{m^2}{\rho^2}\right]+U(\rho)=E\,. \tag{3.147}$$

The physical interpretation of this equation is that the full energy is a sum,

$$E=E_\rho+E_\varphi\,, \tag{3.148}$$

of the radial-motion part

$$E_\rho=-\frac{\hbar^2}{2m}\frac{1}{\rho}\frac{d}{d\rho}\left(\rho\frac{d\mathcal{R}}{d\rho}\right)+U(\rho)\,. \tag{3.149}$$

and the angular-motion part

[63] See, e.g., MA Eq. (10.3) with $\partial/\partial z = 0$.

[64] At this stage, I do not want to mark the particular solution (eigenfunction) ψ and corresponding eigenenergy E with any single index, because based on our experience in Sec. 1.7, we already may expect that in a 2D problem, the role of this index will be played by *two* integers – two quantum numbers.

$$E_\varphi = \frac{\hbar^2 m^2}{2\mathscr{m}\rho^2}. \tag{3.150}$$

Now let us recall that a similar separation exists in classical mechanics, because the total energy of either 2D or 3D particle moving in a central field may be represented as [65]

$$E = \frac{\mathscr{m}}{2}v^2 + U(\rho) = \frac{\mathscr{m}}{2}\left(\dot\rho^2 + \rho^2\dot\varphi^2\right) + U(\rho) \equiv E_\rho + E_\varphi, \tag{3.151}$$

$$\text{with } E_\rho \equiv \frac{p_\rho^2}{2\mathscr{m}} + U(\rho), \qquad \text{and } E_\varphi \equiv \frac{\mathscr{m}}{2}\rho^2\dot\varphi^2 \equiv \frac{p_\varphi^2}{2\mathscr{m}} \equiv \frac{L_z^2}{2\mathscr{m}\rho^2}. \tag{3.152}$$

The comparison of the latter relation with Eqs. (139) and (150) gives us grounds to expect that the quantization rule $L_z = m\hbar$ may be valid not only for this 2D problem but in 3D cases as well. In Sec. 5.6 below, we will see that this is indeed the case.

Returning to Eq. (147), with our 1D wave mechanics experience we may expect that at any fixed m, this ordinary, linear, second-order differential equation should have (for a motion confined to a certain final region of its argument ρ) a discrete energy spectrum described by another integer quantum number – say, n. This means that the eigenfunctions (142) and corresponding eigenenergies (148) and $\mathscr{R}(\rho)$ should be indexed by two quantum numbers, m and n. So, the variable separation is not as "clean" as it was for the rectangular potential well. Normalizing the angular function $\mathscr{f}$ to the full circle, $\Delta\varphi = 2\pi$, we may rewrite Eq. (142) as

$$\psi_{m,n} = \mathscr{R}_{m,n}(\rho)\mathscr{f}_m(\varphi) = \frac{1}{(2\pi)^{1/2}}\mathscr{R}_{m,n}(\rho)e^{im\varphi}. \tag{3.153}$$

A good (and important) example of an analytically solvable problem of this type is a 2D particle whose motion is rigidly confined to a disk of radius R, but otherwise free:

$$U(\rho) = \begin{cases} 0, & \text{for } 0 \le \rho < R, \\ +\infty, & \text{for } R < \rho. \end{cases} \tag{3.154}$$

In this case, the solutions $\mathscr{R}_{m,n}(\rho)$ of Eq. (147) are proportional to the first Bessel functions $J_m(k_n\rho)$, ща epy ашкые лштв[66] with the spectrum of possible values k_n following from the boundary condition $\mathscr{R}_{m,n}(R) = 0$. Let me leave a detailed analysis of this problem for the reader's exercise.

3.6. Spherically symmetric systems: Brute force approach

Now let us proceed to the mathematically more involved, but practically even more important case of the 3D motion in a spherically symmetric potential

$$U(\mathbf{r}) = U(r). \tag{3.155}$$

[65] See, e.g., CM Sec. 3.5.

[66] A summary of the main properties of these functions, including the most important plots and a useful table of values, may be found in EM Sec. 2.7.

Let us start, again, with solving the eigenproblem for a rotor – now a *spherical rotor*, i.e. a particle confined to move on the spherical surface of radius R. The classical rotor's position on the surface is completely described by two coordinates – say, the polar angle θ and the azimuthal angle φ. Its kinetic energy is limited to the angular motion, so for the quantum-mechanical description, in the Laplace operator expressed in spherical coordinates[67] we may keep only those parts, with fixed $r = R$. Because of this, the stationary Schrödinger equation becomes

$$-\frac{\hbar^2}{2mR^2}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right]\psi = E\psi\,. \tag{3.156}$$

(Again, we will attach indices to ψ and E in a minute.) With the natural variable separation,

$$\psi = \Theta(\theta)\mathcal{F}(\varphi)\,, \tag{3.157}$$

Eq. (156), with all terms multiplied by $\sin^2\theta/\Theta\mathcal{F}$, yields

$$-\frac{\hbar^2}{2mR^2}\left[\frac{\sin\theta}{\Theta}\frac{d}{d\theta}\left(\sin\theta\frac{d\Theta}{d\theta}\right)+\frac{1}{\mathcal{F}}\frac{d^2\mathcal{F}}{d^2\varphi}\right] = E\sin^2\theta\,. \tag{3.158}$$

Just as in Eq. (143), the fraction $(d^2\mathcal{F}/dx^2)/\mathcal{F}$ may be a function of φ only and hence has to be constant, giving Eq. (144) for it. So, with the same periodicity condition (145), the azimuthal functions are expressed by (146) again; in the normalized form,

$$\mathcal{F}_m(\varphi) = \frac{1}{(2\pi)^{1/2}}e^{im\varphi}\,. \tag{3.159}$$

With this, the fraction $(d^2\mathcal{F}/d\varphi^2)/\mathcal{F}$ in Eq. (158) equals $(-m^2)$, and after the multiplication of all terms of that equation by $\Theta/\sin^2\theta$, it is reduced to the following ordinary linear differential equation for the polar eigenfunctions $\Theta(\theta)$:

$$-\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{d\Theta}{d\theta}\right)+\frac{m^2}{\sin^2\theta}\Theta = \varepsilon\Theta, \qquad \text{with } \varepsilon \equiv E\Big/\frac{\hbar^2}{2mR^2}\,. \tag{3.160}$$

It is common to recast it into an equation for a new function $P(\xi) \equiv \Theta(\theta)$, with $\xi \equiv \cos\theta$:

$$\frac{d}{d\xi}\left[\left(1-\xi^2\right)\frac{dP}{d\xi}\right]+\left[l(l+1)-\frac{m^2}{1-\xi^2}\right]P = 0\,, \tag{3.161}$$

where a new notation for the normalized energy is introduced: $l(l+1) \equiv \varepsilon$. The motivation for such notation is that, according to the mathematical analysis of Eq. (161) with integer m,[68] it has physically suitable solutions, with P being an either odd or even function of ξ, only if l (called the *orbital quantum number*) is an integer: $l = 0, 1, 2,\ldots$, and only it is not smaller than $|m|$, i.e. if

$$-l \le m \le +l\,. \tag{3.162}$$

[67] See, e.g., MA Eq. (10.9).

[68] This analysis was first carried out by A.-M. Legendre (1752-1833). Just as a historic note: besides many original mathematical achievements, Dr. Legendre had authored a famous textbook, *Éléments de Géométrie,* which dominated teaching geometry through the 19th century.

This fact immediately gives the following spectrum of the spherical rotor's angular[69] energy E – and, as we will see later, that of any spherically symmetric system:

$$E_l = \frac{\hbar^2 l(l+1)}{2mR^2}, \tag{3.163}$$

Angular energy spectrum

so the only effect of the magnetic quantum number m here is imposing the restriction (162) on the non-negative orbital quantum number l. This means, in particular, that each energy level (163) corresponds to $(2l + 1)$ different values of m, i.e. is $(2l + 1)$–degenerate.

To understand the nature of this degeneracy, we need to explore the corresponding eigenfunctions of Eq. (161). They are naturally numbered by two integers, m and l, and are called the a*ssociated Legendre functions* P_l^m. (Note that here m is an upper index, not a power!) For the particular, simplest case $m = 0$, these functions are the so-called *Legendre polynomials* $P_l(\xi) \equiv P_l^0(\xi)$, which may be defined as the solutions of the following *Legendre equation*, resulting from Eq. (161) at $m = 0$:

$$\frac{d}{d\xi}\left[\left(1-\xi^2\right)\frac{d}{d\xi}P\right] + l(l+1)P = 0, \tag{3.164}$$

Legendre equation

and may be calculated explicitly from the following *Rodrigues formula*:[70]

$$P_l(\xi) = \frac{1}{2^l l!}\frac{d^l}{d\xi^l}(\xi^2-1)^l, \quad l = 0, 1, 2, \ldots. \tag{3.165}$$

Legendre polynomials

Using this formula, it is easy to spell out a few lowest Legendre polynomials:

$$P_0(\xi) = 1, \quad P_1(\xi) = \xi, \quad P_2(\xi) = \frac{1}{2}\left(3\xi^2 - 1\right), \quad P_3(\xi) = \frac{1}{2}\left(5\xi^3 - 3\xi\right), \ldots, \tag{3.166}$$

though such explicit expressions become bulkier and bulkier as l is increased. As these expressions (and Fig. 19) show, as the argument ξ is increased, all these functions end up at the same point, $P_l(+1) = +1$, while starting at either at the same point or at the opposite point: $P_l(-1) = (-1)^l$.

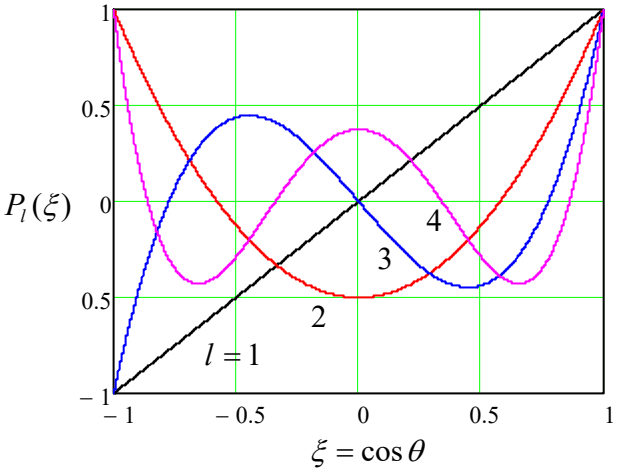


Fig. 3.19. A few lowest Legendre polynomials.

[69] This qualification is important because this E does not include the energy of radial confinement – see Sec. 2.1.

[70] This wonderful formula may be readily proved by plugging it into Eq. (164), but was not so easy to discover! This was done (independently) by B. O. Rodrigues in 1816, J. Ivory in 1824, and C. Jacobi in 1827.

On the way between these two endpoints, the l^{th} polynomial crosses the horizontal axis exactly l times, i.e. Eq. (164) has exactly l roots. (In this behavior, we may readily recognize the "standing wave" pattern typical for all 1D eigenproblems – cf. Figs. 1.8 and 2.35, as well as the discussion of the Sturm oscillation theorem at the end of Sec. 2.9.) It is also easy to use the Rodrigues formula (165) and the integration by parts to show that on the segment $-1 \le \xi \le +1$, th polynomials form a full orthogonal set of functions, with the following normalization rule:

$$\int_{-1}^{+1} P_l(\xi) P_{l'}(\xi) d\xi = \frac{2}{2l+1} \delta_{ll'}. \tag{3.167}$$

For $m > 0$, the associated Legendre functions (now *not* necessarily polynomials!), may be expressed via the Legendre polynomials (165) using the following formula:[71]

Associated Legendre functions

$$P_l^m(\xi) = (-1)^m (1-\xi^2)^{m/2} \frac{d^m}{d\xi^m} P_l(\xi), \tag{3.168}$$

while the functions with a negative magnetic quantum number may be found as

$$P_l^{-m}(\xi) = (-1)^m \frac{(l-m)!}{(l+m)!} P_l^m(\xi), \qquad \text{for } m > 0. \tag{3.169}$$

On the segment $-1 \le \xi \le +1$, the associated Legendre functions with a fixed index m form a full orthogonal set, with the normalization relation

$$\int_{-1}^{+1} P_l^m(\xi) P_{l'}^m(\xi) d\xi = \frac{2}{2l+1} \frac{(l+m)!}{(l-m)!} \delta_{ll'}, \tag{3.170}$$

which is evidently a generalization of Eq. (167) to arbitrary m.

Since the difference between the angles θ and φ is, to a large extent, artificial (due to an arbitrary direction of the polar axis), physicists prefer to use not the functions $\Theta(\theta) \propto P_l^m(\cos\theta)$ and $\mathcal{T}_m(\varphi) \propto e^{im\varphi}$ separately, but normalized products of the type (157), which are called the s*pherical harmonics*:

Spherical harmonics

$$Y_l^m(\theta,\varphi) \equiv \left[\frac{(2l+1)}{4\pi} \frac{(l-m)!}{(l+m)!}\right]^{1/2} P_l^m(\cos\theta) e^{im\varphi}. \tag{3.171}$$

The specific front factor in Eq. (171) is chosen in a way to simplify the following two expressions: the relation of the spherical harmonics with opposite signs of the magnetic quantum number:

$$Y_l^{-m}(\theta,\varphi) = (-1)^m \left[Y_l^m(\theta,\varphi)\right]^*, \tag{3.172}$$

and the following normalization relation:

$$\oint_{4\pi} Y_l^m(\theta,\varphi) \left[Y_{l'}^{m'}(\theta,\varphi)\right]^* d\Omega = \delta_{ll'} \delta_{mm'}, \tag{3.173}$$

[71] Note that some texts use different choices for the front factor (called the *Condon-Shortley phase*) in the functions P_l^m, which do not affect the final results for the spherical harmonics Y_l^m.

with the integration over the whole solid angle. The last formula shows that on a spherical surface, the spherical harmonics form an orthonormal set of functions. This set is also full, so any function defined on the surface may be uniquely represented as a linear combination of Y_l^m.

Despite the somewhat intimidating character of the formulas given above, they yield quite simple expressions for the lowest spherical harmonics, which are most important for applications:

$$l = 0: \quad Y_0^0 = (1/4\pi)^{1/2}, \tag{3.174}$$

$$l = 1: \quad \begin{cases} Y_1^{-1} = (3/8\pi)^{1/2} \sin\theta\, e^{-i\varphi}, \\ Y_1^0 = (3/4\pi)^{1/2} \cos\theta, \\ Y_1^1 = -(3/8\pi)^{1/2} \sin\theta\, e^{i\varphi}, \end{cases} \tag{3.175}$$

$$l = 2: \quad \begin{cases} Y_2^{-2} = (15/32\pi)^{1/2} \sin^2\theta\, e^{-2i\varphi}, \\ Y_2^{-1} = (15/8\pi)^{1/2} \sin\theta\cos\theta\, e^{-i\varphi}, \\ Y_2^0 = (3/16\pi)^{1/2} (3\cos^2\theta - 1), \\ Y_2^1 = -(15/8\pi)^{1/2} \sin\theta\cos\theta\, e^{i\varphi}, \\ Y_2^2 = (15/32\pi)^{1/2} \sin^2\theta\, e^{2i\varphi}. \end{cases} \tag{3.176}$$

It is important to understand the general structure and symmetry of these functions, and in such matters, pictures are invaluable. Since the spherical harmonics with $m \neq 0$ are complex, the most popular way of their graphical representation is to normalize their real and imaginary parts as

$$Y_{lm} \equiv \sqrt{2}(-1)^m \times \begin{cases} \mathrm{Im}\left(Y_l^{|m|}\right) \propto \sin m\varphi, & \text{for } m < 0, \\ \mathrm{Re}\left(Y_l^{|m|}\right) \propto \cos m\varphi, & \text{for } m > 0, \end{cases} \tag{3.177}$$

(for $m = 0$, $Y_{l0} \equiv Y_l^0$), and then plot the magnitude of these real functions[72] in the spherical coordinates as the distance from the origin, while using two colors to show their sign – see Fig. 20.

Let us start from the simplest case $l = 0$. According to Eq. (162), for this lowest orbital quantum number, there may be only one magnetic quantum number: $m = 0$, and according to Eq. (174), the spherical harmonic corresponding to that state is just a constant. Thus the wavefunction of this so-called *s state*[73] is uniformly distributed over the sphere. Since this function has no gradient in any angular direction, it is only natural that the angular kinetic energy (163) of a particle in this state equals zero.

According to the same Eq. (162), for $l = 1$, there are 3 different *p states*, with $m = -1$, $m = 0$, and $m = +1$ – see Eq. (175). As the second row of Fig. 20 shows, these states are essentially identical in structure and are just differently oriented in space, thus readily explaining the 3-fold degeneracy of the kinetic energy (163).

[72] Such real functions Y_{lm}, which also form a full orthonormal set, and are frequently called the *real* (or "tesseral") *spherical harmonics*, are more convenient than the complex harmonics Y_l^m for several applications, especially when the variables of interest are real by definition.

[73] The letter names for the states with various values of l stem from the history of optical spectroscopy – for example, the letter "*s*" (used for states with $l = 0$) originally denoted the "sharp" optical line series, etc. The sequence of the letters is as follows: *s, p, d, f, g*, and then continuing in alphabetical order.

Such a simple explanation, however, is not valid for the 5 different *d states* ($l = 2$), shown in the third row of Fig. 20, as well as the states with higher l: despite their equal energies, they differ not only by their spatial orientation but their structure as well. All states with $m = 0$ have a nonzero gradient only in the θ direction. On the contrary, the states with the ultimate values of m ($\pm l$), change only monotonically (as $\sin^l\theta$) in the polar direction, while oscillating in the azimuthal direction. The states with intermediate values of m provide a gradual transition between these two extremes, oscillating in both directions, stronger and stronger in the azimuthal direction as $|m|$ is increased. Still, the magnetic quantum number, surprisingly, does not affect the angular energy for any l.

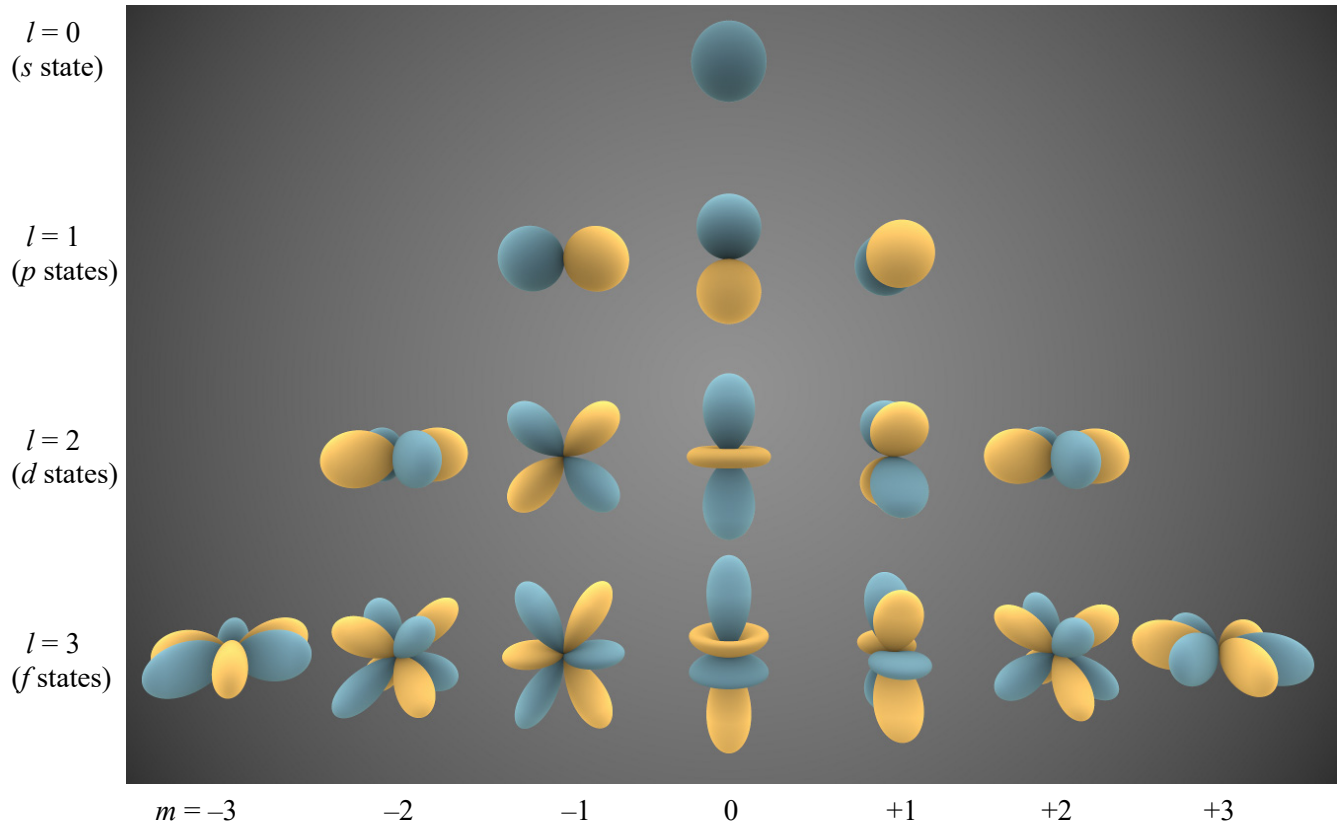


Fig. 3.20. Plots of several lowest real spherical harmonics Y_{lm}. (Adapted from https://en.wikipedia.org/wiki/Spherical_harmonics under the CC BY-SA 3.0 license.)

Another counter-intuitive feature of the spherical harmonics follows from the comparison of Eq. (163) with the second of Eqs. (152), which, in classical mechanics, is valid for the total angular momentum as well. They coincide only if we interpret

$$L^2 \equiv \hbar^2 l(l+1), \tag{3.178}$$

as the value of the total $L^2 \equiv |\mathbf{L}^2|$, including the θ and φ components of the vector $\mathbf{L}$, in the state with eigenfunction Y_l^m. On the other hand, the structure (159) of the azimuthal component $\mathcal{F}(\varphi)$ of the wavefunction is exactly the same as in 2D axially symmetric problems, implying that Eq. (139) still gives correct values $L_z = m\hbar$ for the z-component of the angular momentum. This fact invites a question: why for any state with $l > 0$, $(L_z)^2 = m^2\hbar^2 \le l^2\hbar^2$ is always less than $L^2 = l(l+1)\hbar^2$? In other words, what prevents the angular momentum vector to be fully aligned with the z-axis?

Besides the difficulty of answering this question using the above formulas, this analysis (though mathematically complete), is as intellectually unsatisfactory as the harmonic oscillator analysis in Sec. 2.9. In particular, it does not explain the meaning of the extremely simple relations for the eigenvalues of the energy and the angular momentum, coexisting with rather complicated eigenfunctions.

We will obtain natural answers to all these questions and concerns in Sec. 5.6 below, and now let us proceed to the extension of our wave-mechanical analysis to the 3D motion in an *arbitrary* spherically symmetric potential (155). In this case, we have to use the full form of the Laplace operator in spherical coordinates.[74] The variable separation procedure is an evident generalization of what we have done above, with the particular solutions of the type

$$\psi = \mathcal{R}(r)\Theta(\theta)\mathcal{F}(\varphi), \tag{3.179}$$

whose substitution into the stationary Schrödinger equation yields

$$-\frac{\hbar^2}{2mr^2}\left[\frac{1}{\mathcal{R}}\frac{d}{dr}\left(r^2\frac{d\mathcal{R}}{dr}\right)+\frac{1}{\Theta}\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{d\Theta}{d\theta}\right)+\frac{1}{\sin^2\theta}\frac{1}{\mathcal{F}}\frac{d^2\mathcal{F}}{d\varphi^2}\right]+U(r)=E. \tag{3.180}$$

It is evident that the angular part of the left-hand side (the two last terms in the square brackets) separates from the radial part, and that for the former part, we get Eq. (156) again, with the only change, $R \to r$. This change does not affect the fact that the eigenfunctions of that equation are still the spherical harmonics (171), which obey Eq. (164). As a result, Eq. (180) gives the following equation for the radial function $\mathcal{R}(r)$:

$$-\frac{\hbar^2}{2mr^2}\left[\frac{1}{\mathcal{R}}\frac{d}{dr}\left(r^2\frac{d\mathcal{R}}{dr}\right)-l(l+1)\right]+U(r)=E. \tag{3.181a}$$

Note that no information about the magnetic quantum number m has crept into this radial equation (besides setting the limitation (162) for the possible values of l) so it includes only the orbital quantum number l. The equation may be also rewritten in a form similar to Eq. (148):

$$E = E_r + E_{\theta,\varphi}, \quad \text{with } E_r = -\frac{\hbar^2}{2mr^2}\frac{1}{\mathcal{R}}\frac{d}{dr}\left(r^2\frac{d\mathcal{R}}{dr}\right)+U(r) \text{ and } E_{\theta,\varphi}=\frac{\hbar^2 l(l+1)}{2mr^2}, \tag{3.181b}$$

expressing the same separability of the particle's energy in the central field into the radial and angular components as in classical mechanics – cf. Eqs. (151)-(152), and also Eq. (163). In particular, since the expectation value of the latter component cannot be negative at $l \geq 0$, this means that the ground state of any spherically symmetric system is an *s*-state, with $l = 0$ and hence $m = 0$.

Let us explore the radial equation for the simple case of a 3D particle free to move inside the sphere of radius R – say, confined there by the potential[75]

$$U = \begin{cases} 0, & \text{for } 0 \leq r < R, \\ +\infty, & \text{for } R < r. \end{cases} \tag{3.182}$$

In this case, Eq. (181a) is reduced to

[74] Again, see MA Eq. (10.9).

[75] This problem, besides giving a simple example of the quantization in spherically symmetric systems, is also an important precursor for the discussion of scattering by spherically symmetric potentials in Sec. 8.

$$-\frac{\hbar^2}{2mr^2}\left[\frac{1}{\mathcal{R}}\frac{d}{dr}\left(r^2\frac{d\mathcal{R}}{dr}\right)-l(l+1)\right]=E. \tag{3.183}$$

Multiplying both parts of this equality by $r^2\mathcal{R}$, and introducing the dimensionless argument $\xi \equiv kr$, where k^2 is defined by the usual relation $\hbar^2k^2/2m = E$, we obtain the canonical form of this equation,

$$\xi^2\frac{d^2\mathcal{R}}{d\xi^2}+2\xi\frac{d\mathcal{R}}{d\xi}+\left[\xi^2-l(l+1)\right]\mathcal{R}=0, \tag{3.184}$$

which is satisfied with the so-called *spherical Bessel functions* of the first and second kind, $j_l(\xi)$ and $y_l(\xi)$.[76] These functions are directly related to the Bessel functions of semi-integer order,[77]

$$j_l(\xi)=\left(\frac{\pi}{2\xi}\right)^{1/2}J_{l+\frac{1}{2}}(\xi), \qquad y_l(\xi)=\left(\frac{\pi}{2\xi}\right)^{1/2}Y_{l+\frac{1}{2}}(\xi), \tag{3.185}$$

but are actually much simpler than even the "usual" Bessel functions, such as $J_n(\xi)$ and $Y_n(\xi)$ of an integer order n, because the former ones may be directly expressed via elementary functions:

$$\begin{aligned} &j_0(\xi)=\frac{\sin\xi}{\xi}, && j_1(\xi)=\frac{\sin\xi}{\xi^2}-\frac{\cos\xi}{\xi}, && j_2(\xi)=\left(\frac{3}{\xi^3}-\frac{1}{\xi}\right)\sin\xi-\frac{3}{\xi^2}\cos\xi,...,\\ &y_0(\xi)=-\frac{\cos\xi}{\xi}, && y_1(\xi)=-\frac{\cos\xi}{\xi^2}-\frac{\sin\xi}{\xi}, && y_2(\xi)=-\left(\frac{3}{\xi^3}-\frac{1}{\xi}\right)\cos\xi-\frac{3}{\xi^2}\sin\xi,... \end{aligned} \tag{3.186}$$

A few lowest-order spherical Bessel functions are plotted in Fig. 21.

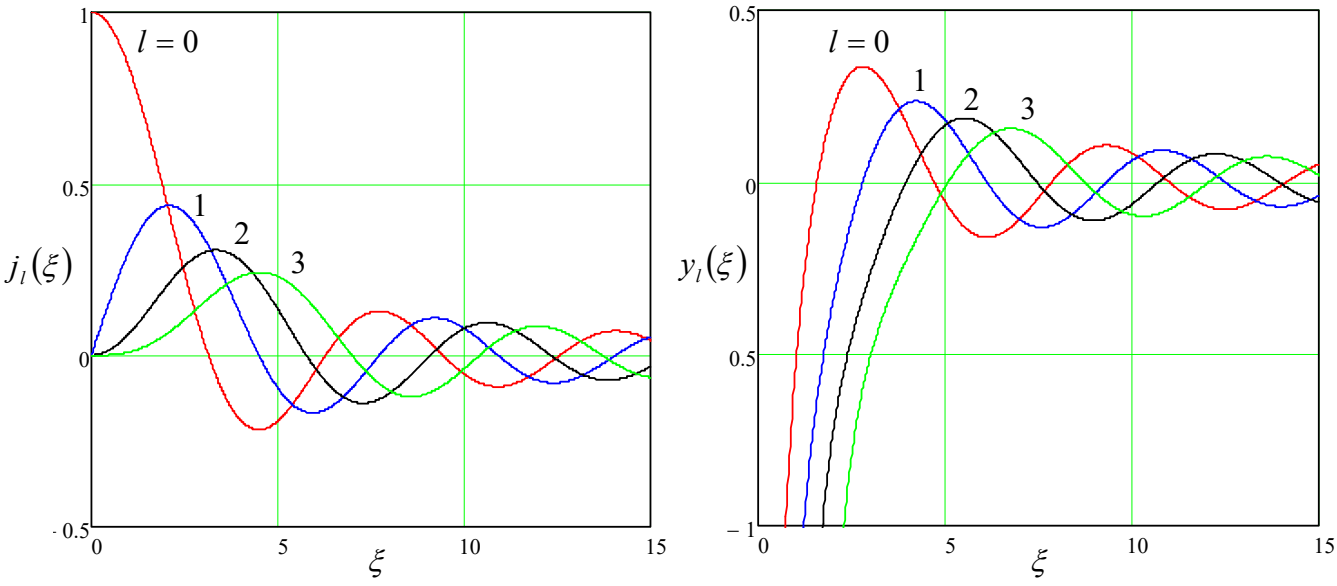


Fig. 3.21. Several lowest-order spherical Bessel functions.

[76] Alternatively, $y_l(\xi)$ are called "spherical Weber functions" or "spherical Neumann functions".

[77] Note that the Bessel functions $J_\nu(\xi)$ and $Y_\nu(\xi)$ of *any* order ν obey the universal recurrence relations and asymptotic formulas (discussed, e.g., in EM Sec. 2.7), so many properties of the functions $j_l(\xi)$ and $y_l(\xi)$ may be readily derived from these relations and Eqs. (185).

As these formulas and plots show, the functions $y_l(\xi)$ are diverging at $\xi \to 0$, and thus cannot be used in the solution of our current problem (182), so we have to take

$$\mathcal{R}_l(r) = \text{const} \times j_l(kr). \tag{3.187}$$

Still, even for these functions, with the sole exception of the simplest function $j_0(\xi)$, the characteristic equation $j_l(kR) = 0$, resulting from the boundary condition $\mathcal{R}(R) = 0$, can be solved only numerically. However, the roots $\xi_{l,n}$ of the equation $j_l(\xi) = 0$, where the integer n (= 1, 2, 3,…) is the root's number, are tabulated in virtually any math handbook, and we may express the eigenvalues we are interested in,

$$k_{l,n} = \frac{\xi_{l,n}}{R}, \qquad E_{l,n} = \frac{\hbar^2 k_{l,n}^2}{2m} \equiv \frac{\hbar^2 \xi_{l,n}^2}{2mR^2}, \tag{3.188}$$

via these tabulated numbers. The table on the right lists several smallest roots, and the corresponding eigenenergies (normalized to their natural unit $E_0 \equiv \hbar^2/2mR^2$), in the order of their growth. It shows a very interesting effect: going up the energy spectrum, the first energies grow due to unit increments of the orbital quantum number l and the corresponding increases of the first roots of the functions $j_l(\xi)$, at the same (lowest) radial quantum number $n = 1$. Then, suddenly, the second root of $j_0(\xi)$, accompanied by a jump to $n = 2$, cuts into this orderly sequence, just to be followed by the first root of $j_3(\xi)$, returning to the initial sequence with $n = 1$. With the further growth of energy, the sequences of l and n become even more entangled.

l	n	$\xi_{l,n}$	$E_{l,n}/E_0 = (\xi_{l,n})^2$
0	1	$\pi \approx 3.1415$	$\pi^2 \approx 9.87$
1	1	4.493	20.19
2	1	5.763	33.21
0	2	$2\pi \approx 6.283$	$4\pi^2 \approx 39.48$
3	1	6.988	48.83

To complete the discussion of our current problem (182), note again that the energy levels listed in the table above are $(2l+1)$-degenerate because each of them corresponds to $(2l + 1)$ different eigenfunctions, each with a specific value of the magnetic quantum number m:

$$\psi_{n,l,m} = C_{l,n} j_l\left(\frac{\xi_{l,n} r}{R}\right) Y_l^m(\theta,\varphi), \qquad \text{with } -l \le m \le +l. \tag{3.189}$$

3.7 Atoms

Now we are ready to discuss atoms, starting from the simplest, exactly solvable *Bohr atom problem*, i.e. that of a single particle's motion in the so-called *attractive Coulomb potential*[78]

$$U(r) = -\frac{C}{r}, \quad \text{with } C > 0. \tag{3.190}$$

Attractive Coulomb potential

The natural scales of E and r in this problem are commonly defined by the requirement of equality of the kinetic and potential energy magnitude scales (dropping all numerical coefficients):

[78] Historically, the solution of this problem in 1928, which reproduced the main results (1.12)-(1.13) of the "old" quantum theory developed by N. Bohr in 1912, but without its phenomenological assumptions, was the decisive step toward the general acceptance of Schrödinger's wave mechanics.

$$E_0 \equiv \frac{\hbar^2}{mr_0^2} \equiv \frac{C}{r_0}, \tag{3.191}$$

similar to its particular case (1.13b). Solving this system of two equations, we get[79]

$$E_0 \equiv \frac{\hbar^2}{mr_0^2} \equiv m\left(\frac{C}{\hbar}\right)^2, \qquad \text{and} \quad r_0 \equiv \frac{\hbar^2}{mC}. \tag{3.192}$$

In the normalized units $\varepsilon \equiv E/E_0$ and $\xi \equiv r/r_0$, Eq. (181) for our case (190) looks relatively simple,

$$\frac{d^2\mathcal{R}}{d\xi^2} + \frac{2}{\xi}\frac{d\mathcal{R}}{d\xi} - l(l+1)\mathcal{R} + 2\left(\varepsilon + \frac{1}{\xi}\right)\mathcal{R} = 0, \tag{3.193}$$

but unfortunately, its eigenfunctions may be called elementary only in the most generous meaning of the word. With the normalization

$$\int_0^\infty \mathcal{R}_{n,l}\mathcal{R}_{n',l} r^2 dr = \delta_{nn'}, \tag{3.194}$$

these (mutually orthogonal) functions may be represented as

Bohr atom: radial functions

$$\mathcal{R}_{n,l}(r) = -\left\{\left(\frac{2}{nr_0}\right)^3 \frac{(n-l-1)!}{2n[(n+l)!]^3}\right\}^{1/2} \exp\left\{-\frac{r}{nr_0}\right\}\left(\frac{2r}{nr_0}\right)^l L_{n-l-1}^{2l+1}\left(\frac{2r}{nr_0}\right). \tag{3.195}$$

Here $L_p^q(\xi)$ are the so-called *associated Laguerre* polynomials, which may be calculated as

Associated Laguerre polynomials

$$L_p^q(\xi) = (-1)^q \frac{d^q}{d\xi^q} L_{p+q}(\xi). \tag{3.196}$$

from the *simple Laguerre polynomials* $L_p(\xi) \equiv L_p^0(\xi)$.[80] In turn, the easiest way to obtain $L_p(\xi)$ is to use the following *Rodrigues formula*:[81]

Rodrigues formula for Laguerre polynomials

$$L_p(\xi) = e^\xi \frac{d^p}{d\xi^p}\left(\xi^p e^{-\xi}\right). \tag{3.197}$$

Note that in contrast with the associated Legendre *functions* P_l^m, participating in the spherical harmonics, all L_p^q are just *polynomials*, and those with small indices p and q are indeed quite simple:

[79] For the most important case of the hydrogen atom, with $C = e^2/4\pi\varepsilon_0$, these scales are reduced, respectively, to the Bohr radius r_B (1.10) and the Hartree energy E_H (1.13a). Note also that according to Eq. (192), for the so-called *hydrogen-like atom* (actually, a positive ion) with $C = Z(e^2/4\pi\varepsilon_0)$, these two key parameters are rescaled as $r_0 = r_B/Z$ and $E_0 = Z^2E_H$.

[80] In Eqs. (196)-(197), p and q are non-negative integers, with no relation whatsoever to the particle's momentum or electric charge. Sorry for this notation, but it is absolutely common, and can hardly result in any confusion.

[81] Named after the same B. O. Rodrigues, and belonging to the same class as his other famous result, Eq. (165) for the Legendre polynomials.

$$\begin{aligned} &L_0^0(\xi)=1, \qquad L_1^0(\xi)=-\xi+1, \qquad L_2^0(\xi)=\xi^2-4\xi+2, \\ &L_0^1(\xi)=1, \qquad L_1^1(\xi)=-2\xi+4, \qquad L_2^1(\xi)=3\xi^2-18\xi+18, \\ &L_0^2(\xi)=2, \qquad L_1^2(\xi)=-6\xi+18, \qquad L_2^2(\xi)=12\xi^2-96\xi+144, \text{ etc.} \end{aligned} \tag{3.198}$$

Returning to Eq. (195), we see that the natural quantization of the radial equation (193) has brought us a new integer quantum number n. To understand its range, we should notice that according to Eq. (197), the highest power of terms in the polynomial L_{p+q} is $(p + q)$, and hence, according to Eq. (196), that of L_p^q is p, so the highest power in the polynomial participating in Eq. (195) is $(n - l - 1)$. Since the power cannot be negative to avoid the unphysical divergence of wavefunctions at $r \to 0$, the radial quantum number n has to obey the restriction $n \geq l + 1$. Since l, as we already know, may take the values $l = 0, 1, 2,\ldots$, we may conclude n may only take the following values:

$$n = 1, 2, 3,... \tag{3.199}$$

What makes this relation very important is the following, most surprising result: the eigenenergies corresponding to the wavefunctions (179), which are indexed with *three* quantum numbers:

$$\psi_{n,l.m} = \mathcal{R}_{n,l}(r)Y_l^m(\theta,\varphi), \tag{3.200}$$

depend only on *one* of them, n:

$$\varepsilon = \varepsilon_n = -\frac{1}{2n^2}, \qquad \text{i.e. } E_n = -\frac{E_0}{2n^2} = -\frac{1}{2n^2}m\left(\frac{C}{\hbar}\right)^2. \tag{3.201}$$

i.e. agree with Bohr's formula (1.12). Because of this reason, n is usually called the *principal quantum number*, and the above relation between it and the "more subordinate" orbital quantum number l is rewritten as

$$l \leq n-1. \tag{3.202}$$

Together with the inequality (162), this gives us the following, very important hierarchy of the three quantum numbers involved in the Bohr atom problem:

Bohr atom: quantum numbers

$$\boxed{1 \leq n \leq \infty \quad \Rightarrow \quad 0 \leq l \leq n-1 \quad \Rightarrow \quad -l \leq m \leq +l.} \tag{3.203}$$

Taking into account the $(2l + 1)$-degeneracy related to m, and using the well-known formula for the arithmetic progression,[82] we see that the n^{th} energy level (201) has the following *orbital degeneracy*:

$$g = \sum_{l=0}^{n-1}(2l+1) \equiv 2\sum_{l=0}^{n-1} l + \sum_{l=0}^{n-1} 1 = 2\frac{n(n-1)}{2} + n \equiv n^2. \tag{3.204}$$

Due to its importance for atoms, let us spell out the hierarchy (203) of a few lowest-energy states, using the traditional state notation in that the value of n is followed by the letter that denotes the value of l:

$$n = 1: \quad l = 0 \quad (\text{one } 1s \text{ state}) \quad m = 0. \tag{3.205}$$

$$\begin{aligned} n = 2: \quad & l = 0 \quad (\text{one } 2s \text{ state}) \quad m = 0, \\ & l = 1 \quad (\text{three } 2p \text{ states}) \quad m = 0, \pm 1. \end{aligned} \tag{3.206}$$

[82] See, e.g., MA Eq. (2.5a).

$$\begin{aligned} n=3:\quad & l=0 \quad (\text{one } 3s \text{ state}) && m=0, \\ & l=1 \quad (\text{three } 3p \text{ states}) && m=0,\pm 1, \\ & l=2 \quad (\text{five } 3d \text{ states}) && m=0,\pm 1,\pm 2. \end{aligned} \tag{3.207}$$

Figure 22 shows plots of the radial functions (195) of the listed states. The most important of them is of course the ground ($1s$) state with $n = 1$ and hence $E = -E_0/2$. According to Eqs. (195) and (198), its radial function is just a simple decaying exponent

Ground state: radial function

$$\mathcal{R}_{1,0}(r) = \frac{2}{r_0^{3/2}} e^{-r/r_0}, \tag{3.208}$$

while its angular distribution is uniform – see Eq. (174). The gap between the ground state energy $E_g = -E_0/2$ and the energy $E = -E_0/8$ of the lowest excited states (with $n = 2$) in a hydrogen atom (in which $E_0 = E_H \approx 27.2$ eV) is as large as ~ 10 eV, so their thermal excitation requires temperatures as high as ~10^5 K, and the overwhelming part of all hydrogen atoms in the visible Universe are in their ground state. Since atomic hydrogen makes up about 75% of the "normal" matter,[83] we are very fortunate that such simple formulas as Eqs. (174) and (208) describe the atomic states prevalent in Mother Nature!

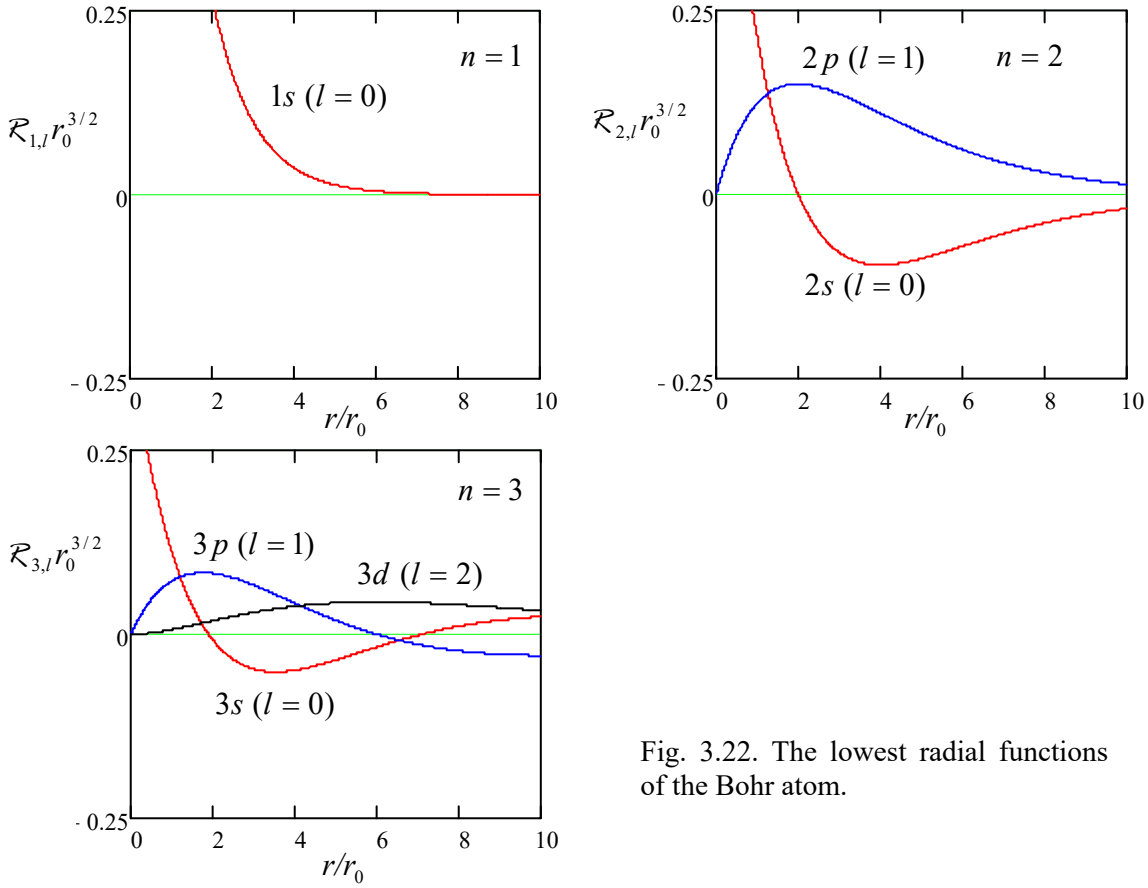


Fig. 3.22. The lowest radial functions of the Bohr atom.

According to Eqs. (195) and (198), the radial functions of the lowest excited states, $2s$ (with $n = 2$ and $l = 0$), and $2p$ (with $n = 2$ and $l = 1$) are also not too complicated:

[83] Excluding the so-far hypothetical *dark matter* and *dark energy*.

$$\mathcal{R}_{2,0}(r) = \frac{1}{(2r_0)^{3/2}}\left(2 - \frac{r}{r_0}\right)e^{-r/2r_0}, \qquad \mathcal{R}_{2,1}(r) = \frac{1}{(2r_0)^{3/2}}\frac{r}{3^{1/2}r_0}e^{-r/2r_0}, \tag{3.209}$$

with the former of these states (2*s*) having a uniform angular distribution, and the three latter (2*p*) states, with different $m = 0, \pm 1$, having simple angular distributions, which differ only by their spatial orientation – see Eq. (175) and the second row of Fig. 20. The most important trend here, clearly visible from the comparison of the two top panels of Fig. 22 as well, is a larger radius of the decay exponent in the radial functions ($2r_0$ for $n = 2$ instead of r_0 for $n = 1$), and hence a larger radial extension of the states. This trend is confirmed by the following general formula:[84]

$$\langle r \rangle_{n,l} = \frac{r_0}{2}\left[3n^2 - l(l+1)\right]. \tag{3.210}$$

The second important trend is that at a fixed n, the orbital quantum number l determines how fast the wavefunction changes with r near the origin, and how much it oscillates in the radial direction at larger values of r. For example, the 2*s* eigenfunction $\mathcal{R}_{2,0}(r)$ is different from zero at $r = 0$, and "makes one wiggle" (has one root) in the radial direction, while the eigenfunctions 2*p* equal zero at $r = 0$ but do not cross the horizontal axis after that. Instead, those wavefunctions oscillate as the functions of an angle – see the second row of Fig. 20. The same trend is clearly visible for $n = 3$ (see the bottom panel of Fig. 22), and continues for the higher values of n.

The states with $l = l_{\max} \equiv n - 1$ may be viewed as crude analogs of the circular motion of a particle in a plane whose orientation defines the quantum number m. On the other hand, the best classical image of the *s*-state ($l = 0$) is a purely radial, spherically symmetric motion of the particle to and from the attracting center. (The latter image is especially imperfect because the motion needs to happen simultaneously in all radial directions.) The classical language becomes reasonable only for the highly degenerate *Rydberg states*, with $n >> 1$, whose linear superpositions may be used to compose wave packets closely following the classical (circular or elliptic) trajectories of the particle – just as was discussed in Sec. 2.2 for the free 1D motion.

Besides Eq. (210), mathematics gives us several other simple relations for the radial functions $\mathcal{R}_{n,l}$ (and, since the spherical harmonics are normalized to 1, for the eigenfunctions as the whole), including those that we will use later in the course:[85]

$$\left\langle \frac{1}{r} \right\rangle_{n,l} = \frac{1}{n^2 r_0}, \quad \left\langle \frac{1}{r^2} \right\rangle_{n,l} = \frac{1}{n^3\left(l+\tfrac{1}{2}\right)r_0^2}, \qquad \left\langle \frac{1}{r^3} \right\rangle_{n,l} = \frac{1}{n^3 l\left(l+\tfrac{1}{2}\right)(l+1)r_0^3}. \tag{3.211}$$

In particular, the first of these formulas means that for *any* eigenfunction $\psi_{n,l,m}$, with all its complicated radial and angular dependencies, there is a simple relation between the potential and full energies:

$$\langle U \rangle_{n,l} = -C\left\langle \frac{1}{r} \right\rangle_{n,l} = -\frac{C}{n^2 r_0} = -\frac{E_0}{n^2} = 2E_n, \tag{3.212}$$

so the average kinetic energy of the particle, $\langle T \rangle_{n,l} = E_n - \langle U \rangle_{n,l}$, is equal to $E_n - 2E_n = |E_n| > 0$.

[84] Note that even at the largest value of l, equal to $(n-1)$, the second term $l(l+1)$ in the square brackets of Eq. (210) is equal to $(n^2 - n)$, and hence cannot over-compensate the first term $3n^2$.

[85] The first of these relations may be proved using the Hellmann-Feynman theorem (see Sec. 1.8); this proof will be offered for the reader's exercise after a more general form of this theorem has been proved in Chapter 6.

As in the several previous cases we have met, simple results (201), (210)-(212) are in sharp contrast with the rather complicated expressions for the corresponding eigenfunctions. Historically, this contrast gave an additional motivation for the development of more general approaches to quantum mechanics, that would replace or at least complement the brute-force (wave-mechanics) analysis. A discussion of such an approach will be the main topic of the next chapter.

Rather strikingly, the above classification of the quantum numbers, with a few steals from the later chapters of this course, allows a semi-quantitative explanation of the whole system of chemical elements. The "only" two additions we need are the following facts:

(i) due to their unavoidable interaction with relatively low-temperature environments, atoms tend to relax into their lowest-energy state, and

(ii) due to the Pauli principle (valid for electrons as the Fermi particles), each *orbital eigenstate* discussed above may house two electrons with opposite spins.

Of course, atomic electrons do interact, so their quantitative description requires quantum mechanics of multiparticle systems, which is rather complex. (Its main concepts will be discussed in Chapter 8.) However, the lion's share of this interaction is reduced to simple electrostatic *screening*, i.e. a partial compensation of the electric charge of the atomic nucleus, as felt by a particular electron, by other electrons of the atom. This screening changes quantitative results (such as the energy scale E_0) much; however, the quantum number hierarchy and hence the state classification, are not affected.

The system of atoms is most often represented as the famous *periodic table of chemical elements*,[86] whose simple version is shown in Fig. 23. (The table in Fig. 24 presents a sequential list of the elements with their electron configurations, following the convention already used in Eqs. (205)-(207), with the additional upper index showing the number of electrons with the indicated values of quantum numbers n and l.) The number in each table's cell, and in the first column of the list, is the so-called *atomic number* Z, which physically is the number of protons in the particular atomic nucleus, and hence the number of electrons in an electrically neutral atom.

The simplest atom, with $Z = 1$, is hydrogen (chemical symbol H) – the only atom for which the theory discussed above is quantitatively correct.[87] According to Eq. (191), the ground state of its only electron corresponds to the quantum number values $n = 1$, $l = 0$, and $m = 0$ – see Eq. (205). In most versions of the periodic table, the cell of H is placed in the top left corner.

In the next atom, helium (symbol He, $Z = 2$), the same orbital quantum state (1s) houses two electrons. As will be discussed in detail in Chapter 8, electrons of the same atom are actually *indistinguishable*, so their quantum states are not independent and may be *entangled*. These factors are important for several properties of helium atoms (and heavier elements as well); however, a bit counter-intuitively, for atom classification purposes, they are not crucial, and we may think about the two electrons of a helium atom just having "opposite spins". Due to the twice higher electric charge of the nucleus of the helium atom, i.e. the twice higher value of the constant C in Eq. (190), resulting in a four-fold increase of the constant E_0 given by Eq. (192), the binding energy of each electron is crudely four times higher than that of the hydrogen atom – though the electron interaction decreases it by about 25%

[86] Also called the *Mendeleev table*, after D. I. Mendeleev who put forward the concept of the quasi-periodicity of chemical element properties as functions of Z phenomenologically in 1869. (The explanation of this periodicity had to wait for 60 more years until the advent of quantum mechanics in the late 1920s.)

[87] Besides very small *fine-structure* and *hyperfine-splitting* corrections – to be discussed, respectively, in Chapters 6 and 8.

– see Sec. 8.2 below. This is why taking one electron away (i.e. the *positive ionization* of a helium atom) requires relatively high energy, ~24.6 eV, which is not available in the usual chemical reactions. On the other hand, a neutral helium atom cannot bind one more electron (i.e. form a *negative ion*) either. As a result, helium, and all other elements with fully completed electron *shells* (the term meaning the sets of states with eigenenergies well separated from higher energy levels) is a chemically inert *noble gas*, thus starting the whole right-most column of the periodic table, allocated for such elements.

1 H	Property legend:																2 He
3 Li	4 Be	alkali metals				transition metals				metalloids		5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	alkali-earth metals				nonmetals				halogens		13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	rare-earth metals	21 Sc	22 Ti	23 V	other metals	25 Mn	26 Fe	27 Co	noble gases	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr		39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	57-71	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn	
87 Fr	88 Ra	89-102	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112 Cn	113 Nh	114 Fl	115 Mc	116 Lv	117 Ts	118 Og	

Lanthanides:	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb	71 Lu
Actinides:	89 Ac	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 Lr

Fig. 3. 23. The periodic table of elements, showing their atomic numbers and chemical symbols, as well as the color-coded basic physical/chemical properties at the so-called *ambient* (meaning usual laboratory) conditions.

The situation changes rather dramatically as we move to the next element, lithium (Li), with $Z = 3$ electrons. Two of them are still accommodated by the inner shell with $n = 1$ (listed in Fig. 24 as the *helium shell* [He]), but the third one has to reside in the next shell with $n = 2$, $l = 0$, and $m = 0$, i.e. in the 2*s* state. According to Eq. (201), the binding energy of this electron is much lower, especially if we take into account that according to Eqs. (210)-(211), the 1*s* electrons of the [He] shell are much closer to the nucleus and almost completely compensate for two-thirds of its electric charge +3*e*. As a result, the 2*s*-state electron is approximately but reasonably described by Eq. (201) with $Z = 1$ and $n = 2$, giving its binding energy close to 3.4 eV (actually, ~5.39 eV), so a lithium atom can give out that electron rather easily – to either an atom/ion of another element to form a chemical compound or to the common conduction band of the solid-state lithium; as a result, at the ambient conditions, this is a typical *alkali metal*. The similarity of chemical properties of lithium and hydrogen, with the *chemical valence* of one,[88] places Li as the starting element of the second *period* (row), with the first period limited to only H and He – see Fig. 23.

[88] Chemical valence (or "valency") is a not very precise term describing the number of the atom's electrons involved in chemical bonding. For the same atom, especially with a large number of electrons in its outer shell, this number may depend on the chemical compound formed. (For example, the valence of iron is two in the *ferrous* oxide, FeO, and three in the *ferric* oxide, Fe_2O_3.)

Atomic number	Atomic symbol	Electron states
Period 1		
1	H	$1s^1$
2	He	$1s^2$
Period 2		[He] shell, plus:
3	Li	$2s^1$
4	Be	$2s^2$
5	B	$2s^2 2p^1$
6	C	$2s^2 2p^2$
7	N	$2s^2 2p^3$
8	O	$2s^2 2p^4$
9	F	$2s^2 2p^5$
10	Ne	$2s^2 2p^6$
Period 3		[Ne] shell, plus:
11	Na	$3s^1$
12	Mg	$3s^2$
13	Al	$3s^2 3p^1$
14	Si	$3s^2 3p^2$
15	P	$3s^2 3p^3$
16	S	$3s^2 3p^4$
17	Cl	$3s^2 3p^5$
18	Ar	$3s^2 3p^6$
Period 4		[Ar] shell, plus:
19	K	$4s^1$
20	Ca	$4s^2$
21	Sc	$3d^1 4s^2$
22	Ti	$3d^2 4s^2$
23	V	$3d^3 4s^2$
24	Cr	$3d^4 4s^2$
25	Mn	$3d^5 4s^2$
26	Fe	$3d^6 4s^2$
27	Co	$3d^7 4s^2$
28	Ni	$3d^8 4s^2$
29	Cu	$3d^9 4s^1$
30	Zn	$3d^{10} 4s^2$
31	Ga	$3d^{10} 4s^2 4p^1$
32	Ge	$3d^{10} 4s^2 4p^2$
33	As	$3d^{10} 4s^2 4p^3$
34	Se	$3d^{10} 4s^2 4p^4$
35	Br	$3d^{10} 4s^2 4p^5$
36	Kr	$3d^{10} 4s^2 4p^6$

Atomic number	Atomic symbol	Electron states
Period 5		[Kr] shell, plus:
37	Rb	$5s^1$
38	Sr	$5s^2$
39	Y	$4d^1 5s^2$
40	Zr	$4d^2 5s^2$
41	Nb	$4d^4 5s^1$
42	Mo	$4d^5 5s^1$
43	Tc	$4d^6 5s^1$
44	Ru	$4d^7 5s^1$
45	Rh	$4d^8 5s^1$
46	Pd	$4d^{10}$
47	Ag	$4d^{10} 5s^1$
48	Cd	$4d^{10} 5s^2$
49	In	$4d^{10} 5s^2 5p^1$
50	Sn	$4d^{10} 5s^2 5p^2$
51	Sb	$4d^{10} 5s^2 5p^3$
52	Te	$4d^{10} 5s^2 5p^4$
53	I	$4d^{10} 5s^2 5p^5$
54	Xe	$4d^{10} 5s^2 5p^6$
Period 6		[Xe] shell, plus:
55	Cs	$6s^1$
56	Ba	$6s^2$
57	La	$5d^1 6s^2$
58	Ce	$4f^1 5d^1 6s^2$
59	Pr	$4f^3 6s^2$
60	Nd	$4f^4 6s^2$
61	Pm	$4f^5 6s^2$
62	Sm	$4f^6 6s^2$
63	Eu	$4f^7 6s^2$
64	Gd	$4f^7 5d^1 6s^2$
65	Tb	$4f^9 6s^2$
66	Dy	$4f^{10} 6s^2$
67	Ho	$4f^{11} 6s^2$
68	Er	$4f^{12} 6s^2$
69	Tm	$4f^{13} 6s^2$
70	Yb	$4f^{14} 6s^2$
71	Lu	$4f^{14} 5d^1 6s^2$
72	Hf	$4f^{14} 5d^2 6s^2$
73	Ta	$4f^{14} 5d^3 6s^2$
74	W	$4f^{14} 5d^4 6s^2$
75	Re	$4f^{14} 5d^5 6s^2$
76	Os	$4f^{14} 5d^6 6s^2$

Atomic number	Atomic symbol	Electron states
77	Ir	$4f^{14} 5d^7 6s^2$
78	Pt	$4f^{14} 5d^9 6s^1$
79	Au	$4f^{14} 5d^{10} 6s^1$
80	Hg	$4f^{14} 5d^{10} 6s^2$
81	Tl	$4f^{14} 5d^{10} 6s^2 6p^1$
82	Pb	$4f^{14} 5d^{10} 6s^2 6p^2$
83	Bi	$4f^{14} 5d^{10} 6s^2 6p^3$
84	Po	$4f^{14} 5d^{10} 6s^2 6p^4$
85	At	$4f^{14} 5d^{10} 6s^2 6p^5$
86	Rn	$4f^{14} 5d^{10} 6s^2 6p^6$
Period 7		[Rn] shell, plus:
87	Fr	$7s^1$
88	Ra	$7s^2$
89	Ac	$6d^1 7s^2$
90	Th	$6d^2 7s^2$
91	Pa	$5f^2 6d^1 7s^2$
92	U	$5f^3 6d^1 7s^2$
93	Np	$5f^4 6d^1 7s^2$
94	Pu	$5f^6 7s^2$
95	Am	$5f^7 7s^2$
96	Cm	$5f^7 6d^1 7s^2$
97	Bk	$5f^9 7s^2$
98	Cf	$5f^{10} 7s^2$
99	Es	$5f^{11} 7s^2$
100	Fm	$5f^{12} 7s^2$
101	Md	$5f^{13} 7s^2$
102	No	$5f^{14} 7s^2$
103	Lr	$5f^{14} 6d^1 7s^2$
104	Rf	$5f^{14} 6d^2 7s^2$
105	Db	$5f^{14} 6d^3 7s^2$
106	Sg	$5f^{14} 6d^4 7s^2$
107	Bh	$5f^{14} 6d^5 7s^2$
108	Hs	$5f^{14} 6d^6 7s^2$
109	Mt	$5f^{14} 6d^7 7s^2$
110	Ds	$5f^{14} 6d^8 7s^2$
111	Rg	$5f^{14} 6d^9 7s^2$
112	Cn	$5f^{14} 6d^{10} 7s^2$
113	Nh	$5f^{14} 6d^{10} 7s^2 7p^1$
114	Fl	$5f^{14} 6d^{10} 7s^2 7p^2$
115	Mc	$5f^{14} 6d^{10} 7s^2 7p^3$
116	Lv	$5f^{14} 6d^{10} 7s^2 7p^4$
117	Ts	$5f^{14} 6d^{10} 7s^2 7p^5$
118	Og	$5f^{14} 6d^{10} 7s^2 7p^6$

Fig. 3.24. Atomic electron configurations. The upper index shows the number of electrons in the states with the indicated quantum numbers n (the first digit) and l (letter-coded as was discussed above).

In the next element, beryllium (symbol Be, $Z = 4$), the $2s$ state ($n = 2$, $l = 0$, $m = 0$) houses one more electron, with the "opposite spin". Due to the higher electric charge of the nucleus, $Q = +4e$, with only half of it compensated by $1s$ electrons of the [He] shell, the binding energy of the $2s$ electrons is somewhat higher than that in lithium, so the ionization energy increases to ~9.32 eV. As a result, beryllium is also chemically active with the valence of two, but not as active as lithium, and is also is metallic in its solid-state phase, but with a lower electric conductivity than lithium.

Moving in this way along the second row of the periodic table (from $Z = 3$ to $Z = 10$), we see a gradual filling of the rest of the total $2n^2 = 2\times2^2 = 8$ different electron states of the $n = 2$ shell (see Eq. (204), with the additional spin degeneracy factor of 2), including two $2s$ states with $m = 0$, and six $2p$ states with $m = 0, \pm1$, with a gradually growing ionization potential (up to ~21.6 eV in Ne with $Z = 10$), i.e. a growing reluctance to either conduct electricity or form *positive* ions. However, the last elements of the row, such as oxygen (O, with $Z = 8$) and especially fluorine (F, with $Z = 9$) can readily pick up extra electrons to fill up their $2p$ states, i.e. form *negative* ions. As a result, these elements are chemically active, with a double valence for oxygen and a single valence for fluorine. However, the final element of this row, neon, has its $n = 2$ shell completely full, and cannot form a stable negative ion. This is why it is a noble gas, like helium. Traditionally, in the periodic table, such elements are placed right under helium (Fig. 23), to emphasize the similarity of their chemical properties. But this necessitates making at least a 6-cell gap in the 1st row. (Actually, the gap is often made larger, to accommodate the next rows – keep reading.)

Period 3, i.e. the 3rd row of the table, starts exactly like period 2, with sodium (Na, with $Z = 11$), also a chemically active alkali metal whose atom features 2+8 = 10 electrons filling the shells with $n = 1$ and $n = 2$ (in Fig. 24, collectively called the neon shell [Ne]), plus one electron in the $3s$ state ($n = 3$, $l = 0$, $m = 0$), which may be again reasonably well described by the hydrogen atom theory – see, e.g., the red curve on the last panel of Fig. 22. Continuing along this row, we could naively expect that, according to Eq. (204), and with the account of double spin degeneracy, this period of the table should have $2n^2 = 2\times3^2 = 18$ elements, with a gradual, sequential filling of first, two $3s$ states, then six $3p$ states, and then ten $3d$ states. However, here we run into a big surprise: after argon (Ar, with $Z = 18$), a relatively inert element with an ionization energy of ~15.7 eV due to the fully filled $3s$ and $3p$ *subshells*, the next element, potassium (K, with $Z = 19$) is an alkali metal again!

The reason for that is the difference of the actual electron energies from those of the hydrogen atom, which is due mostly to electron-electron interactions, and gradually accumulates with the growth of Z. It may be semi-quantitatively understood from the results described in Sec. 6. In hydrogen-like atoms/ions, the electron state energies do not depend on the quantum number l (as well as m) – see Eq. (201). However, the orbital quantum number does affect the wavefunction of an electron. As Fig. 22 shows, the larger l the less the probability for an electron to be close to the nucleus, where its positive charge is less compensated by other electrons. As a result of this effect (and also the relativistic corrections to be discussed in Sec. 6.3), the electron's energy grows with l. Actually, this effect is visible already in period 2 of the table: it manifests itself in the filling order – the p states after the s states. However, for potassium (K, with $Z = 19$) and calcium (Ca, with $Z = 20$), the energies of the $3d$ states become so high that the energies of the two $4s$ states are lower, and the latter states are filled first. As described by Eq. (210), and also by the first of Eqs. (211), the effect of the principal number n on the distance from the nucleus is stronger than that of l, so the $4s$ wavefunctions of K and Ca are relatively far from the nucleus, and determine the chemical valence (equal to 1 and 2, correspondingly) of these elements. The next atoms, from Sc ($Z = 21$) to Zn ($Z = 30$), with the gradually filled "internal" $3d$ states,

are the so-called *transition metals* whose (comparable) ionization energies and chemical properties are determined by the 4*s* electrons.[89]

This fact is the origin of the difference between various forms of the "periodic" table. In its most popular option, shown in Fig. 23, K is used to start the next period 4, and then a new period is started each time and only when the first electron with the next principal quantum number (n) appears.[90] This topology of the table provides a very clear match of the chemical properties of the first element of each period (an alkali metal), as well as its last element (a noble gas). It also automatically means making gaps in all previous rows. Usually, this gap is made between the atoms with completely filled *s* states and with those with the first electron in a *p* state, because here the properties of the elements make a somewhat larger step. (For example, the step from Be to B makes the material an insulator, but the step from Mg to Al makes a smaller difference.) As a result, the elements of the same column have only *approximately* similar chemical valences and physical properties.

In order to accommodate the lower, longer rows, such representation is inconvenient, because the whole table would be too broad. This is why the so-called *rare earth* elements, including *lanthanides* (with Z from 57 to 70, of the 6th row, with a gradual filling of the 4*f* and 5*d* subshells) and *actinides* (Z from 89 to 103, of the 7th row, with a gradual filling of the 5*f* and 6*d* subshells), are usually represented as outlet rows – see Fig. 23. This is quite acceptable for basic chemistry because chemical properties of the elements within each such group are rather close.

To summarize my very short review of this extremely important topic,[91] the "periodic table of elements" is not periodic in the strict sense of the word. Nevertheless, it has had an enormous historic significance for chemistry, as well as atomic and solid-state physics, and is still very convenient for many purposes. For our course, the most important aspect of its discussion is the surprising possibility to describe, at least for classification purposes, such a complex multi-electron system as an atom as a system of quasi-independent electrons in certain quantum states indexed with the same quantum numbers n, l, and m as those of the hydrogen atom. This fact enables the use of various *perturbation theories*, which give a more quantitative description of atomic properties. Some of these techniques will be reviewed in Chapters 6 and 8.

3.8. Spherically symmetric scatterers

The machinery of the Legendre polynomials and the spherical Bessel functions, discussed in Sec. 6, may also be used for the analysis of particle scattering by spherically symmetric potentials (155) beyond the Born approximation (Sec. 3), provided that such a potential $U(r)$ is also localized, i.e. reduces sufficiently fast at $r \to \infty$.[92] Indeed, directing the z-axis along the propagation of the incident plane de Broglie wave ψ_i, and taking its origin in the center of the scatterer, we may expect the scattered wave ψ_s to be axially symmetric, so its expansion in the series over the spherical harmonics includes

[89] The sequence of shell and subshell formation with the atomic number's growth approximately follows the so-called *Madelung rule* (saying that the orbitals with the lowest sum $(n + l)$ are filled first), while the order of state filling inside each subshell closely follows the *Hund rules*, to be discussed in Sec. 8.3.

[90] Another popular option is to return to the first column as soon an atom has one electron in the *s* state (like it is in Cu, Ag, and Au, in addition to the alkali metals).

[91] For a bit more detailed (but still succinct) discussion of the valence and other chemical aspects of atomic structure, I can recommend Chapter 5 of the very clear text by L. Pauling, *General Chemistry*, Dover, 1988.

[92] The quantification of this condition is left for the reader's exercise.

only the terms with $m = 0$. Hence, the solution (64) of the stationary Schrödinger equation (63) in this case may be represented as[93]

$$\psi = \psi_{\rm i} + \psi_{\rm s} = a_{\rm i}\left[e^{ikz} + \sum_{l=0}^{\infty} \mathcal{R}_l(r) P_l(\cos\theta)\right], \tag{3.213}$$

where $k \equiv (2mE)^{1/2}/\hbar$ is defined by the energy E of the incident particle, while the radial functions $\mathcal{R}_l(r)$ have to satisfy Eq. (181), and be finite at $r \to 0$. At large distances $r >> R$, where R is the effective radius of the scatterer, the potential $U(r)$ is negligible, and Eq. (181) is reduced to Eq. (183). In contrast to its analysis in Sec. 6, we should look for its solution using a linear superposition of the spherical Bessel functions of both kinds:

$$\mathcal{R}_l(r) = A_l j_l(kr) + B_l y_l(kr), \qquad \text{at } r >> R, \tag{3.214}$$

because Eq. (183) is now invalid at $r \to 0$, so our former argument for dropping the functions $y_l(kr)$ is no more valid as well. In Eq. (214), A_l and B_l are some complex coefficients, determined by the scattering potential $U(r)$, i.e. by the solution of Eq. (181) at $r \sim R$.

As the explicit expressions (186) show, the spherical Bessel functions $j_l(\xi)$ and $y_l(\xi)$ represent *standing* de Broglie waves, with equal real amplitudes, so their simple linear combinations (called the *spherical Hankel functions of the first and second kind*),

$$h_l^{(1)}(\xi) \equiv j_l(\xi) + iy_l(\xi), \qquad \text{and } h_l^{(2)}(\xi) \equiv j_l(\xi) - iy_l(\xi), \tag{3.215}$$

represent *traveling* spherical waves propagating, respectively, from the origin (i.e. from the center of the scatterer), and toward the origin. In particular, at $\xi >> 1$, l, i.e. at large distances $r >> 1/k$, l/k,[94]

$$h_l^{(1)}(kr) \to \frac{(-i)^{l+1}}{kr} e^{ikr}, \qquad h_l^{(2)}(kr) \to \frac{i^{l+1}}{kr} e^{-ikr}. \tag{3.216}$$

But using the same physical argument as at the beginning of Sec. 1, we may argue that in the case of a localized scatterer, there should be no latter waves at $r >> R$; hence, we have to require the amplitude of the term proportional to $h_l^{(2)}$ to be zero. With the relations reciprocal to Eqs. (215),

$$j_l(\xi) = \frac{1}{2}\left[h_l^{(1)}(\xi) + h_l^{(2)}(\xi)\right], \qquad y_l(\xi) = \frac{1}{2i}\left[h_l^{(1)}(\xi) - h_l^{(2)}(\xi)\right], \tag{3.217}$$

which enable us to rewrite Eq. (214) as

$$\begin{aligned} \mathcal{R}_l(r) &= \frac{A_l}{2}\left[h_l^{(1)}(kr) + h_l^{(2)}(kr)\right] + \frac{B_l}{2i}\left[h_l^{(1)}(kr) - h_l^{(2)}(kr)\right] \\ &\equiv \left(\frac{A_l - iB_l}{2}\right) h_l^{(1)}(kr) + \left(\frac{A_l + iB_l}{2}\right) h_l^{(2)}(kr), \end{aligned} \tag{3.218}$$

this means that the combination $(A_l + iB_l)$ has to be equal zero, i.e. $B_l = iA_l$. Hence we have just one unknown coefficient (say, A_l) for each l,[95] and may rewrite Eq. (218) in an even simpler form:

[93] The particular terms in this sum are frequently called *partial waves*.

[94] For arbitrary l, this result may be confirmed using Eqs. (185) and the asymptotic formulas for the "usual" Bessel functions – see, e.g., EM Eqs. (2.135) and (2.152), valid for an arbitrary (not necessarily integer) index n.

$$\mathcal{R}_l(r) = A_l\left[j_l(kr) + iy_l(kr)\right] \equiv A_l h_l^{(1)}(kr), \qquad \text{at } r >> R, \tag{3.219}$$

and use Eqs. (213) and (216) to write the following expression for the scattered wave at large distances:

$$\psi_s \approx \frac{a_i}{kr} e^{ikr} \sum_{l=0}^{\infty} (-i)^{l+1} A_l P_l(\cos\theta), \qquad \text{for } r >> R, \frac{1}{k}, \frac{l}{k}. \tag{3.220}$$

Comparing this expression with the general Eq. (81), we see that for a spherically symmetric, localized scatterer,

$$f = \frac{1}{k} \sum_{l=0}^{\infty} (-i)^{l+1} A_l P_l(\cos\theta), \tag{3.221}$$

so the differential cross-section (84) is

$$\frac{d\sigma}{d\Omega} = \frac{1}{k^2} \left| \sum_{l=0}^{\infty} (-i)^{l+1} A_l P_l(\cos\theta) \right|^2 \equiv \frac{1}{k^2} \sum_{l,l'=0}^{\infty} i^{l'-l} A_l A_{l'}^* P_l(\cos\theta) P_{l'}(\cos\theta). \tag{3.222}$$

The last expression is more convenient for the calculation of the total cross-section (59):

$$\sigma = \oint_{4\pi} \frac{d\sigma}{d\Omega} d\Omega = 2\pi \int_{-1}^{+1} \frac{d\sigma}{d\Omega} d(\cos\theta) = \frac{2\pi}{k^2} \sum_{l,l'=0}^{\infty} i^{l'-l} A_l A_{l'}^* \int_{-1}^{+1} P_l(\xi) P_{l'}(\xi) d\xi, \tag{3.223}$$

where $\xi \equiv \cos\theta$, because this result may be much simplified by using Eq. (167):[96]

Spherically symmetric scatterer: σ

$$\boxed{\sigma = \sum_{l=0}^{\infty} \sigma_l, \quad \text{with } \sigma_l = \frac{4\pi}{k^2} \frac{|A_l|^2}{2l+1}.} \tag{3.224}$$

Hence the solution of the scattering problem is reduced to the calculation of the partial wave amplitudes A_l defined by Eq. (219) – and for the total cross-section, merely of their magnitudes. This task is much facilitated by using the following formula[97] for the expansion of the incident plane wave into a series over the Legendre polynomials,

$$e^{ikz} \equiv e^{ikr\cos\theta} = \sum_{l=0}^{\infty} i^l (2l+1) j_l(kr) P_l(\cos\theta). \tag{3.225}$$

As the simplest example, let us consider scattering by a completely impenetrable and "hard" (meaning sharp-boundary) sphere, which may be described by the following potential:

$$U(r) = \begin{cases} +\infty, & \text{for } r < R, \\ 0, & \text{for } R < r. \end{cases} \tag{3.226}$$

---

[95] Moreover, using the conservation of the orbital momentum, to be discussed in Sec. 5.6, it is possible to show that this *complex* coefficient may be further reduced to just one *real* parameter, usually recast as the *partial phase shift* δ_l between the l^{th} spherical harmonics of the incident and scattered waves. However, I will not use this notion, because practical calculations are more physically transparent (and not more complex) without it.

[96] Physically, this reduction of the double sum to a single one means that due to the orthogonality of the spherical harmonics, the total scattering probability flows due to each partial wave just add up.

[97] It may be proved by using the Rodrigues formula (165) and integration by parts – the task left for the reader's exercise.

In this case, the total wavefunction has to vanish at $r \le R$, and hence for the external problem ($r \ge R$), the sphere enforces the boundary condition $\psi \equiv \psi_0 + \psi_s = 0$ for all values of θ, at $r = R$. With Eqs. (213), (220), and (225), this condition becomes

$$a_i \sum_{l=0}^{\infty} \left[\mathcal{R}_l(R) + i^l (2l+1) j_l(kR) \right] P_l(\cos\theta) = 0 . \tag{3.227}$$

Due to the orthogonality of the Legendre polynomials, this condition may be satisfied for all angles θ only if all the coefficients before all $P_l(\cos\theta)$ vanish, i.e. if

$$\mathcal{R}_l(R) = -i^l (2l+1) j_l(kR). \tag{3.228}$$

On the other hand, for $r > R$, $U(r) = 0$, so Eq. (183) is valid, and its outward-wave solution (219) has to be valid even at $r \to R$, giving

$$\mathcal{R}_l(R) = A_l \left[j_l(kR) + i y_l(kR) \right]. \tag{3.229}$$

Requiring the two last expressions to give the same result, we get

$$A_l = -i^l (2l+1) \frac{j_l(kR)}{j_l(kR) + i y_l(kR)}, \tag{3.230}$$

so Eqs. (222) and (224) yield:

$$\frac{d\sigma}{d\Omega} = \frac{1}{k^2} \left| \sum_{l=0}^{\infty} (2l+1) \frac{j_l(kR)}{j_l(kR) + i y_l(kR)} P_l(\cos\theta) \right|^2, \qquad \sigma_l = \frac{4\pi(2l+1)}{k^2} \frac{j_l^2(kR)}{j_l^2(kR) + y_l^2(kR)} . \tag{3.231}$$

As Fig. 25a shows, the first of these results gives an angular structure of the scattered de Broglie wave, which is qualitatively similar to that given by the Born approximation – cf. Eq. (98) and Fig. 10.

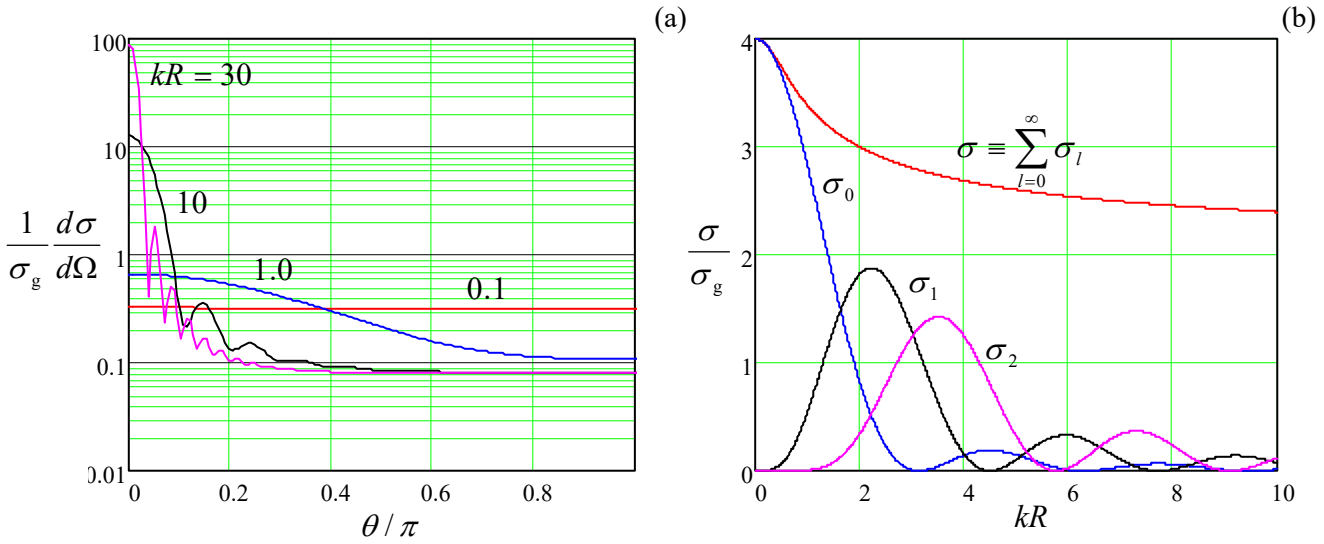


Fig. 3.25. Particle scattering by an impenetrable hard sphere: (a) the differential cross-section normalized to the geometric cross-section $\sigma_g \equiv \pi R^2$ of the sphere, as a function of the scattering angle θ, and (b) the (similarly normalized) total cross-section and its lowest spherical components, as functions of the dimensionless product $kR \propto E^{1/2}$.

Namely, at low particle's energies ($kR << 1$), the scattering is essentially isotropic, while in the opposite, high-energy limit $kR >> 1$, it is mostly confined to small angles $\theta \sim \pi/kR << 1$, and exhibits numerous local destructive-interference minima at angles $\theta_n \sim \pi n/kR$. However, in our current (exact!) theory, these minima are different from zero because the theory describes an effective bending of the de Broglie waves along the back side of the sphere, which smears the interference pattern.

This bending is also responsible for a rather counter-intuitive fact (sometimes called *wave extinction paradox*), described by the second of Eqs. (231) and clearly visible in Fig. 25b: even at $kR \to \infty$, the total cross-section σ of scattering tends to $2\sigma_g \equiv 2\pi R^2$, rather than to the geometric cross-section σ_g as in the purely-classical scattering theory. First discovered for optical waves, this effect is common for all large non-absorbing scatterers. The fact that at $kR << 1$, the cross-section is also larger than σ_g, approaching $4\sigma_g$ at $kR \to 0$, is much less surprising, because in this limit the de Broglie wavelength $\lambda = 2\pi/k$ is much larger than the sphere's radius R, so the sphere effect on the incident wave extends to the area of the order of $\lambda^2 >> \sigma_g \equiv \pi R^2$.

The above analysis may be readily generalized to the case of a step-like (*sharp* but finite) potential (97) – the problem left for the reader's exercise. On the other hand, for a finite and *smooth* scattering potential $U(r)$, by plugging Eq. (225) into Eq. (213) and the result into Eq. (66), and requiring the coefficients before each angular function $P_l(\cos\theta)$ to be balanced, we get the following inhomogeneous generalization of Eq. (181) for the radial functions defined by Eq. (213):

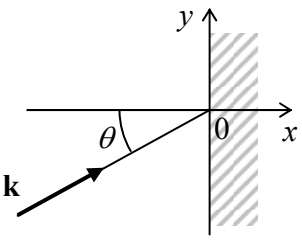
$$\left[E - U(r)\right]\mathcal{R}_l + \frac{\hbar^2}{2mr^2}\left[\frac{d}{dr}\left(r^2\frac{d}{dr}\right) - l(l+1)\right]\mathcal{R}_l(r) = U(r)i^l(2l+1)j_l(kr). \tag{3.232}$$

This differential equation has to be solved in the whole scatterer volume (i.e. for all $r \sim R$) with the boundary conditions for the functions $\mathcal{R}_l(r)$ to be finite at $r \to 0$, and to tend to the asymptotic form (219) at $r >> R$. The last requirement enables the evaluation of the coefficients A_l that are needed for spelling out Eqs. (222) and (224), for any particular potential $U(r)$. Unfortunately, due to the lack of time/space, for particular examples, I have to refer the interested reader to special literature.[98]

3.9. Exercise Problems

3.1. A particle of energy E is incident (in the figure on the right, within the plane of the drawing) on a sharp potential step:

$$U(\mathbf{r}) = \begin{cases} 0, & \text{for } x < 0, \\ U_0, & \text{for } 0 < x. \end{cases}$$



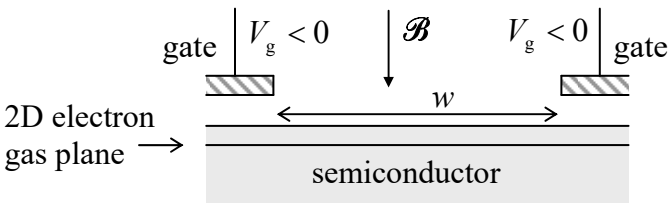
Calculate the particle reflection probability $\mathscr{R}$ as a function of the incidence angle θ, and discuss this function for various magnitudes and signs of U_0.

3.2. For a charged particle moving in a magnetic field $\mathscr{B}$, calculate the commutation relations between Cartesian components of the kinetic ("$m\mathbf{v}$-") momentum operator defined by Eq. (20). Can the result be represented in a vector form?

[98] See, e.g., J. Taylor, *Scattering Theory*, Dover, 2006.

3.3. In the classical mechanics version of the Landau-level problem discussed in Sec. 3.2 of the lecture notes, the geometric center of the particle's orbit is an integral of motion, determined by initial conditions. Calculate the commutation relation between the quantum-mechanical operators corresponding to the Cartesian coordinates of the center.

3.4.* Analyze how are the Landau levels (50) modified by an additional uniform electric field $\mathscr{E}$ directed along the plane of the particle's motion. Contemplate the physical meaning of your result and its implications for the quantum Hall effect in a gate-defined Hall bar. (The area $l \times w$ of such a bar is defined by metallic "gate" electrodes parallel to the 2D electron gas plane – see the figure on the right. The negative voltage V_g applied to the gates squeezes the 2D gas from the area under them into the complementary, Hall-bar part of the plane.)



3.5. Analyze how are the Landau levels (50) modified if a 2D particle is confined in an additional 1D potential well $U(x) = m\omega_0^2 x^2/2$.

3.6. Find the stationary states of a spinless, charged 3D particle moving in "crossed" (mutually perpendicular) uniform electric and magnetic fields, with $\mathscr{E} << c\mathscr{B}$. For such states, calculate the expectation values of the particle's velocity in the direction perpendicular to both fields and compare the result with the solution of the corresponding classical problem.

Hint: You may like to generalize Landau's solution for 2D particles, discussed in Sec. 2, to the 3D case.

3.7. Use the Born approximation to calculate the angular dependence and the total cross-section of scattering of an incident plane wave propagating along the x-axis, by the following pair of similar point inhomogeneities:

$$U(\mathbf{r}) = \mathscr{W}\left[\delta\left(\mathbf{r} - \mathbf{n}_z \frac{a}{2}\right) + \delta\left(\mathbf{r} + \mathbf{n}_z \frac{a}{2}\right)\right].$$

Analyze the results in detail. Derive the condition of the Born approximation's validity for such delta-functional scatterers.

3.8. Use the Born approximation to analyze the scattering of particles of energy E by a very thin, straight, uniform rod of length l, oriented normally to the incident particle's velocity. In particular, calculate the differential and total cross-sections of scattering and analyze the results in the low-energy and high-energy limits.

3.9. Complete the analysis of the Born scattering by a uniform spherical potential (97), started in Sec. 3, by calculation of its total cross-section. Analyze the result in the limits $kR << 1$ and $kR >> 1$.

3.10. Use the Born approximation to calculate the differential cross-section of particle scattering by a very thin spherical shell, whose potential may be approximated as $U(\mathbf{r}) = \mathscr{W}\delta(r - R)$. Analyze the

results in the limits $kR << 1$ and $kR >> 1$, and compare them with those for a uniform sphere considered in Sec. 3.

3.11. Use the Born approximation to calculate the differential and total cross-sections of electron scattering by a screened Coulomb field of a point charge Ze, with the electrostatic potential[99]

$$\phi(\mathbf{r}) = \frac{Ze}{4\pi\varepsilon_0 r} e^{-\lambda r},$$

neglecting spin interaction effects, and analyze the result's dependence on the screening parameter λ. Compare the results with those given by the classical ("Rutherford") formula[100] for the unscreened Coulomb potential ($\lambda \to 0$), and formulate the condition of Born approximation's validity in this limit.

3.12. A quantum particle with electric charge Q is scattered by the field of a localized distributed charge with a spherically symmetric density $\rho(r)$ and zero total charge. Use the Born approximation to calculate the differential cross-section of the *forward scattering* (with the scattering angle $\theta = 0$), and evaluate it for the scattering of electrons by a hydrogen atom in its ground state.

3.13. Prove the optical theorem (99).

Hint: For the general solution (64) of the scattering problem, with ψ_i given by Eq. (6) and ψ_s in the form (81), calculate the full probability current I through a spherical surface of radius $r >> k^{-1}$, and then require that in accordance with the continuity relation (1.48), in this stationary situation, $I = 0$.

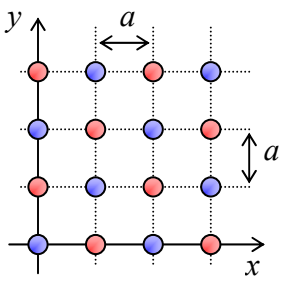
3.14. Reformulate the Born approximation for the 1D case. Use the result to find the scattering and transfer matrices of a "rectangular" (flat-top) scatterer

$$U(x) = \begin{cases} U_0, & \text{for } |x| < d/2, \\ 0, & \text{otherwise.} \end{cases}$$

Compare the results with those of the exact calculations carried out earlier in Chapter 2 and analyze how their relationship changes in the eikonal approximation.

3.15. In the tight-binding approximation, find the lowest stationary states of a particle placed into a system of three similar, isotropic, weakly coupled potential wells located in the vertices of an equilateral triangle.

3.16. The figure on the right shows a fragment of a periodic 2D lattice, with the red and blue points showing the positions of different local potentials.



(i) Find the reciprocal lattice and the 1st Brillouin zone of the system.

(ii) Calculate the wave number k of the monochromatic de Broglie wave incident along the x-axis, at which the lattice creates the lowest-order diffraction peak within the $[x, y]$ plane, and the direction toward this peak.

(iii) Semi-quantitatively, describe the evolution of the intensity of the peak when all local potentials become similar.

[99] This *Yukawa potential* was first suggested in 1935 by H. Yukawa as a model for strong interactions.
[100] See, e.g., CM Sec. 3.5, in particular Eq. (3.73).

Hint: The *order of diffraction* on a multidimensional Bravais lattice is a somewhat ambiguous notion dependent on the lattice type, but the *lowest-order peak* is always that corresponding to the smallest non-zero magnitude of the vector **Q**.

3.17. For the 2D hexagonal lattice (Fig. 12b):

(i) find the reciprocal lattice **Q** and the 1st Brillouin zone;

(ii) use the tight-binding approximation to calculate the dispersion relation $E(\mathbf{q})$ for a 2D particle moving through a potential profile with such periodicity, with an energy close to the eigenenergy of similar isotropic states quasi-localized at the lattice points;

(iii) analyze and sketch/plot the resulting dispersion relation $E(\mathbf{q})$ inside the 1st Brillouin zone.

3.18. Complete the tight-binding-approximation calculation of the band structure of the honeycomb lattice, that was started at the end of Sec. 4. Analyze the results; in particular, prove that the Dirac points $\mathbf{q}_D$ are located in the corners of the 1st Brillouin zone, and express the velocity v_n participating in Eq. (122), in terms of the coupling energy δ_n. Show that the final results do not change if the quasi-localized wavefunctions are not isotropic but are proportional to $\exp\{im\varphi\}$ – as they are, with $m = 1$, for the $2p_z$ electrons of carbon atoms in graphene, which are responsible for its transport properties.

3.19. Examine the basic properties of the so-called *Wannier functions*[101] defined as

$$\phi_{\mathbf{R}}(\mathbf{r}) \equiv C \int_{\text{BZ}} \psi_{\mathbf{q}}(\mathbf{r}) e^{-i\mathbf{q}\cdot\mathbf{R}} d^3q ,$$

where $\psi_{\mathbf{q}}(\mathbf{r})$ is the Bloch wavefunction (108), **R** is any vector of the Bravais lattice, C is a normalization constant, and the integration over the quasimomentum **q** is extended over any (e.g., the first) Brillouin zone.

3.20. Evaluate the long-range interaction (the so-called *London dispersion force*) between two similar, electrically neutral atoms or molecules, modeling each of them as an isotropic 3D harmonic oscillator with the electric dipole moment $\mathbf{d} = q\mathbf{s}$, where **s** is the oscillator's displacement from its equilibrium position.

Hint: You may like to represent the total Hamiltonian of the system as a sum of Hamiltonians of independent 1D harmonic oscillators, and calculate their total ground-state energy as a function of the distance between the dipoles. [102]

3.21. Derive expressions for the stationary wavefunctions and the corresponding energies of a 2D particle of mass m, free to move inside a round disk of radius R. What is the degeneracy of each energy level? Calculate the five lowest energy levels with an accuracy better than 1%.

[101] Named after G. Wannier who introduced these functions in 1939.

[102] This explanation of the interaction between electrically-neutral atoms was put forward in 1930 by F. London, on the background of a prior (1928) work by C. Wang. Note that in some texts this interaction is (rather inappropriately) referred to as the "van der Waals force", though it is only one (long-range) component of the van der Waals model – see, e.g., SM Sec. 4.1.

3.22. Calculate the ground-state energy of a 2D particle of mass m, localized in a very shallow flat-bottom potential well

$$U(\rho)=\begin{cases}-U_0, & \text{for } \rho<R,\\ 0, & \text{for } \rho>R,\end{cases} \quad \text{with } 0<U_0<<\frac{\hbar^2}{mR^2}.$$

3.23. Estimate the energy E of the localized ground state of a 2D particle of mass m, in an axially symmetric potential well of a finite radius R, with an arbitrary but very small potential $U(\rho)$. (Quantify this condition.)

3.24. Spell out the spherical harmonics $Y_4^0(\theta,\varphi)$ and $Y_4^4(\theta,\varphi)$.

3.25. Calculate $\langle x\rangle$ and $\langle x^2\rangle$ in the ground states of the planar and spherical rotors of radius R. What can you say about $\langle p_x\rangle$ and $\langle p_x^2\rangle$?

3.26. A spherical rotor with $r=R=\text{const}$ and mass m is in a state with the following wavefunction: $\psi = C(⅓ + \sin^2\theta)$, where C is a constant. Calculate the energy of its angular motion.

3.27. According to the discussion at the beginning of Sec. 5, stationary wavefunctions of a 3D isotropic harmonic oscillator may be calculated as products of three similar 1D "Cartesian oscillators" – see, in particular Eq. (125), with $d = 3$. However, per the discussion in Sec. 6, the wavefunctions of the type (200), proportional to the spherical harmonics Y_l^m, also describe stationary states of this spherically symmetric system. Represent the wavefunctions (200) of:

(i) the ground state of the oscillator, and

(ii) each of its lowest excited states,

as linear combinations of products of the 1D oscillator's stationary wavefunctions. Also, calculate the degeneracy of the n^{th} energy level of the oscillator.

3.28. A particle of mass m is placed into a spherical, flat-bottom potential well

$$U(\mathbf{r})=\begin{cases}-U_0, & \text{for } r<R,\\ 0, & \text{for } R<r,\end{cases} \quad \text{with } U_0>0.$$

(i) Calculate the smallest U_0 at which the particle has a bound (localized) stationary state.

(ii) Calculate the energy of this state if U_0 is barely larger than that minimum value.

(iii) Does such a localized state exist in a very narrow and deep well that may be described as $U(\mathbf{r}) = -\mathcal{W}\delta(\mathbf{r})$ with a positive and finite $\mathcal{W}$?

3.29. A 3D particle of mass m is placed into a spherically symmetric potential well with $-\infty < U(r) \le U(\infty) = 0$. Relate its ground-state energy to that of a 1D particle of the same mass, moving in the following potential well:

$$U'(x)=\begin{cases}U(x), & \text{for } x\ge 0,\\ +\infty, & \text{for } x\le 0.\end{cases}$$

Use the found relation to:

(i) discuss the origin of the difference between the solutions of Task (i) of the previous problem and of Problem 2.21, and

(ii) calculate the energy spectrum of an electron moving over an impenetrable plane surface of a perfect conductor.

3.30. Calculate the smallest value of the parameter U_0, for that the following spherically symmetric potential well:

$$U(r) = -U_0 e^{-r/R}, \qquad \text{with } U_0, R > 0,$$

has a bound (localized) eigenstate for a particle of mass m.

Hint: You may like to introduce the following new variables: $f \equiv r\mathcal{R}$ and $\xi \equiv Ce^{-r/2R}$, with a proper constant C.

3.31.* A particle of mass m moves in the field of an attractive spherically symmetric potential $U(r) \le U(\infty) \equiv 0$. Find a condition necessary for it to have at least one bound state. Compare the result with those of Problems 28 and 30.

3.32. A particle of mass m, moving in a certain central potential $U(r)$, has a stationary state with the following wavefunction:

$$\psi = Cr^{\alpha} e^{-\beta r} \cos\theta,$$

where C, α, and $\beta > 0$ are constants. Calculate:

(i) the probabilities of all possible values of the quantum numbers m and l, and

(ii) the confining potential and the state's energy.

3.33. For an isotropic 3D harmonic oscillator, calculate:

(i) the energy spectrum resulting from the Bohr quantization of circular classical orbits, and

(ii) the energy spectrum of the s-states in the WKB approximation.

Compare the results with the exact energy spectrum of the oscillator, and comment.

3.34. For a particle of mass m, moving in the spherically symmetric potential $U(\mathbf{r}) = ar^4$:

(i) use the variational method to estimate the ground-state energy,

(ii) calculate the energy spectrum resulting from the Bohr quantization of circular orbits, and

(iii) calculate the energy spectrum of the s-states in the WKB approximation.

Compare the results and comment.

3.35. For a particle of mass m, moving in the attracting Coulomb potential $U(r) = -C/r$ (e.g., the electron in a hydrogen atom):

(i) estimate the ground state energy by using the trial wavefunction $\psi_{\text{trial}} = A/(r + a)^b$, where both $a > 0$ and $b > 1$ are fitting parameters, and

(ii) calculate the energy spectrum of the s-states in the WKB approximation.

Compare the results with the exact energy spectrum of the atom.

3.36. Calculate the energy spectrum of a particle moving in a monotonic but otherwise arbitrary spherically symmetric attractive potential $U(r) < 0$, in the approximation of very large orbital quantum numbers l. Formulate the quantitative condition(s) of validity of your theory. Check that for the Coulomb potential $U(r) = -C/r$, your result agrees with Eq. (201).

Hint: Try to solve Eq. (181) approximately by introducing the same new function $f(r) \equiv r\mathcal{R}(r)$ that was already used in Sec. 1 and in the solutions of a few earlier problems.

3.37. Prove Eq. (210) and the first two of Eqs. (211) for the ground state of a hydrogen-like atom/ion.

3.38. For the ground state of a particle in the Coulomb potential (190), calculate the probability to find it farther from the attracting center than the radius the same particle with the same energy would have on a classical circular orbit.

3.39. For the ground state and the lowest excited states of the hydrogen atom:

(i) calculate the spatial distribution of the electric current flowing around the nucleus,
(ii) evaluate its highest density, and
(iii) calculate and evaluate its magnetic field at the position of the nucleus.

3.40. An electron had been in the ground state of a hydrogen-like atom/ion with nuclear charge Ze when the charge suddenly changed to $(Z + 1)e$.[103] Calculate the probabilities for the electron of the changed system to be:

(i) in its ground state, and
(ii) in one of the lowest excited states.

3.41. Due to a very short pulse of an external force, the nucleus of a hydrogen-like atom/ion, initially at rest in its ground state, starts moving with velocity $\mathbf{v}$. Calculate the probability W_g that the atom remains in its ground state. Evaluate the energy to be given, by the pulse, to a hydrogen atom in order to reduce W_g to 50%.

3.42. Calculate $\langle x^2 \rangle$ and $\langle p_x^2 \rangle$ in the ground state of a hydrogen-like atom/ion. Compare the results with Heisenberg's uncertainty relation. What do these results tell about the electron's velocity in the system?

3.43. Use the Hellmann-Feynman theorem (see Problem 1.7) to prove:

(i) the first of Eqs. (211), and
(ii) the fact that for a spinless particle in an arbitrary spherically symmetric attractive potential $U(r)$, the ground state is always an s-state (with the orbital quantum number $l = 0$).

3.44. For the ground state of a hydrogen atom, calculate:

[103] Such a fast change happens, for example, at the beta-decay, when one of the nucleus' neurons spontaneously turns into a proton, emitting a high-energy electron and a neutrino, which leave the system very fast (instantly on the atomic time scale), and do not affect directly the atom transition's dynamics.

(i) the expectation value of $\mathscr{E}$, where $\boldsymbol{\mathscr{E}}$ is the electric field created by the atom as a whole, and

(ii) the expectation value of $\mathscr{E}^2$ at distances $r >> r_0$ from the nucleus.

Interpret the obtained relation between $\langle \mathscr{E} \rangle^2$ and $\langle \mathscr{E}^2 \rangle$ at distant observation points.

3.45. Find the condition at which a particle of mass m, moving in the field of a very thin spherical shell with $U(\mathbf{r}) = \mathscr{w}\delta(r - R)$ and $\mathscr{w} < 0$, has at least one localized ("bound") stationary state.

3.46. Calculate the lifetime of the lowest metastable state of a particle in the same spherical shell potential as in the previous problem, but now with $\mathscr{w} > 0$, for sufficiently large $\mathscr{w}$. (Quantify this condition.)

3.47. A particle of mass m and energy E is incident on a very thin spherical shell of radius R, whose localized states were the subject of two previous problems, with an arbitrary "weight" $\mathscr{w}$.

(i) Derive general expressions for the differential and total cross-sections of scattering.

(ii) Spell out the contribution σ_0 to the total cross-section σ, given by the spherically symmetric component of the scattered de Broglie wave.

(iii) Analyze the result for σ_0 in the limits of very small and very large magnitudes of $\mathscr{w}$, for both signs of this parameter. In particular, in the limit $\mathscr{w} \to +\infty$, relate the result to the metastable state's lifetime τ calculated in the previous problem.

3.48. Calculate the spherically symmetric contribution σ_0 to the total cross-section of particle scattering by a uniform sphere of radius R, described by the following potential:

$$U(r) = \begin{cases} U_0, & \text{for } r < R, \\ 0, & \text{otherwise,} \end{cases}$$

with an arbitrary constant U_0. Analyze the result in detail, and give an interpretation of its most remarkable features.

3.49. Use the finite difference method with the step $h = a/2$ to calculate as many energy levels as possible, for a particle confined to the interior of:

(i) a square with sides a, and

(ii) a cube with sides a,

with hard walls. For the square, repeat the calculations by using the finer step $h = a/3$. Compare the results for different values of h with each other and with the exact formulas.

Hint: It is advisable to either first solve (or review the solution of) the similar 1D Problem 1.18, or start from reading about the finite difference method.[104] Also: try to exploit the symmetry of the systems.

[104] See, e.g., CM Sec. 8.5 or EM Sec. 2.11.

Chapter 4. Bra-ket Formalism

The objective of this chapter is to describe Dirac's "bra-ket" formalism of quantum mechanics, which not only overcomes some inconveniences of wave mechanics but also enables a natural description of such intrinsic properties of particles as their spin. In the course of the formalism's discussion, I will give only a few simple examples of its application, leaving more involved cases for the following chapters.

4.1. Motivation

As the reader could see from the previous chapters of these notes, wave mechanics gives many results of primary importance. Moreover, it is mostly sufficient for many applications, for example, solid-state electronics and device physics. However, in the course of our survey, we have filed several grievances about this approach. Let me briefly summarize these complaints:

(i) Attempts to analyze the *temporal* evolution of quantum systems, beyond the trivial time behavior of the stationary states, described by Eq. (1.62), run into technical difficulties. For example, we could derive Eq. (2.151) describing the metastable state's decay and Eq. (2.181) describing the quantum oscillations in coupled wells, only for the simplest potential profiles, though it is intuitively clear that these simple results should be common for all problems of this kind. Solving such problems for more complex potential profiles would entangle the time evolution analysis with the calculation of the spatial distribution of the evolving wavefunctions – which (as we could see in Secs. 2.9 and 3.6) may be rather complex even for time-independent potentials. Some separation of the spatial and temporal dependencies is possible using perturbation approaches (to be discussed in Chapter 6) but even those would lead, in the wavefunction language, to very cumbersome formulas.

(ii) The last statement can also be made concerning other issues that are conceptually addressable within the wave mechanics, e.g., the Feynman path integral approach, coupling to the environment, etc. Pursuing them in the wave mechanics language would lead to formulas so bulky that I had postponed their discussion until we would have a more compact formalism on hand.

(iii) In the discussion of several key problems (for example the harmonic oscillator and spherically-symmetric potentials), we have run into rather complicated eigenfunctions coexisting with very simple energy spectra – that infer some simple background physics. It is very important to get this physics revealed.

(iv) In the wave-mechanics postulates formulated in Sec. 1.2, the quantum mechanical operators of the coordinate and momentum are treated rather unequally – see Eqs. (1.26b). However, some key expressions, e.g., for the fundamental eigenfunction of a free particle,

$$\exp\left\{i\frac{\mathbf{p}\cdot\mathbf{r}}{\hbar}\right\}, \tag{4.1}$$

or the harmonic oscillator's Hamiltonian,

$$\hat{H}=\frac{1}{2m}\hat{p}^2+\frac{m\omega_0^2}{2}\hat{r}^2, \tag{4.2}$$

just beg for a similar treatment of coordinates and momenta.

© K. Likharev

However, the strongest motivation for a more general formalism comes from wave mechanics' conceptual inability to describe elementary particles' *spins*[1] and other internal quantum degrees of freedom, such as quarks' flavors. In this context, let us review the basic facts on spin (which is very representative and experimentally the most accessible of all internal quantum numbers), to understand what a more general formalism has to explain – as a minimum.

Figure 1 shows the conceptual scheme of the simplest spin-revealing experiment, first conceived by Otto Stern in 1921 and implemented by Walther Gerlach in 1922. A collimated beam of particles[2] from a natural source, such as a heated cathode, is passed through a gap between the poles of a strong magnet, whose magnetic field $\mathscr{B}$, (in Fig. 1, directed along the z-axis) is nonuniform, so both $\mathscr{B}_z$ and $d\mathscr{B}_z/dz$ are not equal to zero. The experiment shows that even if all particles are in the ground orbital state, the beam splits into two beams of equal intensity.

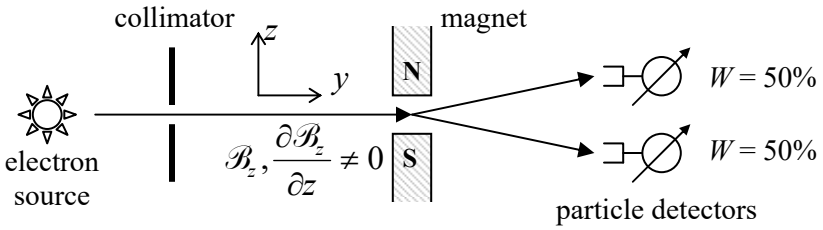


Fig. 4.1. The simplest Stern-Gerlach experiment.

This result may be semi-quantitatively explained on classical (if somewhat phenomenological) grounds by assuming that each particle has an intrinsic, permanent magnetic dipole moment **m**. Indeed, classical electrodynamics tells us[3] that the potential energy U of a magnetic dipole in an external magnetic field $\mathscr{B}$ is equal to $(-\mathbf{m}\cdot\mathscr{B})$, so the force acting on the particle,

$$\mathbf{F} = -\nabla U = -\nabla\left(-\mathbf{m}\cdot\mathscr{B}\right), \tag{4.3}$$

has a non-zero z-component

$$F_z = -\frac{\partial}{\partial z}\left(-m_z\cdot\mathscr{B}_z\right) \equiv m_z\frac{\partial \mathscr{B}_z}{\partial z}. \tag{4.4}$$

Hence if we further assume that the particle's magnetic moment may take only two equally probable discrete values of $m_z = \pm\mu$ (though such discreteness does not follow from any classical model of the particle), this may explain the basic Stern-Gerlach effect qualitatively. The quantitative explanation of the beam splitting angle requires the magnitude of μ to be equal (or very close) to the so-called *Bohr magneton*[4]

$$\mu_{\rm B} \equiv \frac{\hbar e}{2m_{\rm e}} \approx 0.9274\times10^{-23}\,\frac{\rm J}{\rm T}. \tag{4.5}$$

Bohr magneton

[1] Reportedly, the concept of spin as a measure of the internal rotation of a particle was first suggested (though later rejected) by Ralph Kronig, then a 20-year-old student, in January 1925, a few months before two other students, George Uhlenbeck and Samuel Goudsmit, came to this idea independently. The concept was then accepted (first, rather reluctantly) and developed quantitatively by Wolfgang Pauli.

[2] The initial Stern-Gerlach experiments used silver atoms because their larger mass helps to decrease the spit beam widths. However, the discussion below is valid for any spin-½ particles including electrons.

[3] See, e.g., EM Sec. 5.4, in particular Eq. (5.100).

[4] A good mnemonic rule is that it is close to 1 K/T. In the Gaussian units, $\mu_{\rm B} \equiv \hbar e/2m_{\rm e}c \approx 0.9274\times10^{-20}$ erg/G.

However, as we will see below, this value cannot be explained by any internal motion of the particle, say its rotation about the *z*-axis. More importantly, this semi-classical phenomenology cannot explain, even qualitatively, other experimental results, for example those of the set of multistage Stern-Gerlach experiments shown in Fig. 2. In the first of the experiments, the particle beam is first passed through a magnetic field (and its gradient) oriented along the *z*-axis, just as in Fig. 1. Then one of the two resulting beams is absorbed (or removed from the setup in some other way), while the other one is passed through a similar but *x*-oriented field. The experiment shows that this beam is split again into two components of equal intensity. A classical explanation of this experiment would require an even more unnatural additional assumption that the initial particles had random but discrete components of the magnetic moment simultaneously in *two* directions, *z* and *x*.

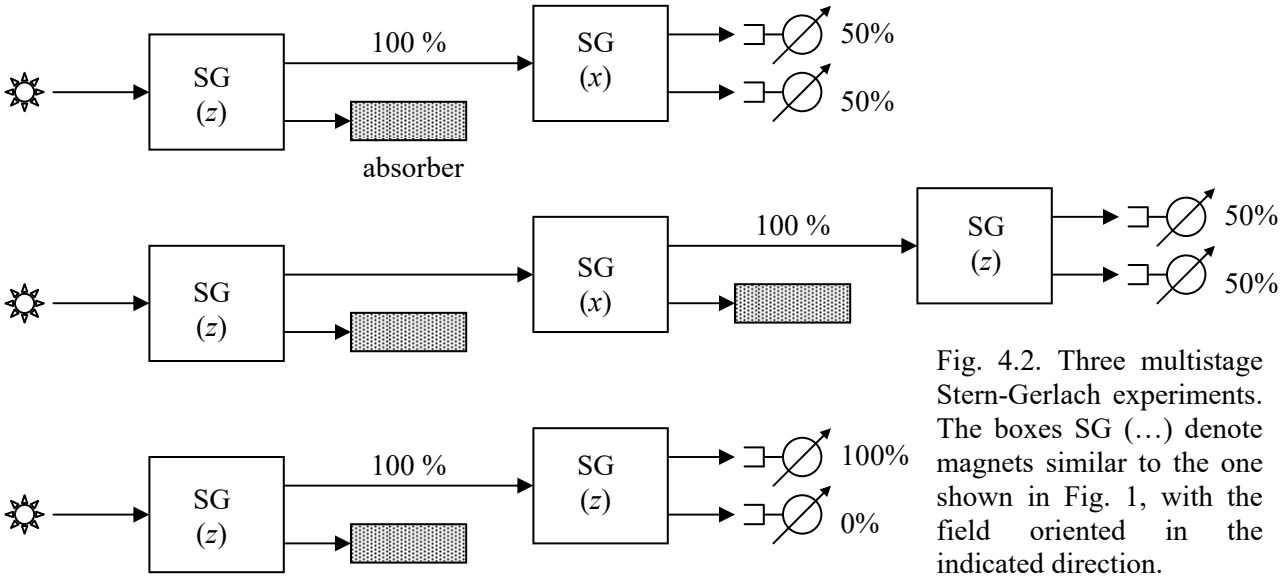


Fig. 4.2. Three multistage Stern-Gerlach experiments. The boxes SG (…) denote magnets similar to the one shown in Fig. 1, with the field oriented in the indicated direction.

However, even this assumption cannot explain the results of the three-stage Stern-Gerlach experiment shown on the middle panel of Fig. 2. Here, the previous two-state setup is complemented with one more absorber and one more magnet, now with the *z*-orientation again. Completely counter-intuitively, it again gives two beams of equal intensity, as if we have not yet filtered out the particles with m_z corresponding to the lower beam, at the first *z*-stage. The only way to save the classical explanation here is to say that maybe, particles *somehow* interact with the magnetic field, so the *x*-polarized beam becomes spontaneously depolarized again somewhere between the two last stages. But any hope for such an explanation is ruined by the control experiment shown on the bottom panel of Fig. 2, whose results indicate that no such depolarization happens.

We will see below that all these (and many more) results find a natural explanation in the so-called *matrix mechanics* pioneered by Werner Heisenberg, Max Born, and Pascual Jordan in 1925. However, the matrix formalism is rather inconvenient for the solution of most problems discussed in Chapters 1-3, and for a short time, it was eclipsed by E. Schrödinger's wave mechanics, which had been put forward just a few months later. However, very soon Paul Adrien Maurice Dirac introduced a more general *bra-ket formalism* of quantum mechanics, which provides a generalization of both approaches and proves their equivalence. Let me describe it, begging for the reader's patience because (in contrast with my usual style), I will not be able to give particular examples of its application for a while – until all the basic notions of the formalism have been introduced.

4.2. States, state vectors, and linear operators

The basic notion of the general formulation of quantum mechanics is the *quantum state* of a system.[5] To get some gut feeling of this notion, if a quantum state α of a particle may be adequately described by wave mechanics, this description is given by the corresponding wavefunction $\Psi_\alpha(\mathbf{r}, t)$. Note, however, that a quantum state as such is *not* a mathematical object,[6] and can participate in mathematical formulas only as a "label" – e.g., the index of the wavefunction Ψ_α. On the other hand, such a wavefunction is *not* a state, but a mathematical object (a complex function of space and time) giving a quantitative description of the state – just as the classical radius vector $\mathbf{r}_\alpha$ and velocity $\mathbf{v}_\alpha$ as real functions of time are mathematical objects describing the motion of the particle in its classical description – see Fig. 3. Similarly, in the Dirac formalism, a certain quantum state α is described by either of two mathematical objects, called the *state vectors*: the *ket-vector* $|\alpha\rangle$ and the *bra-vector* $\langle\alpha|$,[7] whose relationship is close to that between the wavefunction Ψ_α and its complex conjugate Ψ_α^*.

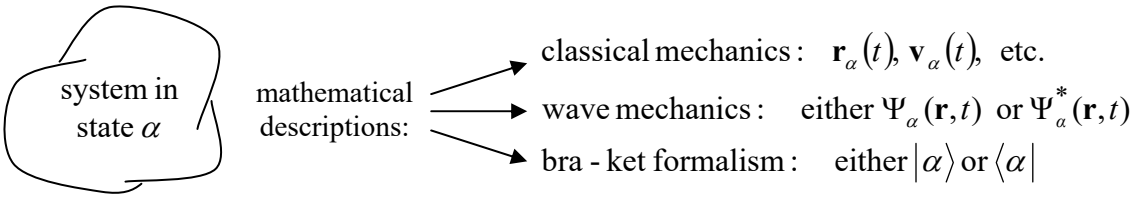


Fig. 4.3. Physical state of a system and its descriptions.

One should be cautious with the term "vector" here. The usual geometric vectors, such as $\mathbf{r}$ and $\mathbf{v}$, are defined in the usual geometric (say, Euclidean) space. In contrast, the bra- and ket-vectors are defined in a more abstract *Hilbert space* – the full set of all possible state vectors of a given system.[8] So, despite certain similarities with the geometric vectors, the bra- and ket-vectors are different mathematical objects, and we need to define the rules of their handling. The primary rules are essentially postulates and are justified only by the correct description of all experimental observations of the rules' corollaries. While there is a general consensus among physicists about what the corollaries are, there are many possible ways to carve from them the different sets of basic postulates. Just as in Sec. 1.2, I will not try too hard to beat the number of the postulates down to the minimum, trying instead to keep their physical meaning transparent.

(i) Ket-vectors. Let us start with *ket-vectors* – sometimes called just *kets* for short. Their most important property is the *linear superposition*. Namely, if several ket-vectors $|\alpha_j\rangle$ describe possible states of a quantum system, numbered by the index j, then any linear combination (*superposition*)

$$|\alpha\rangle = \sum_j c_j |\alpha_j\rangle, \tag{4.6}$$

Linear superposition of ket-vectors

[5] An attentive reader could notice my smuggling the term "system" instead of "particle", which was used in the previous chapters. Indeed, the bra-ket formalism allows the description of quantum systems much more complex than a single spinless particle that is a typical (though not the only possible) subject of wave mechanics.

[6] As was expressed nicely by Asher Peres, one of the pioneers of the quantum information theory, "quantum phenomena do not occur in the Hilbert space, they occur in a laboratory".

[7] The terms *bra* and *ket* were suggested to reflect the fact that the pair $\langle\beta|$ and $|\alpha\rangle$ may be considered as the parts of the combinations like $\langle\beta \mid \alpha\rangle$ (see below), which remind expressions in the usual angle *brackets*.

[8] I have to confess that this is a bit loose definition; it will be refined soon.

where c_j are any (possibly complex) c-numbers, also describes a possible state of the same system.[9] Actually, since ket-vectors are new mathematical objects, the exact meaning of the right-hand side of Eq. (6) becomes clear only after we have postulated the following rules of summation of these vectors,

$$\left|\alpha_j\right\rangle + \left|\alpha_{j'}\right\rangle = \left|\alpha_{j'}\right\rangle + \left|\alpha_j\right\rangle, \tag{4.7}$$

and their multiplication by an arbitrary c-number:

$$c\left|\alpha_j\right\rangle = \left|\alpha_j\right\rangle c\,. \tag{4.8}$$

Note that in the set of wave-mechanics postulates, the statements parallel to Eqs. (7) and (8) were unnecessary because the wavefunctions are the usual (albeit complex) functions of space and time, and we know from the usual algebra that such relations are indeed valid.

As Eq. (6) shows, the coefficient c_j may be interpreted as the "weight" of the state α_j in the linear superposition α. One important particular case is $c_j = 0$, showing that the state α_j does not participate in the superposition α. The corresponding term of the sum (6), i.e. the product

Null-state vector

$$\boxed{0\left|\alpha_j\right\rangle}, \tag{4.9}$$

has a special name: the *null-state* vector. (It is important to avoid confusion between the null state corresponding to vector (9), and the ground state of the system, which is frequently denoted by the ket-vector $|0\rangle$. In some sense, the null state does not exist at all, while the ground state not only does exist but frequently is the most important quantum state of the system.)

(ii) Bra-vectors and inner products. Bra-vectors $\langle\alpha|$, which obey the rules similar to Eqs. (7) and (8), are not new, independent objects: a ket-vector $|\alpha\rangle$ and the corresponding bra-vector $\langle\alpha|$ describe the same state. In other words, there is a *unique dual correspondence* between $|\alpha\rangle$ and $\langle\alpha|$,[10] very similar (though not identical) to that between a wavefunction Ψ and its complex conjugate Ψ^*.[11] The correspondence between these vectors is described by the following rule: if a ket-vector of a linear superposition is described by Eq. (6), then the corresponding bra-vector is

Linear superposition of bra-vectors

$$\boxed{\left\langle\alpha\right| = \sum_j c_j^* \left\langle\alpha_j\right| = \sum_j \left\langle\alpha_j\right| c_j^*}\,. \tag{4.10}$$

The mathematical convenience of using two types of vectors rather than just one becomes clear from the notion of their *inner product* (due to its second, shorthand form, also called the *short bracket*):

Short bracket (inner product)

$$\boxed{\left(\left\langle\beta\right|\right)\left(\left|\alpha\right\rangle\right) \equiv \left\langle\beta\middle|\alpha\right\rangle}, \tag{4.11}$$

which is a scalar c-number, in a certain but limited analogy with the scalar product of the usual geometric vectors. (For one difference, the product (11) may be a complex number.) The main property

[9] One may express the same statement by saying that the vector $|\alpha\rangle$ belongs to the same Hilbert space as all $|\alpha_j\rangle$.

[10] Mathematicians like to say that the ket- and bra-vectors of the same quantum system are defined in two *isomorphic* Hilbert spaces.

[11] This analogy is not occasional: we will see very soon that the wavefunction of a quantum state is just a special ("coordinate") representation of its state vector.

of the inner product is its linearity with respect to any of its component vectors. For example, if a linear superposition α is described by the ket-vector (6), then

$$\langle \beta | \alpha \rangle = \sum_j c_j \langle \beta | \alpha_j \rangle , \tag{4.12}$$

while if Eq. (10) is true, then

$$\langle \alpha | \beta \rangle = \sum_j c_j^* \langle \alpha_j | \beta \rangle . \tag{4.13}$$

In plain English, *c*-number factors may be moved either into or out of the inner products.

The second key property of the inner product is

$$\boxed{\langle \alpha | \beta \rangle = \langle \beta | \alpha \rangle^* .} \tag{4.14}$$

Inner product:: complex conjugate

It is compatible with Eq. (10); indeed, the complex conjugation of both parts of Eq. (12) gives:

$$\langle \beta | \alpha \rangle^* = \sum_j c_j^* \langle \beta | \alpha_j \rangle^* = \sum_j c_j^* \langle \alpha_j | \beta \rangle = \langle \alpha | \beta \rangle . \tag{4.15}$$

Finally, one more rule: the inner product of the bra- and ket-vectors describing the same state (called the *norm squared*) is real and non-negative,

$$\boxed{\| \alpha \|^2 \equiv \langle \alpha | \alpha \rangle \geq 0.} \tag{4.16}$$

State's norm squared

In order to give the reader some feeling about the meaning of this rule: we will see below that if some state α may be described by the corresponding wavefunction $\Psi_\alpha(\mathbf{r}, t)$, then

$$\langle \alpha | \alpha \rangle = \int \Psi_\alpha^* \Psi_\alpha d^3 r \geq 0 . \tag{4.17}$$

Hence the role of the bra- and ket-vectors of the same state is very similar to that of complex-conjugate pairs of its wavefunctions.

(iii) Operators. One more key notion of the Dirac formalism is quantum-mechanical *linear operators*. Just as for the operators discussed in wave mechanics, the function of an operator is to "generate" one state from another: if $|\alpha\rangle$ is a possible ket of the system, and $\hat{A}$ is a legitimate[12] operator, then the following combination,

$$\hat{A} | \alpha \rangle , \tag{4.18}$$

is also a ket-vector describing a possible state of the system, i.e. a ket-vector in the same Hilbert space as the initial vector $|\alpha\rangle$. An alternative formulation of the same rule is the following refinement of the notion of the Hilbert space: for a given set of linear operators of a system, its Hilbert space includes all vectors that may be obtained from each other using the operations of the type (18). In this context, let me note that the operator set, and hence the Hilbert space of a system, usually (if not always) implies its

[12] Here the term "legitimate" means "having a clear sense in the bra-ket formalism". Some examples of "illegitimate" expressions are: $|\alpha\rangle \hat{A}$, $\hat{A} \langle \alpha|$, $|\alpha\rangle|\beta\rangle$, and $\langle \alpha|\langle \beta|$. Note, however, that the last two expressions may be legitimate if α and β are states of different systems, i.e. if their state vectors belong to different Hilbert spaces. We will run into such *direct products* of the bra- and ket-vectors (sometimes denoted, respectively, as $|\alpha\rangle \otimes |\beta\rangle$ and $\langle \alpha| \otimes \langle \beta|$) in Chapters 6-10.

certain approximate model. For example, if the coupling of orbital degrees of freedom of a particle to its spin may be ignored (as it may be for a non-relativistic particle in the absence of an external magnetic field), we may describe the dynamics of the particle using spin operators only. In this case, the set of all possible spin vectors of the particle forms a Hilbert space separate from that of the orbital-state vectors of the same particle.

As the adjective "linear" in the operator definition implies, the main rule governing the operators is their linearity with respect to both any superposition of vectors:

$$\hat{A}\left(\sum_j c_j \left|\alpha_j\right\rangle\right) = \sum_j c_j \hat{A}\left|\alpha_j\right\rangle, \tag{4.19}$$

and any superposition of operators:

$$\left(\sum_j c_j \hat{A}_j\right)\left|\alpha\right\rangle = \sum_j c_j \hat{A}_j\left|\alpha\right\rangle. \tag{4.20}$$

These rules are evidently similar to Eqs. (1.53)-(1.54) of wave mechanics.

The above rules imply that an operator "acts" on the ket-vector on its right; however, a combination of the type $\langle\alpha|\hat{A}$ is also legitimate and represents a new bra-vector. It is important that, generally, this vector does *not* represent the same state as the ket-vector (18); instead, the bra-vector isomorphic to the ket-vector (18) is

Conjugate operator

$$\boxed{\langle\alpha|\hat{A}^\dagger.} \tag{4.21}$$

This statement serves as the definition of the *Hermitian conjugate* (also called "Hermitian adjoint") $\hat{A}^\dagger$ of the initial operator $\hat{A}$. For an important class of operators, called the *Hermitian operators*, the conjugation is inconsequential, i.e. for them

Hermitian operator

$$\boxed{\hat{A}^\dagger = \hat{A}.} \tag{4.22}$$

(This equality, as well as any other operator equation below, means that these operators act similarly on any bra- or ket-vector of the given Hilbert space.) [13]

To proceed further, we need one more additional postulate, sometimes called the *associative axiom of multiplication*: just as an ordinary product of scalars, *any* legitimate bra-ket expression that does not include explicit summations, does not change from an insertion or removal of a pair of parentheses – meaning as usual that the operation inside them has to be performed first. The first two examples of this postulate are given by Eqs. (19) and (20), but the associative axiom is more general and means, for example, that

Long bracket: definition

$$\boxed{\langle\beta|\left(\hat{A}|\alpha\rangle\right) = \left(\langle\beta|\hat{A}\right)|\alpha\rangle \equiv \langle\beta|\hat{A}|\alpha\rangle,} \tag{4.23}$$

[13] If we consider *c*-numbers as a particular type of operators (which is legitimate for any Hilbert space), then according to Eqs. (11) and (21), for them the Hermitian conjugation is equivalent to the simple complex conjugation, so only real *c*-numbers may be considered as a particular type of Hermitian operators (22).

This equality serves as the definition of the last form, called the *long bracket* (evidently, also a scalar), with an operator sandwiched between a bra-vector and a ket-vector. This definition, when combined with the definition of the Hermitian conjugate and Eq. (14), yields an important corollary:

$$\langle\beta|\hat{A}|\alpha\rangle = \langle\beta|\big(\hat{A}|\alpha\rangle\big) = \Big(\big(\langle\alpha|\hat{A}^\dagger\big)|\beta\rangle\Big)^* = \big(\langle\alpha|\hat{A}^\dagger|\beta\rangle\big)^*, \tag{4.24}$$

which is most frequently rewritten as

$$\boxed{\langle\alpha|\hat{A}|\beta\rangle^* = \langle\beta|\hat{A}^\dagger|\alpha\rangle.} \tag{4.25}$$

Long bracket: complex conjugate

The associative axiom also enables us to comprehend the following definition of one more, *outer* product of bra- and ket-vectors:

$$\boxed{|\beta\rangle\langle\alpha|.} \tag{4.26}$$

Outer bra-ket product

In contrast to the inner product (11), which is a *scalar*, this mathematical construct is an *operator*. Indeed, the associative axiom allows us to remove parentheses in the following expression:

$$\big(|\beta\rangle\langle\alpha|\big)|\gamma\rangle = |\beta\rangle\langle\alpha|\gamma\rangle. \tag{4.27}$$

But the last short bracket is just a scalar; hence the mathematical object (26) acting on a ket-vector (in this case, $|\gamma\rangle$) gives a new ket-vector, which is the essence of the operator's action. Very similarly,

$$\langle\delta|\big(|\beta\rangle\langle\alpha|\big) = \langle\delta|\beta\rangle\langle\alpha| \tag{4.28}$$

– again a typical operator's action on a bra-vector. So, Eq. (26) defines an operator.

Now let us perform the following calculation. We may use the parentheses' insertion into the bra-ket equality following from Eq. (14),

$$\langle\gamma|\alpha\rangle\langle\beta|\delta\rangle = \big(\langle\delta|\beta\rangle\langle\alpha|\gamma\rangle\big)^*, \tag{4.29}$$

to transform it into the following form:

$$\langle\gamma|\big(|\alpha\rangle\langle\beta|\big)|\delta\rangle = \big(\langle\delta|\big(|\beta\rangle\langle\alpha|\big)|\gamma\rangle\big)^*. \tag{4.30}$$

Since this equality should be valid for any state vectors $\langle\gamma|$ and $|\delta\rangle$, its comparison with Eq. (25) gives the following operator equality

$$\boxed{\big(|\alpha\rangle\langle\beta|\big)^\dagger = |\beta\rangle\langle\alpha|.} \tag{4.31}$$

Outer product: Hermitian conjugate

This is the conjugate rule for outer products; it reminds Eq. (14) for inner products but involves the Hermitian (rather than the usual complex) conjugation.

The associative axiom is also valid for the *operator multiplication*:

$$\big(\hat{A}\hat{B}\big)|\alpha\rangle = \hat{A}\big(\hat{B}|\alpha\rangle\big), \qquad \langle\beta|\big(\hat{A}\hat{B}\big) = \big(\langle\beta|\hat{A}\big)\hat{B}, \tag{4.32}$$

showing that the action of an operator product on a state vector is nothing more than the sequential action of its operands. However, we have to be careful with the operator products; generally, they do not commute: $\hat{A}\hat{B} \neq \hat{B}\hat{A}$. This is why the *commutator* – the operator defined as

Commutator

$$\left[\hat{A},\hat{B}\right] \equiv \hat{A}\hat{B} - \hat{B}\hat{A}\,, \tag{4.33}$$

is a non-trivial and very useful notion. Another similar notion is the *anticommutator*:[14]

Anti-commutator

$$\left\{\hat{A},\hat{B}\right\} \equiv \hat{A}\hat{B} + \hat{B}\hat{A}\,. \tag{4.34}$$

Finally, the bra-ket formalism broadly uses two special operators. The *null operator* $\hat{0}$ is defined by the following relations:

Null operator

$$\hat{0}\left|\alpha\right\rangle \equiv 0\left|\alpha\right\rangle, \qquad \left\langle\alpha\right|\hat{0} \equiv \left\langle\alpha\right|0\,, \tag{4.35}$$

where α is an arbitrary state; we may say that the null operator "kills" *any* state by turning it into the null state. Another useful notion is the *identity operator*, which is defined by the following action (or rather "inaction" :-) on an arbitrary state vector:

Identity operator

$$\hat{I}\left|\alpha\right\rangle \equiv \left|\alpha\right\rangle, \qquad \left\langle\alpha\right|\hat{I} \equiv \left\langle\alpha\right|. \tag{4.36}$$

These definitions show that the null operator and the identity operator are Hermitian.

4.3. State basis and matrix representation

While some operations in quantum mechanics may be carried out in the general bra-ket formalism outlined above, many calculations are performed for quantum systems that feature a *full and orthonormal* set $\{u\} \equiv \{u_1, u_2, \ldots, u_j, \ldots\}$ of its states u_j, frequently called a *basis*. The former of these terms means that any possible state vector of the system (i.e. any vector of its Hilbert space) may be represented as a unique sum of the type (6) or (10) over its basis vectors:

Expansion over state basis

$$\left|\alpha\right\rangle = \sum_j \alpha_j \left|u_j\right\rangle, \qquad \left\langle\alpha\right| = \sum_j \alpha_j^* \left\langle u_j\right|, \tag{4.37}$$

so, in particular, if α is one of the basis states, say $u_{j'}$, then $\alpha_j = \delta_{jj'}$. The latter term means that

Basis vectors: ortho-normality

$$\left\langle u_j \middle| u_{j'}\right\rangle = \delta_{jj'}\,. \tag{4.38}$$

For the systems that may be described by wave mechanics, examples of the full orthonormal bases are represented by any full and orthonormal set of stationary functions calculated in the previous three chapters of this course – for the simplest example, see Eq. (1.87).

Due to the uniqueness of the expansion (37), the full set of the coefficients α_j involved in the expansion of a state α in certain basis $\{u\}$ gives its complete description – just as the Cartesian components A_x, A_y, and A_z of a usual geometric 3D vector $\mathbf{A}$ in certain reference frame give its complete description. Still, let me emphasize some differences between such representations of the quantum-mechanical state vectors and 3D geometric vectors:

(i) a quantum state basis may have a large or even infinite number of states u_j, and
(ii) the expansion coefficients α_j may be complex.

[14] Another popular notation for the anticommutator (34) is $\left[\hat{A},\hat{B}\right]_+$; it will not be used in these notes.

With these reservations in mind, the analogy with geometric vectors may be pushed further on. Let us inner-multiply both parts of the first of Eqs. (37) by a bra-vector $\langle u_{j'}|$ and then transform the resulting relation using the linearity rules discussed in the previous section, and Eq. (38):

$$\langle u_{j'}|\alpha\rangle = \langle u_{j'}|\sum_j \alpha_j |u_j\rangle = \sum_j \alpha_j \langle u_{j'}|u_j\rangle = \alpha_{j'}. \tag{4.39}$$

Together with Eq. (14), this means that any of the expansion coefficients in Eq. (37) may be represented as an inner product:

$$\boxed{\alpha_j = \langle u_j|\alpha\rangle, \qquad \alpha_j^* = \langle \alpha|u_j\rangle;} \tag{4.40}$$

Expansion coefficients as inner products

these important equalities relations are analogs of equalities $A_j = \mathbf{n}_j \cdot \mathbf{A}$ of the usual vector algebra and will be repeatedly used in this course. With them, the expansions (37) may be rewritten as

$$|\alpha\rangle = \sum_j |u_j\rangle\langle u_j|\alpha\rangle \equiv \sum_j \hat{\Lambda}_j |\alpha\rangle, \qquad \langle\alpha| = \sum_j \langle\alpha|u_j\rangle\langle u_j| \equiv \sum_j \langle\alpha|\hat{\Lambda}_j, \tag{4.41}$$

where

$$\boxed{\hat{\Lambda}_j \equiv |u_j\rangle\langle u_j|\ .} \tag{4.42}$$

Projection operator

Eqs. (41) show that $\hat{\Lambda}_j$ so defined is a legitimate linear operator. This operator, acting on any state vector of the type (37), singles out just one of its components, for example,

$$\hat{\Lambda}_j|\alpha\rangle = |u_j\rangle\langle u_j|\alpha\rangle = \alpha_j|u_j\rangle, \tag{4.43}$$

i.e. "kills" all components of the linear superposition but one. In the geometric analogy, such an operator "projects" the state vector on the j^{th} "direction", hence its name – the *projection operator*. Probably, the most important property of the projection operators, called the *closure* (or "completeness") *relation*, immediately follows from Eq. (41): their sum over the full basis is equivalent to the identity operator

$$\boxed{\sum_j |u_j\rangle\langle u_j| = \hat{I}\ .} \tag{4.44}$$

Closure relation

This means in particular that we may insert the left-hand side of Eq. (44), for any basis, into any bra-ket relation, at any place – the trick that we will use over and over again.

Now let us see how the expansions (37) transform the key notions introduced in the last section, starting with the short bracket (11), i.e. the inner product of two state vectors:

$$\langle\beta|\alpha\rangle = \sum_{j,j'} \langle u_j|\beta_j^*\alpha_{j'}|u_{j'}\rangle = \sum_{j,j'} \beta_j^*\alpha_{j'}\delta_{jj'} = \sum_j \beta_j^*\alpha_j. \tag{4.45}$$

Besides the complex conjugation, this expression is similar to the scalar product of the usual, geometric vectors. Now, let us explore the long bracket (23):

$$\langle\beta|\hat{A}|\alpha\rangle = \sum_{j,j'} \beta_j^* \langle u_j|\hat{A}|u_{j'}\rangle \alpha_{j'} \equiv \sum_{j,j'} \beta_j^* A_{jj'}\alpha_{j'}. \tag{4.46}$$

Here, the last form uses the very important notion of the operator's *matrix elements* defined as

Operator's matrix elements

$$A_{jj'} \equiv \left\langle u_j \middle| \hat{A} \middle| u_{j'} \right\rangle. \tag{4.47}$$

As Eq. (46) shows, the full set of the matrix elements completely characterizes the operator, just as the full set of the expansion coefficients (40) fully characterizes a quantum state. The term "matrix" means, first of all, that it is convenient to represent the full set of $A_{jj'}$ as a square table (*matrix*), with the linear dimension equal to the number of basis states u_j of the system under the consideration. By the way, this number (which may be infinite) is called the *dimensionality* of its Hilbert space.

As two simplest examples, all matrix elements of the null operator, defined by Eqs. (35), are evidently equal to zero (in any basis), and hence it may be represented as a matrix of zeros (called the *null matrix*):

Null matrix

$$0 \equiv \begin{pmatrix} 0 & 0 & \cdots \\ 0 & 0 & \cdots \\ \cdots & \cdots & \cdots \end{pmatrix}, \tag{4.48}$$

while for the identity operator $\hat{I}$ defined by Eqs. (36), we readily get

$$I_{jj'} = \left\langle u_j \middle| \hat{I} \middle| u_{j'} \right\rangle = \left\langle u_j \middle| u_{j'} \right\rangle = \delta_{jj'}, \tag{4.49}$$

i.e. its matrix (naturally called the *identity matrix*) is diagonal – also in any basis:

Identity matrix

$$\mathrm{I} \equiv \begin{pmatrix} 1 & 0 & \ldots \\ 0 & 1 & \ldots \\ \ldots & \ldots & \ldots \end{pmatrix}. \tag{4.50}$$

The convenience of the matrix language extends well beyond the representation of particular operators. For example, let us use the definition (47) to calculate the matrix elements of a product of two operators:

$$(AB)_{jj''} = \left\langle u_j \middle| \hat{A}\hat{B} \middle| u_{j''} \right\rangle. \tag{4.51}$$

Here we may use Eq. (44) for the first (but not the last!) time, inserting the identity operator between the two operators, and then expressing it via the sum of projection operators:

Matrix element of an operator product

$$(AB)_{jj''} = \left\langle u_j \middle| \hat{A}\hat{B} \middle| u_{j''} \right\rangle = \left\langle u_j \middle| \hat{A}\hat{I}\hat{B} \middle| u_{j''} \right\rangle = \sum_{j'} \left\langle u_j \middle| \hat{A} \middle| u_{j'} \right\rangle \left\langle u_{j'} \middle| \hat{B} \middle| u_{j''} \right\rangle = \sum_{j'} A_{jj'} B_{j'j''}. \tag{4.52}$$

This result corresponds to the standard "row by column" rule of calculation of an arbitrary element of the matrix product

$$\mathrm{AB} = \begin{pmatrix} A_{11} & A_{12} & \ldots \\ A_{21} & A_{22} & \ldots \\ \ldots & \ldots & \ldots \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} & \ldots \\ B_{21} & B_{22} & \ldots \\ \ldots & \ldots & \ldots \end{pmatrix}. \tag{4.53}$$

Hence a product of operators may be represented (in a fixed basis!) by that of their matrices (in the same basis).

This is so convenient that the same language is often used to represent not only long brackets,

$$\langle\beta|\hat{A}|\alpha\rangle = \sum_{j'} \beta_j^* A_{jj'} \alpha_{j'} = \left(\beta_1^*, \beta_2^*, \ldots\right) \begin{pmatrix} A_{11} & A_{12} \ldots \\ A_{21} & A_{22} \ldots \\ \ldots & \ldots \ldots \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \ldots \end{pmatrix}, \tag{4.54}$$

Long bracket as a matrix product

but even short brackets:

$$\langle\beta|\alpha\rangle = \sum_{j} \beta_j^* \alpha_j = \left(\beta_1^*, \beta_2^*, \ldots\right) \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \ldots \end{pmatrix}, \tag{4.55}$$

Short bracket as a matrix product

although these equalities require the use of non-square matrices: rows of (complex-conjugate!) expansion coefficients for the representation of bra-vectors, and columns of these coefficients for the representation of ket-vectors. With that, the mapping of quantum states and operators onto matrices becomes completely general.

Now let us have a look at the outer product operator (26). Its matrix elements are just

$$\left(|\alpha\rangle\langle\beta|\right)_{jj'} = \langle u_j|\alpha\rangle\langle\beta|u_{j'}\rangle = \alpha_j \beta_{j'}^*. \tag{4.56}$$

These are the elements of a very special square matrix, whose filling requires the knowledge of just $2N$ scalars (where N is the basis size) rather than N^2 scalars as for an arbitrary operator. However, a simple generalization of such an outer product may represent an arbitrary operator. Indeed, let us insert two identity operators (44), with different summation indices, on both sides of an arbitrary operator:

$$\hat{A} = \hat{I}\hat{A}\hat{I} = \left(\sum_j |u_j\rangle\langle u_j|\right)\hat{A}\left(\sum_{j'} |u_{j'}\rangle\langle u_{j'}|\right), \tag{4.57}$$

and then use the associative axiom to rewrite this expression as

$$\hat{A} = \sum_{j,j'} |u_j\rangle\left(\langle u_j|\hat{A}|u_{j'}\rangle\right)\langle u_{j'}|. \tag{4.58}$$

But the expression in the middle long bracket is just the matrix element (47), so we may write

$$\hat{A} = \sum_{j,j'} |u_j\rangle A_{jj'} \langle u_{j'}|. \tag{4.59}$$

Operator via its matrix elements

The reader should agree that this formula, which is a natural generalization of Eq. (44), is extremely elegant.

The matrix representation is so convenient that it makes sense to extend it to one level lower – from the state vector products to the "bare" state vectors resulting from the operator's action upon a given state. For example, let us use Eq. (59) to represent the ket-vector (18) as

$$|\alpha'\rangle \equiv \hat{A}|\alpha\rangle = \left(\sum_{j,j'} |u_j\rangle A_{jj'} \langle u_{j'}|\right)|\alpha\rangle = \sum_{j,j'} |u_j\rangle A_{jj'} \langle u_{j'}|\alpha\rangle. \tag{4.60}$$

According to Eq. (40), the last short bracket is just $\alpha_{j'}$, so

$$|\alpha'\rangle = \sum_{j,j'} |u_j\rangle A_{jj'} \alpha_{j'} = \sum_j \left(\sum_{j'} A_{jj'} \alpha_{j'}\right)|u_j\rangle \tag{4.61}$$

But the expression in the parentheses is just the coefficient α'_j of the expansion (37) of the resulting ket-vector (60) in the same basis, so

$$\alpha'_j = \sum_{j'} A_{jj'} \alpha_{j'} . \tag{4.62}$$

This result corresponds to the usual rule of multiplication of a matrix by a column, so we may represent any ket-vector by its column matrix, with the operator's action looking like

$$\begin{pmatrix} \alpha'_1 \\ \alpha'_2 \\ \dots \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & \dots \\ A_{21} & A_{22} & \dots \\ \dots & \dots & \dots \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \dots \end{pmatrix} . \tag{4.63}$$

Absolutely similarly, the operator action on the bra-vector (21), represented by its row matrix, is

$$\left(\alpha'^*_1, \alpha'^*_2, \dots\right) = \left(\alpha^*_1, \alpha^*_2, \dots\right) \begin{pmatrix} \left(A^\dagger\right)_{11} & \left(A^\dagger\right)_{12} & \dots \\ \left(A^\dagger\right)_{21} & \left(A^\dagger\right)_{22} & \dots \\ \dots & \dots & \dots \end{pmatrix} . \tag{4.64}$$

By the way, Eq. (64) naturally raises the following question: what are the elements of the matrix on its right-hand side, or more exactly, what is the relation between the matrix elements of an operator and its Hermitian conjugate? The simplest way to answer it is to use Eq. (25) with two arbitrary states (say, u_j and $u_{j'}$) of the same basis in the role of α and β. Together with the orthonormality relation (38), this immediately gives[15]

Hermitian conjugate: matrix elements

$$\boxed{\left(\hat{A}^\dagger\right)_{jj'} = \left(A_{j'j}\right)^* .} \tag{4.65}$$

Thus, the matrix of the Hermitian-conjugate operator is the *complex conjugated and transposed* matrix of the initial operator. This result exposes very clearly the difference between Hermitian and complex conjugation. It also shows that for the Hermitian operators defined by Eq. (22),

$$A_{jj'} = A^*_{j'j} , \tag{4.66}$$

i.e. any pair of their matrix elements, symmetric with respect to the main diagonal, should be the complex conjugate of each other. As a corollary, their main-diagonal elements have to be real:

$$A_{jj} = A^*_{jj} , \quad \text{i.e. } \mathrm{Im}\, A_{jj} = 0. \tag{4.67}$$

[15] For the sake of formula compactness, below I will use the shorthand notation in that the operands of this equality are just $A^\dagger_{jj'}$ and $A^*_{j'j}$. I believe that it leaves little chance for confusion, because the Hermitian conjugation sign † may pertain only to an operator (or its matrix), while the complex conjugation sign *, to a scalar – say a matrix element.

In order to fully appreciate the special role played by Hermitian operators in quantum theory, let us introduce the key notions of *eigenstates* a_j (described by their *eigenvectors* $\langle a_j|$ and $|a_j\rangle$) and *eigenvalues* (*c*-numbers) A_j of an operator $\hat{A}$, both defined by the equation they have to satisfy:[16]

$$\hat{A}\left|a_j\right\rangle = A_j\left|a_j\right\rangle. \tag{4.68}$$

Operator: eigenstates and eigenvalues

Let us prove that the eigenvalues of any Hermitian operator are real,[17]

$$A_j = A_j^*, \quad \text{for } j = 1, 2, ..., N, \tag{4.69}$$

Hermitian operator: eigenvalues

while the eigenstates corresponding to different eigenvalues are orthogonal:

$$\left\langle a_j \middle| a_{j'} \right\rangle = 0, \quad \text{if } A_j \neq A_{j'}. \tag{4.70}$$

Hermitian operator: eigenvectors

The proof of both statements is surprisingly simple. Let us inner-multiply both sides of Eq. (68) by the bra-vector $\langle a_{j'}|$. On the right-hand side of the result, the eigenvalue A_j, as a *c*-number, may be taken out of the bracket, giving

$$\left\langle a_{j'} \middle| \hat{A} \middle| a_j \right\rangle = A_j \left\langle a_{j'} \middle| a_j \right\rangle. \tag{4.71}$$

This equality has to hold for any pair of eigenstates, so we may swap the indices *j* and *j'* in Eq. (71), and write the complex-conjugate of the result:

$$\left\langle a_j \middle| \hat{A} \middle| a_{j'} \right\rangle^* = A_{j'}^* \left\langle a_j \middle| a_{j'} \right\rangle^*. \tag{4.72}$$

Now using Eqs. (14) and (25), together with the Hermitian operator's definition (22), we may transform Eq. (72) into the following form:

$$\left\langle a_{j'} \middle| \hat{A} \middle| a_j \right\rangle = A_{j'}^* \left\langle a_{j'} \middle| a_j \right\rangle. \tag{4.73}$$

Subtracting this equation from Eq. (71), we get

$$0 = \left(A_j - A_{j'}^*\right)\left\langle a_{j'} \middle| a_j \right\rangle. \tag{4.74}$$

There are two possibilities to satisfy this relation. If the indices *j* and *j'* are equal (denote the same eigenstate), then the bracket is the state's norm squared, and cannot be equal to zero. In this case, the left parentheses (with $j = j'$) have to be zero, proving Eq. (69). On the other hand, if *j* and *j'* correspond to different eigenvalues of *A*, the parentheses cannot equal zero (we have just proved that all A_j are real!), and hence the state vectors indexed by *j* and *j'* should be orthogonal, e.g., Eq. (70) is valid.

As will be discussed below, these properties make Hermitian operators suitable, in particular, for the description of physical observables.

[16] This equation should look familiar to the reader – see the stationary Schrödinger equation (1.60), which was the focus of our studies in the first three chapters. We will see soon that that equation is just a particular (coordinate) representation of Eq. (68) for the Hamiltonian as the operator of energy.

[17] The reciprocal statement is also true: if all eigenvalues of an operator are real, it is Hermitian (in any basis). This statement may be readily proved by applying Eq. (93) below to the case when $A_{kk'} = A_k\delta_{kk'}$, with $A_k^* = A_k$.

4.4. Change of basis, and matrix diagonalization

From the discussion of the last section, it may look like the matrix language is fully similar to, and in many instances more convenient than the general bra-ket formalism. In particular, Eqs. (54)-(55) and (63)-(64) show that any part of any bra-ket expression may be directly mapped onto the similar matrix expression, with the only slight inconvenience of using not only columns but also rows (with their elements complex-conjugated), for state vector representation. This invites the question: why do we need the bra-ket language at all? The answer is that the matrix elements depend on the particular choice of the basis set, very much like the Cartesian components of a usual geometric vector depend on the particular choice of reference frame orientation (Fig. 4), and very frequently, at problem solution, it is convenient to use two or more different basis sets for the same system. (Just a bit more patience – numerous examples will follow soon.)

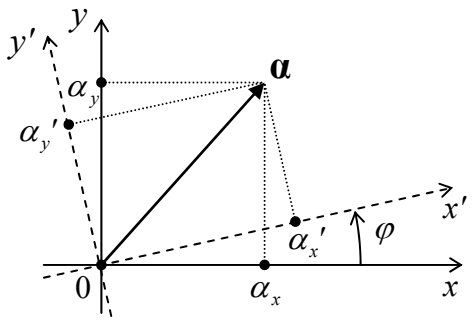


Fig. 4.4. The transformation of components of a 2D vector at a reference frame's rotation.

With this motivation, let us explore what happens at the transform from one basis, $\{u\}$, to another one, $\{v\}$ – both full and orthonormal. First of all, let us prove that for each such pair of bases, and an arbitrary numbering of the states of each base, there exists such an operator $\hat{U}$ that, first,

Unitary operator: definition

$$\left|v_j\right\rangle = \hat{U}\left|u_j\right\rangle, \tag{4.75}$$

and, second,

$$\hat{U}\hat{U}^\dagger = \hat{U}^\dagger\hat{U} = \hat{I}. \tag{4.76}$$

(Due to the last property,[18] $\hat{U}$ is called a *unitary operator*, and Eq. (75), a *unitary transformation*.)

A very simple proof of both statements may be achieved by construction. Indeed, let us take

Unitary operator: construction

$$\hat{U} \equiv \sum_{j'} \left|v_{j'}\right\rangle\left\langle u_{j'}\right|, \tag{4.77}$$

- an evident generalization of Eq. (44). Then, using Eq. (38), we obtain

$$\hat{U}\left|u_j\right\rangle = \sum_{j'} \left|v_{j'}\right\rangle\left\langle u_{j'}\middle|u_j\right\rangle = \sum_{j'} \left|v_{j'}\right\rangle \delta_{j'j} = \left|v_j\right\rangle, \tag{4.78}$$

so Eq. (75) has been proved. Now, applying Eq. (31) to each term of the sum (77), we get

Unitary operator: conjugate

$$\hat{U}^\dagger \equiv \sum_{j'} \left|u_{j'}\right\rangle\left\langle v_{j'}\right|, \tag{4.79}$$

[18] An alternative way to express Eq. (76) is to write $\hat{U}^\dagger = \hat{U}^{-1}$, but I will avoid using this language.

so

$$\hat{U}\hat{U}^{\dagger} = \sum_{j,j'} |v_j\rangle\langle u_j|u_{j'}\rangle\langle v_{j'}| = \sum_{j,j'} |v_j\rangle\delta_{jj'}\langle v_{j'}| = \sum_j |v_j\rangle\langle v_j|. \quad (4.80)$$

But according to the closure relation (44), the last expression is just the identity operator, so one of Eqs. (76) has been proved. (The proof of the second equality is absolutely similar.) As a by-product of our proof, we have also got another important expression – Eq. (79). It implies, in particular, that while, according to Eq. (75), the operator $\hat{U}$ performs the transform from the "old" basis $\{u\}$ to the "new" basis $\{v\}$, its Hermitian adjoint $\hat{U}^{\dagger}$ performs the reciprocal transform:

$$\hat{U}^{\dagger}|v_j\rangle = \sum_{j'} |u_{j'}\rangle\delta_{j'j} = |u_j\rangle. \quad (4.81)$$

Reciprocal basis transform

Now let us see what the matrix elements of the unitary transform operators look like. Generally, as was discussed above, the operator's elements may depend on the basis we calculate them in, so let us be specific – at least initially. For example, let us calculate the desired matrix elements $U_{jj'}$ in the "old" basis $\{u\}$, by using Eq. (77):

$$U_{jj'}|_{\text{in } u} \equiv \langle u_j|\hat{U}|u_{j'}\rangle = \langle u_j|\left(\sum_{j''} |v_{j''}\rangle\langle u_{j''}|\right)|u_{j'}\rangle = \langle u_j|\sum_{j''} |v_{j''}\rangle\delta_{j''j'} = \langle u_j|v_{j'}\rangle. \quad (4.82)$$

Now performing a similar calculation in the "new" basis $\{v\}$, we get

$$U_{jj'}|_{\text{in } v} \equiv \langle v_j|\hat{U}|v_{j'}\rangle = \langle v_j|\left(\sum_{j''} |v_{j''}\rangle\langle u_{j''}|\right)|v_{j'}\rangle = \sum_{j''} \delta_{jj''}\langle u_{j''}\|v_{j'}\rangle = \langle u_j|v_{j'}\rangle. \quad (4.83)$$

Surprisingly, the result is the same! This is of course true for the Hermitian conjugate (79) as well:

$$U^{\dagger}_{jj'}|_{\text{in } u} = U^{\dagger}_{jj'}|_{\text{in } v} = \langle v_j|u_{j'}\rangle. \quad (4.84)$$

These expressions may be used, first of all, to rewrite Eq. (75) in a purely matrix form. Applying the first of Eqs. (41) to any state $v_{j'}$ of the "new" basis, and then Eq. (82), we get

$$|v_{j'}\rangle = \sum_j |u_j\rangle\langle u_j|v_{j'}\rangle = \sum_j U_{jj'}|u_j\rangle. \quad (4.85)$$

Basis transforms: matrix form

Similarly, the reciprocal transform is

$$|u_{j'}\rangle = \sum_j |v_j\rangle\langle v_j|u_{j'}\rangle = \sum_j U^{\dagger}_{jj'}|v_j\rangle. \quad (4.86)$$

These formulas are very convenient for applications; we will use them already in this section.

Next, we may use Eqs. (83)-(84) to express the effect of the unitary transform on the expansion coefficients α_j of the vectors of an *arbitrary* state α, defined by Eq. (37). As a reminder, in the "old" basis $\{u\}$ they are given by Eqs. (40). Similarly, in the "new" basis $\{v\}$,

$$\alpha_j|_{\text{in } v} = \langle v_j|\alpha\rangle. \quad (4.87)$$

Again inserting the identity operator in its closure form (44) with the internal index j', and then using Eqs. (84) and (40), we get

$$\alpha_j\big|_{\text{in }v} = \langle v_j|\left(\sum_{j'}|u_{j'}\rangle\langle u_{j'}|\right)|\alpha\rangle = \sum_{j'}\langle v_j|u_{j'}\rangle\langle u_{j'}|\alpha\rangle = \sum_{j'}U^{\dagger}_{jj'}\langle u_{j'}|\alpha\rangle = \sum_{j'}U^{\dagger}_{jj'}\alpha_{j'}\big|_{\text{in }u}\,. \quad (4.88)$$

The reciprocal transform is performed by matrix elements of the operator $\hat{U}$:

$$\alpha_j\big|_{\text{in }u} = \sum_{j'}U_{jj'}\alpha_{j'}\big|_{\text{in }v}\,. \quad (4.89)$$

Per Eqs. (82)-(84), the matrix elements $U_{jj'}$ and $U^{\dagger}_{jj'}$ are the same in the bases $\{u\}$ and $\{v\}$, so Eqs. (88)-(89) may be rewritten in the following compact matrix form:

$$|\alpha\rangle_{\text{in }v} = \mathrm{U}^{\dagger}|\alpha\rangle_{\text{in }u}, \qquad |\alpha\rangle_{\text{in }u} = \mathrm{U}|\alpha\rangle_{\text{in }v}, \quad (4.90)$$

even though the reader should remember that these relations are different from the usual matrix formulas, which use the same basis for all its components.

So, if the transform (75) from the "old" basis $\{u\}$ to the "new" basis $\{v\}$ is performed by a unitary operator, the change (88) of state vector components at this transformation requires its Hermitian conjugate. This fact is similar to the transformation of components of a usual vector at coordinate frame rotation. For example, for a 2D vector whose actual position in space is fixed (Fig. 4):

$$\begin{pmatrix}\alpha_x'\\ \alpha_y'\end{pmatrix} = \begin{pmatrix}\cos\varphi & \sin\varphi\\ -\sin\varphi & \cos\varphi\end{pmatrix}\begin{pmatrix}\alpha_x\\ \alpha_y\end{pmatrix}, \quad (4.91)$$

but the reciprocal transform is performed by a different matrix, which may be obtained from that participating in Eq. (91) by the replacement $\varphi \to -\varphi$. This replacement has a clear geometric sense: if the "new" reference frame $\{x', y'\}$ is obtained from the "old" frame $\{x, y\}$ by a *counterclockwise* rotation by angle φ, the reciprocal transformation requires such rotation with angle $-\varphi$. (In this analogy, the unitary property (76) of the unitary transform operators corresponds to the equality of the determinants of both rotation matrices to 1.)

Now let us use the same trick of identity operator insertion, repeated twice, to find the transformation rule for matrix elements of an arbitrary operator:

$$A_{jj'}\big|_{\text{in }v} \equiv \langle v_j|\hat{A}|v_{j'}\rangle = \langle v_j|\left(\sum_{k}|u_k\rangle\langle u_k|\right)\hat{A}\left(\sum_{k'}|u_{k'}\rangle\langle u_{k'}|\right)|v_{j'}\rangle = \sum_{k,k'}U^{\dagger}_{jk}A_{kk'}\big|_{\text{in }u}U_{k'j'}\,; \quad (4.92)$$

Matrix elements' transforms

absolutely similarly, we may also get

$$A_{jj'}\big|_{\text{in }u} \equiv \sum_{k,k'}U_{jk}A_{kk'}\big|_{\text{in }v}U^{\dagger}_{k'j'}\,. \quad (4.93)$$

In the spirit of Eq. (90), we may represent these results in the similar matrix form:

$$\mathrm{A}\big|_{\text{in }v} = \mathrm{U}^{\dagger}\mathrm{A}\big|_{\text{in }u}\mathrm{U}, \qquad \mathrm{A}\big|_{\text{in }u} = \mathrm{U}\mathrm{A}\big|_{\text{in }v}\mathrm{U}^{\dagger}, \quad (4.94)$$

where, again, the matrix elements of U and $\mathrm{U}^{\dagger}$ may be calculated in any of the bases $\{u\}$ and $\{v\}$ – but not in an arbitrary basis!

As a sanity check, let us apply Eq. (93) to the identity operator:

$$\hat{I}\big|_{\text{in } v} = \left(\hat{U}^{\dagger}\hat{I}\hat{U}\right)_{\text{in } u} = \left(\hat{U}^{\dagger}\hat{U}\right)_{\text{in } u} = \hat{I}\big|_{\text{in } u} \tag{4.95}$$

– just as it should be. One more invariant of the basis change is the *trace* of any operator, defined as the sum of the diagonal terms of its matrix:

$$\boxed{\text{Tr } \hat{A} \equiv \text{Tr A} \equiv \sum_j A_{jj} .} \tag{4.96}$$

Operator/ matrix: trace

The (easy) proof of this fact, using previous relations, is left for the reader's exercise.

So far, I have implied that both state bases $\{u\}$ and $\{v\}$ are known, and the natural question is where this information comes from in the quantum mechanics of actual physical systems. To get a partial answer to this question, let us return to Eq. (68), which defines the eigenstates and the eigenvalues of an operator. Let us assume that the eigenstates a_j of a certain operator $\hat{A}$ form a full and orthonormal set, and calculate the matrix elements of the operator in the basis $\{a\}$ of these states, at their arbitrary numbering. For that, it is sufficient to inner-multiply both sides of Eq. (68), written for some eigenstate $a_{j'}$, by the bra-vector of an arbitrary state a_j of the same set:

$$\left\langle a_j \middle| \hat{A} \middle| a_{j'} \right\rangle = \left\langle a_j \middle| A_{j'} \middle| a_{j'} \right\rangle . \tag{4.97}$$

The left-hand side of this equality is the matrix element $A_{jj'}$ we are looking for, while its right-hand side is just $A_{j'}\delta_{jj'}$. As a result, we see that the matrix is diagonal, with the diagonal consisting of the operator's eigenvalues:

$$\boxed{A_{jj'} = A_j \delta_{jj'} .} \tag{4.98}$$

Matrix elements in eigenstate basis

In particular, in the eigenstate basis (but not necessarily in an arbitrary basis!), A_{jj} means the same as A_j. Thus the important problem of finding the eigenvalues and eigenstates of an operator is equivalent to the *diagonalization* of its matrix,[19] i.e. finding the basis in which the operator's matrix acquires the diagonal form (98); then the diagonal elements are the eigenvalues, and the basis itself is the desirable set of eigenstates.

To see how this is done in practice, let us inner-multiply Eq. (68) by a bra-vector of the basis (say, $\{u\}$) in that we have happened to know the matrix elements $A_{jj'}$:

$$\left\langle u_k \middle| \hat{A} \middle| a_j \right\rangle = \left\langle u_k \middle| A_j \middle| a_j \right\rangle . \tag{4.99}$$

On the left-hand side, we can (as usual :-) insert the identity operator between the operator $\hat{A}$ and the ket-vector, and then use the closure relation (44) in the same basis $\{u\}$, while on the right-hand side, we can move the eigenvalue A_j (a c-number) out of the bracket, and then insert a summation over the same index as in the closure, compensating it with the proper Kronecker delta symbol:

$$\left\langle u_k \middle| \hat{A} \sum_{k'} \middle| u_{k'} \right\rangle \left\langle u_{k'} \middle| a_j \right\rangle = A_j \sum_{k'} \left\langle u_{k'} \middle| a_j \right\rangle \delta_{kk'} . \tag{4.100}$$

Moving out the signs of summation over k', and using the definition (47) of the matrix elements, we get

[19] Note that the expression "matrix diagonalization" is a very common but dangerous jargon. Formally, a matrix is just a table, an ordered set of c-numbers, and cannot be "diagonalized". It is OK to use this jargon (I will do this) if you remember clearly what it actually means – see the definition above.

$$\sum_{k'} \left(A_{kk'} - A_j \delta_{kk'}\right) \langle u_{k'} | a_j \rangle = 0 . \tag{4.101}$$

But the set of such equalities, for all N possible values of the index k, is just a system of homogeneous linear equations for unknown c-numbers $\langle u_{k'}|a_j\rangle$. According to Eqs. (82)-(84), these numbers are nothing else than the matrix elements $U_{k'j}$ of a unitary matrix providing the required transformation from the initial basis $\{u\}$ to the basis $\{a\}$ that diagonalizes the matrix A. This system may be represented in the matrix form:

Matrix diagonalization

$$\begin{pmatrix} A_{11} - A_j & A_{12} & \dots \\ A_{21} & A_{22} - A_j & \dots \\ \dots & \dots & \dots \end{pmatrix} \begin{pmatrix} U_{1j} \\ U_{2j} \\ \dots \end{pmatrix} = 0 , \tag{4.102}$$

and the condition of its consistency,

Characteristic equation for eigenvalues

$$\begin{vmatrix} A_{11} - A_j & A_{12} & \dots \\ A_{21} & A_{22} - A_j & \dots \\ \dots & \dots & \dots \end{vmatrix} = 0, \tag{4.103}$$

plays the role of the characteristic equation of the system. This equation has N roots A_j – the eigenvalues of the operator $\hat{A}$; after they have been calculated, plugging any of them back into the system (102), we can use it to find N matrix elements U_{kj} (k = 1, 2, …N) corresponding to this particular eigenvalue. However, since the equations (102) are homogeneous, they allow finding U_{kj} only to a constant multiplier. To ensure their normalization, i.e. enforce the unitary character of the matrix U, we may use the requirement for all eigenvectors to be normalized (just as the basis vectors are):

$$\langle a_j | a_j \rangle \equiv \sum_k \langle a_j | u_k \rangle \langle u_k | a_j \rangle \equiv \sum_k \left| U_{kj} \right|^2 = 1, \tag{4.104}$$

for each j. This normalization completes the diagonalization.[20]

Now (at last!) I can give the reader some examples. As a simple but very important case, let us diagonalize each of the operators described (in a certain two-function basis $\{u\}$, i.e. in two-dimensional Hilbert space) by the so-called *Pauli matrices*

Pauli matrices

$$\sigma_x \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_y \equiv \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_z \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{4.105}$$

Though introduced by a physicist, with a specific purpose to describe the electron's spin, these matrices have a general mathematical significance, because together with the 2×2 identity matrix, they provide a full, linearly-independent system – meaning that an arbitrary 2×2 matrix may be represented as

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = b\mathrm{I} + c_x \sigma_x + c_y \sigma_y + c_z \sigma_z , \tag{4.106}$$

[20] A possible slight complication here is that the characteristic equation may give equal eigenvalues for certain groups of different eigenvectors. In such cases, the requirement of the mutual orthogonality of these *degenerate states* should be additionally enforced.

with a unique set of four c-number coefficients b, c_x, c_y, and c_z.

Since the matrix σ_z is already diagonal, with the evident eigenvalues ±1, let us start with diagonalizing the matrix σ_x. For it, the characteristic equation (103) is evidently

$$\begin{vmatrix} -A_j & 1 \\ 1 & -A_j \end{vmatrix} = 0, \qquad \text{i.e. } A_j^2 - 1 = 0, \tag{4.107}$$

and has two roots, $A_{1,2} = \pm 1$. (Again, the state numbering is arbitrary!) So the eigenvalues of the matrix σ_x are the same as those of the matrix σ_z. (The reader may readily check that the eigenvalues of the matrix σ_y are also the same.) However, the eigenvectors of the operators corresponding to these three matrices are different. To find them for σ_x, let us plug its first eigenvalue, $A_1 = +1$, back into equations (101) spelled out for this particular case ($j = 1$; $k, k' = 1,2$):

$$\begin{aligned} -\langle u_1 | a_1 \rangle + \langle u_2 | a_1 \rangle = 0, \\ \langle u_1 | a_1 \rangle - \langle u_2 | a_1 \rangle = 0. \end{aligned} \tag{4.108}$$

These two equations are compatible (of course, because the used eigenvalue $A_1 = +1$ satisfies the characteristic equation), and any of them gives

$$\langle u_1 | a_1 \rangle = \langle u_2 | a_1 \rangle, \quad \text{i.e.} \quad U_{11} = U_{21}. \tag{4.109}$$

With that, the normalization condition (104) yields

$$|U_{11}|^2 = |U_{21}|^2 = \frac{1}{2}. \tag{4.110}$$

Although the normalization is insensitive to the simultaneous multiplication of U_{11} and U_{21} by the same phase factor $\exp\{i\varphi\}$ with any real φ, it is convenient to keep the coefficients real, for example taking $\varphi = 0$, to get

$$U_{11} = U_{21} = \frac{1}{\sqrt{2}}. \tag{4.111}$$

Performing an absolutely similar calculation for the second characteristic value, $A_2 = -1$, we get $U_{12} = -U_{22}$, and we may choose the common phase to have

$$U_{12} = -U_{22} = \frac{1}{\sqrt{2}}, \tag{4.112}$$

so the whole unitary matrix for diagonalization of the operator corresponding to σ_x is[21]

$$\mathrm{U}_x = \mathrm{U}_x^\dagger = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \tag{4.113}$$

Unitary matrix diagonalizing σ_x

For what follows, it will be convenient to have this result expressed in the ket-relation form – see Eqs. (85)-(86):

$$|a_1\rangle = U_{11}|u_1\rangle + U_{21}|u_2\rangle = \frac{1}{\sqrt{2}}\left(|u_1\rangle + |u_2\rangle\right), \qquad |a_2\rangle = U_{12}|u_1\rangle + U_{22}|u_2\rangle = \frac{1}{\sqrt{2}}\left(|u_1\rangle - |u_2\rangle\right), \tag{4.114a}$$

[21] Though this particular unitary matrix U_x is Hermitian, this is not true for an arbitrary choice of the phases φ.

$$|u_1\rangle = U_{11}^{\dagger}|a_1\rangle + U_{21}^{\dagger}|a_2\rangle = \frac{1}{\sqrt{2}}\left(|a_1\rangle + |a_2\rangle\right), \qquad |u_2\rangle = U_{12}^{\dagger}|a_1\rangle + U_{22}^{\dagger}|a_2\rangle = \frac{1}{\sqrt{2}}\left(|a_1\rangle - |a_2\rangle\right). \quad (4.114b)$$

Now let me show that these results are already sufficient to understand the Stern-Gerlach experiments described in Sec. 1 – but with two additional postulates. The first of them is that the interaction of a particle with the external magnetic field, besides that due to its orbital motion, may be described by the following *operator vector* of its spin dipole magnetic moment:[22]

Spin magnetic moment

$$\hat{\mathbf{m}} = \gamma \hat{\mathbf{S}}, \quad (4.115a)$$

where the constant coefficient γ, specific for every particle type, is called the *gyromagnetic ratio*,[23] and $\hat{\mathbf{S}}$ is the *operator vector*[24] *of spin*, with three Cartesian components:

Spin vector operator

$$\hat{\mathbf{S}} = \mathbf{n}_x \hat{S}_x + \mathbf{n}_y \hat{S}_y + \mathbf{n}_z \hat{S}_z. \quad (4.115b)$$

Here $\mathbf{n}_{x,y,z}$ are the usual Cartesian unit vectors in the 3D geometric space (in the quantum-mechanics sense, they are just *c*-numbers, or rather "*c*-vectors"), while $\hat{S}_{x,y,z}$ are the "usual" (scalar) operators. For the so-called *spin-½ particles* (including the electron),[25] these components may be simply, as

$$\hat{S}_{x,y,z} = \frac{\hbar}{2}\hat{\sigma}_{x,y,z}, \quad (4.116a)$$

Pauli vector

expressed via those of the *Pauli vector* $\hat{\boldsymbol{\sigma}} \equiv \mathbf{n}_x \hat{\sigma}_x + \mathbf{n}_y \hat{\sigma}_y + \mathbf{n}_z \hat{\sigma}_z$, so we may also write

$$\hat{\mathbf{S}} = \frac{\hbar}{2}\hat{\boldsymbol{\sigma}}. \quad (4.116b)$$

In turn, in the so-called *z-basis*, each Cartesian component of the latter operator is just the corresponding Pauli matrix (105), so it may be also convenient to use the following 3D vector of these matrices:[26]

Pauli vector's matrix

$$\boldsymbol{\sigma} \equiv \mathbf{n}_x \sigma_x + \mathbf{n}_y \sigma_y + \mathbf{n}_z \sigma_z \equiv \begin{pmatrix} \mathbf{n}_z & \mathbf{n}_x - i\mathbf{n}_y \\ \mathbf{n}_x + i\mathbf{n}_y & -\mathbf{n}_z \end{pmatrix}. \quad (4.117)$$

The *z*-basis, in which such matrix representation of $\hat{\boldsymbol{\sigma}}$ is valid, is *defined* as an orthonormal basis of certain two states, commonly denoted ↑ ("spin up") an ↓ ("spin down"). In this basis, the matrix of the operator $\hat{\sigma}_z$ is diagonal, with eigenvalues, respectively, + 1 and –1, and hence the matrix $S_z \equiv (\hbar/2)\sigma_z$ of $\hat{S}_z$ is also diagonal with the eigenvalues $+\hbar/2$ and $-\hbar/2$ – see the last of Eqs. (105). Note that

[22] This was the key point in the electron spin's description, developed by W. Pauli in 1925-1927.

[23] For the electron, with its negative charge $q = -e$, the gyromagnetic ratio is negative: $\gamma_e = -g_e e/2m_e$, where $g_e \approx 2$ is the electron's dimensionless *g-factor*. Due to quantum-electrodynamic (relativistic) effects, this *g*-factor is slightly higher than 2: $g_e = 2(1 + \alpha/2\pi + \ldots) \approx 2.002319304\ldots$, where $\alpha \equiv e^2/4\pi\varepsilon_0\hbar c \equiv (E_H/m_e c^2)^{1/2} \approx 1/137$ is the so-called *fine structure constant*. (The origin of its name will be clear from the discussion in Sec. 6.3.)

[24] The basic rule of dealing with operator vectors is to perform all vector operations just as with the usual geometric vectors. (The vector ∇ is a good example – see the formulas for in MA Secs. 8-12.)

[25] At this point, the adjective "spin-½ " should be understood as just a name. The physical sense of this term and the generalization of the theory to other values of spin will be discussed in Sec. 5.7.

[26] Note that is some texts, the term "Pauli vector" is used for this matrix $\boldsymbol{\sigma}$ rather than for the operator $\hat{\boldsymbol{\sigma}}$.

we do not "understand" what exactly the states ↑ and ↓ are,[27] but loosely associate them with some internal rotation of a spin-½ particle about the z-axis, with either positive or negative angular momentum component S_z. However, attempts to use such classical interpretation for quantitative predictions run into fundamental difficulties – see Sec. 6 below.

The second necessary postulate describes the general relation between the bra-ket formalism and experiment. Namely, in quantum mechanics, each real observable A is represented by a Hermitian operator $\hat{A} = \hat{A}^\dagger$, and the result of its measurement,[28] in a quantum state α described by a linear superposition of the eigenstates a_j of the operator,

$$|\alpha\rangle = \sum_j \alpha_j |a_j\rangle, \quad \text{with } \alpha_j = \langle a_j|\alpha\rangle, \tag{4.118}$$

may be only one of the corresponding eigenvalues A_j.[29] Specifically, if the ket (118) and all eigenkets $|a_j\rangle$ are normalized to 1,

$$\langle\alpha|\alpha\rangle = 1, \qquad \langle a_j|a_j\rangle = 1, \tag{4.119}$$

then the probability of a certain measurement outcome A_j is[30]

$$W_j = |\alpha_j|^2 \equiv \alpha_j^* \alpha_j \equiv \langle\alpha|a_j\rangle\langle a_j|\alpha\rangle, \tag{4.120}$$

Quantum measurement postulate

This relation is evidently a generalization of Eq. (1.22) in wave mechanics. As a sanity check, let us assume that the set of the eigenstates a_j is full, and calculate the sum of the probabilities to find the system in each of these states:

$$\sum_j W_j = \sum_j \langle\alpha|a_j\rangle\langle a_j|\alpha\rangle = \langle\alpha|\hat{I}|\alpha\rangle = 1. \tag{4.121}$$

Now returning to the Stern-Gerlach experiment, conceptually the description of the first (z-oriented) experiment shown in Fig. 1 is formally the hardest for us, because the statistical ensemble describing the unpolarized particle beam at its input is *mixed* ("incoherent"), and cannot be described by a *pure* ("coherent") superposition of the type (6) that have been the subject of our studies so far. (We will discuss mixed ensembles in Chapter 7.) However, it is intuitively clear that its results are compatible with the description of the two output beams as sets of particles in the pure states ↑ and ↓, respectively. The absorber following that first stage (Fig. 2) just takes all spin-down particles out of the picture, producing an output beam of polarized particles in the definite ↑ state. For such a beam, the

[27] If you think about it, the word "understand" typically means that we can express a new notion in terms of those discussed earlier and thus considered "known". (For example, in our current case, we cannot describe the spin states by any wavefunction $\psi(\mathbf{r})$, or any other mathematical notion discussed in the previous three chapters and hence considered "known".) The bra-ket formalism was invented exactly to enable mathematical analyses of such "new" quantum states we do not initially "understand". Gradually, as we learn more and more about their properties and get *accustomed* to these notions, we start treating them as "known" ones.

[28] Here again, just like in Sec. 1.2, the statement implies the abstract notion of "ideal experiments", deferring the discussion of real (physical) measurements until Chapter 10.

[29] As a reminder, at the end of Sec. 3 we have already proved that such eigenstates corresponding to different values A_j are orthogonal. If any of these values is degenerate, i.e. corresponds to several different eigenstates, they should be also selected orthogonal, in order for Eq. (118) to be valid.

[30] This relation, in particular, explains the most common term for the (generally, complex) coefficients α_j, which was already mentioned several times earlier: the *probability amplitudes*.

probabilities (120) are $W_\uparrow = 1$ and $W_\downarrow = 0$. This is certainly compatible with the result of the "control" experiment shown on the bottom panel of Fig. 2: the repeated SG (z) stage does not split such a beam, keeping the probabilities the same.

Now let us discuss the double Stern-Gerlach experiment shown on the top panel of Fig. 2. For that, let us represent the z-polarized beam in another basis – of the two states (I will denote them as $\rightarrow$ and $\leftarrow$) in that, by definition, the matrix S_x is diagonal. But this is exactly the set we called $a_{1,2}$ in the σ_x matrix diagonalization problem solved above. On the other hand, the states $\uparrow$ and $\downarrow$ are exactly what we called $u_{1,2}$ in that problem because in this basis, we know the matrix σ explicitly – see Eq. (117). Hence, in the application to the particle spin problem, we may rewrite Eqs. (114) as

Relation between eigenvectors of S_x and S_z

$$|\rightarrow\rangle = \frac{1}{\sqrt{2}}\Big(|\uparrow\rangle + |\downarrow\rangle\Big), \qquad |\leftarrow\rangle = \frac{1}{\sqrt{2}}\Big(|\uparrow\rangle - |\downarrow\rangle\Big), \tag{4.122}$$

$$|\uparrow\rangle = \frac{1}{\sqrt{2}}\Big(|\rightarrow\rangle + |\leftarrow\rangle\Big), \qquad |\downarrow\rangle = \frac{1}{\sqrt{2}}\Big(|\rightarrow\rangle - |\leftarrow\rangle\Big), \tag{4.123}$$

Currently for us the first of Eqs. (123) is most important, because it shows that the quantum state of particles entering the SG (x) stage may be represented as a coherent superposition of particles with $S_x = +\hbar/2$ and $S_x = -\hbar/2$. Notice that the beams have equal probability amplitude moduli, so according to Eq. (120), the split beams $\rightarrow$ and $\leftarrow$ have equal intensities, in accordance with experimental results.

Now, let us discuss the most mysterious (from the classical point of view) multistage SG experiment shown on the middle panel of Fig. 2. After the second absorber has taken out all particles in, say, the $\leftarrow$ state, the remaining particles, all in the state $\rightarrow$, are passed to the final, SG (z), stage. But according to the first of Eqs. (122), this state may be represented as a (coherent) linear superposition of the $\uparrow$ and $\downarrow$ states, with equal probability amplitudes. The final stage separates particles in these two states into separate beams, with equal probabilities $W_\uparrow = W_\downarrow = ½$ to find an particle in each of them, thus explaining the experimental results.

To conclude our discussion of the multistage Stern-Gerlach experiment, let me note that though it cannot be explained in terms of wave mechanics (which operates with *scalar* de Broglie waves), it has an analogy in classical theories of *vector* fields, such as the classical electrodynamics. Indeed, let a plane electromagnetic wave propagate normally to the plane of the drawing in Fig. 5, and pass through the linear polarizer 1.

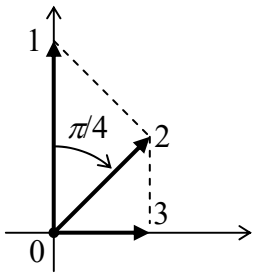


Fig. 4.5. A light polarization sequence similar to the three-stage Stern-Gerlach experiment shown on the middle panel of Fig. 2.

Similarly to the output of the initial SG (z) stages (including the absorbers) shown in Fig. 2, the output wave is linearly polarized in one direction – the vertical direction in Fig. 5. Now its electric field vector has no horizontal component – as may be revealed by the wave's full absorption in a

perpendicular polarizer 3. However, let us pass the wave through polarizer 2 first. In this case, the output wave does acquire a horizontal component, as can be, again, revealed by passing it through polarizer 3. If the angles between the polarization directions 1 and 2, and between 2 and 3, are both equal to $\pi/4$, each polarizer reduces the wave amplitude by a factor of $\sqrt{2}$, and hence the intensity by a factor of 2, exactly like in the multistage SG experiment, with the polarizer 2 playing the role of the SG (x) stage. The "only" difference is that the necessary angle between the polarizer orientations is $\pi/4$, rather than $\pi/2$ for the Stern-Gerlach experiment. In quantum electrodynamics (see Chapter 9 below), which confirms classical predictions for this experiment, this difference may be explained by that between the *integer* spin of electromagnetic field quanta (photons) and the *half-integer* spin of electrons.

4.5. Observables: Expectation values and uncertainties

After this particular (and hopefully inspiring) example, let us discuss the general relation between the Dirac formalism and experiment in more detail. The expectation value of an observable over *any* statistical ensemble (not necessarily a coherent one) may be always calculated using the general statistical rule (1.37). For the particular case of a coherent superposition (118), we can combine that rule with Eq. (120) and the second of Eqs. (118):

$$\langle A\rangle_\alpha = \sum_j A_j W_j = \sum_j \alpha_j^* A_j \alpha_j = \sum_j \langle \alpha | a_j \rangle A_j \langle a_j | \alpha \rangle \equiv \langle \alpha | \left(\sum_j |a_j\rangle A_j \langle a_j| \right) |\alpha\rangle . \tag{4.124}$$

Now using Eq. (59) for the particular case of the eigenstate basis $\{a\}$, for which Eq. (98) is valid, we arrive at a very simple and important formula[31]

$$\boxed{\langle A\rangle_\alpha = \langle \alpha | \hat{A} | \alpha \rangle .} \tag{4.125}$$

Expectation value as a long bracket

This is a clear analog of the wave-mechanics formula (1.23) – and as we will see soon, may be used to derive it.[32] A great convenience of Eq. (125) is that it does not explicitly involve the eigenvector set of the corresponding operator, and allows the calculation to be performed in any convenient basis.

For example, let us consider an arbitrary coherent state α of spin-½,[33] and calculate the expectation values of its components. The calculations are easier in the z-basis because we know the matrix elements of the spin operator components in that basis. Representing the ket- and bra-vectors of the given state as linear superpositions of the corresponding vectors of the basis states $\uparrow$ and $\downarrow$,

$$|\alpha\rangle = \alpha_\uparrow |\uparrow\rangle + \alpha_\downarrow |\downarrow\rangle, \qquad \langle\alpha| = \langle\uparrow|\alpha_\uparrow^* + \langle\downarrow|\alpha_\downarrow^* . \tag{4.126}$$

and plugging these expressions into Eq. (125) written for the observable S_z, we get

---

[31] This equality reveals the full beauty of Dirac's notation. Indeed, initially in this chapter, the quantum-mechanical brackets just *reminded* the angular brackets used for statistical averaging. Now we see that in this particular (but most important) case, the angular brackets of these two types may be indeed *equal* to each other!

[32] Note also that Eq. (120) may be rewritten in a form similar to Eq. (125): $W_j = \langle \alpha | \hat{\Lambda}_j | \alpha \rangle$, where $\hat{\Lambda}_j$ is the operator (42) of the state's projection upon the j^{th} eigenstate a_j.

[33] For clarity, the noun "spin-½" is used, here and below, to denote the spin degree of freedom of a spin-½ particle, independent of its orbital motion.

$$\begin{aligned}\langle S_z\rangle &= \left(\langle\uparrow|\alpha_\uparrow^* + \langle\downarrow|\alpha_\downarrow^*\right)\hat{S}_z\left(\alpha_\uparrow|\uparrow\rangle + \alpha_\downarrow|\downarrow\rangle\right)\\ &= \alpha_\uparrow\alpha_\uparrow^*\langle\uparrow|\hat{S}_z|\uparrow\rangle + \alpha_\downarrow\alpha_\downarrow^*\langle\downarrow|\hat{S}_z|\downarrow\rangle + \alpha_\uparrow\alpha_\downarrow^*\langle\downarrow|\hat{S}_z|\uparrow\rangle + \alpha_\downarrow\alpha_\uparrow^*\langle\uparrow|\hat{S}_z|\downarrow\rangle.\end{aligned} \tag{4.127}$$

Now there are two equivalent ways (both very simple) to calculate the long brackets in this expression. The first one is to represent each of them in the matrix form in the *z*-basis, in which the bra- and ket-vectors of states ↑ and ↓ are the matrix rows (1, 0) and (0, 1), or similar matrix columns – the exercise highly recommended to the reader. Another (perhaps more elegant) way is to use the general Eq. (59), in the *z*-basis, together with the spin-½-specific Eqs. (116a) and (105) to write

Spin-½ component operators

$$\hat{S}_x = \frac{\hbar}{2}\left(|\uparrow\rangle\langle\downarrow| + |\downarrow\rangle\langle\uparrow|\right), \quad \hat{S}_y = -i\frac{\hbar}{2}\left(|\uparrow\rangle\langle\downarrow| - |\downarrow\rangle\langle\uparrow|\right), \quad \hat{S}_z = \frac{\hbar}{2}\left(|\uparrow\rangle\langle\uparrow| - |\downarrow\rangle\langle\downarrow|\right). \tag{4.128}$$

For our particular calculation, we may plug the last of these expressions into Eq. (127), and use the orthonormality conditions (38):

$$\langle\uparrow|\uparrow\rangle = \langle\downarrow|\downarrow\rangle = 1, \quad \langle\uparrow|\downarrow\rangle = \langle\downarrow|\uparrow\rangle = 0. \tag{4.129}$$

Both approaches give (of course) the same result:

$$\langle S_z\rangle = \frac{\hbar}{2}\left(\alpha_\uparrow\alpha_\uparrow^* - \alpha_\downarrow\alpha_\downarrow^*\right). \tag{4.130}$$

This particular result might be also obtained using Eq. (120) for the probabilities $W_\uparrow = \alpha_\uparrow\alpha_\uparrow^*$ and $W_\downarrow = \alpha_\downarrow\alpha_\downarrow^*$, namely:

$$\langle S_z\rangle = W_\uparrow\left(+\frac{\hbar}{2}\right) + W_\downarrow\left(-\frac{\hbar}{2}\right) = \alpha_\uparrow\alpha_\uparrow^*\left(+\frac{\hbar}{2}\right) + \alpha_\downarrow\alpha_\downarrow^*\left(-\frac{\hbar}{2}\right). \tag{4.131}$$

The formal way (127), based on the general Eq. (125), has, however, the advantage of being applicable to finding the observables whose operators are *not* diagonal in the *z*-basis, as well. In particular, absolutely similar calculations give

$$\langle S_x\rangle = \alpha_\uparrow\alpha_\uparrow^*\langle\uparrow|\hat{S}_x|\uparrow\rangle + \alpha_\downarrow\alpha_\downarrow^*\langle\downarrow|\hat{S}_x|\downarrow\rangle + \alpha_\uparrow\alpha_\downarrow^*\langle\downarrow|\hat{S}_x|\uparrow\rangle + \alpha_\downarrow\alpha_\uparrow^*\langle\uparrow|\hat{S}_x|\downarrow\rangle = \frac{\hbar}{2}\left(\alpha_\uparrow\alpha_\downarrow^* + \alpha_\downarrow\alpha_\uparrow^*\right), \tag{4.132}$$

$$\langle S_y\rangle = \alpha_\uparrow\alpha_\uparrow^*\langle\uparrow|\hat{S}_y|\uparrow\rangle + \alpha_\downarrow\alpha_\downarrow^*\langle\downarrow|\hat{S}_y|\downarrow\rangle + \alpha_\uparrow\alpha_\downarrow^*\langle\downarrow|\hat{S}_y|\uparrow\rangle + \alpha_\downarrow\alpha_\uparrow^*\langle\uparrow|\hat{S}_y|\downarrow\rangle = i\frac{\hbar}{2}\left(\alpha_\uparrow\alpha_\downarrow^* - \alpha_\downarrow\alpha_\uparrow^*\right), \tag{4.133}$$

Let us have a good look at a particular spin state, for example the spin-up state ↑. According to Eq. (126), in this state $\alpha_\uparrow = 1$ and $\alpha_\downarrow = 0$, so Eqs. (130)-(133) yield:

$$\langle S_z\rangle = \frac{\hbar}{2}, \qquad \langle S_x\rangle = \langle S_y\rangle = 0. \tag{4.134}$$

Now let us use the same Eq. (125) to calculate the spin component uncertainties. According to Eqs. (105) and (116)-(117), the operator of each spin component squared is equal to $(\hbar/2)^2\hat{I}$, so the general Eq. (1.33) yields

$$(\delta S_z)^2 = \langle S_z^2 \rangle - \langle S_z \rangle^2 = \langle \uparrow | \hat{S}_z^2 | \uparrow \rangle - \left(\frac{\hbar}{2}\right)^2 = \left(\frac{\hbar}{2}\right)^2 \langle \uparrow | \hat{I} | \uparrow \rangle - \left(\frac{\hbar}{2}\right)^2 = 0, \tag{4.135a}$$

$$(\delta S_x)^2 = \langle S_x^2 \rangle - \langle S_x \rangle^2 = \langle \uparrow | \hat{S}_x^2 | \uparrow \rangle - 0 = \left(\frac{\hbar}{2}\right)^2 \langle \uparrow | \hat{I} | \uparrow \rangle = \left(\frac{\hbar}{2}\right)^2, \tag{4.135b}$$

$$(\delta S_y)^2 = \langle S_y^2 \rangle - \langle S_y \rangle^2 = \langle \uparrow | \hat{S}_y^2 | \uparrow \rangle - 0 = \left(\frac{\hbar}{2}\right)^2 \langle \uparrow | \hat{I} | \uparrow \rangle = \left(\frac{\hbar}{2}\right)^2. \tag{4.135c}$$

While Eqs. (134) and (135a) are compatible with the classical notion of the angular momentum of magnitude $\hbar/2$ being directed exactly along the z-axis, this correspondence should not be overstretched, because such a classical picture cannot explain Eqs. (135b) and (135c). The best (but still imprecise!) classical image I can offer is the spin vector **S** oriented, on average, in the z-direction, but still having its x- and y-components strongly "wobbling" (fluctuating) about their zero average values.

It is straightforward to verify that in the x-polarized and y-polarized states, the situation is similar, with the corresponding change of axis indices. Thus, in neither of these states, all three spin components have definite values. Let me show that this is not just an occasional fact, but reflects one of the most profound properties of quantum mechanics, the *uncertainty relations*. For that, let us consider two measurable observables, A and B, of the same quantum system. There are two possibilities here. If the operators corresponding to these observables commute,

$$\left[\hat{A}, \hat{B}\right] = 0, \tag{4.136}$$

then all matrix elements of the commutator in any orthogonal basis (in particular, in the basis of eigenstates a_j of the operator $\hat{A}$) have to equal zero:

$$\langle a_j | \left[\hat{A}, \hat{B}\right] | a_{j'} \rangle \equiv \langle a_j | \hat{A}\hat{B} | a_{j'} \rangle - \langle a_j | \hat{B}\hat{A} | a_{j'} \rangle = 0. \tag{4.137}$$

In the first bracket of the middle expression, let us act by the (Hermitian!) operator $\hat{A}$ on the bra-vector, while in the second one, on the ket-vector. According to Eq. (68), such action turns the operators into the corresponding eigenvalues, which may be taken out of the long brackets, so we get

$$A_j \langle a_j | \hat{B} | a_{j'} \rangle - A_{j'} \langle a_j | \hat{B} | a_{j'} \rangle \equiv \left(A_j - A_{j'} \right) \langle a_j | \hat{B} | a_{j'} \rangle = 0. \tag{4.138}$$

This means that if all eigenstates of the operator $\hat{A}$ are non-degenerate (i.e. $A_j \neq A_{j'}$ if $j \neq j'$), the matrix of the operator $\hat{B}$ has to be diagonal in the basis $\{a\}$, i.e., the operators $\hat{A}$ and $\hat{B}$ have common eigenstates. Such pairs of observables (and their operators) that can share their eigenstates are called *compatible*. For example, in the wave mechanics of a particle, its momentum (1.26) and kinetic energy (1.27) are compatible, sharing their eigenfunctions (1.29). Now we see that this is not occasional, because each Cartesian component of the kinetic energy is proportional to the square of the corresponding component of the momentum, and any operator commutes with an arbitrary integer power of itself:

$$\left[\hat{A}, \hat{A}^n\right] \equiv \left[\hat{A}, \underbrace{\hat{A}\hat{A}...\hat{A}}_{n}\right] = \hat{A}\underbrace{\hat{A}\hat{A}...\hat{A}}_{n} - \underbrace{\hat{A}\hat{A}...\hat{A}}_{n}\hat{A} = 0. \tag{4.139}$$

Now, what if the operators $\hat{A}$ and $\hat{B}$ do *not* commute? Then the following *general uncertainty relation* is valid:

General uncertainty relation

$$\delta A\,\delta B \geq \frac{1}{2}\left|\left\langle\left[\hat{A},\hat{B}\right]\right\rangle\right|, \tag{4.140}$$

where all expectation values are for the same but arbitrary state of the system. The proof of Eq. (140) may be divided into two steps, the first one proving the so-called *Schwartz inequality* for any two possible states, say α and β:[34]

Schwartz inequality

$$\langle\alpha|\alpha\rangle\langle\beta|\beta\rangle \geq \left|\langle\alpha|\beta\rangle\right|^2. \tag{4.141}$$

Its proof may be readily achieved by applying the postulate (16) – that the norm of any legitimate state of the system cannot be negative – to the state with the following ket-vector:

$$|\delta\rangle \equiv |\alpha\rangle - \frac{\langle\beta|\alpha\rangle}{\langle\beta|\beta\rangle}|\beta\rangle, \tag{4.142}$$

where α and β are possible, non-null states of the system, so the denominator in Eq. (142) is not equal to zero. For this case, Eq. (16) gives

$$\left(\langle\alpha| - \frac{\langle\alpha|\beta\rangle}{\langle\beta|\beta\rangle}\langle\beta|\right)\left(|\alpha\rangle - \frac{\langle\beta|\alpha\rangle}{\langle\beta|\beta\rangle}|\beta\rangle\right) \geq 0. \tag{4.143}$$

Opening the parentheses, we get

$$\langle\alpha|\alpha\rangle - \frac{\langle\alpha|\beta\rangle}{\langle\beta|\beta\rangle}\langle\beta|\alpha\rangle - \frac{\langle\beta|\alpha\rangle}{\langle\beta|\beta\rangle}\langle\alpha|\beta\rangle + \frac{\langle\alpha|\beta\rangle\langle\beta|\alpha\rangle}{\langle\beta|\beta\rangle^2}\langle\beta|\beta\rangle \geq 0. \tag{4.144}$$

After the cancellation of one inner product $\langle\beta|\beta\rangle$ in the numerator and the denominator of the last term, it cancels with the 2nd (or the 3rd) term. What remains is the Schwartz inequality (141).

Now let us apply this inequality to states

$$|\alpha\rangle \equiv \tilde{\hat{A}}|\gamma\rangle \quad \text{and} \quad |\beta\rangle \equiv \tilde{\hat{B}}|\gamma\rangle, \tag{4.145}$$

where, in both relations, γ is the same possible state of the system, and the deviation operators are defined similarly to the deviations of the observables (see Sec. 1.2):

$$\tilde{\hat{A}} \equiv \hat{A} - \langle A\rangle, \qquad \tilde{\hat{B}} \equiv \hat{B} - \langle B\rangle. \tag{4.146}$$

With this substitution, and taking into account again that the observable operators $\hat{A}$ and $\hat{B}$ are Hermitian, Eq. (141) yields

$$\langle\gamma|\tilde{\hat{A}}^2|\gamma\rangle\langle\gamma|\tilde{\hat{B}}^2|\gamma\rangle \geq \left|\langle\gamma|\tilde{\hat{A}}\tilde{\hat{B}}|\gamma\rangle\right|^2. \tag{4.147}$$

Since the state γ is arbitrary, we may use Eq. (125) to rewrite this relation as an operator inequality:

[34] This inequality is the quantum-mechanical analog of the usual vector algebra's result $\alpha^2\beta^2 \geq |\boldsymbol{\alpha}\cdot\boldsymbol{\beta}|^2$.

$$\delta A\,\delta B \geq \left|\left\langle \hat{\tilde{A}}\hat{\tilde{B}} \right\rangle\right|. \tag{4.148}$$

Actually, this is already an uncertainty relation, even "better" (stronger) than its standard form (140); moreover, it is more convenient in some cases. To prove Eq. (140), we need a couple of more steps. First, let us notice that the operator product participating in Eq. (148) may be recast as

$$\hat{\tilde{A}}\hat{\tilde{B}} = \frac{1}{2}\left\{\hat{\tilde{A}},\hat{\tilde{B}}\right\} - \frac{i}{2}\hat{C}, \quad \text{where } \hat{C} \equiv i\left[\hat{\tilde{A}},\hat{\tilde{B}}\right]. \tag{4.149}$$

Any anticommutator of Hermitian operators, including that in Eq. (149), is a Hermitian operator, and its eigenvalues are purely real, so its expectation value (in any state) is also purely real. On the other hand, the commutator part of Eq. (149) is just

$$\hat{C} \equiv i\left[\hat{\tilde{A}},\hat{\tilde{B}}\right] \equiv i\left(\hat{A}-\langle A\rangle\right)\left(\hat{B}-\langle B\rangle\right) - i\left(\hat{B}-\langle B\rangle\right)\left(\hat{A}-\langle A\rangle\right) \equiv i\left(\hat{A}\hat{B}-\hat{B}\hat{A}\right) \equiv i\left[\hat{A},\hat{B}\right]. \tag{4.150}$$

Second, according to Eqs. (52) and (65), the Hermitian conjugate of any product of the Hermitian operators $\hat{A}$ and $\hat{B}$ is just the product of these operators swapped. Using this fact, we may write

$$\hat{C}^\dagger = \left(i\left[\hat{A},\hat{B}\right]\right)^\dagger = -i\left(\hat{A}\hat{B}\right)^\dagger + i\left(\hat{B}\hat{A}\right)^\dagger = -i\hat{B}\hat{A} + i\hat{A}\hat{B} = i\left[\hat{A},\hat{B}\right] = \hat{C}, \tag{4.151}$$

so the operator $\hat{C}$ is also Hermitian, i.e. its eigenvalues are also real, and thus its expectation value is purely real as well. As a result, the square of the expectation value of the operator product (149) may be represented as

$$\left\langle \hat{\tilde{A}}\hat{\tilde{B}} \right\rangle^2 = \left\langle \frac{1}{2}\left\{\hat{\tilde{A}},\hat{\tilde{B}}\right\} \right\rangle^2 + \left\langle \frac{1}{2}\hat{C} \right\rangle^2. \tag{4.152}$$

Since the first term on the right-hand side of this equality cannot be negative, we may write

$$\left\langle \hat{\tilde{A}}\hat{\tilde{B}} \right\rangle^2 \geq \left\langle \frac{1}{2}\hat{C} \right\rangle^2 \equiv \left\langle \frac{i}{2}\left[\hat{A},\hat{B}\right] \right\rangle^2, \tag{4.153}$$

and hence continue Eq. (148) as

$$\delta A\,\delta B \geq \left|\left\langle \hat{\tilde{A}}\hat{\tilde{B}} \right\rangle\right| \geq \frac{1}{2}\left|\left\langle \left[\hat{A},\hat{B}\right] \right\rangle\right|, \tag{4.154}$$

thus proving Eq. (140).

For the particular case of operators $\hat{x}$ and $\hat{p}_x$ (or a similar pair of operators for another Cartesian coordinate), we may readily combine Eq. (140) with Eq. (2.14b) to prove the original Heisenberg's uncertainty relation (2.13). For the spin-½ operators defined by Eq. (116)-(117), it is very simple (and highly recommended to the reader) to show that

$$\left[\hat{\sigma}_j,\hat{\sigma}_{j'}\right] = 2i\sum_{j''=1}^{3}\varepsilon_{jj'j''}\hat{\sigma}_{j''}, \qquad \text{i.e. } \left[\hat{S}_j,\hat{S}_{j'}\right] = i\hbar\sum_{j''=1}^{3}\varepsilon_{jj'j''}\hat{S}_{j''}, \tag{4.155}$$

Spin- ½: commutation relations

where $\varepsilon_{jj'j''}$ is the Levi-Civita permutation symbol.[35] As a result, the uncertainty relations (140) for all Cartesian components of spin-½ systems are similar, for example

[35] See, e.g., MA Eq. (13.2).

Spin-½: uncertainty relations

$$\delta S_x \delta S_y \geq \frac{\hbar}{2}\left|\langle S_z \rangle\right|, \text{ etc.} \tag{4.156}$$

In particular, as we already know, in the ↑ state the right-hand side of this relation equals $(\hbar/2)^2 > 0$, so neither of the uncertainties δS_x, δS_y can equal zero. As a reminder, our direct calculation earlier in this section has shown that each of these uncertainties is equal to $\hbar/2$, i.e. their product is equal to the lowest value allowed by the uncertainty relation (156) – just as the Gaussian wave packets (2.16) provide the lowest possible value of the product $\delta x \delta p_x$, allowed by the Heisenberg relation (2.13).

4.6. Quantum dynamics: Three pictures

So far in this chapter, I shied away from the discussion of the system's dynamics, implying that the bra- and ket-vectors were just their "snapshots" at a certain instant t. Now we are sufficiently prepared to examine their evolution in time. One of the most beautiful features of quantum mechanics is that this evolution may be described using either of three alternatives (called *pictures*), giving exactly the same final results for the expectation values of all observables.

From the standpoint of our wave-mechanics experience, the *Schrödinger picture* is the most natural one. In this picture, the operators corresponding to time-independent observables (e.g., to the Hamiltonian function H of an isolated system) are also constant in time, while the bra- and ket-vectors evolve in time as

$$\langle \alpha(t)| = \langle \alpha(t_0)|\, \hat{u}^\dagger(t,t_0), \qquad |\alpha(t)\rangle = \hat{u}(t,t_0)\,|\alpha(t_0)\rangle. \tag{4.157a}$$

Here $\hat{u}(t,t_0)$ is the *time-evolution operator*, which obeys the following differential equation:

$$i\hbar \frac{\partial}{\partial t}\hat{u} = \hat{H}\hat{u}, \tag{4.157b}$$

where $\hat{H}$ is the Hamiltonian operator of the system – which is always Hermitian: $\hat{H}^\dagger = \hat{H}$, and t_0 is the initial moment of time. (Note that Eqs. (157) remain valid even if the Hamiltonian depends on time explicitly.) Differentiating the second of Eqs. (157a) over time t, and then using Eq. (157b) twice, we can merge these two relations into a single equation, without explicit use of the time-evolution operator:

Schrödinger equation

$$i\hbar \frac{\partial}{\partial t}|\alpha(t)\rangle = \hat{H}|\alpha(t)\rangle, \tag{4.158}$$

which is frequently more convenient. (However, for some purposes the notion of the time-evolution operator, together with Eq. (157b), are useful – as we will see in a minute.) While Eq. (158) is a very natural generalization of the wave-mechanical equation (1.25), and is also frequently called the *Schrödinger equation*,[36] it still should be considered as a new, more general postulate, which finds its final justification (as it is usual in physics) in the agreement of its corollaries with experiment – more exactly, in the absence of a single credible contradiction to an experiment.

Starting the discussion of Eq. (158), let us first consider the case of a time-independent Hamiltonian, whose eigenstates a_n and eigenvalues E_n obey Eq. (68) for this operator:[37]

[36] Moreover, we will be able to *derive* Eq. (1.25) from Eq. (158) – see below.

[37] I have switched the state index notation from j to n, which was used for numbering stationary states in Chapter 1, to emphasize the special role played by the stationary states a_n in quantum dynamics.

$$\hat{H}|a_n\rangle = E_n|a_n\rangle, \tag{4.159}$$

and hence are also time-independent. (Similarly to the wavefunctions ψ_n defined by Eq. (1.60), a_n are called the *stationary states* of the system.) Let us use Eqs. (158)-(159) to calculate the law of time evolution of the expansion coefficients α_n (i.e. the probability amplitudes) defined by Eq. (118), in a stationary state basis, using Eq. (158):

$$\dot{\alpha}_n(t) = \frac{d}{dt}\langle a_n|\alpha(t)\rangle = \langle a_n|\frac{d}{dt}|\alpha(t)\rangle = \langle a_n|\frac{1}{i\hbar}\hat{H}|\alpha(t)\rangle = \frac{E_n}{i\hbar}\langle a_n|\alpha(t)\rangle = -\frac{i}{\hbar}E_n\alpha_n. \tag{4.160}$$

This is the same simple equation as Eq. (1.61), and its integration, with the initial moment t_0 taken for 0, yields a similar result – cf. Eq. (1.62):

$$\boxed{\alpha_n(t) = \alpha_n(0)\exp\left\{-\frac{i}{\hbar}E_n t\right\}.} \tag{4.161}$$

Time evolution of probability amplitudes

In order to illustrate how this result works, let us consider the dynamics of a spin-½ in a time-independent, uniform external magnetic field $\mathscr{B}$. To construct the system's Hamiltonian, we may apply the correspondence principle to the classical expression for the energy of a magnetic moment **m** in the external magnetic field $\mathscr{B}$, [38]

$$U = -\mathbf{m}\cdot\mathscr{B}. \tag{4.162}$$

In quantum mechanics, the operator corresponding to the moment **m** is given by Eq. (115) (suggested by W. Pauli), so the spin-field interaction is described by the so-called *Pauli Hamiltonian*, which may be, due to Eqs. (116)-(117), represented in several equivalent forms:

$$\boxed{\hat{H} = -\hat{\mathbf{m}}\cdot\mathscr{B} \equiv -\gamma\hat{\mathbf{S}}\cdot\mathscr{B} = -\gamma\frac{\hbar}{2}\hat{\boldsymbol{\sigma}}\cdot\mathscr{B}.} \tag{4.163a}$$

Pauli Hamiltonian: operator

If the z-axis is aligned with the field's direction, this expression is reduced to

$$\hat{H} = -\gamma\mathscr{B}\hat{S}_z \equiv -\gamma\mathscr{B}\frac{\hbar}{2}\hat{\sigma}_z. \tag{4.163b}$$

According to Eq. (117), in the z-basis of the spin states ↑ and ↓, the matrix of the operator (163b) is

$$\boxed{\mathrm{H} = -\frac{\gamma\hbar\mathscr{B}}{2}\sigma_z \equiv \frac{\hbar\Omega}{2}\sigma_z, \quad \text{where } \Omega \equiv -\gamma\mathscr{B}.} \tag{4.164}$$

Pauli Hamiltonian: z-basis matrix

The constant Ω so defined coincides with the classical frequency of the precession, about the z-axis, of an axially-symmetric rigid body (the so-called *symmetric top*), with an angular momentum **S** and the magnetic moment $\mathbf{m} = \gamma\mathbf{S}$, induced by the external torque $\boldsymbol{\tau} = \mathbf{m}\times\mathscr{B}$.[39] (For an electron, with its negative gyromagnetic ratio $\gamma_e = -g_e e/2m_e$, neglecting the tiny difference of the g_e-factor from 2, we get

$$\Omega = \frac{e}{m_e}\mathscr{B}, \tag{4.165}$$

so according to Eq. (3.48), the frequency Ω coincides with the electron's cyclotron frequency ω_c.)

[38] See, e.g., EM Eq. (5.100). As a reminder, we have already used this expression for the derivation of Eq. (3).
[39] See, e.g., CM Sec. 4.5, in particular Eq. (4.72), and EM Sec. 5.5, in particular Eq. (5.114) and its discussion.

In order to apply the general Eq. (161) to this case, we need to find the eigenstates a_n and eigenenergies E_n of our Hamiltonian. However, with our (smart :-) choice of the z-axis, the Hamiltonian matrix is already diagonal:

$$\hat{H} = \frac{\hbar\Omega}{2}\sigma_z \equiv \frac{\hbar\Omega}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{4.166}$$

meaning that the states ↑ and ↓ are the eigenstates of this system, with the eigenenergies, respectively, [40]

Spin-½ in magnetic field: eigenenergies

$$E_\uparrow = +\frac{\hbar\Omega}{2} \quad \text{and} \quad E_\downarrow = -\frac{\hbar\Omega}{2}. \tag{4.167}$$

Note that their difference,

$$\Delta E \equiv \left|E_\uparrow - E_\downarrow\right| = \hbar\left|\Omega\right| = \hbar\left|\gamma\mathscr{B}\right|, \tag{4.168}$$

corresponds to the classical energy $2|m\mathscr{B}|$ of flipping a magnetic dipole with the moment's magnitude $m = \gamma\hbar/2$, oriented along the direction of the field $\mathscr{B}$. Note also that if the product $\gamma\mathscr{B}$ is positive, then Ω is negative, so $E_\uparrow$ is negative, while $E_\downarrow$ is positive. This is in agreement with the classical picture of a magnetic dipole **m** having negative potential energy when it is aligned with the external magnetic field $\mathscr{B}$ – see Eq. (162) again.

So, for the time evolution of the probability amplitudes of these states, Eq. (161) immediately yields the following expressions:

$$\alpha_\uparrow(t) = \alpha_\uparrow(0)\exp\left\{-\frac{i}{2}\Omega t\right\}, \qquad \alpha_\downarrow(t) = \alpha_\downarrow(0)\exp\left\{+\frac{i}{2}\Omega t\right\}, \tag{4.169}$$

allowing a ready calculation of the time evolution of the expectation values of any observable. In particular, we can calculate the expectation value of S_z as a function of time by applying Eq. (130) to the (arbitrary) time moment t:

$$\langle S_z\rangle(t) = \frac{\hbar}{2}\left[\alpha_\uparrow(t)\alpha_\uparrow^*(t) - \alpha_\downarrow(t)\alpha_\downarrow^*(t)\right] = \frac{\hbar}{2}\left[\alpha_\uparrow(0)\alpha_\uparrow^*(0) - \alpha_\downarrow(0)\alpha_\downarrow^*(0)\right] = \langle S_z\rangle(0). \tag{4.170}$$

Thus the expectation value of the spin component parallel to the applied magnetic field remains constant in time, regardless of the initial state of the system. However, this is not true for the components perpendicular to the field. For example, Eq. (132), applied to the moment t, gives

$$\langle S_x\rangle(t) = \frac{\hbar}{2}\left[\alpha_\uparrow(t)\alpha_\downarrow^*(t) + \alpha_\downarrow(t)\alpha_\uparrow^*(t)\right] = \frac{\hbar}{2}\left[\alpha_\uparrow(0)\alpha_\downarrow^*(0)e^{-i\Omega t} + \alpha_\downarrow(0)\alpha_\uparrow^*(0)e^{+i\Omega t}\right]. \tag{4.171}$$

Clearly, this expression describes sinusoidal oscillations with frequency (164). The amplitude and the phase of these oscillations depend on initial conditions. Indeed, solving Eqs. (132)-(133) for the probability amplitude products, we get the following relations:

$$\hbar\alpha_\downarrow(t)\alpha_\uparrow^*(t) = \langle S_x\rangle(t) + i\langle S_y\rangle(t), \qquad \hbar\alpha_\uparrow(t)\alpha_\downarrow^*(t) = \langle S_x\rangle(t) - i\langle S_y\rangle(t), \tag{4.172}$$

valid for any time t. Plugging their values for $t = 0$ into Eq. (171), we get

[40] So, spin-½ gives one more example of two-level systems whose discussion was started in Sec. 2.6. The fact that all quantum two-level systems are isomorphic (see Sec. 5.1) adds importance to our current discussion.

$$\begin{aligned}\langle S_x\rangle(t) &= \frac{1}{2}\left[\langle S_x\rangle(0)+i\langle S_y\rangle(0)\right]e^{+i\Omega t}+\frac{1}{2}\left[\langle S_x\rangle(0)-i\langle S_y\rangle(0)\right]e^{-i\Omega t}\\ &\equiv \langle S_x\rangle(0)\cos\Omega t-\langle S_y\rangle(0)\sin\Omega t\,.\end{aligned} \tag{4.173}$$

An absolutely similar calculation using Eq. (133) gives

$$\langle S_y\rangle(t)=\langle S_y\rangle(0)\cos\Omega t+\langle S_x\rangle(0)\sin\Omega t\,. \tag{4.174}$$

These formulas show, for example, that if at moment $t = 0$ the spin's state was ↑, i.e. $\langle S_x\rangle(0) = \langle S_y\rangle(0) = 0$, then the oscillation amplitudes of both "lateral" components of the spin vanish. On the other hand, if the spin was initially in the state →, i.e. had the definite, largest possible value of S_x equal to $\hbar/2$ (in classics, we would say "the spin-½ was oriented in the x-direction"), then both expectation values $\langle S_x\rangle$ and $\langle S_y\rangle$ oscillate in time[41] with this amplitude, and with the phase shift $\pi/2$ between them.

So, the quantum-mechanical results for the expectation values of the Cartesian components of spin-½ are indistinguishable from the classical results for the precession, with the frequency $\Omega = -\gamma\mathscr{B}$,[42] of a symmetric top with the angular momentum $\mathbf{L}$ of magnitude $\hbar/2$, about the field's direction (our axis z), under the effect of an external torque $\boldsymbol{\tau} = \mathbf{m}\times\mathscr{B}$ exerted by the field $\mathscr{B}$ on the magnetic moment $\mathbf{m} = \gamma\mathbf{L}$. Note, however, that the classical language does not describe the large quantum-mechanical uncertainties of the components, obeying Eqs. (156), which are absent in the classical picture – at least when the precession starts from a definite orientation of the angular momentum vector.[43]

Recall also that at the stationary orbital motion of a particle, the component L_z of its angular momentum is always a multiple of $\hbar$ – see, e.g., Eq. (3.139). As a result, the angular momentum of a spin-½ particle, with its stationary values $S_z = \pm\hbar/2$, cannot be explained by the summation of orbital moments of its hypothetical components, i.e. by any internal rotation of the particle about its axis.

After this illustration, let us return to the discussion of the general Schrödinger equation (157b) and prove the following fascinating fact: it is possible to write the general solution of this *operator* equation. In the easiest case when the Hamiltonian is time-independent, this solution turns out to be an exact analog of Eq. (161),

$$\hat{u}(t,t_0)=\hat{u}(t_0,t_0)\exp\left\{-\frac{i}{\hbar}\hat{H}(t-t_0)\right\}=\exp\left\{-\frac{i}{\hbar}\hat{H}(t-t_0)\right\}. \tag{4.175}$$

To start its proof we should, first of all, understand what a function (in this particular case, the exponent) of an operator means. In the operator (and matrix) algebra, such nonlinear functions are *defined* by their Taylor expansions; in particular, Eq. (175) means that

---

[41] This is one more (hopefully, redundant :-) illustration of the difference between the averaging over the statistical ensemble and that over time: in Eqs. (170), (173)-(174), and also in quite a few relations below, only the former averaging has been performed, so the results are still functions of time.

[42] Note that according to this relation, the gyromagnetic ratio γ may be interpreted as the angular frequency of the spin precession in a unit magnetic field – hence the name. In particular, for electrons, $|\gamma_e| \approx 1.761\times10^{11}$ s^{-1}T^{-1}; for protons, the ratio is much smaller, $\gamma_p \equiv g_p e/2m_p \approx 2.675\times10^{8}$ s^{-1}T^{-1} – mostly because of their larger mass m_p, at a g-factor of the same order as for the electron: $g_p \approx 5.586$. For heavier spin-½ particles, e.g., atomic nuclei with such spin, the values of γ are correspondingly smaller – e.g., $\gamma \approx 8.681\times10^{6}$ s^{-1}T^{-1} for the ^{57}Fe nucleus.

[43] If the initial conditions are random, the classical motion is stochastic even if its laws are deterministic.

$$\hat{u}(t,t_0) = \hat{I} + \sum_{k=1}^{\infty} \frac{1}{k!}\left[-\frac{i}{\hbar}\hat{H}(t-t_0)\right]^k$$
$$\equiv \hat{I} + \frac{1}{1!}\left(-\frac{i}{\hbar}\right)\hat{H}(t-t_0) + \frac{1}{2!}\left(-\frac{i}{\hbar}\right)^2 \hat{H}^2(t-t_0)^2 + \frac{1}{3!}\left(-\frac{i}{\hbar}\right)^3 \hat{H}^3(t-t_0)^3 + ..., \quad (4.176)$$

where $\hat{H}^2 \equiv \hat{H}\hat{H}$, $\hat{H}^3 \equiv \hat{H}\hat{H}\hat{H}$, etc. Working with such a series of operator products is not as hard as one could imagine, due to their regular structure. For example, let us differentiate both sides of Eq. (176) over t, at constant t_0, at the last step using this equality again – that time, backward:

$$\frac{\partial}{\partial t}\hat{u}(t,t_0) = \hat{0} + \frac{1}{1!}\left(-\frac{i}{\hbar}\right)\hat{H} + \frac{1}{2!}\left(-\frac{i}{\hbar}\right)^2 \hat{H}^2 2(t-t_0) + \frac{1}{3!}\left(-\frac{i}{\hbar}\right)^2 \hat{H}^3 3(t-t_0)^2 + ...$$
$$\equiv \left(-\frac{i}{\hbar}\right)\hat{H}\left[\hat{I} + \frac{1}{1!}\left(-\frac{i}{\hbar}\right)\hat{H}(t-t_0) + \frac{1}{2!}\left(-\frac{i}{\hbar}\right)^2 \hat{H}^2(t-t_0)^2\right] + ... \equiv -\frac{i}{\hbar}\hat{H}\hat{u}(t,t_0), \quad (4.177)$$

so the differential equation (158) is indeed satisfied. On the other hand, Eq. (175) also satisfies the initial condition

$$\hat{u}(t_0,t_0) = \hat{u}^{\dagger}(t_0,t_0) = \hat{I} \quad (4.178)$$

that immediately follows from the definition (157a) of the evolution operator. Thus, Eq. (175) indeed gives the (unique) solution for the time evolution operator – in the Schrödinger picture.

Now let us allow the operator $\hat{H}$ to be a function of time, but with the condition that its "values" (in fact, operators) at different instants commute with each other:

$$\left[\hat{H}(t'), \hat{H}(t'')\right] = 0, \quad \text{for any } t', t''. \quad (4.179)$$

(A good example is the Pauli Hamiltonian (4.163) for a spin in a classical magnetic field $\mathscr{B}$ even if it depends on time. Indeed, the spin operator $\hat{\mathbf{S}}$ does not depend explicitly on time and hence commutes with itself as well as with the c-numbers $\mathscr{B}(t')$ and $\mathscr{B}(t'')$. Note, however, that a similar operator describing the effect of a classical position-independent force $\mathbf{F}(t)$ on the orbital motion of a particle,

$$\hat{H}_F = -\mathbf{F}(t)\cdot\hat{\mathbf{r}}, \quad (4.180)$$

may be deceiving: though it satisfies Eq. (179), this relation is invalid for the particle's *full* Hamiltonian including its kinetic energy.) In this case, it is sufficient to replace, in all the above formulas, the product $\hat{H}(t-t_0)$ with the corresponding integral over time; in particular, Eq. (175) is generalized as

Evolution operator: explicit expression

$$\hat{u}(t,t_0) = \exp\left\{-\frac{i}{\hbar}\int_{t_0}^{t}\hat{H}(t')dt'\right\}. \quad (4.181)$$

This replacement means that the first form of Eq. (176) should be replaced with

$$\hat{u}(t,t_0) = \hat{I} + \sum_{k=1}^{\infty}\frac{1}{k!}\left(-\frac{i}{\hbar}\right)^k\left(\int_{t_0}^{t}\hat{H}(t')dt'\right)^k \equiv \hat{I} + \sum_{k=1}^{\infty}\frac{1}{k!}\left(-\frac{i}{\hbar}\right)^k\int_{t_0}^{t}dt_1\int_{t_0}^{t}dt_2...\int_{t_0}^{t}dt_k\hat{H}(t_1)\hat{H}(t_2)...\hat{H}(t_k). \quad (4.182)$$

The proof that Eq. (182) satisfies Eq. (158) is absolutely similar to the one carried out above.

We may now use Eq. (181) to show that the time-evolution operator remains unitary at any moment, even for a time-dependent Hamiltonian, if it satisfies Eq. (179). Indeed, Eq. (181) yields

$$\hat{u}(t,t_0)\hat{u}^{\dagger}(t,t_0)=\exp\left\{-\frac{i}{\hbar}\int_{t_0}^{t}\hat{H}(t')dt'\right\}\exp\left\{+\frac{i}{\hbar}\int_{t_0}^{t}\hat{H}(t'')dt''\right\}. \tag{4.183}$$

Since each of these exponents may be represented with the Taylor series (182), and, thanks to Eq. (179), different components of these sums may be swapped at will, the expression (183) may be manipulated exactly as the product of *c*-number exponents, for example rewritten as

$$\hat{u}(t,t_0)\hat{u}^{\dagger}(t,t_0)=\exp\left\{-\frac{i}{\hbar}\left[\int_{t_0}^{t}\hat{H}(t')dt'-\int_{t_0}^{t}\hat{H}(t'')dt''\right]\right\}=\exp\{\hat{0}\}=\hat{I}. \tag{4.184}$$

This property ensures, in particular, that the system state's normalization does not depend on time:

$$\langle\alpha(t)|\alpha(t)\rangle=\langle\alpha(t_0)|\hat{u}^{\dagger}(t,t_0)\hat{u}(t,t_0)|\alpha(t_0)\rangle=\langle\alpha(t_0)|\alpha(t_0)\rangle. \tag{4.185}$$

The most difficult cases for the explicit solution of Eq. (158) are those where Eq. (179) is violated.[44] It may be proved that in these cases, Eqs. (181)-(182) should be replaced with the following *Dyson series* using the so-called *time-ordering operator* $\hat{\mathcal{T}}$:

$$\hat{u}(t,t_0)=\hat{\mathcal{T}}\exp\left\{-\frac{i}{\hbar}\int_{t_0}^{t}\hat{H}(t')dt'\right\}=\hat{I}+\sum_{k=1}^{\infty}\frac{1}{k!}\left(-\frac{i}{\hbar}\right)^{k}\int_{t_0}^{t}dt_1\int_{t_0}^{t}dt_2\ldots\int_{t_0}^{t}dt_k\,\hat{\mathcal{T}}\left[\hat{H}(t_1)\hat{H}(t_2)\ldots\hat{H}(t_k)\right],$$
$$\text{where }\hat{\mathcal{T}}\left[\hat{H}(t_1)\hat{H}(t_2)\right]\equiv\begin{cases}\hat{H}(t_1)\hat{H}(t_2), & \text{for } t_2\le t_1,\\ \hat{H}(t_2)\hat{H}(t_1), & \text{for } t_1\le t_2.\end{cases} \tag{4.186}$$

Since we would not have time/space to use this relation in this course, I will skip its proof.[45]

Let me now return to the general discussion of quantum dynamics to outline its alternative, the *Heisenberg picture*. For its introduction, let us recall that according to Eq. (125), in quantum mechanics the expectation value of any observable A is a long bracket. Let us explore the even more general form of such a bracket:

$$\langle\alpha|\hat{A}|\beta\rangle, \tag{4.187}$$

because in some applications, the states α and β may be different. As was discussed above, in the Schrödinger picture the bra- and ket-vectors of the states evolve in time, while the operators of observables remain time-independent (if they do not explicitly depend on time). As a result, Eq. (187) applied to the moment t, may be represented as

$$\langle\alpha(t)|\hat{A}_{\rm S}|\beta(t)\rangle, \tag{4.188}$$

where the index "S" is added to emphasize the Schrödinger picture. Let us apply the evolution law (157a) to the bra- and ket-vectors in this expression:

$$\langle\alpha(t)|\hat{A}_{\rm S}|\beta(t)\rangle=\langle\alpha(t_0)|\hat{u}^{\dagger}(t,t_0)\hat{A}_{\rm S}\hat{u}(t,t_0)|\beta(t_0)\rangle. \tag{4.189}$$

[44] We will run into such situations in Chapter 7, but will not need to apply Eq. (186) there.

[45] It may be found, for example, in Chapter 5 of J. Sakurai's textbook – see *References*.

This equality means that if we form a long bracket with bra- and ket-vectors of the initial-time states, together with the following time-dependent *Heisenberg operator*[46]

Heisenberg operator

$$\hat{A}_{\rm H}(t) \equiv \hat{u}^{\dagger}(t,t_0)\hat{A}_{\rm S}\hat{u}(t,t_0) = \hat{u}^{\dagger}(t,t_0)\hat{A}_{\rm H}(t_0)\hat{u}(t,t_0)\,, \tag{4.190}$$

all experimentally measurable results will remain the same as in the Schrödinger picture:

Heisenberg picture

$$\langle\alpha(t)|\hat{A}|\beta(t)\rangle = \langle\alpha(t_0)|\hat{A}_{\rm H}(t,t_0)|\beta(t_0)\rangle. \tag{4.191}$$

For full clarity, let us see how the Heisenberg picture works for the same simple (but very important!) problem of the spin-½ precession in a *z*-oriented magnetic field, described (in the *z*-basis) by the Hamiltonian matrix (164). In that basis, Eq. (157b) for the time-evolution operator becomes

$$i\hbar\frac{\partial}{\partial t}\begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22}\end{pmatrix} = \frac{\hbar\Omega}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1\end{pmatrix}\begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22}\end{pmatrix} \equiv \frac{\hbar\Omega}{2}\begin{pmatrix} u_{11} & u_{12} \\ -u_{21} & -u_{22}\end{pmatrix}. \tag{4.192}$$

We see that in this simple case, the differential equations for different matrix elements of the evolution operator matrix are decoupled, and readily solvable by using the universal initial conditions (178):[47]

$$\mathrm{u}(t,0) = \begin{pmatrix} e^{-i\Omega t/2} & 0 \\ 0 & e^{i\Omega t/2}\end{pmatrix} \equiv \mathrm{I}\cos\frac{\Omega t}{2} - i\sigma_z\sin\frac{\Omega t}{2}. \tag{4.193}$$

Now let us use them in Eq. (190) to calculate the Heisenberg-picture operators of spin components – still in the *z*-basis. Dropping the index "H" for the notation brevity (the Heisenberg-picture operators are clearly marked by their dependence on time anyway), we get

$$\begin{aligned} \mathrm{S}_x(t) &= \mathrm{u}^{\dagger}(t,0)\mathrm{S}_x(0)\mathrm{u}(t,0) = \frac{\hbar}{2}\mathrm{u}^{\dagger}(t,0)\sigma_x\mathrm{u}(t,0) \\ &= \frac{\hbar}{2}\begin{pmatrix} e^{i\Omega t/2} & 0 \\ 0 & e^{-i\Omega t/2}\end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0\end{pmatrix}\begin{pmatrix} e^{-i\Omega t/2} & 0 \\ 0 & e^{i\Omega t/2}\end{pmatrix} \\ &= \frac{\hbar}{2}\begin{pmatrix} 0 & e^{i\Omega t} \\ e^{-i\Omega t} & 0\end{pmatrix} = \frac{\hbar}{2}\left(\sigma_x\cos\Omega t - \sigma_y\sin\Omega t\right) \equiv \mathrm{S}_x(0)\cos\Omega t - \mathrm{S}_y(0)\sin\Omega t\,. \end{aligned} \tag{4.194}$$

Absolutely similar calculations of the other spin components yield

$$\mathrm{S}_y(t) = \frac{\hbar}{2}\begin{pmatrix} 0 & -ie^{i\Omega t} \\ ie^{-i\Omega t} & 0\end{pmatrix} = \frac{\hbar}{2}\left(\sigma_y\cos\Omega t + \sigma_x\sin\Omega t\right) \equiv \mathrm{S}_y(0)\cos\Omega t + \mathrm{S}_x(0)\sin\Omega t\,, \tag{4.195}$$

[46] Note that this relation is similar in structure to the first of Eqs. (94), with the state bases $\{v\}$ and $\{u\}$ loosely associated with the time moments, respectively, t and t_0.

[47] We could of course use this solution, together with Eq. (157), to obtain all the above results for this system within the Schrödinger picture. In our simple case, the use of Eqs. (161) for this purpose was more straightforward, but in some cases, e.g., for some time-dependent Hamiltonians, an explicit calculation of the time-evolution matrix may be the best (or even the only practicable) way to proceed.

$$S_z(t) = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \frac{\hbar}{2}\sigma_z = S_z(0). \tag{4.196}$$

One practical advantage of these formulas is that they describe the system's evolution for *arbitrary* initial conditions, thus making the analysis of initial state effects very simple. Indeed, since in the Heisenberg picture, the expectation values of observables are calculated using Eq. (191) (with $\beta = \alpha$), with time-independent bra- and ket-vectors, such averaging of Eqs. (194)-(196) immediately returns us to Eqs. (170), (173), and (174), which were obtained above in the Schrödinger picture. Moreover, these equations for the Heisenberg operators formally coincide with the classical equations of the torque-induced precession for *c*-number variables. (Below we will see that the same exact correspondence is valid for the Heisenberg picture of the orbital motion.)

In order to see that the last fact is by no means a coincidence, let us combine Eqs. (157b) and (190) to form an explicit differential equation of the Heisenberg operator's evolution. For that, let us differentiate Eq. (190) over time:

$$\frac{d}{dt}\hat{A}_{\rm H} = \frac{\partial \hat{u}^\dagger}{\partial t}\hat{A}_{\rm S}\hat{u} + \hat{u}^\dagger\frac{\partial \hat{A}_{\rm S}}{\partial t}\hat{u} + \hat{u}^\dagger\hat{A}_{\rm S}\frac{\partial \hat{u}}{\partial t}. \tag{4.197}$$

Plugging in the derivatives of the time evolution operator from Eq. (157b) and its Hermitian conjugate, and multiplying both sides of the equation by $i\hbar$, we get

$$i\hbar\frac{d}{dt}\hat{A}_{\rm H} = -\hat{u}^\dagger\hat{H}\hat{A}_{\rm S}\hat{u} + i\hbar\hat{u}^\dagger\frac{\partial \hat{A}_{\rm S}}{\partial t}\hat{u} + \hat{u}^\dagger\hat{A}_{\rm S}\hat{H}\hat{u}. \tag{4.198a}$$

If for the Schrödinger-picture's Hamiltonian, the condition (179) is satisfied, then, according to Eqs. (177) or (182), the Hamiltonian commutes with the time evolution operator and its Hermitian conjugate, and may be swapped with any of them.[48] Hence, we may rewrite Eq. (198a) as

$$i\hbar\frac{d}{dt}\hat{A}_{\rm H} = -\hat{H}\hat{u}^\dagger\hat{A}_{\rm S}\hat{u} + i\hbar\hat{u}^\dagger\frac{\partial \hat{A}_{\rm S}}{\partial t}\hat{u} + \hat{u}^\dagger\hat{A}_{\rm S}\hat{u}\hat{H} \equiv i\hbar\hat{u}^\dagger\frac{\partial \hat{A}_{\rm S}}{\partial t}\hat{u} + \left[\hat{u}^\dagger\hat{A}_{\rm S}\hat{u}, \hat{H}\right]. \tag{4.198b}$$

Now using the definition (190) again, for both terms on the right-hand side, we may write

$$\boxed{i\hbar\frac{d}{dt}\hat{A}_{\rm H} = i\hbar\left(\frac{\partial \hat{A}}{\partial t}\right)_{\rm H} + \left[\hat{A}_{\rm H}, \hat{H}\right].} \tag{4.199}$$

Heisenberg equation of motion

This is the so-called *Heisenberg equation of motion*.

Let us see how this equation looks for the same problem of the spin-½ precession in a *z*-oriented, time-independent magnetic field described in the *z*-basis by the Hamiltonian matrix (164), which does not depend on time. In this basis, Eq. (199) for the operator vector of spin reads[49]

[48] Due to the same reason, $\hat{H}_{\rm H} \equiv \hat{u}^\dagger\hat{H}_{\rm S}\hat{u} = \hat{u}^\dagger\hat{u}\hat{H}_{\rm S} = \hat{H}_{\rm S}$; this is why the Hamiltonian operator's index may be dropped in Eqs. (198)-(199).

[49] Using the commutation relations (155), this equation may be readily generalized to the case of an arbitrary magnetic field $\mathscr{B}(t)$ and an arbitrary state basis – the exercise highly recommended to the reader.

$$i\hbar\begin{pmatrix}\dot{\mathbf{S}}_{11} & \dot{\mathbf{S}}_{12}\\ \dot{\mathbf{S}}_{21} & \dot{\mathbf{S}}_{22}\end{pmatrix}=\frac{\hbar\Omega}{2}\left[\begin{pmatrix}\mathbf{S}_{11} & \mathbf{S}_{12}\\ \mathbf{S}_{21} & \mathbf{S}_{22}\end{pmatrix},\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\right]=\hbar\Omega\begin{pmatrix}0 & -\mathbf{S}_{12}\\ \mathbf{S}_{21} & 0\end{pmatrix}. \tag{4.200}$$

Once again, the equations for different matrix elements are decoupled, and their solution is elementary:

$$\begin{aligned}&\mathbf{S}_{11}(t)=\mathbf{S}_{11}(0)=\text{const}, \quad &&\mathbf{S}_{22}(t)=\mathbf{S}_{22}(0)=\text{const},\\ &\mathbf{S}_{12}(t)=\mathbf{S}_{12}(0)e^{+i\Omega t}, \quad &&\mathbf{S}_{21}(t)=\mathbf{S}_{21}(0)e^{-i\Omega t}.\end{aligned} \tag{4.201}$$

According to Eq. (190), the initial values of the Heisenberg-picture matrix elements are just the Schrödinger-picture ones, so using Eq. (117) we may rewrite this solution in either of two forms:

$$\begin{aligned}\mathbf{S}(t)&=\frac{\hbar}{2}\left[\mathbf{n}_x\begin{pmatrix}0 & e^{+i\Omega t}\\ e^{-i\Omega t} & 0\end{pmatrix}+\mathbf{n}_y\begin{pmatrix}0 & -ie^{+i\Omega t}\\ ie^{-i\Omega t} & 0\end{pmatrix}+\mathbf{n}_z\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\right]\\ &\equiv\frac{\hbar}{2}\begin{pmatrix}\mathbf{n}_z & \mathbf{n}_-e^{+i\Omega t}\\ \mathbf{n}_+e^{-i\Omega t} & -\mathbf{n}_z\end{pmatrix}, \qquad \text{where } \mathbf{n}_\pm\equiv\mathbf{n}_x\pm i\mathbf{n}_y\ .\end{aligned} \tag{4.202}$$

The simplicity of the last expression is spectacular. (Remember, it covers *any* initial conditions and *all* three spatial components of spin!) On the other hand, for some purposes the previous form may be more convenient; in particular, its Cartesian components give our earlier results (194)-(196).[50]

One of the advantages of the Heisenberg picture is that it provides a more clear link between classical and quantum mechanics, found by P. Dirac. Indeed, analytical classical mechanics may be used to derive the following equation of time evolution of an arbitrary function $A(q_j, p_j, t)$ of the generalized coordinates q_j and momenta p_j of the system, and time t: [51]

$$\frac{dA}{dt}=\frac{\partial A}{\partial t}-\{A,H\}_{\rm P}, \tag{4.203}$$

where H is the classical Hamiltonian function of the system, and $\{..,..\}_{\rm P}$ is the so-called *Poisson bracket* defined, for two arbitrary functions $A(q_j, p_j, t)$ and $B(q_j, p_j, t)$, as

Poisson bracket

$$\{A,B\}_{\rm P}\equiv\sum_j\left(\frac{\partial A}{\partial p_j}\frac{\partial B}{\partial q_j}-\frac{\partial A}{\partial q_j}\frac{\partial B}{\partial p_j}\right). \tag{4.204}$$

Comparing Eq. (203) with Eq. (199), we see that the correspondence between the classical and quantum mechanics (in the Heisenberg picture) is provided by the following symbolic relation

[50] Note that the "values" of the same Heisenberg operator at different moments of time may or may not commute. For example, consider a free 1D particle, with the time-independent Hamiltonian $\hat{H}=\hat{p}^2/2m$. In this case, Eq. (199) yields the following equations: $i\hbar\dot{\hat{x}}=[\hat{x},\hat{H}]=i\hbar\hat{p}/m$ and $i\hbar\dot{\hat{p}}=[\hat{p},\hat{H}]=0$, with simple solutions (similar to those for the classical motion): $\hat{p}(t)=\text{const}=\hat{p}(0)$ and $\hat{x}(t)=\hat{x}(0)+\hat{p}(0)t/m$, so $[\hat{x}(0),\hat{x}(t)]=[\hat{x}(0),\hat{p}(0)]t/m\equiv[\hat{x}_{\rm S},\hat{p}_{\rm S}]t/m=i\hbar t/m\neq 0$, for $t\neq 0$.

[51] See, e.g., CM Eq. (10.17). The notation there does not use the subscript "P" that is employed in Eqs. (203)-(205) to distinguish the classical Poisson bracket (204) from the quantum anticommutator (34).

$$\{A,B\}_{\rm P} \leftrightarrow \frac{i}{\hbar}\left[\hat{A},\hat{B}\right]. \tag{4.205}$$

Classical vs. quantum mechanics

This relation may be used, in particular, for finding appropriate operators for some observables, if their form is not immediately evident from the correspondence principle.

Finally, let us discuss one more alternative picture of quantum dynamics. It is attributed to P. A. M. Dirac, and is called either the "Dirac picture", or (more frequently) the *interaction picture*. The last name stems from the fact that this picture is very useful for *perturbative* (approximate) approaches to systems whose Hamiltonians may be partitioned into two parts,

$$\hat{H} = \hat{H}_0 + \hat{H}_{\rm int}, \tag{4.206}$$

where $\hat{H}_0$ is the sum of relatively simple Hamiltonians of the component subsystems, while the second term in Eq. (206) represents their weak interaction.[52] (Note, however, that all relations in the balance of this section are exact and not directly based on the interaction weakness.) In this case, it is natural to consider, together with the full operator $\hat{u}(t,t_0)$ of the system's evolution, which obeys Eq. (157b), a similarly defined unitary operator $\hat{u}_0(t,t_0)$ of the "unperturbed" evolution described by $\hat{H}_0$ alone:

$$i\hbar\frac{\partial}{\partial t}\hat{u}_0 = \hat{H}_0\hat{u}_0, \tag{4.207}$$

and also the following *interaction evolution operator*,

$$\hat{u}_{\rm I} \equiv \hat{u}_0^\dagger\hat{u}\,. \tag{4.208}$$

Interaction evolution operator

The motivation for these definitions becomes more clear if we insert the reciprocal relation,

$$\hat{u} \equiv \hat{u}_0\,\hat{u}_0^\dagger\hat{u} = \hat{u}_0\hat{u}_{\rm I}, \tag{4.209}$$

and its Hermitian conjugate,

$$\hat{u}^\dagger = \left(\hat{u}_0\hat{u}_{\rm I}\right)^\dagger = \hat{u}_{\rm I}^\dagger\hat{u}_0^\dagger, \tag{4.210}$$

into the basic Eq. (189):

$$\begin{aligned}\left\langle\alpha\middle|\hat{A}\middle|\beta\right\rangle &= \left\langle\alpha(t_0)\middle|\hat{u}^\dagger(t,t_0)\hat{A}_{\rm S}\hat{u}(t,t_0)\middle|\beta(t_0)\right\rangle \\ &= \left\langle\alpha(t_0)\middle|\hat{u}_{\rm I}^\dagger(t,t_0)\hat{u}_0^\dagger(t,t_0)\hat{A}_{\rm S}\hat{u}_0(t,t_0)\hat{u}_{\rm I}(t,t_0)\middle|\beta(t_0)\right\rangle.\end{aligned} \tag{4.211}$$

This relation shows that any long bracket (187), i.e. any experimentally verifiable result of quantum mechanics, may be expressed as

$$\left\langle\alpha\middle|\hat{A}\middle|\beta\right\rangle = \left\langle\alpha_{\rm I}(t)\middle|\hat{A}_{\rm I}(t)\middle|\beta_{\rm I}(t)\right\rangle, \tag{4.212}$$

if we assume that both the state vectors and the operators depend on time, with the vectors evolving only due to the *interaction operator* $\hat{u}_{\rm I}$,

$$\left\langle\alpha_{\rm I}(t)\right| \equiv \left\langle\alpha(t_0)\right|\hat{u}_{\rm I}^\dagger(t,t_0), \qquad \left|\beta_{\rm I}(t)\right\rangle \equiv \hat{u}_{\rm I}(t,t_0)\left|\beta(t_0)\right\rangle, \tag{4.213}$$

Interaction picture: state vectors

[52] This picture may also useful in more standard problems of the perturbation theory (see Ch. 6 below) where $\hat{H}_{\rm int}$ describes a weak perturbation of a *single* system described by a relatively simple Hamiltonian $\hat{H}_0$.

while the operators' evolution being governed by the *unperturbed operator* $\hat{u}_0$:

Interaction picture: operators

$$\hat{A}_{\rm I}(t) \equiv \hat{u}_0^\dagger(t,t_0)\hat{A}_{\rm S}\hat{u}_0(t,t_0). \tag{4.214}$$

These relations describe the interaction picture of quantum dynamics. Let me defer an example of its use until the perturbative analysis of open quantum systems in Sec. 7.6, and end this section with proof that the interaction evolution operator (208) satisfies the following natural equation,

$$i\hbar\frac{\partial}{\partial t}\hat{u}_{\rm I} = \hat{H}_{\rm I}\hat{u}_{\rm I}, \tag{4.215}$$

where $\hat{H}_{\rm I}$ is the interaction Hamiltonian formed from $\hat{H}_{\rm int}$ in accordance with the same rule (214):

$$\hat{H}_{\rm I}(t) \equiv \hat{u}_0^\dagger(t,t_0)\hat{H}_{\rm int}\hat{u}_0(t,t_0). \tag{4.216}$$

The proof is very straightforward: first using the definition (208), and then Eqs. (157b) and the Hermitian conjugate of Eq. (207), we may write

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}\hat{u}_{\rm I} &\equiv i\hbar\frac{\partial}{\partial t}\left(\hat{u}_0^\dagger\hat{u}\right) \equiv i\hbar\frac{\partial\hat{u}_0^\dagger}{\partial t}\hat{u} + \hat{u}_0^\dagger i\hbar\frac{\partial\hat{u}}{\partial t} = -\hat{H}_0\hat{u}_0^\dagger\hat{u} + \hat{u}_0^\dagger\hat{H}\hat{u} = -\hat{H}_0\hat{u}_0^\dagger\hat{u} + \hat{u}_0^\dagger\left(\hat{H}_0 + \hat{H}_{\rm int}\right)\hat{u} \\ &\equiv -\hat{H}_0\hat{u}_0^\dagger\hat{u} + \hat{u}_0^\dagger\hat{H}_0\hat{u} + \hat{u}_0^\dagger\hat{H}_{\rm int}\hat{u} \equiv \left(-\hat{H}_0\hat{u}_0^\dagger + \hat{u}_0^\dagger\hat{H}_0\right)\hat{u} + \hat{u}_0^\dagger\hat{H}_{\rm int}\hat{u}. \end{aligned} \tag{4.217}$$

Since $\hat{u}_0^\dagger$ may be represented as an integral of an exponent of $\hat{H}_0$ over time (similar to Eq. (181) relating $\hat{u}$ and $\hat{H}$), these operators commute, so the parentheses in the last form of Eq. (217) vanish. Now plugging $\hat{u}$ from the last form of Eq. (209), we get the equation,

$$i\hbar\frac{\partial}{\partial t}\hat{u}_{\rm I} = \hat{u}_0^\dagger\hat{H}_{\rm int}\hat{u}_0 u_{\rm I} \equiv \left(\hat{u}_0^\dagger\hat{H}_{\rm int}\hat{u}_0\right)\hat{u}_{\rm I}, \tag{4.218}$$

which is clearly equivalent to the combination of Eqs. (215) and (216).

As Eq. (215) shows, if the energy scale of the interaction $H_{\rm int}$ is much smaller than that of the background Hamiltonian H_0, the interaction evolution operators $\hat{u}_{\rm I}$ and $\hat{u}_{\rm I}^\dagger$, and hence the state vectors (213) evolve relatively slowly, without fast background oscillations. This is very convenient for the perturbative approaches to complex interacting systems, in particular to the "open" quantum systems that weakly interact with their environment – see Sec. 7.6.

4.7. Coordinate and momentum representations

Now let me show that in application to the orbital motion of a particle, the bra-ket formalism naturally reduces to the notions and postulates of wave mechanics, which were discussed in Chapter 1. For that, we first have to modify some of the above formulas for the case of a basis with a continuous spectrum of eigenvalues. In that case, it is more appropriate to replace discrete indices, such as *j*, *j'*, etc. broadly used above, with the corresponding eigenvalue – just as it was done earlier for functions of the wave vector – see, e.g., Eqs. (1.88), (2.20), etc. For example, the key Eq. (68), defining the eigenkets and eigenvalues of an operator, may be conveniently rewritten in the form

$$\hat{A}|a_A\rangle = A|a_A\rangle. \tag{4.219}$$

More substantially, all sums over such continuous eigenstate sets should be replaced with integrals. For example, for a full and orthonormal set of the continuous eigenstates $|a_A\rangle$, the closure relation (44) should be replaced with

Continuous spectrum: closure relation

$$\int dA\,|a_A\rangle\langle a_A| = \hat{I}, \tag{4.220}$$

where the integral is over the whole interval of possible eigenvalues of the observable A.[53] Applying this relation to the ket-vector of an arbitrary state α, we get the following replacement of Eq. (37):

$$|\alpha\rangle \equiv \hat{I}|\alpha\rangle = \int dA\,|a_A\rangle\langle a_A|\alpha\rangle = \int dA\,\langle a_A|\alpha\rangle|a_A\rangle. \tag{4.221}$$

For the particular case when $|\alpha\rangle = |a_{A'}\rangle$, this relation requires that

Continuous spectrum: state orthonormality

$$\langle a_A|a_{A'}\rangle = \delta(A - A'); \tag{4.222}$$

this formula replaces the orthonormality condition (38).

According to Eq. (221), in the continuous case the bracket $\langle a_A|\alpha\rangle$ still plays the role of probability amplitude, i.e. a complex c-number whose modulus squared determines the state a_A's probability – see the last form of Eq. (120). However, for a continuous observable, the probability of finding the system exactly in a particular state is infinitesimal; instead (as was already discussed in Sec. 1.2), we should speak about the probability $dW = w(A)dA$ of finding the observable within a small interval $dA << A$ near the value A, with *probability density* $w(A) \propto |\langle a_A|\alpha\rangle|^2$. The coefficient of proportionality in this relation may be found by making a similar change from the summation to integration in the normalization condition (121):

$$\int dA\langle\alpha|a_A\rangle\langle a_A|\alpha\rangle = 1. \tag{4.223}$$

Since the total probability of the system being in *some* state should be equal to $\int w(A)dA$, this means that

Continuous spectrum: probability density

$$w(A) = \langle\alpha|a_A\rangle\langle a_A|\alpha\rangle = |\langle\alpha|a_A\rangle|^2. \tag{4.224}$$

Now let us see how we can calculate the expectation values of continuous observables, i.e. their ensemble averages. If we speak about the same observable A whose eigenstates are used as the continuous basis (or any compatible observable), everything is simple. Indeed, inserting Eq. (224) into the general statistical relation

$$\langle A\rangle = \int w(A)AdA \tag{4.225}$$

that is the obvious continuous version of Eq. (1.37), we get

$$\langle A\rangle = \int\langle\alpha|a_A\rangle A\langle a_A|\alpha\rangle dA. \tag{4.226}$$

Inserting a delta function to represent this expression formally as a double integral,

$$\langle A\rangle = \int dA\int dA'\langle\alpha|a_A\rangle A\delta(A - A')\langle a_{A'}|\alpha\rangle, \tag{4.227}$$

[53] The generalization to cases when the eigenvalue spectrum consists of both a continuous interval plus some set of discrete values, is straightforward, though leads to somewhat bulky formulas.

and using the continuous-spectrum version of Eq. (98),

$$\langle a_A|\hat{A}|a_{A'}\rangle = A\delta(A-A'), \tag{4.228}$$

we may write

$$\langle A\rangle = \int dA\int dA'\langle\alpha|a_A\rangle\langle a_A|\hat{A}|a_{A'}\rangle\langle a_{A'}|\alpha\rangle \equiv \langle\alpha|\hat{A}|\alpha\rangle, \tag{4.229}$$

so Eq. (4.125) remains valid in the continuous-spectrum case without any changes. This formula is very convenient for applications because it does not require the calculation of the eigenstates a_A, and its matrix form is valid in any basis.

Now we are ready for a discussion of the relationship between the bra-ket formalism and wave mechanics. (For the notation simplicity I will discuss its 1D version; its generalization to 2D and 3D cases is straightforward.) Let us start with postulating the (intuitively, almost evident) existence of a quantum state basis, whose ket-vectors will be called $|x\rangle$, corresponding to a certain definite value x of the particle's coordinate. Writing the trivial identity $x|x\rangle = x|x\rangle$ and comparing it with Eq. (219), we see that they do not contradict each other if we assume that x on the left-hand side of this relation is the Hermitian operator $\hat{x}$ of the particle's coordinate, in a specific representation when its action on a ket- (or bra-) vector is just the multiplication by the *c*-number x:

$$\hat{x}|x\rangle = x|x\rangle. \tag{4.230}$$

In this way, we consider vectors $|x\rangle$ to be the eigenstates of the operator $\hat{x}$. (This looks like a proof, but is actually a separate, independent postulate, no matter how plausible.)

Let me hope that the reader will excuse me if I do not pursue here strict proof that the set of all x-states is full and orthogonal,[54] so we may apply Eq. (222) to it:

$$\langle x|x'\rangle = \delta(x-x'). \tag{4.231}$$

Using this basis is called the *coordinate representation* – the term which was already mentioned several times in this course, but without explanation. In the basis of the x-states, the inner product $\langle a_A|\alpha(t)\rangle$ becomes $\langle x|\alpha(t)\rangle$, and Eq. (223) takes the following form:

$$w(x,t) = \langle\alpha(t)|x\rangle\langle x|\alpha(t)\rangle \equiv \langle x|\alpha(t)\rangle^*\langle x|\alpha(t)\rangle. \tag{4.232}$$

Comparing this formula with the basic postulate (1.22) of wave mechanics, we see that they coincide if the wavefunction of a time-dependent state α is identified with that short bracket:[55]

Wave-function as inner product

$$\boxed{\Psi_\alpha(x,t) \equiv \langle x|\alpha(t)\rangle}. \tag{4.233}$$

This key formula provides the desired connection between the bra-ket formalism and the wave mechanics, and should not be too surprising for the (thoughtful :-) reader. Indeed, Eq. (45) shows that any inner product of two state vectors describing two states is a measure of their similarity – just as the scalar product of two geometric vectors is; the orthonormality condition (38) is a particular

[54] Such proof is rather involved mathematically, but physically this fact should be evident.

[55] I do not quite like expressions like $\langle x|\Psi\rangle$ used in some papers and even textbooks. Of course, one is free to replace α with any other letter (Ψ including) to denote a quantum state, but then it is better not to use the same letter to denote the wavefunction, i.e. an inner product of two state vectors, to avoid confusion.

manifestation of this fact. In this language, the particular value (233) of a wavefunction Ψ_α at some point x and moment t characterizes "how much of a particular coordinate x" the state α contains at time t. (Of course, this informal language is too crude to reflect the fact that $\Psi_\alpha(x, t)$ is a complex function, which has not only a modulus but also an argument – the quantum-mechanical phase.)

Now let us rewrite the most important formulas of the bra-ket formalism in the wave mechanics notation. Inner-multiplying both parts of Eq. (219), written for an arbitrary operator, by the ket-vector $\langle x|$, and then inserting into the left-hand side of that relation the identity operator in the form (220) for coordinate x', we get

$$\int dx' \langle x|\hat{A}|x'\rangle\langle x'|a_A\rangle = A\langle x|a_A\rangle, \tag{4.234}$$

i.e., using the wavefunction's definition (233),

$$\int dx' \langle x|\hat{A}|x'\rangle \Psi_A(x') = A\,\Psi_A(x), \tag{4.235}$$

where, for the notation brevity, the time dependence of the wavefunction is just implied (with the capital Ψ serving as a reminder of this fact), and will be restored when needed. For a general operator, we would have to stop here, because if it does not commute with the coordinate operator, its matrix in the x-basis is not diagonal, and the integral on the left-hand side of Eq. (235) cannot be worked out explicitly. However, virtually all quantum-mechanical operators discussed in this course[56] are (*space*-) *local*: they depend on only one spatial coordinate, say x. For such an operator, we may define its coordinate representation by the following equality (valid for an arbitrary wavefunction, not only Ψ_A):

$$\hat{A}\big|_{\text{in } x}\Psi(x) \equiv \int \langle x|\hat{A}|x'\rangle\Psi(x')dx'. \tag{4.236}$$

Operator: coordinate representation

The explicit form of the coordinate representation still needs to be determined for each operator type. Let us consider, for example, the 1D version of the Hamiltonian (1.41),

$$\hat{H} = \frac{\hat{p}_x^2}{2m} + U(\hat{x}), \tag{4.237}$$

which was the basis of all our discussions in Chapter 2. Its potential-energy part U (even if it is time-dependent as well) commutes with the operator $\hat{x}$, i.e. its matrix in the x-basis is diagonal. For such an operator, the long bracket in Eq. (236) may be transformed using Eq. (231): $\langle x|U|x'\rangle = U(x)\delta(x-x')$, so the right-hand part of this equality becomes just $U(x)\Psi(x)$. Comparing it with the left-hand part, we see that the coordinate representation of such an operator is given merely by the c-number function $U(x)$. (Eq. (230) may be viewed as just a particular manifestation of this rule.)

The situation with the momentum operator $\hat{p}_x$ (and hence the kinetic energy $\hat{p}_x^2/2m$), which do not commute with $\hat{x}$, is less evident. Let me show that its coordinate representation is given by the 1D version of Eq. (1.26), if we *postulate* that the commutation relation (2.14),

$$[\hat{x}, \hat{p}] = i\hbar\hat{I}, \qquad \text{i.e.} \quad \hat{x}\hat{p}_x - \hat{p}_x\hat{x} = i\hbar\hat{I}, \tag{4.238}$$

[56] The only substantial exception is the statistical operator $\hat{w}(x, x')$, to be discussed separately in Chapter 7.

is valid in *any* representation.[57] For that, let us consider the following matrix element: $\langle x|\hat{x}\hat{p}_x - \hat{p}_x\hat{x}|x'\rangle$. On one hand, we may use Eq. (238), and then Eq. (231), to write

$$\langle x|\hat{x}\hat{p}_x - \hat{p}_x\hat{x}|x'\rangle = \langle x|i\hbar\hat{I}|x'\rangle = i\hbar\langle x|x'\rangle = i\hbar\delta(x-x'). \tag{4.239}$$

On the other hand, since $\hat{x}|x'\rangle = x'|x'\rangle$ and $\langle x|\hat{x} = \langle x|x$, we may represent the same matrix element as

$$\langle x|\hat{x}\hat{p}_x - \hat{p}_x\hat{x}|x'\rangle = \langle x|x\hat{p}_x - \hat{p}_x x'|x'\rangle = (x-x')\langle x|\hat{p}_x|x'\rangle. \tag{4.240}$$

Comparing Eqs. (239) and (240), we get

$$\langle x|\hat{p}_x|x'\rangle = i\hbar\frac{\delta(x-x')}{x-x'}. \tag{4.241}$$

As it follows from the definition of the delta function,[58] all expressions involving it acquire final sense only at their integration, in our current case, as described by Eq. (236). Plugging Eq. (241) into the right-hand side of that relation, we get

$$\int\langle x|\hat{p}_x|x'\rangle\Psi(x')dx' = i\hbar\int\frac{\delta(x-x')}{x-x'}\Psi(x')dx'. \tag{4.242}$$

Since the right-hand-part integral is contributed only by an infinitesimal vicinity of the point $x' = x$, we may calculate it by expanding the continuous wavefunction $\Psi(x')$ into the Taylor series in small $(x' - x)$, and keeping only two leading terms of the series, so Eq. (242) is reduced to

$$\int\langle x|\hat{p}_x|x'\rangle\Psi(x')dx' = i\hbar\left[\Psi(x)\int\frac{\delta(x-x')}{x-x'}dx' - \int\delta(x-x')\frac{\partial\Psi(x')}{\partial x'}\Big|_{x'=x}dx'\right]. \tag{4.243}$$

Since the delta function may be always understood as an even function of its argument, in our case of $(x - x')$, the first term on the right-hand side is proportional to an integral of an odd function in symmetric limits and is equal to zero, and we get[59]

$$\int\langle x|\hat{p}_x|x'\rangle\Psi(x')dx' = -i\hbar\frac{\partial\Psi}{\partial x}. \tag{4.244}$$

Comparing this expression with the left-hand side of Eq. (236) with $\hat{A} = \hat{p}_x$, we see that in the coordinate representation, we indeed get the 1D version of Eq. (1.26), which was used so much in Chapter 2,[60]

$$\hat{p}_x\big|_{\text{in } x} = -i\hbar\frac{\partial}{\partial x}. \tag{4.245}$$

---

[57] Another possible approach to the axiomatics of wave mechanics is to *derive* Eg. (238) by *postulating* the form, $\hat{\mathcal{T}}_X = \exp\{-i\hat{p}_x X/\hbar\}$, of the operator that shifts any wavefunction by distance X along the x-axis. In my approach, this expression will be *derived* when we need it (in Sec. 5.5), while Eq. (238) is *postulated*.

[58] If necessary, please revisit MA Sec. 14.

[59] One more useful expression of this type, which may be proved similarly, is $(\partial/\partial x)\delta(x - x') = \delta(x - x')\partial/\partial x'$.

[60] This means, in particular, that in the sense of Eq. (236), the operator of differentiation is local, despite the fact that its action on a function f may be interpreted as the limit of the fraction $\Delta f/\Delta x$, involving *two* points. (In some axiomatic systems, local operators are *defined* as arbitrary polynomials of functions and their derivatives.)

It is virtually evident (and straightforward to prove by using the Taylor expansion just as in Sec. 6) that the coordinate representation of any operator function $f(\hat{p}_x)$ is

$$f\left(-i\hbar\frac{\partial}{\partial x}\right). \tag{4.246}$$

In particular, this pertains to the kinetic energy operator in Eq. (237), so the coordinate representation of this Hamiltonian also takes the very familiar form:

$$\hat{H}\big|_{\text{in } x} = \frac{1}{2m}\left(-i\hbar\frac{\partial}{\partial x}\right)^2 + U(x,t) \equiv -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + U(x,t). \tag{4.247}$$

Now returning to the discussion of the general Eq. (235), and comparing its last form with that of Eq. (236), we see that for a local operator in the coordinate representation, the eigenproblem (219) takes the form

$$\hat{A}\big|_{\text{in } x}\Psi_A(x) = A\Psi_A(x), \tag{4.248}$$

Eigenproblem in *x*-representation

even if the operator $\hat{A}$ does not commute with the operator $\hat{x}$. The most important case of this coordinate-representation form of the eigenproblem (68) is the familiar Eq. (1.60) for the eigenvalues E_n of the energy of a system with a time-independent Hamiltonian.

The operator locality also simplifies the expression for its expectation value. Indeed, plugging the closure relation in the form (231) into the general Eq. (125) twice (written in the first case for x and in the second case for x'), we get

$$\langle A\rangle = \int dx \int dx' \langle\alpha(t)|x\rangle\langle x|\hat{A}|x'\rangle\langle x'|\alpha(t)\rangle = \int dx \int dx' \Psi_\alpha^*(x,t)\langle x|\hat{A}|x'\rangle\Psi_\alpha(x',t). \tag{4.249}$$

Now, Eq. (236) reduces this result to just

$$\langle A\rangle = \int dx \int dx' \Psi_\alpha^*(x,t)\hat{A}\big|_{\text{in } x}\Psi_\alpha(x,t)\delta(x-x') \equiv \int \Psi_\alpha^*(x,t)\hat{A}\big|_{\text{in } x}\Psi_\alpha(x,t)dx. \tag{4.250}$$

i.e. to Eq. (1.23), which had to be postulated in Chapter 1 where the x-representation of the operators was just implied.

Finally, let us discuss the time evolution of the wavefunction, in the Schrödinger picture. For that, we may use Eq. (233) to calculate the (partial) time derivative of the wavefunction of some state α:

$$i\hbar\frac{\partial\Psi_\alpha}{\partial t} = i\hbar\frac{\partial}{\partial t}\langle x|\alpha(t)\rangle. \tag{4.251}$$

Since the coordinate operator $\hat{x}$ does not depend on time explicitly, its eigenstates x are stationary, and we can swap the time derivative and the time-independent bra-vector $\langle x|$. Now using the Schrödinger-picture equation (158), and then inserting the identity operator in the continuous form (220) of the closure relation, written for the coordinate eigenstates,

$$\int dx'|x'\rangle\langle x'| = \hat{I}, \tag{4.252}$$

we may continue to develop the right-hand side of Eq. (251) as

$$\langle x|i\hbar\frac{\partial}{\partial t}|\alpha(t)\rangle = \langle x|\hat{H}|\alpha(t)\rangle = \int dx'\langle x|\hat{H}|x'\rangle\langle x'|\alpha(t)\rangle = \int dx'\,\langle x|\hat{H}|x'\rangle\Psi_\alpha(x'), \tag{4.253}$$

If the Hamiltonian operator is local, we may apply Eq. (236) to the last expression, to get the familiar form (1.28) of the Schrödinger equation:

$$i\hbar\frac{\partial\Psi_\alpha}{\partial t} = \hat{H}\big|_{\text{in }x}\Psi_\alpha. \tag{4.254}$$

So, for the local operators that obey Eq. (236), we have been able to derive all the basic notions and postulates of the wave mechanics from the bra-ket formalism. Moreover, the formalism has allowed us to get a very useful equation (248) for an arbitrary local operator, which will be repeatedly used below. (In the first three chapters of this course, we have only used its particular case (1.60) for the Hamiltonian operator.)

Now let me deliver on my promise to develop a more balanced view of the de Broglie wave (4.1), which would be more respectful to the evident **r** ↔ **p** symmetry of the coordinate and momentum. Let us discuss the 1D case when the wave may be represented as

$$\psi_p(x) = a_p\exp\left\{i\frac{px}{\hbar}\right\}, \qquad \text{for all} -\infty < x < +\infty. \tag{4.255}$$

(For the sake of brevity, from this point to the end of the section, I am dropping the index x in the notation of the momentum – just as it was done in Chapter 2.) Let us have a good look at this function. Since it satisfies Eq. (248) for the 1D momentum operator (245),

$$\hat{p}\big|_{\text{in }x}\psi_p = p\psi_p, \tag{4.256}$$

ψ_p is an eigenfunction of that operator. But this means that we can also write Eq. (219) for the corresponding ket-vector:

$$\hat{p}|p\rangle = p|p\rangle, \tag{4.257}$$

and according to Eq. (233), the wavefunction (255) may be represented as

$$\psi_p(x) = \langle x|p\rangle, \qquad \text{so } \psi_p^*(x) = \langle p|x\rangle. \tag{4.258}$$

These expressions are quite remarkable in their $x \leftrightarrow p$ symmetry – which may be pursued further on. Before doing that, however, we have to discuss the normalization of such wavefunctions. Indeed, in this case, the probability density $w(x)$ of the wave (255) is constant, so its integral

$$\int_{-\infty}^{+\infty} w(x)dx = \int_{-\infty}^{+\infty}\psi_p(x)\psi_p^*(x)dx \tag{4.259}$$

diverges if $a_p \neq 0$. Earlier in the course, we discussed two ways to avoid this divergence. One is to use a very large but finite integration volume – see Eq. (1.31). Another way is to work with wave packets of the type (2.20), possibly of a very large length and hence a very narrow spread of the momentum values. Then the integral (259) may be required to equal 1 without any conceptual problem.

However, both these methods, while being convenient for the solution of many particular problems, violate the $x \leftrightarrow p$ symmetry and hence are unfit for our current conceptual discussion.

Instead, let us continue to identify the eigenvectors $\langle p|$ and $|p\rangle$ of the momentum with the bra- and ket-vectors $\langle a_A|$ and $|a_A\rangle$ of the general theory described at the beginning of this section. Then the normalization condition (222) becomes

$$\langle p|p'\rangle = \delta(p-p'). \tag{4.260}$$

Inserting the identity operator in the form (252), with the integration variable x' replaced by x, into the left-hand side of this equation, and using Eq. (258), we can translate this normalization rule to the wavefunction language:

$$\int dx \langle p|x\rangle\langle x|p'\rangle \equiv \int dx \psi_p^*(x)\psi_{p'}(x) = \delta(p-p'). \tag{4.261}$$

For the particular wavefunction (255), this requirement turns into the following condition:

$$a_p^* a_{p'} \int_{-\infty}^{+\infty} \exp\left\{i\frac{(p'-p)x}{\hbar}\right\}dx \equiv |a_p|^2 2\pi\hbar\delta(p-p') = \delta(p-p'), \tag{4.262}$$

so, finally, $a_p = e^{i\phi}/(2\pi\hbar)^{1/2}$, where ϕ is an arbitrary (real) phase, and Eq. (255) becomes[61]

$$\psi_p(x) = \langle x|p\rangle = \frac{1}{(2\pi\hbar)^{1/2}}\exp\left\{i\left(\frac{px}{\hbar}+\phi\right)\right\}. \tag{4.263}$$

Now let us represent an arbitrary wavefunction $\psi(x)$ as a wave packet of the type (2.20), based on the wavefunctions (263), taking $\phi = 0$ for the notation brevity, because the phase may be incorporated into the (generally, complex) envelope function $\varphi(p)$:

$$\psi(x) = \frac{1}{(2\pi\hbar)^{1/2}}\int \varphi(p)\exp\left\{i\frac{px}{\hbar}\right\}dp\,. \tag{4.264}$$

x-representation: wavefunctions

From the mathematical point of view, this is just a 1D Fourier spatial transform, and its reciprocal is

$$\varphi(p) \equiv \frac{1}{(2\pi\hbar)^{1/2}}\int \psi(x)\exp\left\{-i\frac{px}{\hbar}\right\}dx\,. \tag{4.265}$$

p-representation: wavefunctions

These expressions are completely symmetric, and represent the same wave packet; this is why the functions $\psi(x)$ and $\varphi(p)$ are frequently called the *reciprocal representations* of a quantum state of the particle: respectively, its coordinate (x-) and momentum (p-) representations. Using Eq. (258), and Eq. (263) with $\phi = 0$, they may be recast into simpler forms,

$$\psi(x) = \int \varphi(p)\langle x|p\rangle dp, \qquad \varphi(p) = \int \psi(x)\langle p|x\rangle dx\,, \tag{4.266}$$

in which the inner products satisfy the basic postulate (14) of the bra-ket formalism:

$$\langle p|x\rangle = \frac{1}{(2\pi\hbar)^{1/2}}\exp\left\{-i\frac{px}{\hbar}\right\} = \langle x|p\rangle^*. \tag{4.267}$$

[61] Repeating such calculation for each Cartesian component of a plane monochromatic wave of arbitrary dimensionality d, we get $\psi_{\mathbf{p}} = (2\pi\hbar)^{-d/2}\exp\{i(\mathbf{p}\cdot\mathbf{r}/\hbar + \varphi)\}$.

Next, we already know that in the x-representation, i.e. in the usual wave mechanics, the coordinate operator $\hat{x}$ is reduced to the multiplication by x, and the momentum operator is proportional to the partial derivative over the coordinate:

x-representation: operators

$$\hat{x}\Big|_{\text{in } x} = x, \qquad \hat{p}\Big|_{\text{in } x} = -i\hbar\frac{\partial}{\partial x}. \tag{4.268}$$

It is natural to guess that in the p-representation, the expressions for operators would be reciprocal:

p-representation: operators

$$\hat{x}\Big|_{\text{in } p} = +i\hbar\frac{\partial}{\partial p}, \qquad \hat{p}\Big|_{\text{in } p} = p, \tag{4.269}$$

with the only difference of one sign, which is due to the opposite signs of the Fourier exponents in Eqs. (264) and (265). The proofs of Eqs. (269) are straightforward; for example, acting by the momentum operator on the arbitrary wavefunction (264), we get

$$\hat{p}\psi(x) = -i\hbar\frac{\partial}{\partial x}\psi(x) = \frac{1}{(2\pi\hbar)^{1/2}}\int\varphi(p)\left(-i\hbar\frac{\partial}{\partial x}\exp\left\{i\frac{px}{\hbar}\right\}\right)dp = \frac{1}{(2\pi\hbar)^{1/2}}\int p\varphi(p)\exp\left\{i\frac{px}{\hbar}\right\}dp, \tag{4.270}$$

and similarly for the operator $\hat{x}$ acting on the function $\varphi(p)$. Comparing the final form of Eq. (270) with the initial Eq. (264), we see that the action of the operators (268) on the wavefunction ψ (i.e. the state's x-representation) gives the same results as the action of the operators (269) on the function φ (i.e. its p-representation).

It is also illuminating to have a different look at this coordinate-momentum duality. For that, notice that according to Eqs. (82)-(84), we may consider the bracket $\langle x|p\rangle$ as an element of the (infinite-size) matrix U_{xp} of the unitary transform from the x-basis to the p-basis. Let us use this fact to derive the general operator transform rule that would be a continuous version of Eq. (92). Say, we want to calculate the general matrix element of some operator known in the x-representation, in the p-representation:

$$\langle p|\hat{A}|p'\rangle. \tag{4.271}$$

Inserting two identity operators (252) written for x and x' into this bracket, and then using Eq. (258) and its complex conjugate, and also Eq. (236) (again, valid only for space-local operators!), we get

$$\begin{aligned}\langle p|\hat{A}|p'\rangle &= \int dx\int dx'\langle p|x\rangle\langle x|\hat{A}|x'\rangle\langle x'|p'\rangle = \int dx\int dx'\psi_p^*(x)\langle x|\hat{A}|x'\rangle\psi_{p'}(x') \\ &= \frac{1}{2\pi\hbar}\int dx\exp\left\{-i\frac{px}{\hbar}\right\}\hat{A}\Big|_{\text{in } x}\exp\left\{+i\frac{p'x}{\hbar}\right\}.\end{aligned} \tag{4.272}$$

As a sanity check, for the momentum operator itself, this relation yields:

$$\langle p|\hat{p}|p'\rangle = \frac{1}{2\pi\hbar}\int dx\exp\left\{-i\frac{px}{\hbar}\right\}\left(-i\hbar\frac{\partial}{\partial x}\right)\exp\left\{i\frac{p'x}{\hbar}\right\} = \frac{p'}{2\pi\hbar}\int_{-\infty}^{+\infty}\exp\left\{i\frac{(p'-p)x}{\hbar}\right\}dx = p'\delta(p'-p). \tag{4.273}$$

Due to Eq. (257), this result is equivalent to the second of Eqs. (269).

From a thoughtful reader, I anticipate the following natural question: why is the momentum representation used much less often than the coordinate representation – i.e. wave mechanics? The answer is purely practical: with an important exception of the 1D harmonic oscillator (to be revisited in

Sec. 5.4), in most systems, the orbital-motion Hamiltonian (237) is not $x \leftrightarrow p$ symmetric, with the potential energy $U(\mathbf{r})$ typically being a more complex function than the kinetic energy $p^2/2m$. Because of that, it is easier to analyze such systems treating the potential energy operator just as a *c*-number multiplier, as it is in the coordinate representation – and as this was done in Chapters 1-3.

The most significant exception from this practice is the motion in a periodic potential in the presence of a coordinate-independent external force $\mathbf{F}(t)$. As was discussed in Secs. 2.7 and 3.4, in such periodic systems the eigenenergies $E_n(\mathbf{q})$, playing the role of the effective kinetic energy of the particle, may be rather involved functions of its quasimomentum $\hbar\mathbf{q}$, while its effective potential energy $U_{\text{ef}} = -\mathbf{F}(t)\cdot\mathbf{r}$ due to the additional force $\mathbf{F}(t)$ is a very simple function of coordinates. This is why detailed analyses of the quantum effects that were briefly discussed in Sec. 2.8 (the Bloch oscillations, etc.) and also such statistical phenomena as drift, diffusion, etc.[62] in solid-state theory are typically based on the momentum (or rather quasimomentum) representation.

4.8. Exercise problems

4.1. Prove that if $\hat{A}$ and $\hat{B}$ are linear operators, and C is a *c*-number, then:

(i) $\left(\hat{A}^\dagger\right)^\dagger = \hat{A}$; (ii) $\left(C\hat{A}\right)^\dagger = C^*\hat{A}^\dagger$; (iii) $\left(\hat{A}\hat{B}\right)^\dagger = \hat{B}^\dagger\hat{A}^\dagger$;

(iv) the operators $\hat{A}\hat{A}^\dagger$ and $\hat{A}^\dagger\hat{A}$ are Hermitian.

4.2. Prove that for any linear operators $\hat{A}, \hat{B}, \hat{C}$, and $\hat{D}$,

$$\left[\hat{A}\hat{B}, \hat{C}\hat{D}\right] = \hat{A}\left\{\hat{B}, \hat{C}\right\}\hat{D} - \hat{A}\hat{C}\left\{\hat{B}, \hat{D}\right\} + \left\{\hat{A}, \hat{C}\right\}\hat{D}\hat{B} - \hat{C}\left\{\hat{A}, \hat{D}\right\}\hat{B}.$$

4.3. Calculate all possible binary products $\sigma_j\sigma_{j'}$ (for $j, j' = x, y, z$) of the Pauli matrices defined by Eqs. (105), and their commutators and anticommutators (defined similarly to those of the corresponding operators). Summarize the results by using the Kronecker delta and Levi-Civita permutation symbols.[63]

4.4. Calculate the following expressions,

(i) $(\mathbf{c}\cdot\boldsymbol{\sigma})^n$, and then

(ii) $(b\mathrm{I} + \mathbf{c}\cdot\boldsymbol{\sigma})^n$,

for the scalar product $\mathbf{c}\cdot\boldsymbol{\sigma}$ of the Pauli vector's matrix $\boldsymbol{\sigma} \equiv \mathbf{n}_x\sigma_x + \mathbf{n}_y\sigma_y + \mathbf{n}_z\sigma_z$ by an arbitrary *c*-number geometric vector $\mathbf{c}$, where n is a non-negative integer *c*-number and b is an arbitrary scalar *c*-number.

Hint: For Task (ii), you may like to use the binomial theorem[64] and then transform the result to a form enabling you to use the same theorem backward.

4.5. Use the solution of the previous problem to derive Eqs. (2.191) for the transparency $\mathscr{T}$ of the Dirac comb – a system of N similar, equidistant, delta-functional potential barriers.

---

[62] In this series, a brief discussion of these effects may be found in SM Chapter 6.

[63] See, e.g., MA Eqs. (13.1) and (13.2).

[64] See, e.g. MA Eq. (2.9).

4.6. Use the solution of Problem 4(i) to spell out the following matrix: $\exp\{i\theta\,\mathbf{n}\cdot\boldsymbol{\sigma}\}$, where $\boldsymbol{\sigma}$ is the 3D vector (117) of the Pauli matrices, $\mathbf{n}$ is a *c*-number geometric vector of unit length, and θ is a *c*-number scalar.

4.7. Use the solution of Problem 4(ii) to calculate $\exp\{\mathrm{A}\}$, where A is an arbitrary 2×2 matrix.

4.8. Express all elements of the matrix $\mathrm{B} \equiv \exp\{\mathrm{A}\}$ explicitly via those of the 2×2 matrix A. Spell out your result for the following matrices:

$$\mathrm{A} = \begin{pmatrix} a & a \\ a & a \end{pmatrix}, \qquad \mathrm{A}' = \begin{pmatrix} i\varphi & i\varphi \\ i\varphi & i\varphi \end{pmatrix},$$

with real a and φ.

4.9. Prove that for arbitrary square matrices A and B,

$$\mathrm{Tr}\,(\mathrm{AB}) = \mathrm{Tr}\,(\mathrm{BA})\,.$$

Is each diagonal element $(AB)_{jj}$ necessarily equal to $(BA)_{jj}$?

4.10. Calculate the trace of the following 2×2 matrix:

$$\mathrm{A} \equiv (\mathbf{a}\cdot\boldsymbol{\sigma})(\mathbf{b}\cdot\boldsymbol{\sigma})(\mathbf{c}\cdot\boldsymbol{\sigma}),$$

where $\boldsymbol{\sigma}$ is the Pauli vector's matrix, while $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ are arbitrary *c*-number vectors.

4.11. Prove that the matrix trace of an arbitrary operator does not change at its unitary transformation.

4.12. Prove that for any two full and orthonormal bases $\{u\}$ and $\{v\}$ of the same Hilbert space,

$$\mathrm{Tr}\left(|u_j\rangle\langle v_{j'}|\right) = \langle v_{j'}|u_j\rangle.$$

4.13. Is the 1D scattering matrix S, defined by Eq. (2.124), unitary? What about the 1D transfer matrix T defined by Eq. (2.125)?

4.14. Calculate the trace of the following matrix:

$$\exp\{i\mathbf{a}\cdot\boldsymbol{\sigma}\}\exp\{i\mathbf{b}\cdot\boldsymbol{\sigma}\},$$

where $\boldsymbol{\sigma}$ is the Pauli vector's matrix, while $\mathbf{a}$ and $\mathbf{b}$ are *c*-number geometric vectors.

4.15. Prove the following operator-vector identity:

$$(\boldsymbol{\sigma}\cdot\hat{\mathbf{r}})(\boldsymbol{\sigma}\cdot\hat{\mathbf{p}}) = \mathrm{I}\,\hat{\mathbf{r}}\cdot\hat{\mathbf{p}} + i\boldsymbol{\sigma}\cdot(\hat{\mathbf{r}}\times\hat{\mathbf{p}}),$$

where $\boldsymbol{\sigma}$ is the Pauli vector's matrix, and I is the 2×2 identity matrix.

Hint: Take into account that the operator vectors $\hat{\mathbf{r}}$ and $\hat{\mathbf{p}}$ are defined in the orbital-motion Hilbert space, different from that of the Pauli vector $\hat{\boldsymbol{\sigma}}$, and hence commute with it – even though they do not commute with each other.

4.16. Let A_j be the eigenvalues of some operator $\hat{A}$. Express the following two sums,

$$\Sigma_1 \equiv \sum_j A_j \quad \text{and} \quad \Sigma_2 \equiv \sum_j A_j^2 ,$$

via the matrix elements $A_{jj'}$ of this operator in an arbitrary basis.

4.17. Calculate $\langle \sigma_z \rangle$ of a spin–½ in the quantum state with the following ket-vector:

$$|\alpha\rangle = \text{const} \times \left(|\uparrow\rangle + |\downarrow\rangle + |\rightarrow\rangle + |\leftarrow\rangle \right),$$

where $(\uparrow, \downarrow)$ and $(\rightarrow, \leftarrow)$ are the eigenstates of the Pauli matrices σ_z and σ_x, respectively.

Hint: Double-check whether your solution is general.

4.18. A spin-½ is fully polarized in the positive z-direction. Calculate the probabilities of the alternative outcomes of a perfect Stern-Gerlach experiment with the magnetic field oriented in an arbitrarily different direction.

4.19. In a certain basis, the Hamiltonian of a two-level system is described by the matrix

$$\mathrm{H} = \begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix}, \qquad \text{with } E_1 \neq E_2 ,$$

while the operator of some observable A of this system, by the matrix

$$\mathrm{A} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

For the system's state with the energy definitely equal to E_1, find the possible results of measurements of the observable A and the probabilities of the corresponding measurement outcomes.

4.20. Three states $u_{1,2,3}$ form a full and orthonormal basis of a system with the following Hamiltonian

$$\hat{H} = -\delta \left(|u_1\rangle\langle u_2| + |u_2\rangle\langle u_3| + |u_3\rangle\langle u_1| \right) + \text{h.c.},$$

where δ is a real constant, while h.c. means the Hermitian conjugate of the previous expression. Calculate its stationary states and energy levels. Can you relate this system to any other(s) discussed earlier in the course?

4.21. Guided by Eq. (2.203), and by the solutions of the previous problem and also of Problem 3.15, suggest a Hamiltonian describing particle's dynamics in an infinite 1D chain of similar potential wells within the tight-binding approximation, in the bra-ket formalism. Verify that its eigenstates and eigenvalues correspond to those discussed in Sec. 2.7.

4.22. In a certain full and orthonormal basis of three states $u_{1,2,3}$, operators $\hat{A}$ and $\hat{B}$ are defined by the following equalities:

$$\hat{A}|u_1\rangle = |u_3\rangle, \quad \hat{A}|u_2\rangle = |u_2\rangle, \quad \hat{A}|u_3\rangle = |u_1\rangle; \qquad \hat{B}|u_1\rangle = |u_1\rangle, \quad \hat{B}|u_2\rangle = 0, \quad \hat{B}|u_3\rangle = -|u_3\rangle .$$

(i) Prove that the operators $\hat{A}^2$ and $\hat{B}$ commute and form an orthonormal basis of their common eigenstates.

(ii) Give the most general expression for the matrix (in the *u*-basis) of an operator that would commute with $\hat{B}$.

4.23. Calculate the eigenvectors and eigenvalues of the following matrices:

$$\mathrm{A} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad \mathrm{B} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

4.24. A certain state γ is an eigenstate of each of two operators, $\hat{A}$ and $\hat{B}$. What can be said about the corresponding eigenvalues *a* and *b*, if the operators anticommute?

4.25. An operator $\hat{A}$ commutes with each of two other operators $\hat{B}$ and $\hat{C}$, but these two operators do not commute: $\left[\hat{B},\hat{C}\right] \neq 0$. Prove that the full set of eigenvalues of the operator $\hat{A}$ includes some degenerate ones.

4.26. Derive the differential equation for the time evolution of the expectation value of an observable, by using (i) the Schrödinger picture and (ii) the Heisenberg picture of quantum dynamics.

4.27. At $t = 0$, a spin-½ whose interaction with an external field is described by the Hamiltonian

$$\hat{H} = \mathbf{c}\cdot\hat{\boldsymbol{\sigma}} \equiv c_x\hat{\sigma}_x + c_y\hat{\sigma}_y + c_z\hat{\sigma}_z$$

(where $c_{x,y,z}$ are real *c*-number constants, and $\hat{\sigma}_{x,y,z}$ are the Pauli operators) was in the state $\uparrow$, one of the two eigenstates of $\hat{\sigma}_z$. In the Schrödinger picture, calculate the time evolution of:

(i) the ket-vector $|\alpha\rangle$ of the spin (in any time-independent basis you like),
(ii) the probabilities to find the spin in the states $\uparrow$ and $\downarrow$, and
(iii) the expectation values of all three Cartesian components of the spin vector.

Analyze and interpret the results for the particular case $c_y = c_z = 0$.

Hint: Think about the best basis to use for the solution.

4.28. For the same system as in the previous problem, use the Heisenberg picture to calculate the time evolution of:

(i) all three Cartesian components of the spin operator $\hat{\mathbf{S}}_{\mathrm{H}}(t)$, and
(ii) the expectation values of the spin components.

Compare the latter results with those of the previous problem.

4.29. For the same system as in the two previous problems, calculate the matrix elements of the operator $\hat{\sigma}_z$ in the basis of the stationary states of the system.

4.30. In the Schrödinger picture of quantum dynamics, certain three operators satisfy the following commutation relation:

$$\left[\hat{A},\hat{B}\right]=\hat{C}\,.$$

What is their relation in the Heisenberg picture, at a certain time instant t?

4.31. Prove the Bloch theorem given by either Eq. (3.107) or Eq. (3.108), where **R** is an arbitrary vector of the Bravais lattice (3.106).

Hint: Analyze the commutation properties of the so-called *translation operator* $\hat{\mathcal{T}}_{\mathbf{R}}$, defined by the following result of its action on an arbitrary function $f(\mathbf{r})$:

$$\hat{\mathcal{T}}_{\mathbf{R}} f(\mathbf{r}) = f(\mathbf{r}+\mathbf{R}),$$

and apply them to an eigenfunction $\psi(\mathbf{r})$ of the stationary Schrödinger equation for a particle moving in the periodic potential described by Eq. (3.105).

4.32. A constant force F is applied to an (otherwise free) 1D particle of mass m. Calculate the stationary wavefunctions of the particle in:

(i) the coordinate representation, and
(ii) the momentum representation.

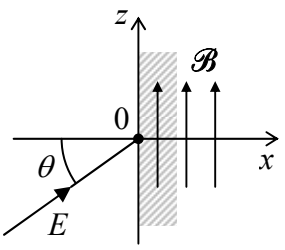
Discuss the relation between the results.

4.33. Use the momentum representation to re-solve the problem discussed at the beginning of Sec. 2.6, i.e. calculate the eigenenergy of a 1D particle of mass m, localized in a very short potential well of "weight" $\mathcal{W}$.

4.34. The momentum representation of a certain operator of orbital 1D motion is p^{-1}. Use two different approaches to find its coordinate representation.

4.35.* For a particle moving in a 3D periodic potential, develop the bra-ket formalism for the **q**-representation, in which a complex amplitude similar to a_q in Eq. (2.234) (but generalized to 3D and all energy bands) plays the role of the wavefunction. In particular, calculate the operators **r** and **v** in this representation, and use the result to prove Eq. (2.237) for the 1D case in the low-field limit.

4.36. A uniform, time-independent magnetic field $\mathcal{B} = \mathbf{n}_z\mathcal{B}$ is induced in one semi-space, while the other semi-space is field-free, with a sharp plane boundary $x = 0$ between these two regions – see figure on the right. A monochromatic beam of non-relativistic, electrically-neutral spin-½ particles with a gyromagnetic ratio $\gamma \neq 0$,[65] in a certain spin state and with a kinetic energy E, propagating within the $[x, z]$ plane, is incident on this boundary from the field-free side under angle θ. Calculate the coefficient of particle reflection from the boundary.



[65] The fact that γ may be different from zero even for electrically-neutral particles such as neutrons, is explained by the Standard Model of the elementary particles, in which a neutron "consists" (in a broad sense of this word) of three electrically-charged quarks with zero net charge.

Chapter 5. Some Exactly Solvable Problems

The objective of this chapter is to describe several relatively simple but important applications of the bra-ket formalism, including a few core problems of wave mechanics we have already started to discuss in Chapters 2 and 3.

5.1. Two-level systems

The discussion of the bra-ket formalism in the previous chapter was peppered with numerous illustrations of its main concepts on the examples of "spin-½-like" systems with the smallest non-trivial (two-dimensional) Hilbert space. In such a system, the bra- and ket-vectors of an arbitrary quantum state α may be represented as a linear superposition of just two basis vectors, for example

$$|\alpha\rangle = \alpha_{\uparrow}|\uparrow\rangle + \alpha_{\downarrow}|\downarrow\rangle, \tag{5.1}$$

where the states $\uparrow$ and $\downarrow$ are defined as the eigenstates of the Pauli matrix σ_z – see Eq. (4.105). For the genuine spin-½ particles (such as electrons) placed in a z-oriented time-independent magnetic field, these states are the stationary "spin-up" and "spin-down" stationary states of the Pauli Hamiltonian (4.163), with the corresponding two energy levels (4.167).

However, an approximate but reasonable quantum description of some other important systems may also be given in such Hilbert space. For example, as was discussed in Sec. 2.6, two weakly coupled space-localized orbital states of a spin-free particle are sufficient for an approximate description of its quantum oscillations between two potential wells. A similar coupling of two traveling waves explains the energy band splitting in the weak-potential approximation of the band theory – Sec. 2.7. As will be shown in the next chapter, in systems with time-independent Hamiltonians, such a situation almost unavoidably appears each time when two energy levels are much closer to each other than to other levels. Moreover, as will be shown in Sec. 6.5, a similar truncated description is adequate even in cases when two levels E_n and $E_{n'}$ of an unperturbed system are not close to each other, but the corresponding states become coupled by an applied ac field of a frequency ω very close to the difference $(E_n - E_{n'})/\hbar$. Such *two-level systems* are nowadays the focus of additional attention in the view of prospects of their use for quantum information processing and encryption.[1] This is why I will spend a bit more time reviewing the main properties of an arbitrary two-level system.

The most general form of the Hamiltonian of a two-level system is represented, in an arbitrary basis, by a 2×2 matrix

$$\mathrm{H} = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix}. \tag{5.2}$$

According to the discussion in Secs. 4.3-4.5, since the Hamiltonian operator has to be Hermitian, the diagonal elements of the matrix H have to be real, and its off-diagonal elements have to be complex

[1] In the last context, to be discussed in Sec. 8.5, the two-level systems are usually called *qubits*.

conjugate: $H_{21} = H_{12}^*$. As a result, we may not only represent H as a linear combination (4.106) of the identity matrix and the Pauli matrices but also reduce it to a more specific form:

$$\mathrm{H} = b\mathrm{I} + \mathbf{c}\cdot\boldsymbol{\sigma} = \begin{pmatrix} b + c_z & c_x - ic_y \\ c_x + ic_y & b - c_z \end{pmatrix} \equiv \begin{pmatrix} b + c_z & c_- \\ c_+ & b - c_z \end{pmatrix}, \qquad \text{with } c_\pm \equiv c_x \pm ic_y, \tag{5.3}$$

where the scalar b and the Cartesian components of the vector $\mathbf{c}$ are *real* c-number coefficients:

$$b = \frac{H_{11} + H_{22}}{2}, \quad c_x = \frac{H_{12} + H_{21}}{2} \equiv \mathrm{Re}\, H_{21}, \quad c_y = \frac{H_{21} - H_{12}}{2i} \equiv \mathrm{Im}\, H_{21}, \quad c_z = \frac{H_{11} - H_{22}}{2}. \tag{5.4}$$

If such a Hamiltonian does not depend on time, the corresponding characteristic equation (4.103) for the system's energy levels $E_\pm$,

$$\begin{vmatrix} b + c_z - E & c_- \\ c_+ & b - c_z - E \end{vmatrix} = 0, \tag{5.5}$$

is a simple quadratic equation, with the following roots:

$$E_\pm = b \pm c \equiv b \pm \left(c_+ c_- + c_z^2\right)^{1/2} \equiv b \pm \left(c_x^2 + c_y^2 + c_z^2\right)^{1/2} \equiv \frac{H_{11} + H_{22}}{2} \pm \left[\left(\frac{H_{11} - H_{22}}{2}\right)^2 + \left|H_{21}\right|^2\right]^{1/2}. \tag{5.6}$$

The parameter $b \equiv (H_{11} + H_{22})/2$ evidently gives the average energy $E^{(0)}$ of the system, which does not contribute to the level splitting

$$\Delta E \equiv E_+ - E_- = 2c \equiv 2\left(c_x^2 + c_y^2 + c_z^2\right)^{1/2} \equiv \left[\left(H_{11} - H_{22}\right)^2 + 4\left|H_{21}\right|^2\right]^{1/2}. \tag{5.7}$$

So, the splitting is a hyperbolic function of the coefficient $c_z \equiv (H_{11} - H_{22})/2$. A plot of this function is the famous level-anticrossing diagram (Fig. 1), which has already been discussed in Sec. 2.7 in the particular context of the weak-potential limit of the 1D band theory.

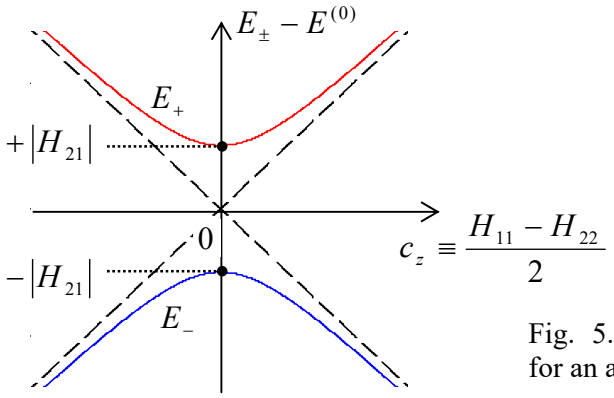


Fig. 5.1. The level-anticrossing diagram for an arbitrary two-level system.

The physics of the diagram becomes especially clear if the two states of the basis used to spell out the matrix (2) may be interpreted as the stationary states of two potentially independent subsystems, with the energies, respectively, H_{11} and H_{22}. (For example, in the case of two weakly coupled potential wells discussed in Sec. 2.6, these are the ground-state energies of two very distant wells.) Then the off-diagonal elements $c_- \equiv H_{12}$ and $c_+ \equiv H_{21} = H_{12}^*$ describe the subsystem coupling, and the anticrossing

diagram shows how do the eigenenergies of the coupled system depend (at fixed coupling) on the difference of the subsystem energies. As was already discussed in Sec. 2.7, the most striking feature of the diagram is that any non-zero coupling $|c_\pm| \equiv (c_x^2 + c_y^2)^{1/2}$ changes the topology of the eigenstate energies, creating a gap of the width ΔE.

As it follows from our discussions of particular two-level systems in Secs. 2.6 and 4.6, the dynamics of such systems also has a general feature – the quantum oscillations. Namely, if we put any two-level system into any initial state different from one of its eigenstates ±, and then let it evolve on its own, the probability of its finding the system in any of the "partial" states exhibits oscillations with the frequency

$$\Omega = \frac{2c}{\hbar} = \frac{\Delta E}{\hbar} \equiv \frac{E_+ - E_-}{\hbar}, \tag{5.8}$$

lowest at the exact subsystem symmetry ($c_z = 0$, i.e. $H_{11} = H_{22}$), when it is proportional to the coupling strength: $\Omega_{min} = 2|c_\pm|/\hbar \equiv 2|H_{12}|/\hbar = 2|H_{21}|/\hbar$.

In the case discussed in Sec. 2.6, these are the oscillations of a particle between the two coupled potential wells (or rather of the probabilities to find it in either well) – see, e.g., Eqs. (2.181). On the other hand, for a spin-½ particle in an external magnetic field, these oscillations take the form of spin precession in the plane normal to the field, with periodic oscillations of its Cartesian components (or rather their expectation values) – see, e.g., Eqs. (4.173)-(4.174). Some other examples of the quantum oscillations in two-level systems may be rather unexpected; for example, the ammonium molecule NH_3 (Fig. 2) has two similar states that differ only by the inversion of the nitrogen atom relative to the common plane of the three hydrogen atoms. These states are weakly coupled due to the quantum-mechanical tunneling of the nitrogen atom through this plane.[2] Since for this particular molecule, in the absence of external fields, the level splitting ΔE corresponds to an experimentally convenient frequency $\Omega/2\pi \approx 24$ GHz, it played an important historic role in the initial development of the atomic frequency standards and microwave quantum generators (*masers*) in the early 1950s,[3] which paved the way for laser technology.

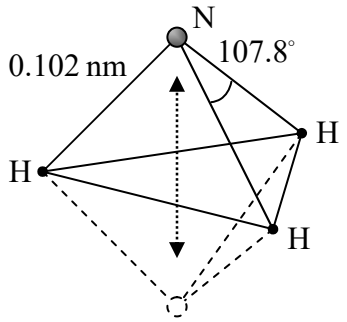


Fig. 5.2. An ammonia molecule and its inversion.

Now let us now discuss a very convenient geometric representation of an arbitrary quantum state α of any two-level system. As Eq. (1) shows, such a state is completely described by two complex

[2] Since the hydrogen atoms are much lighter, it may be fairer to speak about the tunneling of their triangle around the (nearly immobile) nitrogen atom.

[3] In particular, these molecules were used in the demonstration of the first maser by C. Townes' group in 1954.

coefficients (c-numbers) – say, $\alpha_\uparrow$ and $\alpha_\downarrow$. If the vectors of the basis states $\uparrow$ and $\downarrow$ are normalized, then these coefficients must obey the following restriction:

$$W_\Sigma = \langle\alpha|\alpha\rangle = \Big(\langle\uparrow|\alpha_\uparrow^* + \langle\downarrow|\alpha_\downarrow^*\Big)\Big(\alpha_\uparrow|\uparrow\rangle + \alpha_\downarrow|\downarrow\rangle\Big) = \alpha_\uparrow^*\alpha_\uparrow + \alpha_\downarrow^*\alpha_\downarrow = |\alpha_\uparrow|^2 + |\alpha_\downarrow|^2 = 1. \quad (5.9)$$

This requirement is automatically satisfied if we take the moduli of $\alpha_\uparrow$ and $\alpha_\downarrow$ equal to the sine and cosine of the same real angle. Thus we may write, for example,

$$\alpha_\uparrow = \cos\frac{\theta}{2}e^{i\gamma}, \quad \alpha_\downarrow = \sin\frac{\theta}{2}e^{i(\gamma+\varphi)}. \quad (5.10)$$

Moreover, according to the general Eq. (4.125), if we deal with just one two-level system,[4] the common phase factor $\exp\{i\gamma\}$ drops out of the calculation of any expectation value, so we may take $\gamma = 0$, and Eq. (10) is reduced to

$$\boxed{\alpha_\uparrow = \cos\frac{\theta}{2}, \quad \alpha_\downarrow = \sin\frac{\theta}{2}e^{i\varphi}.} \quad (5.11)$$

Bloch sphere representation

The reason why the argument of these sine and cosine functions is usually taken in the form $\theta/2$, is clear from Fig. 3a: Eq. (11) conveniently maps each state α of a two-level system onto a certain *representation point* on a unit-radius *Bloch sphere*,[5] with the polar angle θ and the azimuthal angle φ.

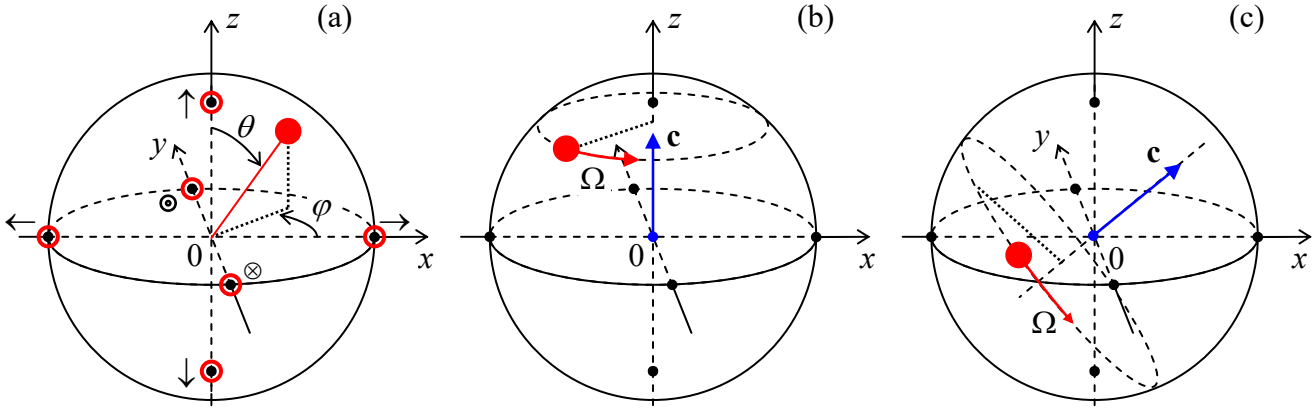


Fig. 5.3. The Bloch sphere: (a) the representation of an arbitrary state (solid red point) and the eigenstates of the Pauli matrices (black-dotted points), and (b, c) the two-level system's evolution: (b) in a constant "field" **c** directed along the z-axis, and (c) in an arbitrarily orientated field.

In particular, the basis state $\uparrow$, described by Eq. (1) with $\alpha_\uparrow = 1$ and $\alpha_\downarrow = 0$, corresponds to the North Pole of the sphere ($\theta = 0$), while the opposite state $\downarrow$, with $\alpha_\uparrow = 0$ and $\alpha_\downarrow = 1$, to its South Pole ($\theta = \pi$). Similarly, the eigenstates $\rightarrow$ and $\leftarrow$ of the matrix σ_x, described by Eqs. (4.122), i.e. having $\alpha_\uparrow =$

[4] If you need a reminder of why this condition is crucial, please revisit the discussion at the end of Sec. 1.6. Note also that the mutual phase shifts between different two-level systems are important, in particular, for quantum information processing (see Sec. 8.5 below), so most discussions of these applications have to start from Eq. (10) rather than Eq. (11).

[5] This representation was suggested in 1946 by the same Felix Bloch who pioneered the energy band theory discussed in Chapters 2-3.

$1/\sqrt{2}$ and $\alpha_\downarrow = \pm 1/\sqrt{2}$, correspond to the equator ($\theta = \pi/2$) points with, respectively, $\varphi = 0$ and $\varphi = \pi$. Two more special points (denoted in Fig. 3a as ⊙ and ⊗) are also located on the sphere's equator, at $\theta = \pi/2$ and $\varphi = \pm\pi/2$; it is easy to check that they correspond to the eigenstates of the matrix σ_y (in the same z-basis).

To understand why this mutually perpendicular location of these three special point pairs on the Bloch sphere is not occasional, let us plug Eqs. (11) into Eqs. (4.131)-(4.133) for the expectation values of the spin-½ components. In terms of the Pauli vector operator (4.117), $\hat{\boldsymbol{\sigma}} \equiv \hat{\mathbf{S}}/(\hbar/2)$, the result is

$$\langle \sigma_x \rangle = \sin\theta\cos\varphi, \qquad \langle \sigma_y \rangle = \sin\theta\sin\varphi, \qquad \langle \sigma_z \rangle = \cos\theta, \tag{5.12}$$

showing that the radius vector of any representation point is just the expectation value $\langle \boldsymbol{\sigma} \rangle$.

Now let us use Eq. (3) to see how the representation point moves in various cases, ignoring the term $b\mathrm{I}$ – which, again, describes the offset of the total energy of the system relative to some reference level, and does not affect its dynamics. First of all, according to Eq. (4.158), if $\mathbf{c} = 0$ (when the Hamiltonian operator turns to zero, and hence the state vectors do not depend on time) the point does not move at all, and its position is determined by initial conditions, i.e. by the system's preparation. If $\mathbf{c} \neq 0$, we may re-use some results of Sec. 4.6, obtained for the Pauli Hamiltonian (4.163a), which coincides with Eq. (3) if[6]

$$\mathbf{c} = -\gamma \mathscr{B} \frac{\hbar}{2}. \tag{5.13}$$

In particular, if the field $\mathscr{B}$, and hence the vector $\mathbf{c}$, is directed along the z-axis and is time-independent, Eqs. (4.170) and (4.173)-(4.174) show that the representation point $\langle \boldsymbol{\sigma} \rangle$ on the Bloch sphere rotates within a plane normal to this axis (see Fig. 3b) with the angular velocity

$$\frac{d\varphi}{dt} \equiv \Omega = -\gamma \mathscr{B}_z \equiv \frac{2c_z}{\hbar}. \tag{5.14}$$

Almost evidently, since the selection of the coordinate axes is arbitrary, this picture should remain valid for any orientation of the vector $\mathbf{c}$, with the representation point rotating, on the Bloch sphere, around its direction, with the angular speed $|\Omega| = 2c/\hbar$ – see Fig. 3c. This fact may be proved using any picture of the quantum dynamics, discussed in Sec. 4.6. Actually, the reader may already have done that by solving Problems 4.27-4.28, just to see that even for the particular, simple initial state of the system (↑), the final results for the Cartesian components of the vector $\langle \boldsymbol{\sigma} \rangle$ are somewhat bulky. However, this description may be readily simplified, even for an arbitrary time dependence of the "field" vector $\mathbf{c}(t)$ in Eq. (3), by using the geometric vector language.

For that, let us rewrite Eq. (3) (again, with $b = 0$) in the operator form,

$$\hat{H} = \mathbf{c}(t) \cdot \hat{\boldsymbol{\sigma}}, \tag{5.15}$$

valid in an arbitrary basis. According to Eq. (4.199), the corresponding Heisenberg equation of motion for the j^{th} Cartesian components of the vector-operator $\hat{\boldsymbol{\sigma}}$ (which does not depend on time explicitly, $\partial\hat{\sigma}/\partial t = 0$) is

[6] This correspondence justifies using the use of the term "field" for the vector $\mathbf{c}$.

$$i\hbar\dot{\hat{\sigma}}_j = \left[\hat{\sigma}_j, \hat{H}\right] \equiv \left[\hat{\sigma}_j, \mathbf{c}(t)\cdot\hat{\boldsymbol{\sigma}}\right] \equiv \left[\hat{\sigma}_j, \sum_{j'=1}^{3} c_{j'}(t)\hat{\sigma}_{j'}\right] \equiv \sum_{j'=1}^{3} c_{j'}(t)\left[\hat{\sigma}_j, \hat{\sigma}_{j'}\right]. \tag{5.16}$$

Now using the commutation relations (4.155), which remain valid in any basis and in any picture of time evolution,[7] we get

$$i\hbar\dot{\hat{\sigma}}_j = 2i\sum_{j',j''=1}^{3} c_{j'}(t)\hat{\sigma}_{j''}\varepsilon_{jj'j''}\,, \tag{5.17}$$

where $\varepsilon_{jj'j''}$ is the Levi-Civita symbol. But it is straightforward to verify that the usual vector product of two 3D vectors may be represented in a similar Cartesian-component form:

$$\left(\mathbf{a}\times\mathbf{b}\right)_j = \begin{vmatrix} \mathbf{n}_1 & \mathbf{n}_2 & \mathbf{n}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}_j = \sum_{j',j''=1}^{3} a_{j'}b_{j''}\varepsilon_{jj'j''}\,, \tag{5.18}$$

As a result, the right-hand side of Eq. (17) may be rewritten as $2i\left[\mathbf{c}(t)\times\hat{\boldsymbol{\sigma}}\right]_j$, and that relation may be recast in a vector form – or rather several equivalent forms:

$$i\hbar\dot{\hat{\boldsymbol{\sigma}}} = 2i\mathbf{c}(t)\times\hat{\boldsymbol{\sigma}}, \qquad \text{or } \dot{\hat{\boldsymbol{\sigma}}} = \frac{2}{\hbar}\mathbf{c}(t)\times\hat{\boldsymbol{\sigma}}, \qquad \text{or } \dot{\hat{\boldsymbol{\sigma}}} = \boldsymbol{\Omega}(t)\times\hat{\boldsymbol{\sigma}}, \tag{5.19}$$

where the vector $\boldsymbol{\Omega}$ is defined as

$$\frac{\hbar}{2}\boldsymbol{\Omega}(t) \equiv \mathbf{c}(t) \tag{5.20}$$

– an evident vector generalization of Eq. (14).[8] As we have seen in Sec. 4.6, any linear relation between two Heisenberg operators is also valid for the expectation values of the corresponding observables, so the last form of Eq. (19) yields:

$$\langle\dot{\boldsymbol{\sigma}}\rangle = \boldsymbol{\Omega}(t)\times\langle\boldsymbol{\sigma}\rangle. \tag{5.21}$$

But this is the well-known kinematic formula[9] for the rotation of a constant-length classical 3D vector $\langle\boldsymbol{\sigma}\rangle$ around the instantaneous direction of the vector $\boldsymbol{\Omega}(t)$, with the instantaneous angular velocity $\Omega(t)$. So, the time evolution of the representation point on the Bloch sphere is quite simple, especially in the case of a time-independent $\mathbf{c}$, and hence $\boldsymbol{\Omega}$ – see Fig. 3c.[10] Note that it is sufficient to turn off the field to stop the precession instantly. (Since Eq. (21) is the first-order differential equation, the representation point has no effective inertia.[11]) Hence, changing the direction and the magnitude of the

---

[7] Indeed, if some three operators in the Schrödinger picture are related as $[\hat{A}_S, \hat{B}_S] = \hat{C}_S$, then according to Eq. (4.190), in the Heisenberg picture:

$$[\hat{A}_H, \hat{B}_H] = [\hat{u}^\dagger\hat{A}_H\hat{u}, \hat{u}^\dagger\hat{B}_H\hat{u}] \equiv \hat{u}^\dagger\hat{A}_H\hat{u}\hat{u}^\dagger\hat{B}_H\hat{u} - \hat{u}^\dagger\hat{B}_H\hat{u}\hat{u}^\dagger\hat{A}_H\hat{u} \equiv \hat{u}^\dagger[\hat{A}_S, \hat{B}_S]\hat{u} \equiv \hat{u}^\dagger\hat{C}_S\hat{u} = \hat{C}_H.$$

[8] It is also easy to verify that in the particular case $\boldsymbol{\Omega} = \Omega\mathbf{n}_z$, Eqs. (19) are reduced, in the z-basis, to Eqs. (4.200) for the spin-½ vector matrix $\mathbf{S} = (\hbar/2)\boldsymbol{\sigma}$.

[9] See, e.g., CM Sec. 4.1, in particular Eq. (4.8).

[10] The bulkiness of the solutions of Problems 4.27-4.28 (which were offered just as useful exercises in quantum dynamic formalisms) reflects the awkward expression of the resulting simple circular motion of the vector $\langle\boldsymbol{\sigma}\rangle$ (see Fig. 3c) via its Cartesian components.

[11] This is also true for the classical angular momentum $\mathbf{L}$ at its torque-induced precession – see, e.g., CM Sec. 4.5.

effective external field, it is possible to drive the representation point of a two-level system from any initial position to any final position on the Bloch sphere, i.e. make the system take any of its possible quantum states.

In the particular case of a spin-½ in a magnetic field $\mathscr{B}(t)$, it is more customary to use Eqs. (13) and (20) to rewrite Eq. (21) as the following equation for the expectation value of the spin vector $\mathbf{S} = (\hbar/2)\boldsymbol{\sigma}$:

$$\left\langle \dot{\mathbf{S}} \right\rangle \equiv \gamma \left\langle \mathbf{S} \right\rangle \times \mathscr{B}(t). \tag{5.22}$$

As we know from the discussion in Chapter 4, such a classical description of the spin's evolution does not give a full picture of the quantum reality; in particular, it does not describe the possible large uncertainties of its components – see, e.g., Eqs. (4.135). The situation, however, is different for a collection of $N >> 1$ similar, non-interacting spins, initially prepared to be in the same state – for example by polarizing all spins with a strong external field $\mathscr{B}_0$, at relatively low temperatures T, with $k_B T << \gamma \mathscr{B}_0 \hbar$. (A practically important example of such a collection is a set of nuclear spins in macroscopic condensed-matter samples, where the spin interaction with each other and the environment is typically very small.) For such a collection, Eq. (22) is still valid, while the relative uncertainty of the resulting sample's magnetization $\mathbf{M} = n\langle \mathbf{m} \rangle = n\gamma\langle \mathbf{S} \rangle$ (where $n \equiv N/V$ is the spin density) is proportional to $1/N^{1/2} << 1$. Thus, the evolution of magnetization may be described, with good precision, by the essentially classical equation:

$$\dot{\mathbf{M}} = \gamma \mathbf{M} \times \mathscr{B}(t). \tag{5.23}$$

This equation, or the equivalent set of three *Bloch equations*[12] for its Cartesian components, with the right-hand side augmented with small terms describing the effects of dephasing and relaxation (to be discussed in Chapter 7), is used, in particular, to describe the *magnetic resonance*, taking place when the frequency (4.164) of the magnetization's precession in a strong dc magnetic field approaches the frequency of an additionally applied (and usually weak) ac/rf field. Two species of this effect, the electron paramagnetic resonance (EPR) and the nuclear magnetic resonance (NMR) are broadly used in material science, chemistry, and medicine. Unfortunately, I will not have time to discuss the related technical issues and methods (in particular, interesting ac/rf pulsing techniques, including the so-called *spin echo* and *Ramsey interferometry*) in detail, and have to refer the reader to special literature.[13]

5.2. The Ehrenfest theorem

In Sec. 4.7, we have derived all the basic relations of wave mechanics from the bra-ket formalism, which will also enable us to get some important additional results in that area. One of them is a pair of very interesting relations, together called the *Ehrenfest theorem*. To derive them, for the simplest case of 1D orbital motion, let us calculate the following commutator:

[12] They were introduced by F. Bloch in the same 1946 paper as the Bloch-sphere representation. In the 1950s when the value of Eq. (21) for quantum optics became recognized, this equation and its open-system generalizations became known as *optical Bloch equations*. Currently, the term 'Bloch equations' is frequently used for any two-level systems, regardless of the physical origin of the Hamiltonian (15).

[13] For introductions see, e.g., J. Wertz and J. Bolton, *Electron Spin Resonance*, 2nd ed., Wiley, 2007; J. Keeler, *Understanding NMR Spectroscopy,* 2nd ed., Wiley, 2010.

$$\left[\hat{x},\hat{p}_x^2\right] \equiv \hat{x}\hat{p}_x\hat{p}_x - \hat{p}_x\hat{p}_x\hat{x}. \tag{5.24}$$

Let us apply the commutation relation (4.238), in the following form:

$$\hat{x}\hat{p}_x = \hat{p}_x\hat{x} + i\hbar\hat{I}, \tag{5.25}$$

to the first term of the right-hand side of Eq. (24) twice, with the goal to chase the coordinate operator into the rightmost position:

$$\hat{x}\hat{p}_x\hat{p}_x = \left(\hat{p}_x\hat{x} + i\hbar\hat{I}\right)\hat{p}_x \equiv \hat{p}_x\hat{x}\hat{p}_x + i\hbar\hat{p}_x = \hat{p}_x\left(\hat{p}_x\hat{x} + i\hbar\hat{I}\right) + i\hbar\hat{p}_x \equiv \hat{p}_x\hat{p}_x\hat{x} + 2i\hbar\hat{p}_x. \tag{5.26}$$

The first term of this result cancels with the last term of Eq. (24), so the commutator becomes quite simple:

$$\left[\hat{x},\hat{p}_x^2\right] = 2i\hbar\hat{p}_x. \tag{5.27}$$

Let us use this equality to calculate the Heisenberg-picture equation of motion of the operator $\hat{x}$, by applying the general Heisenberg equation (4.199) to the 1D orbital motion described by the Hamiltonian (4.237), but possibly with a more general, time-dependent potential energy U:

$$\frac{d\hat{x}}{dt} = \frac{1}{i\hbar}\left[\hat{x},\hat{H}\right] = \frac{1}{i\hbar}\left[\hat{x},\frac{\hat{p}_x^2}{2m} + U(\hat{x},t)\right]. \tag{5.28}$$

The potential energy operator is a function of the coordinate operator and hence, as we know, commutes with it. Thus, the right-hand side of Eq. (28) is proportional to the commutator (27), and we get

$$\boxed{\frac{d\hat{x}}{dt} = \frac{\hat{p}_x}{m}}. \tag{5.29}$$

Heisenberg equation for coordinate

In this operator equation, we readily recognize the full analog of the classical relation between the particle's momentum and its velocity.

Now let us see what a similar procedure gives for the momentum's derivative:

$$\frac{d\hat{p}_x}{dt} = \frac{1}{i\hbar}\left[\hat{p}_x,\hat{H}\right] = \frac{1}{i\hbar}\left[\hat{p}_x,\frac{\hat{p}_x^2}{2m} + U(\hat{x},t)\right]. \tag{5.30}$$

The kinetic energy operator commutes with the momentum operator and hence drops from the right-hand side of this equation. To calculate the remaining commutator of the momentum and the potential energy, let us use the fact that any smooth (infinitely differentiable) function may be represented by its Taylor expansion:

$$U(\hat{x},t) = \sum_{k=0}^{\infty}\frac{1}{k!}\frac{\partial^k U}{\partial\hat{x}^k}\hat{x}^k, \tag{5.31}$$

where the derivatives of U may be understood as c-numbers (evaluated at $x = 0$, and the given time t), so we may write

$$\left[\hat{p}_x,U(\hat{x},t)\right] = \sum_{k=0}^{\infty}\frac{1}{k!}\frac{\partial^k U}{\partial\hat{x}^k}\left[\hat{p}_x,\hat{x}^k\right] = \sum_{k=0}^{\infty}\frac{1}{k!}\frac{\partial^k U}{\partial\hat{x}^k}\left(\hat{p}_x\underbrace{\hat{x}\hat{x}..\hat{x}}_{k\text{ times}} - \underbrace{\hat{x}\hat{x}...\hat{x}}_{k\text{ times}}\hat{p}_x\right). \tag{5.32a}$$

Applying Eq. (25) k times to the last term in the parentheses, exactly as we did in Eq. (26), we get

$$\left[\hat{p}_x, U(\hat{x},t)\right] = -\sum_{k=1}^{\infty}\frac{1}{k!}\frac{\partial^k U}{\partial \hat{x}^k} ik\hbar \hat{x}^{k-1} \equiv -i\hbar\sum_{k=1}^{\infty}\frac{1}{(k-1)!}\frac{\partial^k U}{\partial \hat{x}^k}\hat{x}^{k-1}. \tag{5.32b}$$

But the last sum is just the Taylor expansion of the derivative $\partial U/\partial x$. Indeed,

$$\frac{\partial U}{\partial \hat{x}} = \sum_{k'=0}^{\infty}\frac{1}{k'!}\frac{\partial^{k'}}{\partial \hat{x}^{k'}}\left(\frac{\partial U}{\partial \hat{x}}\right)\hat{x}^{k'} = \sum_{k'=0}^{\infty}\frac{1}{k'!}\frac{\partial^{k'+1} U}{\partial \hat{x}^{k'+1}}\hat{x}^{k'} = \sum_{k=1}^{\infty}\frac{1}{(k-1)!}\frac{\partial^k U}{\partial \hat{x}^k}\hat{x}^{k-1}, \tag{5.33}$$

where at the last step, the summation index was changed from k' to $k - 1$. As a result, we may recast Eq. (5.32b) as

$$\left[\hat{p}_x, U(\hat{x},t)\right] = -i\hbar\frac{\partial}{\partial \hat{x}}U(\hat{x},t), \tag{5.34}$$

so Eq. (30) yields:

Heisenberg equation for momentum

$$\boxed{\frac{d\hat{p}_x}{dt} = -\frac{\partial}{\partial \hat{x}}U(\hat{x},t).} \tag{5.35}$$

This equation also coincides with the classical equation of motion! Moreover, averaging Eqs. (29) and (35) over the initial state (as Eq. (4.191) prescribes), we get similar results for the expectation values:[14]

Ehrenfest theorem

$$\boxed{\frac{d\langle x\rangle}{dt} = \frac{\langle p_x\rangle}{m}, \qquad \frac{d\langle p_x\rangle}{dt} = -\left\langle\frac{\partial U}{\partial x}\right\rangle.} \tag{5.36}$$

However, it is important to remember that the similarity of these quantum-mechanical equations and their classical mechanics analogs is superficial, and the degree of difference between the two mechanics very much depends on the problem. As one extreme, let us consider the case when a particle's state, at any moment between t_0 and t, may be accurately represented by one, relatively p_x-narrow wave packet. Then we may interpret Eqs. (36) as the equations of the essentially classical motion of the wave packet's center, and consider this fact as a manifestation of the correspondence principle. However, even in this case, it is important to remember the purely quantum mechanical effects of non-zero wave packet broadening and its spread in time, which were discussed in Sec. 2.2.

As an opposite extreme, let us revisit the "leaky" potential well discussed in Sec. 2.5 – see Fig. 2.15. Since both the potential $U(x)$ and the initial wavefunction of that system are symmetric relative to point $x = 0$ at all times, the right-hand sides of both Eqs. (36) identically equal zero, and hence they predict that the average values of the coordinate and the momentum stay equal to zero at all times. Of course, this prediction is correct, but it does not tell us much about the rich dynamics of the system: the finite lifetime of the metastable state, the formation of two wave packets, their waveform and propagation speed (see Fig. 2.17), and about the insights the full solution gives for the quantum measurement theory and the system's irreversibility. Another similar example is the energy band theory (Sec. 2.7), with its purely quantum effect of the allowed energy bands and forbidden energy gaps, of which Eqs. (36) give no clue.

To summarize, the Ehrenfest theorem is useful as an illustration of the correspondence principle and as the sanity check of quantum-mechanical calculation results, but its predictive power should not be exaggerated.

[14] The equation set (36) constitutes the *Ehrenfest theorem*, named after its author, Paul Ehrenfest.

5.3. The Feynman path integral

As has been already mentioned, even within the realm of wave mechanics, the bra-ket language may simplify some calculations that would be very bulky using the notation used in Chapters 1-3. Probably the best example is the famous alternative, *path-integral* formulation of quantum mechanics.[15] I will review this important concept, cutting one math corner for the sake of brevity.[16] (This shortcut will be clearly marked below.)

Let us inner-multiply both parts of Eq. (4.157a), which is essentially the definition of the time-evolution operator, by the bra-vector of state x,

$$\langle x|\alpha(t)\rangle = \langle x|\hat{u}(t,t_0)|\alpha(t_0)\rangle, \tag{5.37}$$

insert the identity operator before the ket-vector on the right-hand side, and then use the closure condition in the form of Eq. (4.252), with x' replaced with x_0:

$$\langle x|\alpha(t)\rangle = \int dx_0 \langle x|\hat{u}(t,t_0)|x_0\rangle\langle x_0|\alpha(t_0)\rangle. \tag{5.38}$$

According to Eq. (4.233), this equality may be represented as

$$\Psi_\alpha(x,t) = \int dx_0 \langle x|\hat{u}(t,t_0)|x_0\rangle \Psi_\alpha(x_0,t_0). \tag{5.39}$$

Comparing this expression with Eq. (2.44), we see that the long bracket in this relation is nothing other than the 1D propagator that was discussed in Sec. 2.2, i.e.

$$G(x,t;x_0,t_0) = \langle x|\hat{u}(t,t_0)|x_0\rangle. \tag{5.40}$$

Let me hope that the reader sees that this equality corresponds to the physical sense of the propagator.

Now let us break the time segment $[t_0, t]$ into N (for the time being, not necessarily equal) parts by inserting $(N-1)$ intermediate points (Fig. 4) with

$$t_0 < t_1 < \ldots < t_k < \ldots < t_{N-1} < t, \tag{5.41}$$

and use the definition (4.157) of the time evolution operator to write

$$\hat{u}(t,t_0) = \hat{u}(t,t_{N-1})\hat{u}(t_{N-1},t_{N-2})\ldots\hat{u}(t_2,t_1)\hat{u}(t_1,t_0). \tag{5.42}$$

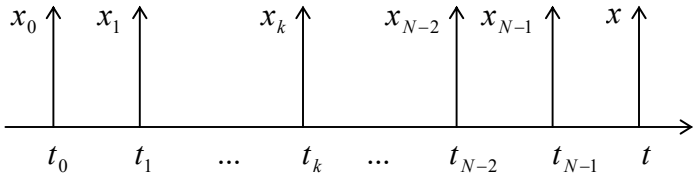


Fig. 5.4. Time partition and coordinate notation at the initial stage of the Feynman path integral's derivation.

[15] This formulation was developed in 1948 by Richard P. Feynman. (According to his memories, this work was motivated by a "mysterious" remark by P. Dirac in his pioneering 1930 textbook on quantum mechanics.)

[16] A more thorough discussion of the path-integral approach may be found in the famous text by R. Feynman and A. Hibbs, *Quantum Mechanics and Path Integrals*, first published in 1965. (For its latest edition by Dover in 2010, the book was emended by D. Styler.) For a more recent monograph, which reviews more applications, see L. Schulman, *Techniques and Applications of Path Integration*, Wiley, 1981.

After plugging Eq. (42) into Eq. (40), let us insert the identity operator, again in the closure form (4.252), but written for x_k rather than x', between each two partial evolution operators including the time argument t_k. The result is

$$G(x,t;x_0,t_0)=\int dx_{N-1}\int dx_{N-2}\ldots\int dx_1\langle x|\hat{u}(t,t_{N-1})|x_{N-1}\rangle\langle x_{N-1}|\hat{u}(t_{N-1},t_{N-2})|x_{N-2}\rangle\ldots\langle x_1|\hat{u}(t_1,t_0)|x_0\rangle. \quad (5.43)$$

The physical sense of each integration variable x_k is the wavefunction's argument at time t_k – see Fig. 4.

The key Feynman's step was the realization that if all intervals are taken similar and sufficiently small, $t_k - t_{k-1} = d\tau \to 0$, all the partial brackets participating in Eq. (43) may be expressed via the free-particle's propagator given by Eq. (2.49), even if the particle is not free, but moves in a stationary potential profile $U(x)$. To show that, let us use either Eq. (4.175) or Eq. (4.181), which, for a small time interval $d\tau$, give the same result:

$$\hat{u}(\tau+d\tau,\tau)=\exp\left\{-\frac{i}{\hbar}\hat{H}d\tau\right\}=\exp\left\{-\frac{i}{\hbar}\left(\frac{\hat{p}^2}{2m}d\tau+U(\hat{x})d\tau\right)\right\}. \quad (5.44)$$

Generally, an exponent of a sum of two operators may be treated as that of c-number arguments, and in particular factored into a product of two exponents, only if the operators commute. (In that case, we can use all the standard algebra for the exponents of c-number arguments.) In our case, this is not so because the operator $\hat{p}^2/2m$ does not commute with $\hat{x}$, and hence with $U(\hat{x})$. However, it may be shown[17] that for an infinitesimal time interval $d\tau$, the non-zero commutator

$$\left[\frac{\hat{p}^2}{2m}d\tau, U(\hat{x})d\tau\right]\neq 0, \quad (5.45)$$

proportional to $(d\tau)^2$, may be ignored in the first, linear approximation in $d\tau$. As a result, we may factor the right-hand side in Eq. (44) by writing

$$\hat{u}(\tau+d\tau,\tau)_{d\tau\to 0}\to\exp\left\{-\frac{i}{\hbar}\frac{\hat{p}^2}{2m}d\tau\right\}\exp\left\{-\frac{i}{\hbar}U(\hat{x})d\tau\right\}. \quad (5.46)$$

(This approximation is very much similar in spirit to the trapezoidal-rule approximation in the usual 1D integration,[18] which is also asymptotically impeachable.)

Since the second exponential function on the right-hand side of Eq. (46) commutes with the coordinate operator, we may move it out of each partial bracket participating in Eq. (43), with $U(x)$ turning into a c-number function:

$$\langle x_{\tau+d\tau}|\hat{u}(\tau+d\tau,\tau)|x_\tau\rangle=\langle x_{\tau+d\tau}|\exp\left\{-\frac{i}{\hbar}\frac{\hat{p}^2}{2m}d\tau\right\}|x_\tau\rangle\exp\left\{-\frac{i}{\hbar}U(x)d\tau\right\}. \quad (5.47)$$

But the remaining bracket is just the propagator of a free particle, so for it, we may use Eq. (2.49):

[17] This is exactly the corner I am going to cut because a strict mathematical proof of this (intuitively evident) statement would take more time/space than I can afford.

[18] See, e.g., MA Eq. (5.2).

$$\left\langle x_{\tau+d\tau}\left|\exp\left\{-\frac{i}{\hbar}\frac{\hat{p}^2}{2m}d\tau\right\}\right|x_\tau\right\rangle=\left(\frac{m}{2\pi i\hbar d\tau}\right)^{1/2}\exp\left\{i\frac{m(dx)^2}{2\hbar d\tau}\right\}. \tag{5.48}$$

As the result, the full propagator (43) takes the form

$$G(x,t;x_0,t_0)=\lim_{\substack{d\tau\to 0\\ N\to\infty}}\int dx_{N-1}\int dx_{N-2}..\int dx_1\left(\frac{m}{2\pi i\hbar d\tau}\right)^{N/2}\exp\left\{\sum_{k=1}^{N}\left[i\frac{m(dx)^2}{2\hbar d\tau}-i\frac{U(x)}{\hbar}d\tau\right]_{\tau=t_k}\right\}. \tag{5.49}$$

At $N\to\infty$ and hence $d\tau\equiv(t-t_0)/N\to 0$, the sum under the exponent in this expression may be approximated with the corresponding integral:

$$\sum_{k=1}^{N}\frac{i}{\hbar}\left[\frac{m}{2}\left(\frac{dx}{d\tau}\right)^2-U(x)\right]_{\tau=t_k}d\tau\to\frac{i}{\hbar}\int_{t_0}^{t}\left[\frac{m}{2}\left(\frac{dx}{d\tau}\right)^2-U(x)\right]d\tau, \tag{5.50}$$

and the expression in the square brackets is just the particle's Lagrangian function $\mathscr{L}$.[19] The integral of this function over time is the classical action $\mathcal{S}$ calculated along a particular "path" $x(\tau)$.[20] As a result, defining the (1D) *path integral* as

$$\int(\cdots)D[x(\tau)]\equiv\lim_{\substack{d\tau\to 0\\ N\to\infty}}\left(\frac{m}{2\pi i\hbar d\tau}\right)^{N/2}\int dx_{N-1}\int dx_{N-2}\ldots\int dx_1(\cdots), \tag{5.51a}$$

1D path integral: definition

we can bring our result to the following (superficially simple) form:

$$G(x,t;x_0,t_0)=\int\exp\left\{\frac{i}{\hbar}\mathcal{S}[x(\tau)]\right\}D[x(\tau)]. \tag{5.51b}$$

1D propagator via path integral

The name "path integral" for the mathematical construct (51a) may be readily explained if we keep the number N of time intervals large but still finite, and also approximate each of the enclosed integrals with a sum over $M >> 1$ discrete points along the coordinate axis – see Fig. 5a.

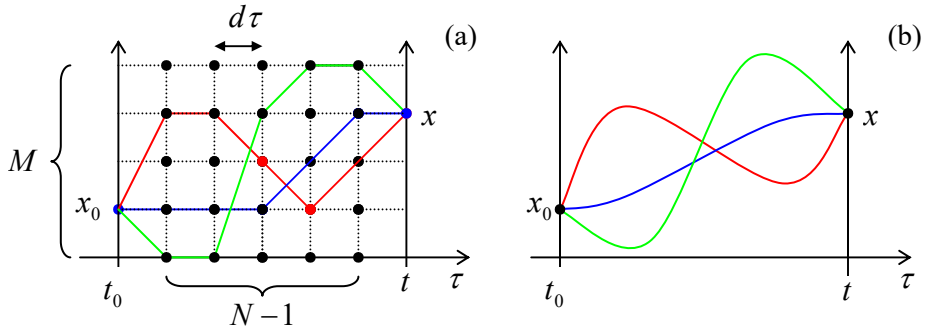


Fig. 5.5. Several 1D classical paths: (a) in the discrete approximation and (b) in the continuous limit.

Then the path integral (51a) is the product of $(N-1)$ sums corresponding to different values of time τ, each of them with M terms, each of those representing the function under the integral at a particular spatial point. Multiplying those $(N-1)$ sums, we get a sum of $(N-1)M$ terms, each

[19] See, e.g., CM Sec. 2.1.
[20] See, e.g., CM Sec. 10.3.

evaluating the function at a specific spatial-temporal point [x, τ]. These terms may be now grouped to represent all possible different continuous classical paths $x[\tau]$ from the initial point [x_0, t_0] to the finite point [x, t]. It is evident that the last interpretation remains true even in the continuous limit $N, M \to \infty$ (see Fig. 5b).

Why does such a path representation of the sum make sense? This is because in the classical limit, the particle follows just a certain path, corresponding to the minimum of the action $\mathcal{S}_{\rm cl}$. As a result, for all close trajectories, the difference $(\mathcal{S} - \mathcal{S}_{\rm cl})$ is proportional to the square of the deviation from the classical trajectory. Hence, for a quasiclassical motion, with $\mathcal{S}_{\rm cl} >> \hbar$, there is a bunch of close trajectories, with $(\mathcal{S} - \mathcal{S}_{\rm cl}) << \hbar$, that give substantial contributions to the path integral. On the other hand, strongly non-classical trajectories, with $(\mathcal{S} - \mathcal{S}_{\rm cl}) >> \hbar$, give phases $\mathcal{S}/\hbar$ rapidly oscillating from one trajectory to the next one, and their contributions to the path integral are averaged out.[21] As a result, for a quasi-classical motion, the propagator's exponent may be evaluated on the classical path only:

$$G_{\rm cl} \propto \exp\left\{\frac{i}{\hbar}\mathcal{S}_{\rm cl}\right\} = \exp\left\{\frac{i}{\hbar}\int_{t_0}^{t}\left[\frac{m}{2}\left(\frac{dx}{d\tau}\right)^2 - U(x)\right]d\tau\right\}. \tag{5.52}$$

The sum of the kinetic and potential energies is the full energy E of the particle, which remains constant for motion in a stationary potential $U(x)$, so we may rewrite the expression under the last integral as[22]

$$\left[\frac{m}{2}\left(\frac{dx}{d\tau}\right)^2 - U(x)\right]d\tau = \left[m\left(\frac{dx}{d\tau}\right)^2 - E\right]d\tau = m\frac{dx}{d\tau}dx - Ed\tau. \tag{5.53}$$

With this replacement, Eq. (52) yields

$$G_{\rm cl} \propto \exp\left\{\frac{i}{\hbar}\int_{x_0}^{x} m\frac{dx}{d\tau}dx\right\}\exp\left\{-\frac{i}{\hbar}E(t-t_0)\right\} = \exp\left\{\frac{i}{\hbar}\int_{x_0}^{x} p(x)dx\right\}\exp\left\{-\frac{i}{\hbar}E(t-t_0)\right\}, \tag{5.54}$$

where p is the classical momentum of the particle. But (at least, leaving the pre-exponential factor alone) this is the WKB approximation result that was derived and studied in detail in Chapter 2!

One may question the value of such a complicated calculation, which yields results that could be readily obtained from Schrödinger's wave mechanics. Feynman's approach is indeed not used too often, but it has its merits. First, it has an important philosophical (and hence heuristic) value. Indeed, Eq. (51) may be interpreted by saying that the essence of quantum mechanics is the exploration, by the system, of all possible paths $x(\tau)$, each of them classical-like, in the sense that the particle's coordinate x and velocity $dx/d\tau$ are exactly defined simultaneously at each point. The resulting contributions to the path integral are added up coherently to form the actual propagator G, and via it, the final probability $W \propto |G|^2$ of the particle's propagation from [x_0, t_0] to [x, t]. As the scale of the action $\mathcal{S}$ of the motion

[21] This fact may be proved by expanding the difference $(\mathcal{S} - \mathcal{S}_{\rm cl})$ in the Taylor series in the path variation (leaving only the leading quadratic terms) and working out the resulting Gaussian integrals. This integration, together with the pre-exponential coefficient in Eq. (51a), gives exactly the pre-exponential factor that we have already found refining the WKB approximation in Sec. 2.4.

[22] The same trick is often used in analytical classical mechanics – say, for proving the Hamilton principle, and for the derivation of the Hamilton-Jacobi equations – see, e.g., CM Secs. 10.3-4.

decreases and becomes comparable to $\hbar$, more and more paths produce substantial contributions to this sum, and hence to W, providing a larger and larger difference between the quantum and classical properties of the system.

Second, the path integral provides a justification for some simple explanations of quantum phenomena. A typical example is the quantum interference effects discussed in Sec. 3.1 – see, e.g., Fig. 3.1 and the corresponding text. In that discussion, we used the Huygens principle to argue that at the two-slit interference, the WKB approximation might be restricted to contributions from two paths that pass through different slits, but otherwise consist of straight-line segments. To have another look at that assumption, let us generalize the path integral to multi-dimensional geometries. Fortunately, the simple structure of Eq. (51b) makes such generalization virtually evident:

$$G(\mathbf{r},t;\mathbf{r}_0,t_0)=\int \exp\left\{\frac{i}{\hbar}\mathcal{S}[\mathbf{r}(\tau)]\right\}\mathcal{D}[\mathbf{r}(\tau)],\qquad \mathcal{S}\equiv\int_{t_0}^{t}\mathscr{L}\left(\mathbf{r},\frac{d\mathbf{r}}{d\tau}\right)d\tau=\int_{t_0}^{t}\left[\frac{m}{2}\left(\frac{d\mathbf{r}}{d\tau}\right)^2-U(\mathbf{r})\right]d\tau. \quad (5.55)$$

3D propagator as a path integral

where the definition (51a) of the path integral should be also modified correspondingly. (I will not go into these technical details.) For the Young-type experiment (Fig. 3.1), where a classical particle could reach the detector only after passing through one of the slits, the classical paths dominating the contribution from each slit are the straight-line segments shown in Fig. 3.1, and if they are much longer than the de Broglie wavelength, the propagator may be well approximated by the sum of two integrals of $\mathscr{L}d\tau = i\mathbf{p}(\mathbf{r})\cdot d\mathbf{r}/\hbar$ – just as this was done in Sec. 3.1.

Last but not least, the path integral allows simple solutions to some problems that would be hard to obtain by other methods. As the simplest example, let us consider the problem of tunneling in multi-dimensional space, sketched in Fig. 6 for the 2D case – just for the graphics' simplicity. Here, the potential profile $U(x, y)$ has a saddle-like shape. (Another helpful image is a mountain path between two summits, in Fig. 6 located on the top and at the bottom of the shown region.) A particle of energy E may move classically in the left and right regions with $U(x, y) < E$, but if E is not sufficiently high, it can pass from one of these regions to another one only via the quantum-mechanical tunneling under the pass. Let us calculate the transparency of this potential barrier in the WKB approximation, ignoring the possible pre-exponential factor. [23]

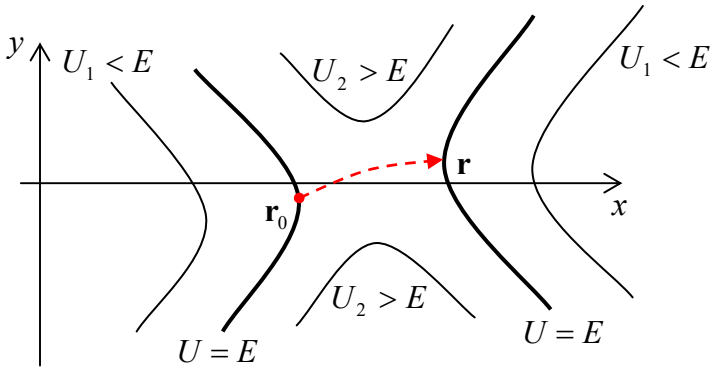


Fig. 5.6. A saddle-type 2D potential profile and the instanton trajectory of a particle of energy E (schematically).

[23] Actually, one can argue that the pre-exponential factor should be close to 1, just like in Eq. (2.117), especially if the potential is smooth, in the sense of Eq. (2.107), in all spatial directions. (Let me remind the reader that for most practical applications of quantum tunneling, the pre-exponential factor is of relatively minor importance.)

According to the evident multi-dimensional generalization Eq. (54), for the classically forbidden region, where $E < U(x, y)$, and hence $\mathbf{p}(\mathbf{r})/\hbar = i\boldsymbol{\kappa}(\mathbf{r})$, the contributions to the propagator (55) are proportional to

$$e^{-I} \exp\left\{-\frac{i}{\hbar}E(t-t_0)\right\}, \qquad \text{where } I \equiv \int_{\mathbf{r}_0}^{\mathbf{r}} \boldsymbol{\kappa}(\mathbf{r}) \cdot d\mathbf{r}, \tag{5.56}$$

where $\kappa \equiv |\boldsymbol{\kappa}|$ may be calculated just in the 1D case – cf. Eq. (2.97):

$$\frac{\hbar^2 \kappa^2(\mathbf{r})}{2m} = U(\mathbf{r}) - E. \tag{5.57}$$

Hence the path integral in this region is much simpler than in the classically allowed region because the spatial exponents are purely real and there is no complex interference between them. Due to the minus sign before I in the exponent (56), the *largest* contribution to G evidently comes from the trajectory (or a narrow bundle of close trajectories) for which the integral I has the *smallest* value, so the barrier transparency may be calculated as

3D tunneling in WKB limit

$$\mathcal{T} \approx \left|G\right|^2 \approx e^{-2I} \equiv \exp\left\{-2\int_{\mathbf{r}_0}^{\mathbf{r}} \boldsymbol{\kappa}(\mathbf{r}') \cdot d\mathbf{r}'\right\}, \tag{5.58}$$

where $\mathbf{r}$ and $\mathbf{r}_0$ are certain points on the opposite classical turning-point surfaces: $U(\mathbf{r}) = U(\mathbf{r}_0) = E$ – see Fig. 6.

Thus the barrier transparency problem is reduced to finding the trajectory (including the points $\mathbf{r}$ and $\mathbf{r}_0$) that connects these two surfaces and minimizes the functional I. This is of course a well-known problem of the calculus of variations,[24] but it is interesting that the path integral provides a simple alternative way of solving it. Let us consider an auxiliary problem of particle's motion in the potential profile $U_{\text{inv}}(\mathbf{r})$ that is *inverted* relative to the particle's energy E, i.e. is defined by the following equality:

$$U_{\text{inv}}(\mathbf{r}) - E \equiv E - U(\mathbf{r}). \tag{5.59}$$

As was discussed above, at fixed energy E, the path integral for the WKB motion in the classically allowed region of potential $U_{\text{inv}}(x, y)$ (that coincides with the classically forbidden region of the original problem) is dominated by the classical trajectory corresponding to the minimum of

$$\mathcal{S}_{\text{inv}} = \int_{\mathbf{r}_0}^{\mathbf{r}} \mathbf{p}_{\text{inv}}(\mathbf{r}') \cdot d\mathbf{r}' = \hbar \int_{\mathbf{r}_0}^{\mathbf{r}} \mathbf{k}_{\text{inv}}(\mathbf{r}') \cdot d\mathbf{r}, \tag{5.60}$$

where $\mathbf{k}_{\text{inv}}$ should be determined from the WKB relation

$$\frac{\hbar^2 k_{\text{inv}}^2(\mathbf{r})}{2m} \equiv E - U_{\text{inv}}(\mathbf{r}). \tag{5.61}$$

But comparing Eqs. (57), (59), and (61), we see that $\mathbf{k}_{\text{inv}} = \boldsymbol{\kappa}$ at each point! This means that in the WKB limit, the tunneling path corresponds to the classical (so-called *instanton*[25]) trajectory of the same

[24] For a concise introduction to the field see, e.g., I. Gelfand and S. Fomin, *Calculus of Variations*, Dover, 2000, or L. Elsgolc, *Calculus of Variations*, Dover, 2007.

[25] In the quantum field theory, the instanton concept may be formulated somewhat differently and has more complex applications – see, e.g. R. Rajaraman, *Solitons and Instantons*, North-Holland, 1987.

particle moving in the inverted potential $U_{inv}(\mathbf{r})$. If the initial point $\mathbf{r}_0$ is fixed, this trajectory may be readily found by means of classical mechanics. (Note that the initial kinetic energy, and hence the initial velocity of the instanton launched from point $\mathbf{r}_0$ should be zero because by the classical turning point definition, $U_{inv}(\mathbf{r}_0) = U(\mathbf{r}_0) = E$.) Thus the problem is further reduced to a simpler task of maximizing the transparency (58) by choosing the optimal position of $\mathbf{r}_0$ on the equipotential surface $U(\mathbf{r}_0) = E$ – see Fig. 6. Moreover, for many symmetric potentials, the position of this point may be readily guessed even without calculations – as it is in Problems 6 and 7, left for the reader's exercise.

Note that besides the calculation of the potential barrier's transparency, the instanton trajectory has one more important implication: the so-called *traversal time* τ_t of the classical motion along it, from the point $\mathbf{r}_0$ to the point $\mathbf{r}$, in the inverted potential defined by Eq. (59), plays the role of the most important (though not the only one) time scale of the particle's tunneling under the barrier.[26]

5.4. Revisiting harmonic oscillator

Now let us return to the 1D harmonic oscillator, which may be understood as any system, regardless of its physical nature, described by the Hamiltonian (4.237) with the potential energy (2.111):

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{m\omega_0^2 \hat{x}^2}{2}. \tag{5.62}$$

Harmonic oscillator: Hamiltonian

In Sec. 2.9 we have used a "brute-force" (wave-mechanics) approach to analyze the eigenfunctions $\psi_n(x)$ and eigenvalues E_n of this Hamiltonian, and found that, unfortunately, this approach required relatively complex mathematics, which does not enable an easy calculation of the key characteristics of the system. Fortunately, the bra-ket formalism helps to make such calculations.

First, by introducing the normalized (dimensionless) operators of coordinates and momentum:[27]

$$\hat{\xi} \equiv \frac{\hat{x}}{x_0}, \qquad \hat{\zeta} \equiv \frac{\hat{p}}{m\omega_0 x_0}, \tag{5.63}$$

where $x_0 \equiv (\hbar/m\omega_0)^{1/2}$ is the natural coordinate scale discussed in detail in Sec. 2.9, we can represent the Hamiltonian (62) in a very simple and $x \leftrightarrow p$ symmetric form:

$$\hat{H} = \frac{\hbar\omega_0}{2}\left(\hat{\xi}^2 + \hat{\zeta}^2\right). \tag{5.64}$$

This symmetry, as well as our discussion of the very similar coordinate and momentum representations in Sec. 4.7, hints that much may be gained by treating the operators $\hat{\xi}$ and $\hat{\zeta}$ on equal footing. Inspired by this clue, let us introduce a new operator

$$\hat{a} \equiv \frac{\hat{\xi} + i\hat{\zeta}}{\sqrt{2}} \equiv \left(\frac{m\omega_0}{2\hbar}\right)^{1/2}\left(\hat{x} + i\frac{\hat{p}}{m\omega_0}\right). \tag{5.65a}$$

Annihilation operator: definition

[26] For more on this interesting issue see, e.g., M. Buttiker and R. Landauer, *Phys. Rev. Lett.* **49**, 1739 (1982), and references therein.

[27] This normalization is not really necessary, it just makes the following calculations less bulky – and thus more aesthetically appealing.

Since both operators $\hat{\xi}$ and $\hat{\zeta}$ correspond to real observables, i.e. have real eigenvalues and hence are Hermitian (self-adjoint), the Hermitian conjugate of the operator $\hat{a}$ is simply its complex conjugate:

Creation operator: definition

$$\hat{a}^\dagger \equiv \frac{\hat{\xi} - i\hat{\zeta}}{\sqrt{2}} \equiv \left(\frac{m\omega_0}{2\hbar}\right)^{1/2}\left(\hat{x} - i\frac{\hat{p}}{m\omega_0}\right). \tag{5.65b}$$

Because of the reason that will be clear very soon, $\hat{a}^\dagger$ and $\hat{a}$ (in this order!) are called the *creation and annihilation operators*.

Now solving the simple system of two linear equations (65) for $\hat{\xi}$ and $\hat{\zeta}$, we get the following reciprocal relations:

$$\hat{\xi} = \frac{\hat{a} + \hat{a}^\dagger}{\sqrt{2}}, \quad \hat{\zeta} = \frac{\hat{a} - \hat{a}^\dagger}{\sqrt{2}\,i}, \qquad \text{i.e. } \hat{x} = \left(\frac{\hbar}{m\omega_0}\right)^{1/2}\frac{\hat{a} + \hat{a}^\dagger}{\sqrt{2}}, \quad \hat{p} = (\hbar m\omega_0)^{1/2}\frac{\hat{a} - \hat{a}^\dagger}{\sqrt{2}\,i}. \tag{5.66}$$

Our Hamiltonian (64) includes only squares of these operators. Calculating them, we have to be careful to avoid swapping the new operators, because they do not commute. Indeed, for the normalized operators (63), Eq. (2.14) gives

$$\left[\hat{\xi}, \hat{\zeta}\right] \equiv \frac{1}{x_0^2 m\omega_0}\left[\hat{x}, \hat{p}\right] = i\hat{I}, \tag{5.67}$$

so Eqs. (65) yield

Creation-annihilation operators: commutation relation

$$\left[\hat{a}, \hat{a}^\dagger\right] = \frac{1}{2}\left[\left(\hat{\xi} + i\hat{\zeta}\right), \left(\hat{\xi} - i\hat{\zeta}\right)\right] = -\frac{i}{2}\left(\left[\hat{\xi}, \hat{\zeta}\right] - \left[\hat{\zeta}, \hat{\xi}\right]\right) = \hat{I}\,. \tag{5.68}$$

With such due caution, Eq. (66) gives

$$\hat{\xi}^2 = \frac{1}{2}\left(\hat{a}^2 + \hat{a}^{\dagger 2} + \hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a}\right), \qquad \hat{\zeta}^2 = -\frac{1}{2}\left(\hat{a}^2 + \hat{a}^{\dagger 2} - \hat{a}\hat{a}^\dagger - \hat{a}^\dagger\hat{a}\right). \tag{5.69}$$

Plugging these expressions back into Eq. (64), we get

$$\hat{H} = \frac{\hbar\omega_0}{2}\left(\hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a}\right). \tag{5.70}$$

This expression is elegant enough but may be recast into an even more convenient form. For that, let us rewrite the commutation relation (68) as

$$\hat{a}\hat{a}^\dagger = \hat{a}^\dagger\hat{a} + \hat{I}\,, \tag{5.71}$$

and plug it into Eq. (70). The result is

$$\hat{H} = \frac{\hbar\omega_0}{2}\left(2\hat{a}^\dagger\hat{a} + \hat{I}\right) \equiv \hbar\omega_0\left(\hat{N} + \frac{1}{2}\hat{I}\right), \tag{5.72}$$

where, in the last form, one more (evidently, Hermitian) operator

Number operator: definition

$$\hat{N} \equiv \hat{a}^\dagger\hat{a} \tag{5.73}$$

has been introduced. Since, according to Eq. (72), the operators $\hat{H}$ and $\hat{N}$ differ only by the addition of the identity operator and multiplication by a *c*-number, these operators commute. Hence, according to

the general arguments of Sec. 4.5, they share a set of stationary eigenstates n (they are frequently called the *Fock states*), and we can write the standard eigenproblem (4.68) for the new operator as

$$\hat{N}|n\rangle = N_n|n\rangle, \tag{5.74}$$

where N_n are some eigenvalues that, according to Eq. (72), determine also the energy spectrum of the oscillator:

$$E_n = \hbar\omega_0\left(N_n + \frac{1}{2}\right). \tag{5.75}$$

So far, we know only that all eigenvalues N_n are real; to calculate them, let us carry out the following calculation – splendid in its simplicity and efficiency. Consider the result of the action of the operator $\hat{N}$ on the ket-vector $\hat{a}^\dagger|n\rangle$. Using the definition (73) and then the associative rule of the bra-ket formalism, we may write

$$\hat{N}\left(\hat{a}^\dagger|n\rangle\right) \equiv \left(\hat{a}^\dagger\hat{a}\right)\left(\hat{a}^\dagger|n\rangle\right) = \hat{a}^\dagger\left(\hat{a}\hat{a}^\dagger\right)|n\rangle. \tag{5.76}$$

Now using the commutation relation (71), and then Eq. (74), we may continue as

$$\hat{a}^\dagger\left(\hat{a}\hat{a}^\dagger\right)|n\rangle = \hat{a}^\dagger\left(\hat{a}^\dagger\hat{a} + \hat{I}\right)|n\rangle = \hat{a}^\dagger\left(\hat{N} + \hat{I}\right)|n\rangle = \hat{a}^\dagger\left(N_n + 1\right)|n\rangle = \left(N_n + 1\right)\left(\hat{a}^\dagger|n\rangle\right). \tag{5.77}$$

For clarity, let us summarize the result of this calculation:

$$\hat{N}\left(\hat{a}^\dagger|n\rangle\right) = \left(N_n + 1\right)\left(\hat{a}^\dagger|n\rangle\right). \tag{5.78}$$

Performing a similar calculation for the operator $\hat{a}$, we get a similar formula, but with a different sign:

$$\hat{N}\left(\hat{a}|n\rangle\right) = \left(N_n - 1\right)\left(\hat{a}|n\rangle\right). \tag{5.79}$$

It is time to stop the calculations for a minute, and translate their results into plain English: if $|n\rangle$ is the eigenket of the operator $\hat{N}$ with an eigenvalue N_n, then $\hat{a}^\dagger|n\rangle$ and $\hat{a}|n\rangle$ are also eigenkets of that operator, with the eigenvalues $(N_n + 1)$ and $(N_n - 1)$, respectively. This statement may be vividly represented by the so-called *ladder diagram* shown in Fig. 7.

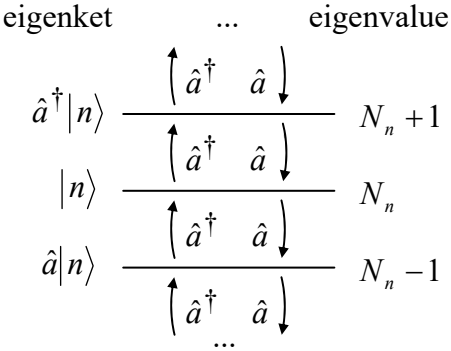


Fig. 5.7. The "ladder diagram" of eigenstates of a 1D harmonic oscillator. Arrows show the actions of the creation and annihilation operators on the eigenstates.

The operator $\hat{a}^\dagger$ moves the system one step up this ladder, while the operator $\hat{a}$ brings it one step down. In other words, the former operator creates a new *excitation* of the system,[28] while the latter operator kills ("annihilates") such excitation.[29] On the other hand, according to Eq. (74) inner-multiplied by the bra-vector $\langle n|$, the operator $\hat{N}$ does not change the state of the system, but "counts" its position on the ladder:

$$\langle n|\hat{N}|n\rangle = \langle n|N_n|n\rangle = N_n. \tag{5.80}$$

This is why $\hat{N}$ is called the *number operator*, in our current context meaning the number of the elementary excitations of the oscillator.

This calculation still needs completion. Indeed, we still do not know whether the ladder shown in Fig. 7 shows *all* eigenstates of the oscillator, and what exactly the numbers N_n are. Fascinating enough, both questions may be answered by exploring just one paradox. Let us start with some state n (read: a step of the ladder), and keep going down the ladder, applying the operator $\hat{a}$ again and again. According to Eq. (79), at each step, the eigenvalue N_n is decreased by one, so eventually, it should become negative. However, this cannot happen because any actual eigenstate, including the states represented by kets $|d\rangle \equiv \hat{a}\,|n\rangle$ and $|n\rangle$, should have a positive norm – see Eq. (4.16). Comparing the norms,

$$\|n\|^2 = \langle n|n\rangle, \qquad \|d\|^2 = \langle n|\hat{a}^\dagger a|n\rangle = \langle n|\hat{N}|n\rangle = N_n\langle n|n\rangle, \tag{5.81}$$

we see that both of them cannot be positive simultaneously if N_n is negative.

To resolve this paradox let us notice that the action of the creation and annihilation operators on the stationary states n may consist of not only their promotion to an adjacent step of the ladder diagram but also by their multiplication by some c-numbers:

$$\hat{a}|n\rangle = A_n|n-1\rangle, \qquad \hat{a}^\dagger|n\rangle = A'_n|n+1\rangle. \tag{5.82}$$

(The linear relations (78)-(79) clearly allow that.) Let us calculate the coefficients A_n assuming, for convenience, that all eigenstates, including the states n and $(n-1)$, are normalized:

$$\langle n|n\rangle = 1, \qquad \langle n-1|n-1\rangle = \langle n|\frac{\hat{a}^\dagger}{A_n^*}\frac{\hat{a}}{A_n}|n\rangle = \frac{1}{A_n^* A_n}\langle n|\hat{N}|n\rangle = \frac{N_n}{A_n^* A_n}\langle n|n\rangle = 1. \tag{5.83}$$

From here, we get $|A_n| = (N_n)^{1/2}$, i.e.

$$\hat{a}|n\rangle = N_n^{1/2} e^{i\varphi_n}|n-1\rangle, \tag{5.84}$$

where φ_n is an arbitrary real phase. Now let us consider what happens if all numbers N_n are integers. (Because of the definition of N_n, given by Eq. (74), it is convenient to call these integers n, i.e. to use the same letter as for the corresponding eigenstate.) Then when we have come down to the state with $n = 0$, an attempt to make one more step down gives

$$\hat{a}|0\rangle = 0|-1\rangle. \tag{5.85}$$

[28] For electromagnetic field oscillators, such excitations are called *photons*; for mechanical wave oscillators, *phonons*, etc.

[29] This is exactly why $\hat{a}^\dagger$ is called the *creation* operator, and $\hat{a}$, the *annihilation* operator.

But according to Eq. (4.9), the state on the right-hand side of this equation is the "null state", i.e. does not exist.[30] This gives the (only known :-) resolution of the state ladder paradox: the ladder has the lowest step with $N_n = n = 0$.

As a by-product of our discussion, we have obtained a very important relation $N_n = n$, which means, in particular, that the state ladder shown in Fig. 7 includes *all* eigenstates of the oscillator. Plugging this relation into Eq. (75), we see that the *full* spectrum of eigenenergies of the harmonic oscillator is described by the simple formula

$$E_n = \hbar\omega_0\left(n + \frac{1}{2}\right), \quad n = 0, 1, 2\ldots, \tag{5.86}$$

which was already discussed in Sec. 2.9. It is rather remarkable that the bra-ket formalism has allowed us to derive it without calculating the corresponding (rather cumbersome) wavefunctions $\psi_n(x)$ – see Eqs. (2.284).

Moreover, this formalism may be also used to calculate virtually any matrix element of the oscillator, without using $\psi_n(x)$. However, to do that, we should first calculate the coefficient A'_n participating in the second of Eqs. (82). This may be done similarly to the above calculation of A_n; alternatively, since we already know that $|A_n| = (N_n)^{1/2} = n^{1/2}$, we may notice that according to Eqs. (73) and (82), the eigenproblem (74), which in our new notation for N_n becomes

$$\hat{N}|n\rangle = n|n\rangle, \tag{5.87}$$

may be rewritten as

$$\hat{N}|n\rangle \equiv \hat{a}^\dagger\hat{a}|n\rangle = \hat{a}^\dagger A_n|n-1\rangle = A_n A'_{n-1}|n\rangle. \tag{5.88}$$

Comparing the right-hand sides of Eqs. (87) and (88), we see that $|A'_{n-1}| = n/|A_n| = n^{1/2}$, i.e. $A'_n = (n+1)^{1/2}\exp(i\varphi_n')$. Taking all phases φ_n and φ_n' equal to zero for simplicity, we may spell out Eqs. (82) as[31]

$$\hat{a}^\dagger|n\rangle = (n+1)^{1/2}|n+1\rangle, \qquad \hat{a}|n\rangle = n^{1/2}|n-1\rangle. \tag{5.89}$$

Fock state ladder

Now we can use these formulas to calculate, for example, the matrix elements of the operator $\hat{x}$ in the Fock state basis:

$$\begin{aligned}\langle n'|\hat{x}|n\rangle \equiv x_0\langle n'|\hat{\xi}|n\rangle &= \frac{x_0}{\sqrt{2}}\langle n'|\left(\hat{a}+\hat{a}^\dagger\right)|n\rangle \equiv \frac{x_0}{\sqrt{2}}\left(\langle n'|\hat{a}|n\rangle + \langle n'|\hat{a}^\dagger|n\rangle\right)\\ &= \frac{x_0}{\sqrt{2}}\left[n^{1/2}\langle n'|n-1\rangle + (n+1)^{1/2}\langle n'|n+1\rangle\right].\end{aligned} \tag{5.90}$$

Taking into account the Fock state orthonormality:

$$\langle n'|n\rangle = \delta_{n'n}, \tag{5.91}$$

this result becomes

[30] Please note again the radical difference between the null state on the right-hand side of Eq. (85) and the state described by the ket-vector $|0\rangle$ on the left-hand side of that relation. The latter state *does* exist and, moreover, represents the most important, ground state of the system, with $n = 0$ – see Eqs. (2.274)-(2.275).

[31] A useful mnemonic rule for these key relations is that the *c*-number coefficient in any of them is equal to the square root of the *largest* number of the two states it relates.

Coordinate's matrix elements

$$\langle n'|\hat{x}|n\rangle = \frac{x_0}{\sqrt{2}}\left[n^{1/2}\delta_{n',n-1} + (n+1)^{1/2}\delta_{n',n+1}\right] \equiv \left(\frac{\hbar}{2m\omega_0}\right)^{1/2}\left[n^{1/2}\delta_{n',n-1} + (n+1)^{1/2}\delta_{n',n+1}\right]. \quad (5.92)$$

Acting absolutely similarly, for the momentum's matrix elements we get a similar expression:

$$\langle n'|\hat{p}|n\rangle = i\left(\frac{\hbar m\omega_0}{2}\right)^{1/2}\left[-n^{1/2}\delta_{n',n-1} + (n+1)^{1/2}\delta_{n',n+1}\right] \quad (5.93)$$

Hence the matrices of both operators in the Fock-state basis have only two diagonals adjacent to the main diagonal; all other elements (including the main-diagonal ones) are zeros.

The matrix elements of higher powers of these operators, as well as their products, may be handled similarly, though the higher the power, the bulkier the result. For example,

$$\begin{aligned}\langle n'|\hat{x}^2|n\rangle &= \langle n'|\hat{x}\hat{x}|n\rangle = \sum_{n''=0}^{\infty}\langle n'|\hat{x}|n''\rangle\langle n''|\hat{x}|n\rangle \\ &= \frac{x_0^2}{2}\sum_{n''=0}^{\infty}\left[(n'')^{1/2}\delta_{n',n''-1} + (n''+1)^{1/2}\delta_{n',n''+1}\right]\left[n^{1/2}\delta_{n'',n-1} + (n+1)^{1/2}\delta_{n'',n+1}\right] \\ &= \frac{x_0^2}{2}\left\{\left[n(n-1)\right]^{1/2}\delta_{n',n-2} + \left[(n+1)(n+2)\right]^{1/2}\delta_{n',n+2} + (2n+1)\delta_{n',n}\right\}.\end{aligned} \quad (5.94)$$

For applications, the most important of these matrix elements are those on its main diagonal:

$$\langle x^2\rangle \equiv \langle n|\hat{x}^2|n\rangle = \frac{x_0^2}{2}(2n+1). \quad (5.95)$$

This expression shows, in particular, that the expectation value of the oscillator's potential energy in the n^{th} Fock state is

$$\langle U\rangle \equiv \frac{m\omega_0^2}{2}\langle x^2\rangle = \frac{m\omega_0^2 x_0^2}{2}\left(n+\frac{1}{2}\right) \equiv \frac{\hbar\omega_0}{2}\left(n+\frac{1}{2}\right). \quad (5.96)$$

This is exactly one-half of the total energy (86) of the oscillator. As a sanity check, an absolutely similar calculation for the momentum squared, and hence for the kinetic energy $p^2/2m$, yields

$$\langle p^2\rangle = \langle n|\hat{p}^2|n\rangle = (m\omega_0 x_0)^2\left(n+\frac{1}{2}\right) \equiv \hbar m\omega_0\left(n+\frac{1}{2}\right), \quad \text{so} \quad \left\langle\frac{p^2}{2m}\right\rangle = \frac{\hbar\omega_0}{2}\left(n+\frac{1}{2}\right), \quad (5.97)$$

i.e. both partial energies are equal to $E_n/2$, just as in a classical oscillator.[32]

Note that according to Eqs. (92) and (93), the expectation values of both x and p in any Fock state are equal to zero:

$$\langle x\rangle \equiv \langle n|\hat{x}|n\rangle = 0, \qquad \langle p\rangle \equiv \langle n|\hat{p}|n\rangle = 0, \quad (5.98)$$

[32] Still note that operators of the partial (potential and kinetic) energies do *not* commute with either each other or with the full-energy (Hamiltonian) operator, so the Fock states n are *not* their eigenstates. This fact maps onto the well-known oscillations of these partial energies (with the frequency $2\omega_0$) in a classical oscillator, at the full energy staying constant.

This is why, according to the general Eqs. (1.33)-(1.34), the results (95) and (97) also give the variances of the coordinate and the momentum, i.e. the squares of their uncertainties, $(\delta x)^2$ and $(\delta p)^2$. In particular, for the ground state ($n = 0$), these uncertainties are

$$\delta x = \frac{x_0}{\sqrt{2}} \equiv \left(\frac{\hbar}{2m\omega_0} \right)^{1/2}, \qquad \delta p = \frac{m\omega_0 x_0}{\sqrt{2}} \equiv \left(\frac{\hbar m \omega_0}{2} \right)^{1/2}. \tag{5.99}$$

In the theory of precise measurements (to be reviewed in brief in Chapter 10), these expressions are often called the *standard quantum limit*.

5.5. Glauber states and squeezed states

There is a huge difference between a quantum stationary (Fock) state of the oscillator and its classical state. Indeed, let us write the well-known classical equations of motion of the oscillator (using capital letters to distinguish classical variables from the arguments of quantum wavefunctions): [33]

$$\dot{X} = \frac{P}{m}, \qquad \dot{P} = -\frac{\partial U}{\partial x} = -m\omega_0^2 X. \tag{5.100}$$

The simplest method to solve these equations is to introduce the dimensionless complex variable

$$\alpha(t) \equiv \frac{1}{\sqrt{2}x_0} \left[X(t) + i \frac{P(t)}{m\omega_0} \right], \tag{5.101}$$

With this definition, Eqs. (100) are conveniently merged into one equation,

$$\dot{\alpha} = -i\omega_0 \alpha, \tag{5.102}$$

with an evident, very simple solution

$$\alpha(t) = \alpha(0) \exp\{-i\omega_0 t\}, \tag{5.103}$$

so per Eq. (102):

$$X(t) = \sqrt{2}x_0 \, \mathrm{Re}\left[\alpha(0) \exp\{-i\omega_0 t\}\right], \qquad P(t) = \sqrt{2} m\omega_0 x_0 \, \mathrm{Im}\left[\alpha(0) \exp\{-i\omega_0 t\}\right],$$

where the constant $\alpha(0)$ is just the (normalized) classical complex amplitude of the oscillations, so their real amplitude is $A = \sqrt{2}x_0|\alpha(t)| = \sqrt{2}x_0|\alpha(0)|$.[34] By the appropriate choice of the time origin, the complex amplitude may be always made real; then $X \propto \cos\omega_0 t$ and $P \propto -\sin\omega_0 t$.

On the so-called *phase plane*, with the Cartesian coordinates x and p, this solution describes a clockwise rotation of the representation point $\{X(t), P(t)\}$ along an elliptic trajectory starting from the initial point $\{X(0), P(0)\}$. The normalization of the momentum by $m\omega_0$, similar to the one performed by the second of Eqs. (63), makes this trajectory pleasingly circular, with a constant radius equal to the oscillation amplitude A, corresponding to the constant full energy

$$E = \frac{m\omega_0^2}{2} A^2, \qquad \text{with } A^2 = \left[X(t)\right]^2 + \left[\frac{P(t)}{m\omega_0}\right]^2 = \text{const} = \left[X(0)\right]^2 + \left[\frac{P(0)}{m\omega_0}\right]^2, \tag{5.104}$$

[33] If Eqs. (100) are not evident, please consult a classical mechanics course – e.g., CM Sec. 3.2 and/or Sec. 10.1.
[34] See, e.g., CM Chapter 5, especially Eqs. (5.4).

determined by the initial conditions – see Fig. 8.)

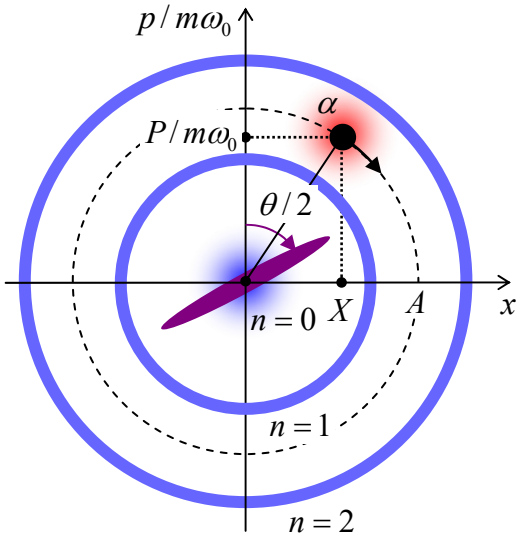


Fig. 5.8. Representations of various states of a harmonic oscillator on the phase plane. The bold black point represents a classical state with displacement amplitude A, with the dashed line showing its trajectory. The (very imperfect) classical images of the Fock states with $n = 0$, 1, and 2 are shown in blue. The blurred red spot is the (equally schematic) image of the Glauber state α, with $|\alpha| = A/\sqrt{2}x_0$. Finally, the magenta elliptical spot is a classical image of a *squeezed* ground state – see below. Arrows show the direction of the states' evolution in time.

On the other hand, according to the basic Eq. (4.161), the time dependence of a Fock state, as of a stationary state of the oscillator, is limited to the phase factor $\exp\{-iE_nt/\hbar\}$. This factor drops out at the averaging (4.125) for any observable. As a result, in this state the expectation values of x, p, or any function thereof are time-independent; moreover, as Eqs. (98) show, $\langle x\rangle = \langle p\rangle = 0$. Taking into account Eqs. (96)-(97), the closest (though very imperfect) geometric image[35] of such a state on the phase plane is a static circle of the radius $A_n = x_0(2n + 1)^{1/2}$, along which the wavefunction is uniformly spread – see the blue rings in Fig. 8. For the ground state ($n = 0$), with the wavefunction (2.275), a better image may be a blurred round spot, of a radius $\sim x_0$, at the origin. (It is easy to criticize such blurring, intended to represent the non-vanishing spreads (99), because it fails to reflect the fact that the total energy of the oscillator in the state, $E_0 = \hbar\omega_0/2$ is definite, without any uncertainty.)

So, the difference between a classical state of the oscillator and its Fock state n is very profound; it is much similar to the difference between the classical picture of a freely moving 1D particle and the traveling de Broglie wave (1.88). However, the Fock states are not the only possible quantum states of the oscillator: according to the basic Eq. (4.6), any state described by the ket-vector

$$|\alpha\rangle = \sum_{n=0}^{\infty} \alpha_n |n\rangle \tag{5.105}$$

with an arbitrary set of (complex) c-numbers α_n, is also its legitimate state, subject only to the normalization condition $\langle\alpha|\alpha\rangle = 1$, giving

$$\sum_{n=0}^{\infty} |\alpha_n|^2 = 1. \tag{5.106}$$

[35] I have to confess that such geometric mapping of a quantum state onto the phase plane $[x, p]$ is not exactly defined; you may think about colored areas in Fig. 8 as the regions of the observable pairs $\{x, p\}$ most probably obtained in measurements. A quantitative definition of such a mapping will be given in Sec. 7.3 using the Wigner function, though, as we will see there, even such imaging has certain internal contradictions. Still, such cartoons as Fig. 8 have a substantial heuristic value, provided that their limitations are kept in mind.

It is natural to ask: could we select the coefficients α_n in such a special way that the state properties would be closer to the classical one; in particular the expectation values $\langle x \rangle$ and $\langle p \rangle$ of the coordinate and momentum would evolve in time as the classical values $X(t)$ and $P(t)$, while the uncertainties of these observables would be, just as in the ground state, given by Eqs. (99), and hence have the smallest possible uncertainty product, $\delta x \delta p = \hbar/2$. As early as 1926, E. Schrödinger showed that the answer was positive. In particular, by using special properties of the Hermite polynomials (2.281), he showed that the corresponding wavefunction, in the coordinate representation, is

$$\Psi_\alpha(x,t) = \left(\frac{m\omega_0}{\pi\hbar}\right)^{1/4} \exp\left\{-\frac{m\omega_0}{2\hbar}[x - X(t)]^2 + i\frac{P(t)x}{\hbar} - i\phi(t)\right\}, \tag{5.107}$$

Glauber state: wavefunction

where

$$\phi(t) = \frac{\omega_0 t}{2} + \frac{X(t)P(t)}{2m\omega_0^2 x_0^2} + \text{const.} \tag{5.108}$$

This solution, whose validity may be readily verified by its substitution to the full Schrödinger equation for the oscillator's Hamiltonian (2.271) with the account of Eqs. (100), shows that such a *Glauber state*[36] is essentially the ground state but with its center shifted from the phase plane's origin to the classical oscillation point $\{X(t), P(t)\}$ – see the blurred red spot in Fig. 8. Moreover, it clearly shows that the coordinate's uncertainty, which is not affected by the x-independent phase shift $\phi(t)$, does not change with time, i.e. that in the harmonic oscillator, the Gaussian wave packet (107), once formed, does not spread with time. (As we have seen in Sec. 2.2, for a free particle, this is impossible.)

Moreover, a similar (though bulkier) calculation shows that the wavefunction (107), with the appropriately modified phase $\phi(t)$, also satisfies the Schrödinger equation of an oscillator under the effect of a pulse of a classical force $F(t)$, provided that the oscillator initially was in its ground state and that the classical evolution law $\{X(t), P(t)\}$ takes this force into account.[37] Since for many experimental implementations of the harmonic oscillator, the ground state may be readily formed (for example, by letting the oscillator relax via its weak coupling to a low-temperature environment), the Glauber state is usually easier to form than any Fock state with $n > 0$. This is why the Glauber states are so important and deserve a thorough discussion.

However, for such a discussion, the usual methods of wave mechanics and even the expansion (105) are rather inconvenient, because of the bulky coordinate representation (2.284) of the Fock states n. Instead, the needed calculations may be more readily done in the bra-ket formalism.

Let us start by expressing the double shift of the ground state (by X and by P), which is so evident in Eq. (107), in the operator language. Forgetting about the P for a minute, let us find the *translation operator* $\hat{\mathcal{T}}_X$ that would produce the desired shift of an arbitrary wavefunction $\psi(x)$ by a c-number distance X along the coordinate argument x. This means that

[36] Named after R. J. Glauber who studied these states in detail in the mid-1960s using operator methods – see below. Another popular adjective, "coherent", for the Glauber states is very misleading, because *all* quantum states of *all* systems we have studied in this course so far, including the Fock states of the harmonic oscillator, may be represented as coherent (pure) superpositions of the basis states. This is why I will not use this term for the Glauber states.

[37] To find it, it is sufficient to integrate Eqs. (100) with $F(t)$ added to the right-hand side of the second of these equations. For their solution for an arbitrary $F(t)$, see, e.g., CM Eqs. (5.27) and (5.34) with $\delta = 0$.

$$\hat{\mathcal{T}}_X \psi(x) \equiv \psi(x - X). \tag{5.109}$$

Representing the wavefunction ψ as the standard wave packet (4.264), we see that

$$\hat{\mathcal{T}}_X \psi(x) = \frac{1}{(2\pi\hbar)^{1/2}} \int \varphi(p) \exp\left\{ i \frac{p(x-X)}{\hbar} \right\} dp \equiv \frac{1}{(2\pi\hbar)^{1/2}} \int \left[\varphi(p) \exp\left\{ -i \frac{pX}{\hbar} \right\} \right] \exp\left\{ i \frac{px}{\hbar} \right\} dp. \tag{5.110}$$

Hence, the shift may be achieved by the multiplication of each Fourier component of the packet, with the momentum p, by $\exp\{-ipX/\hbar\}$. This gives us a hint that the general form of the translation operator, valid in *any* representation, should be

$$\hat{\mathcal{T}}_X = \exp\left\{ -i \frac{\hat{p}X}{\hbar} \right\}. \tag{5.111}$$

The proof of this formula is provided merely by the fact that, as we know from Chapter 4, any operator is uniquely determined by the set of its matrix elements in any full and orthogonal basis, in particular, the basis of the momentum states p. According to Eq. (110), the analog of Eq. (4.235) for the p-representation, applied to the translation operator (which is evidently local), is

$$\int dp \langle p | \hat{\mathcal{T}}_X | p' \rangle \varphi(p') = \exp\left\{ -i \frac{pX}{\hbar} \right\} \varphi(p), \tag{5.112}$$

so the operator (111) does exactly the job we need it to.

The operator that provides the shift of momentum by a c-number P is absolutely similar in structure – with the opposite sign under the exponent, due to the opposite sign of the exponent in the reciprocal Fourier transform, so the simultaneous shift by both X and P may be achieved by the following translation operator:

Translation operator

$$\hat{\mathcal{T}}_\alpha = \exp\left\{ i \frac{P\hat{x} - \hat{p}X}{\hbar} \right\}. \tag{5.113}$$

As we already know, for a harmonic oscillator, the creation-annihilation operators are more natural, so we may use Eqs. (66) to recast Eq. (113) as

$$\hat{\mathcal{T}}_\alpha = \exp\left\{ \alpha \hat{a}^\dagger - \alpha^* \hat{a} \right\}, \qquad \text{so } \hat{\mathcal{T}}_\alpha{}^\dagger = \exp\left\{ \alpha^* \hat{a} - \alpha \hat{a}^\dagger \right\}, \tag{5.114}$$

where α (which, generally, may be a function of time) is the c-number defined by Eq. (101). Now, getting clues from Eq. (107), we may form the Glauber state's ket-vector just as

$$|\alpha\rangle = \hat{\mathcal{T}}_\alpha |0\rangle. \tag{5.115}$$

This formula, valid in any representation, is very elegant, but using it for practical calculations (say, of the expectation values of observables) is not too easy because of the exponent-of-operators form of the translation operator (113). Fortunately, it turns out that a much simpler representation of the Glauber state is possible. To show this, let us start with the following general (and very useful) property of exponential functions of an operator argument: if

$$\left[\hat{A}, \hat{B} \right] = \mu \hat{I}, \tag{5.116}$$

(where $\hat{A}$ and $\hat{B}$ are arbitrary linear operators, and μ is a c-number), then

$$\exp\{+\hat{A}\}\hat{B}\exp\{-\hat{A}\} = \hat{B} + \mu\hat{I}. \tag{5.117}$$

This relation may be readily proved by expanding the operator $\hat{f}(\lambda) \equiv \exp\{+\lambda\hat{A}\}\hat{B}\exp\{-\lambda\hat{A}\}$ in the Taylor series with respect to the *c*-number parameter λ, and then evaluating the result for $\lambda = 1$. (This simple exercise is left for the reader.)

Let us apply Eqs. (116)-(117) to two cases, both with

$$\hat{A} = \alpha^*\hat{a} - \alpha\hat{a}^\dagger, \quad \text{so} \quad \exp\{+\hat{A}\} = \hat{\mathcal{T}}_\alpha{}^\dagger, \quad \exp\{-\hat{A}\} = \hat{\mathcal{T}}_\alpha. \tag{5.118}$$

First, let us take $\hat{B} = \hat{I}$; then Eq. (116) is valid with $\mu = 0$, and Eq. (117) yields

$$\hat{\mathcal{T}}_\alpha{}^\dagger\hat{\mathcal{T}}_\alpha = \hat{I}, \tag{5.119}$$

This equality means that the translation operator is unitary – not a big surprise, because if we shift a classical point on the phase plane by a complex number ($+\alpha$) and then by ($-\alpha$), we certainly must come back to the initial position. Eq. (119) means merely that this fact is true for any quantum state as well.

Second, let us take $\hat{B} = \hat{a}$; in order to find the corresponding parameter μ, we must calculate the commutator on the left-hand side of Eq. (116) for this case. By using, at the due step of the calculation, Eq. (68), we get

$$\left[\hat{A}, \hat{B}\right] = \left[\alpha^*\hat{a} - \alpha\hat{a}^\dagger, \hat{a}\right] = -\alpha\left[\hat{a}^\dagger, \hat{a}\right] = \alpha\hat{I}, \tag{5.120}$$

so in this case $\mu = \alpha$, and Eq. (117) yields

$$\hat{\mathcal{T}}_\alpha{}^\dagger\hat{a}\hat{\mathcal{T}}_\alpha = \hat{a} + \alpha\hat{I}. \tag{5.121}$$

We have approached the summit of this beautiful calculation. Let us consider the following operator:

$$\hat{\mathcal{T}}_\alpha\hat{\mathcal{T}}_\alpha{}^\dagger\hat{a}\hat{\mathcal{T}}_\alpha. \tag{5.122}$$

Using Eq. (119), we may reduce this product to $\hat{a}\hat{\mathcal{T}}_\alpha$, while the application of Eq. (121) to the same expression (122) yields $\hat{\mathcal{T}}_\alpha\hat{a} + \alpha\hat{\mathcal{T}}_\alpha$. Hence, we get the following operator equality:

$$\hat{a}\hat{\mathcal{T}}_\alpha = \hat{\mathcal{T}}_\alpha\hat{a} + \alpha\hat{\mathcal{T}}_\alpha, \tag{5.123}$$

which may be applied to any state. Now acting by both sides on the ground state's ket $|0\rangle$, and using the fact that $\hat{a}|0\rangle$ is the null state (while per Eq. (115), $\hat{\mathcal{T}}_\alpha|0\rangle \equiv |\alpha\rangle$), we finally get a very simple and elegant result:[38]

$$\boxed{\hat{a}|\alpha\rangle = \alpha|\alpha\rangle.} \tag{5.124}$$

Glauber state as eigenstate

[38] This result is also somewhat counterintuitive. Indeed, according to Eq. (89), the annihilation operator $\hat{a}$, acting upon a Fock state *n*, "beats it down" to the lower-energy state ($n - 1$). However, according to Eq. (124), the action of the same operator on a Glauber state α does *not* lead to the state change and hence to any energy change! The resolution of this paradox is given by the representation of the Glauber state as a series of Fock states – see Eq. (134) below. The operator $\hat{a}$ indeed transfers each Fock component of this series to a lower-energy state, but it also re-weighs each term of the series, so the complete energy of the Glauber state remains constant.

Thus any Glauber state α is one of the eigenstates of the annihilation operator, namely the one with the eigenvalue equal to the *c*-number parameter α of the state, i.e. to the complex representation (101) of the classical point which is the center of the Glauber state's wavefunction.[39] This fact makes the calculations of all Glauber state properties much simpler. As an example, let us calculate $\langle x \rangle$ in the Glauber state with some *c*-number α:

$$\langle x \rangle = \langle \alpha | \hat{x} | \alpha \rangle = \frac{x_0}{\sqrt{2}} \langle \alpha | \left(\hat{a} + \hat{a}^\dagger \right) | \alpha \rangle \equiv \frac{x_0}{\sqrt{2}} \left(\langle \alpha | \hat{a} | \alpha \rangle + \langle \alpha | \hat{a}^\dagger | \alpha \rangle \right). \tag{5.125}$$

In the first term in the parentheses, we can apply Eq. (124) directly, while in the second term, we can use the bra-counterpart of that relation, $\langle \alpha | \hat{a}^\dagger = \langle \alpha | \alpha^*$. Now assuming that the Glauber state is normalized, $\langle \alpha | \alpha \rangle = 1$, and using Eq. (101), we get

$$\langle x \rangle = \frac{x_0}{\sqrt{2}} \left(\langle \alpha | \alpha | \alpha \rangle + \langle \alpha | \ldots \alpha^* | \alpha \rangle \right) = \frac{x_0}{\sqrt{2}} \left(\alpha + \alpha^* \right) = X, \tag{5.126}$$

Acting absolutely similarly, we may verify that $\langle p \rangle = P$, and that δx and δp do indeed obey Eqs. (99), i.e. do not depend on the shift α. (This simple exercise is highly recommended to the reader.)

As the last sanity check, let us use Eq. (124) to re-calculate the Glauber state's wavefunction (107). Inner-multiplying both sides of that relation by the bra-vector $\langle x|$, and using the definition (65a) of the annihilation operator, we get

$$\frac{1}{\sqrt{2} x_0} \langle x | \left(\hat{x} + i \frac{\hat{p}}{m \omega_0} \right) | \alpha \rangle = \alpha \langle x | \alpha \rangle. \tag{5.127}$$

Since $\langle x|$ is the bra-vector of the eigenstate of the Hermitian operator $\hat{x}$, they may be swapped, with the operator giving its eigenvalue *x*; acting on that bra-vector by the (local!) operator of momentum, we have to use it in the coordinate representation – see Eq. (4.245). As a result, we get

$$\frac{1}{\sqrt{2} x_0} \left(x \langle x | \alpha \rangle + \frac{\hbar}{m \omega_0} \frac{\partial}{\partial x} \langle x | \alpha \rangle \right) = \alpha \langle x | \alpha \rangle. \tag{5.128}$$

But $\langle x|\alpha\rangle$ is nothing else than the Glauber state's wavefunction Ψ_α, so Eq. (128) gives a first-order differential equation:

$$\frac{1}{\sqrt{2} x_0} \left(x \Psi_\alpha + \frac{\hbar}{m \omega_0} \frac{\partial}{\partial x} \Psi_\alpha \right) = \alpha \Psi_\alpha. \tag{5.129}$$

Chasing Ψ_α and *x* to the opposite sides of the equation, and using the definition (101) of the parameter α, we can bring this equation to the following form (valid at fixed *t*, and hence fixed *X* and *P*):

$$\frac{d\Psi_\alpha}{\Psi_\alpha} = \frac{m \omega_0}{\hbar} \left[-x + \left(X + i \frac{P}{m \omega_0} \right) \right] dx. \tag{5.130}$$

Integrating both parts, we return to Eq. (107).

[39] This fact means that the spectrum of eigenvalues α in Eq. (124), viewed as an eigenproblem, is continuous – it may be *any* complex number.

Now we can use Eq. (124) for finding the coefficients α_n in the expansion (105) of the Glauber state α in the series over the Fock states n. Plugging Eq. (105) into both sides of Eq. (124), using the second of Eqs. (89) on the left-hand side, and requiring the coefficients at each ket-vector $|n\rangle$ in both parts of the resulting relation to be equal, we get the following recurrence relation:

$$\alpha_{n+1} = \frac{\alpha}{(n+1)^{1/2}}\alpha_n. \tag{5.131}$$

Applying this relation sequentially for $n = 0, 1, 2$, etc., we get

$$\alpha_n = \frac{\alpha^n}{(n!)^{1/2}}\alpha_0. \tag{5.132}$$

Now we can find α_0 from the normalization requirement (106), getting

$$|\alpha_0|^2 \sum_{n=0}^{\infty} \frac{|\alpha|^{2n}}{n!} = 1. \tag{5.133}$$

In this sum, we may readily recognize the Taylor expansion of the function $\exp\{|\alpha|^2\}$, so the final result (besides an arbitrary common phase multiplier) is

$$|\alpha\rangle = \exp\left\{-\frac{|\alpha|^2}{2}\right\}\sum_{n=0}^{\infty}\frac{\alpha^n}{(n!)^{1/2}}|n\rangle. \tag{5.134}$$

Glauber state vs Fock states

Hence, if the oscillator is in the Glauber state α, the probabilities $W_n \equiv \alpha_n\alpha_n^*$ of finding the system on the n^{th} energy level (86) obey the well-known *Poisson distribution* (Fig. 9):

$$W_n = \frac{\langle n\rangle^n}{n!}e^{-\langle n\rangle}, \tag{5.135}$$

Poisson distribution

where $\langle n\rangle$ is the statistical average of n – see Eq. (1.37):

$$\langle n\rangle = \sum_{n=0}^{\infty} n W_n. \tag{5.136}$$

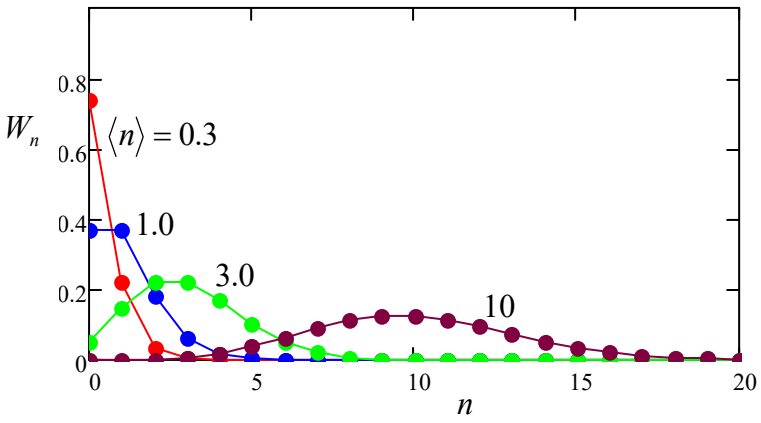


Fig. 5.9. The Poisson distribution (135) for several values of $\langle n\rangle$. Note that W_n are defined only for integer values of n, so the lines are only guides for the eye.

Note that the result of such summation is not necessarily an integer; in our particular case,

$$\langle n \rangle = |\alpha|^2. \tag{5.137}$$

For applications, perhaps the most important property of this distribution is that for any $\langle n \rangle$,

Glauber state: r.m.s. uncertainty

$$\left\langle \tilde{n}^2 \right\rangle \equiv \left\langle \left(n - \langle n \rangle\right)^2 \right\rangle = \langle n \rangle, \qquad \text{so that } \delta n \equiv \left\langle \tilde{n}^2 \right\rangle^{1/2} = \langle n \rangle^{1/2}. \tag{5.138}$$

Another important property is that at $\langle n \rangle >> 1$, the Poisson distribution approaches the Gaussian one, with W_n peaking at $n = \langle n \rangle = |\alpha|^2$, and a small relative r.m.s. uncertainty: $\delta n/\langle n \rangle << 1$ – see Fig. 9.

Now let us discuss the Glauber state's evolution in time. In the wave-mechanics language, it is completely described by the dynamics (100) of the *c*-number shifts $X(t)$ and $P(t)$ participating in the wavefunction (107). An alternative and equivalent way of dynamics description is to use the Heisenberg equation of motion. As Eqs. (29) and (35) tell us, such equations for the Heisenberg operators of coordinate and momentum have to be similar to the classical equations (100):

$$\dot{\hat{x}}_{\rm H} = \frac{\hat{p}_{\rm H}}{m}, \qquad \dot{\hat{p}}_{\rm H} = -m\omega_0^2 \hat{x}_{\rm H}. \tag{5.139}$$

Now using Eqs. (66), for the Heisenberg-picture creation and annihilation operators we get the equations

$$\dot{\hat{a}}_{\rm H} = -i\omega_0 \hat{a}_{\rm H}, \qquad \dot{\hat{a}}^\dagger_{\rm H} = +i\omega_0 \hat{a}^\dagger_{\rm H}, \tag{5.140}$$

which are completely similar to the classical equation (102) for the *c*-number parameter α and its complex conjugate, and hence have the solutions identical to Eq. (103):

$$\hat{a}_{\rm H}(t) = \hat{a}_{\rm H}(0) e^{-i\omega_0 t}, \qquad \hat{a}^\dagger_{\rm H}(t) = \hat{a}^\dagger_{\rm H}(0) e^{i\omega_0 t}. \tag{5.141}$$

As was discussed in Sec. 4.6, such equations are very convenient because they enable simple calculation of time evolution of observables for any initial state of the oscillator (Fock, Glauber, or any other) by using Eq. (4.191). In particular, Eq. (141), without any calculations, shows that regardless of its initial state, the oscillator always returns to it *exactly* with the period $2\pi/\omega_0$.[40]

Applied to the particular case of the ground state of the oscillator, Eq. (141) confirms that the Gaussian wave packet of the special width (99) does not spread in time at all – even temporarily. At this point, I have to notice that there exist other ground-like states whose initial wave packets are still Gaussian but have different widths, say $\delta x < x_0/\sqrt{2}$. As we already know from Sec. 2.2, the momentum spread δp has to be correspondingly larger, but the uncertainty product may still be the smallest: $\delta x \delta p = \hbar/2$. Such *squeezed ground states* ζ, with zero expectation values of x and p, may be generated from the Fock/Glauber ground state:

Squeezed ground state

$$|\zeta\rangle = \hat{S}_\zeta |0\rangle, \tag{5.142a}$$

by using the so-called *squeezing operator*:

[40] Actually, this fact is also evident from the Schrödinger picture of the oscillator's time evolution: due to the exactly equal distances $\hbar\omega_0$ between the eigenenergies (86), the time functions $a_n(t)$ in the fundamental expansion (1.69) of its wavefunction oscillate with frequencies $n\omega_0$, and hence they all share the basic time period $2\pi/\omega_0$.

$$\hat{\mathcal{S}}_\zeta \equiv \exp\left\{\frac{1}{2}\left(\zeta^* \hat{a}\hat{a} - \zeta\, \hat{a}^\dagger \hat{a}^\dagger\right)\right\}, \tag{5.142b}$$

Squeezing operator

which depends on the complex c-number parameter $\zeta = re^{i\theta}$, where r and θ are real. The parameter's modulus r determines the squeezing degree; if ζ is real (i.e. $\theta = 0$), then

$$\delta x = \frac{x_0}{\sqrt{2}} e^{-r}, \qquad \delta p = \frac{m\omega_0 x_0}{\sqrt{2}} e^{r}, \qquad \text{so} \quad \delta x \delta p = \frac{m\omega_0 x_0^2}{2} \equiv \frac{\hbar}{2}. \tag{5.143}$$

On the phase plane (Fig. 8), this state, with $r > 0$, may be represented by an oval spot squeezed along one of two mutually perpendicular axes (hence the state's name), and stretched by the same factor e^r along the counterpart axis; the same formulas but with $r < 0$ describe squeezing along the other axis. On the other hand, the phase θ of the squeezing parameter ζ determines the angle $\theta/2$ of the squeezing/stretching axes about the phase plane origin – see the magenta ellipse in Fig. 8. If $\theta \neq 0$, Eqs. (143) are valid for the variables $\{x', p'\}$ obtained from $\{x, p\}$ via clockwise rotation by that angle. For any of such origin-centered squeezed ground states, the time evolution is reduced to an increase of the angle with the rate ω_0, i.e. to the clockwise rotation of the ellipse, without its deformation, with the angular velocity ω_0 – see the magenta arrows in Fig. 8. As a result, the uncertainties δx and δp oscillate in time with the double frequency $2\omega_0$. Such squeezed ground states may be formed, for example, by a parametric excitation of the oscillator,[41] with a parameter modulation depth close to, but still below the threshold of the excitation of degenerate parametric oscillations.

By action of an additional external force (or by appropriate initial state preparation), the center of a squeezed state may be displaced from the origin to an arbitrary point $\{X, P\}$. Such a displaced squeezed state may be described by the action of the translation operator (113) upon the ground squeezed state, i.e. by the action of the operator product $\hat{\mathcal{T}}_\alpha \hat{\mathcal{S}}_\zeta$ on the usual (Fock/Glauber, i.e. non-squeezed) ground state. Calculations similar to those that led us from Eq. (114) to Eq. (124), show that the displaced squeezed state is an eigenstate of the following mixed operator:

$$\hat{b} \equiv \hat{a} \cosh r + \hat{a}^\dagger e^{i\theta} \sinh r, \tag{5.144}$$

with the same parameters r and θ, with the eigenvalue

$$\beta = \alpha \cosh r + \alpha^* e^{i\theta} \sinh r, \tag{5.145}$$

thus generalizing Eq. (124), which corresponds to $r = 0$. For the particular case $\alpha = 0$, Eq. (145) yields $\beta = 0$, i.e. the action of the operator (144) on the squeezed ground state ζ yields the null state. Just as Eq. (124) in the case of the Glauber states, Eqs. (144)-(145) make the calculation of the basic properties of the squeezed states (for example, the proof of Eqs. (143) for the case $\alpha = \theta = 0$) very straightforward.

Unfortunately, I do not have more time/space for a further discussion of the squeezed states in this chapter (besides a few problems given for the reader's exercise), but their importance for precise quantum measurements will be discussed in Sec. 10.2 below.[42]

---

[41] For a discussion and classical theory of this effect, see, e.g., CM Sec. 5.5.

[42] For more on the squeezed states see, e.g., Chapter 7 in the monograph by C. Gerry and P. Knight, *Introductory Quantum Optics*, Cambridge U. Press, 2005. Also, note the spectacular measurements of the Glauber and squeezed states of electromagnetic (optical) oscillators by G. Breitenbach *et al.*, *Nature* **387**, 471 (1997), a large

5.6. Orbital angular momentum

One more blank spot to fill has been left by our study, in Sec. 3.6, of wave mechanics of particle motion in spherically symmetric 3D potentials. Indeed, while the azimuthal components of the eigenfunctions (the spherical harmonics) of such systems are very simple,

$$\psi_m = (2\pi)^{-1/2} e^{im\varphi}, \quad \text{with } m = 0, \pm 1, \pm 2, \ldots, \tag{5.146}$$

their polar components include the associated Legendre functions $P_l^m(\cos\theta)$, which may be expressed via elementary functions only indirectly – see Eqs. (3.165) and (3.168). This makes all the calculations less than transparent and, in particular, does not allow a clear insight into the origin of the very simple energy spectrum of such systems – see, e.g., Eq. (3.163). The bra-ket formalism, applied to the angular momentum operator, not only enables such insight and produces a very convenient tool for many calculations involving spherically symmetric potentials but also opens a clear way toward the unification of the orbital momentum with the particle's spin – the latter task to be addressed in the next section.

Let us start by using the correspondence principle to spell out the quantum-mechanical vector operator of the orbital angular momentum $\mathbf{L} \equiv \mathbf{r}\times\mathbf{p}$ of a point particle:

Angular momentum operator

$$\hat{\mathbf{L}} \equiv \hat{\mathbf{r}} \times \hat{\mathbf{p}} = \begin{vmatrix} \mathbf{n}_x & \mathbf{n}_y & \mathbf{n}_z \\ \hat{r}_1 & \hat{r}_2 & \hat{r}_3 \\ \hat{p}_1 & \hat{p}_2 & \hat{p}_3 \end{vmatrix}, \qquad \text{i.e. } \hat{L}_j \equiv \sum_{j',j''=1}^{3} \hat{r}_{j'} \hat{p}_{j''} \varepsilon_{jj'j''}, \tag{5.147}$$

where $\varepsilon_{jj'j''}$ is the Levi-Civita permutation symbol, which we have already used in Sec. 4.5, and also in Sec. 1 of this chapter in similar expressions (17)-(18). From this definition, we can readily calculate the commutation relations for all Cartesian components of the vector operators of **L**, **r**, and **p**, for example,

$$\left[\hat{L}_j, \hat{r}_{j'}\right] = \left[\sum_{k,j''=1}^{3} \hat{r}_k \hat{p}_{j''} \varepsilon_{jkj''}, \hat{r}_{j'}\right] \equiv -\sum_{k,j''=1}^{3} \hat{r}_k \left[\hat{r}_{j'}, \hat{p}_{j''}\right] \varepsilon_{jkj''} = -i\hbar \sum_{k,j''=1}^{3} \hat{r}_k \delta_{j'j''} \varepsilon_{jkj''} = i\hbar \sum_{j''=1}^{3} \hat{r}_{j''} \varepsilon_{jj'j''}. \tag{5.148}$$

The summary of all these calculations may be represented in similar forms:

Key commutation relations

$$\left[\hat{L}_j, \hat{r}_{j'}\right] = i\hbar \sum_{j''=1}^{3} \hat{r}_{j''} \varepsilon_{jj'j''}, \qquad \left[\hat{L}_j, \hat{p}_{j'}\right] = i\hbar \sum_{j''=1}^{3} \hat{p}_{j''} \varepsilon_{jj'j''}, \qquad \left[\hat{L}_j, \hat{L}_{j'}\right] = i\hbar \sum_{j''=1}^{3} \hat{L}_{j''} \varepsilon_{jj'j''}; \tag{5.149}$$

the last of them shows that the commutator of two different Cartesian components of $\hat{\mathbf{L}}$ is proportional to its complementary component.

Also introducing, in a natural way, the (scalar!) operator of the observable $L^2 \equiv |\mathbf{L}|^2$,

Operator of L^2

$$\hat{L}^2 \equiv \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2 \equiv \sum_{j=1}^{3} \hat{L}_j^2, \tag{5.150}$$

it is straightforward to check that this operator commutes with each of the Cartesian components:

(ten-fold) squeezing achieved in such oscillators by H. Vahlbruch *et al.*, *Phys. Rev. Lett.* **100**, 033602 (2008), and the first results on the ground state squeezing in micromechanical oscillators, with resonance frequencies $\omega_0/2\pi$ as low as a few MHz, by using their parametric coupling to microwave electromagnetic oscillators – see, e.g., E. Wollman *et al.*, *Science* **349**, 952 (2015) and/or J.-M. Pirkkalainen *et al.*, *Phys. Rev. Lett.* **115**, 243601 (2015).

$$\left[\hat{L}^2, \hat{L}_j\right] = 0. \tag{5.151}$$

This result, at first sight, may seem to contradict the last of Eqs. (149). Indeed, haven't we learned in Sec. 4.5 that commuting operators (e.g., $\hat{L}^2$ and any of $\hat{L}_j$) share their eigenstate sets? If yes, shouldn't this set has to be common for all four angular momentum operators? The resolution in this paradox may be found in the condition that was mentioned just after Eq. (4.138), but (sorry!) was not sufficiently emphasized there. According to that relation, if an operator has *degenerate* eigenstates (i.e. if some $A_j = A_{j'}$ even for $j \neq j'$), they should not be necessarily all shared by another compatible operator.

This is exactly the situation with the orbital angular momentum operators, which may be schematically represented by the *Venn diagram*[43] shown in Fig. 10: the eigenstates of the operator $\hat{L}^2$ are highly degenerate,[44] and their set is broader than those of any component operator $\hat{L}_j$ (that, as will be shown below, are non-degenerate – until we consider the particle's spin).

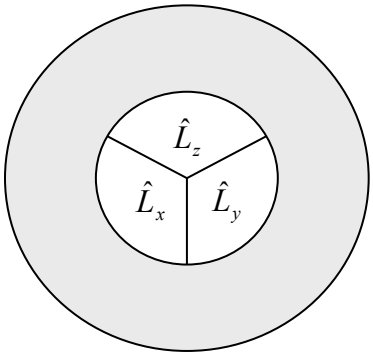


Fig. 5.10. The Venn diagram showing the partitioning of the set of eigenstates of the operator $\hat{L}^2$. Each inner sector corresponds to the states shared with one of the Cartesian component operators $\hat{L}_j$, while the outer (shaded) ring represents the eigenstates of $\hat{L}^2$ that are not shared with either of $\hat{L}_j$ – for example, all linear combinations of the eigenstates of different component operators.

Let us focus on just one of these three joint sets of eigenstates – by tradition, of the operators $\hat{L}^2$ and $\hat{L}_z$. (This tradition stems from the canonical form of the spherical coordinates, in which the polar angle is measured from the *z*-axis. Indeed, in the coordinate representation, we may write

$$\hat{L}_z \equiv \hat{x}\hat{p}_y - \hat{y}\hat{p}_x = x\left(-i\hbar\frac{\partial}{\partial y}\right) - y\left(-i\hbar\frac{\partial}{\partial x}\right) = -i\hbar\frac{\partial}{\partial \varphi}\ . \tag{5.152}$$

Writing the standard eigenproblem for the operator in this representation, $\hat{L}_z\psi_m = L_z\psi_m$, we see that it is satisfied by the eigenfunctions (146), with eigenvalues $L_z = \hbar m$ – the fact that was already conjectured in Sec. 3.5.) More specifically, let us consider a set of eigenstates $\{l, m\}$ corresponding to a certain degenerate eigenvalue of the operator $\hat{L}^2$, and all possible eigenvalues of the operator $\hat{L}_z$, i.e. all possible quantum numbers *m*. (At this point, *l* is just a label of the eigenvalue of the operator $\hat{L}^2$; it will

[43] This is just a particular example of the Venn diagrams (introduced in the 1880s by John Venn) that show possible relations (such as intersections, unions, complements, etc.) between various sets of objects, and are a very useful notion of the general set theory.

[44] Note that this particular result is consistent with the classical picture of the angular momentum vector: even when its length is fixed, the vector may be oriented in various directions, corresponding to different values of its Cartesian components. However, in the classical picture, all these components may have exactly fixed values simultaneously, while in the quantum picture, this is not true.

be defined more explicitly in a minute.) To analyze this set, it is instrumental to introduce the so-called *ladder* (also called, respectively, "raising" and "lowering") *operators*[45]

Ladder operators

$$\hat{L}_{\pm} \equiv \hat{L}_x \pm i\hat{L}_y . \tag{5.153}$$

It is simple (and hence left for the reader's exercise) to use this definition and the last of Eqs. (149) to calculate the following commutators:

Important commutation relations

$$\left[\hat{L}_+, \hat{L}_-\right] = 2\hbar\hat{L}_z, \quad \text{and} \quad \left[\hat{L}_z, \hat{L}_\pm\right] = \pm\hbar\hat{L}_\pm, \tag{5.154}$$

and also to use Eqs. (149)-(150) to prove two other important operator relations:

$$\hat{L}^2 = \hat{L}_z^2 + \hat{L}_+\hat{L}_- - \hbar\hat{L}_z, \qquad \hat{L}^2 = \hat{L}_z^2 + \hat{L}_-\hat{L}_+ + \hbar\hat{L}_z . \tag{5.155}$$

Now let us rewrite the last of Eqs. (154) as

$$\hat{L}_z\hat{L}_\pm = \hat{L}_\pm\hat{L}_z \pm \hbar\hat{L}_\pm , \tag{5.156}$$

and act by both its sides upon the ket-vector $|l, m\rangle$ of an arbitrary common eigenstate:

$$\hat{L}_z\hat{L}_\pm|l,m\rangle = \hat{L}_\pm\hat{L}_z|l,m\rangle \pm \hbar\hat{L}_\pm|l,m\rangle. \tag{5.157}$$

Since the eigenvalues of the operator $\hat{L}_z$ are equal to $\hbar m$, in the first term of the right-hand side of Eq. (157) we may write

$$\hat{L}_z|l,m\rangle = \hbar m|l,m\rangle. \tag{5.158}$$

With that, Eq. (157) may be recast as

$$\hat{L}_z\left(\hat{L}_\pm|l,m\rangle\right) = \hbar(m\pm1)\left(\hat{L}_\pm|l,m\rangle\right). \tag{5.159}$$

In a spectacular similarity with Eqs. (78)-(79) for the harmonic oscillator, Eq. (159) means that the states $\hat{L}_\pm|l,m\rangle$ are also eigenstates of the operator $\hat{L}_z$, corresponding to eigenvalues $\hbar(m \pm 1)$. Thus the ladder operators work exactly as the creation and annihilation operators of a harmonic oscillator, moving the system up or down a ladder of eigenstates – see Fig. 11.

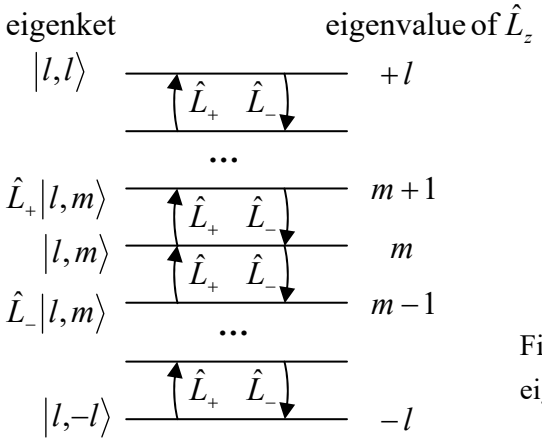


Fig. 5.11. The ladder diagram of the common eigenstates of the operators $\hat{L}^2$ and $\hat{L}_z$.

[45] Note a substantial similarity between this definition and Eqs. (65) for the creation/annihilation operators.

The most significant difference is that now the state ladder must end in both directions, because an infinite increase of $|m|$, with whichever sign of m, would cause the expectation values of the operator

$$\hat{L}_x^2 + \hat{L}_y^2 \equiv \hat{L}^2 - \hat{L}_z^2, \tag{5.160}$$

which corresponds to a non-negative observable, becoming negative. Hence there have to be two states at both ends of the ladder, with such ket-vectors $|l, m_{max}\rangle$ and $|l, m_{min}\rangle$ that

$$\hat{L}_+ \left| l, m_{max} \right\rangle = 0, \qquad \hat{L}_- \left| l, m_{min} \right\rangle = 0. \tag{5.161}$$

Due to the symmetry of the whole problem with respect to the replacement $m \to -m$, we should have $m_{min} = -m_{max}$. This m_{max} is exactly the quantum number traditionally called l, i.e.

$$-l \le m \le +l. \tag{5.162}$$

Relation between m and l

This relation of the quantum numbers m and l is semi-quantitatively compatible with the classical image of the angular momentum vector **L**, of the same length L, pointing in various directions, thus affecting the value of its component L_z. In this classical picture, however, L^2 would be equal to the square of $(L_z)_{max}$, i.e. to $(\hbar l)^2$; however, in quantum mechanics, this is not so. Indeed, applying both parts of the second of the operator equalities (155) to the top state's vector $|l, m_{max}\rangle \equiv |l, l\rangle$, we get

$$\begin{aligned} \hat{L}^2 |l,l\rangle &= \hbar \hat{L}_z |l,l\rangle + \hat{L}_z^2 |l,l\rangle + \hat{L}_- \hat{L}_+ |l,l\rangle = \hbar^2 l |l,l\rangle + \hbar^2 l^2 |l,l\rangle + 0 \\ &\equiv \hbar^2 l(l+1) |l,l\rangle. \end{aligned} \tag{5.163}$$

Since by our initial assumption, all eigenvectors $|l, m\rangle$ correspond to the same eigenvalue of $\hat{L}^2$, this result means that *all* these eigenvalues are equal to $\hbar^2 l(l + 1)$. Just as in the case of the spin-½ vector operators discussed in Sec. 4.5, the deviation of this result from $\hbar^2 l^2$ may be interpreted as a result of unavoidable uncertainties ("fluctuations") of the x- and y-components of the angular momentum, which give non-zero positive contributions to $\langle L_x^2 \rangle$ and $\langle L_y^2 \rangle$, and hence to $\langle L^2 \rangle$, even if the angular momentum vector is aligned with the z-axis in the best possible way.[46]

(For applications, one more relation, in one of its two equivalent forms, may be convenient:

$$\hat{L}_\pm \left| l, m \right\rangle = \hbar \left[l(l+1) - m(m \pm 1) \right]^{1/2} \left| l, m \pm 1 \right\rangle \equiv \hbar \left[(l \pm m + 1)(l \mp m) \right]^{1/2} \left| l, m \pm 1 \right\rangle. \tag{5.164}$$

This equality, valid to the multiplier $e^{i\varphi}$ with an arbitrary real phase φ, may be readily proved from the above relations in the same way as the parallel Eqs. (89) for the harmonic-oscillator operators (65) were proved in Sec. 4; due to this similarity, the proof is also left for the reader's exercise.[47])

---

[46] Curiously, a similar formula $\langle L^2 \rangle = \hbar^2 l(l + 1)$ may be also obtained by assuming that all $(2l + 1)$ values $L_z = \hbar m$ of a system with fixed l have equal probability. (Let me leave the proof for the reader's exercise.)

[47] The reader is also challenged to use the commutation relations discussed above to prove one more important property of the common eigenstates of the operators $\hat{L}_z$ and $\hat{L}^2$:

$$\left\langle l, m \left| \hat{r}_j \right| l', m' \right\rangle = 0, \qquad \text{unless } l' = l \pm 1 \text{ and } m' = \text{either } m \pm 1 \text{ or } m.$$

This property gives the *selection rule* for the orbital electric-dipole quantum transitions, to be discussed later in the course, especially in Sec. 9.3. (The final selection rules at these transitions may be affected by the particle's spin – see the next section.)

Note that the formulas discussed in this section, with the sole exception of Eq. (146), are not conditioned by a particular Hamiltonian of the system under analysis. However, they (as well as those discussed in the next section) are especially important for particles moving in spherically-symmetric potentials, which were discussed in Sec. 3.6. It is easy (and hence is also left for the reader's exercise) to prove that in this case, the particle's Hamiltonian operator commutes with that of the angular momentum, so according to Eq. (4.199), in the Heisenberg picture of quantum dynamics, the Cartesian components $\hat{L}_j$ as well as $\hat{L}^2$ do not depend on time, and hence their expectation values are integrals of motion.

By using the expression of Cartesian coordinates via the spherical ones exactly as this was done in Eq. (152), we get the following expressions for the ladder operators (153) in the coordinate representation:

$$\hat{L}_{\pm} = \hbar e^{\pm i\varphi}\left(\pm\frac{\partial}{\partial\theta} + i\,\mathrm{cotan}\,\theta\frac{\partial}{\partial\varphi}\right). \tag{5.165}$$

Angular momentum operators: coordinate representation

Now plugging this relation, together with Eq. (152), into any of Eqs. (155), we get

$$\hat{L}^2 = -\hbar^2\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right]. \tag{5.166}$$

But this is exactly the operator (besides its division by the constant parameter $2mR^2$) that stands on the left-hand side of Eq. (3.156). Hence that equation, which was explored by the "brute-force" (wave-mechanical) approach in Sec. 3.6, may be understood as the eigenproblem for the operator $\hat{L}^2$ in the coordinate representation, with the eigenfunctions $Y_l^m(\theta,\varphi)$ corresponding to the eigenkets $|l, m\rangle$, and the eigenvalues $L^2 = 2mR^2E$. As a reminder, the main result of that, rather involved analysis was expressed by Eq. (3.163), which now may be rewritten as

$$L_l^2 \equiv 2mR^2E_l = \hbar^2 l(l+1), \tag{5.167}$$

in full agreement with Eq. (163), which was obtained by much more efficient means based on the bra-ket formalism. In particular, it is fascinating to see how easy it is to operate with the eigenvectors $|l, m\rangle$, while the coordinate representations of these vectors, the spherical harmonics $Y_l^m(\theta,\varphi)$, may be only expressed by rather complicated functions – please have one more look at Eq. (3.171) and Fig. 3.20.

5.7. Spin and total angular momentum

The theory described in the last section is useful for much more than orbital motion analysis. In particular, it helps to generalize the spin-½ results discussed in Chapter 4 to other values of spin s – the parameter still to be quantitatively defined. For that, let us notice that the commutation relations (4.155) for spin-½, which were derived from the Pauli matrix properties, may be rewritten in exactly the same form as Eqs. (149) and (151) for the orbital momentum:

Spin operators: commutation relations

$$\left[\hat{S}_j, \hat{S}_{j'}\right] = i\hbar\sum_{j''=1}^{3}\hat{S}_{j''}\varepsilon_{jj'j''}, \qquad \left[\hat{S}^2, \hat{S}_j\right] = 0. \tag{5.168}$$

It had been postulated (and then confirmed by numerous experiments) that these relations hold for quantum particles with any spin. Now notice that all the calculations of the last section have been based *almost* exclusively on such relations – the only exception will be discussed imminently. Hence, we may repeat them for the spin operators, and get the relations similar to Eqs. (158) and (163):

$$\hat{S}_z|s,m_s\rangle = \hbar m_s|s,m_s\rangle, \quad \hat{S}^2|s,m_s\rangle = \hbar^2 s(s+1)|s,m_s\rangle, \quad 0 \le s, \quad -s \le m_s \le +s, \tag{5.169}$$

Spin operators: eigenstates and eigenvalues

where m_s is a quantum number parallel to the orbital magnetic number m, and the non-negative constant s is defined as the maximum value of $|m_s|$. The c-number s is exactly what is called the *particle's spin*.

Now let us return to the only part of our orbital moment calculations that has *not* been derived from the commutation relations. This was the fact, based on the solution (146) of the orbital motion problems, that the quantum number m (the analog of m_s) may be only an integer. For spin, we do not have such a solution, so the spectrum of numbers m_s (and hence its limits $\pm s$) should be found from the more loose requirement that the eigenstate ladder, extending from $-s$ to $+s$, has an integer number of steps. Hence, $2s$ has to be an integer, i.e. the spin s of a quantum particle may be either *integer* (as it is, for example, for photons, gluons, and massive bosons $W^{\pm}$ and Z^0), or *half-integer* (e.g., for all quarks and leptons, notably including electrons).[48] For $s = ½$, this picture yields all the properties of the spin-½ that were derived in Chapter 4 from Eqs. (4.115)-(4.117). In particular, the operators $\hat{S}^2$ and $\hat{S}_z$ have two common eigenstates (↑ and ↓), with $S_z = \hbar m_s = \pm\hbar/2$, both with $S^2 = s(s+1)\hbar^2 = (3/4)\hbar^2$.

Note that this analogy with the angular momentum sheds new light on the symmetry properties of spin-½. Indeed, the fact that m in Eq. (146) is an integer was derived in Sec. 3.5 from the requirement that making a full circle around the z-axis, we should find a similar final value of the wavefunction ψ_m, which may differ from the initial one only by an inconsequential factor $\exp\{2\pi i m\} = +1$. With the replacement $m \to m_s = \pm ½$, such an operation would multiply the wavefunction by $\exp\{\pm\pi i\} = -1$, i.e. reverse its sign. Of course, spin properties cannot be described by a usual wavefunction, but this *odd parity* of electrons, shared by all other spin-½ particles, is clearly revealed in properties of multiparticle systems (see Chapter 8 below), and as a result, in their statistics (see, e.g., SM Chapter 2).

Now we are sufficiently equipped to analyze the situations in which a particle has both the orbital momentum and the spin – as an electron inside an atom. In classical mechanics, such an object, with the spin **S** interpreted as the angular moment of its internal rotation, would be characterized by the *total angular momentum* vector **J** = **L** + **S.** Following the correspondence principle, we may assume that quantum-mechanical properties of this observable may be described by the similarly defined vector operator:

$$\hat{\mathbf{J}} \equiv \hat{\mathbf{L}} + \hat{\mathbf{S}}, \tag{5.170}$$

Total angular momentum

with Cartesian components

$$\hat{J}_z \equiv \hat{L}_z + \hat{S}_z, \text{ etc.}, \tag{5.171}$$

and the magnitude squared equal to

$$\hat{J}^2 \equiv \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2. \tag{5.172}$$

[48] As a reminder, in the Standard Model of particle physics, such hadrons as mesons and baryons (notably including protons and neutrons) are essentially composite particles. However, at non-relativistic energies, protons and neutrons may be considered fundamental particles with $s = ½$.

Let us examine the key properties of this vector operator. Since its two components (170) describe different degrees of freedom of the particle, i.e. belong to different Hilbert spaces, they have to be completely commuting:

$$\left[\hat{L}_j,\hat{S}_{j'}\right]=0,\qquad \left[\hat{L}^2,\hat{S}^2\right]=0,\qquad \left[\hat{L}_j,\hat{S}^2\right]=0,\qquad \left[\hat{L}^2,\hat{S}_j\right]=0. \tag{5.173}$$

The above formulas are sufficient to derive the commutation relations for the operator $\hat{\mathbf{J}}$, and unsurprisingly, they turn out to be absolutely similar to those of its orbital and spin components:

Total momentum: commutation relations

$$\boxed{\left[\hat{J}_j,\hat{J}_{j'}\right]=i\hbar\sum_{j''=1}^{3}\hat{J}_{j''}\varepsilon_{jj'j''},\qquad \left[\hat{J}^2,\hat{J}_j\right]=0\,.} \tag{5.174}$$

Now by repeating all the arguments of the last section, we may derive the following expressions for the common eigenstates of the operators $\hat{J}^2$ and $\hat{J}_z$:

Total momentum: eigenstates, and eigenvalues

$$\boxed{J_z\left|j,m_j\right\rangle=\hbar m_j\left|j,m_j\right\rangle,\quad \hat{J}^2\left|j,m_j\right\rangle=\hbar^2 j(j+1)\left|j,m_j\right\rangle,\quad 0\le j,\quad -j\le m_j\le+j,} \tag{5.175}$$

where j and m_j are new quantum numbers.[49] Repeating the arguments just made for s and m_s, we may conclude that j and m_j may be either integers or half-integers.

Before we proceed, one remark on notation: it is very convenient to use the same letter m for numbering eigenstates of all momentum components participating in Eq. (171), with corresponding indices (j, l, and s), in particular, to replace what we called m with m_l. With this replacement, the main results of the last section may be summarized in a form similar to Eqs. (168), (169), (174), and (175):

Orbital momentum: basic properties (new notation)

$$\boxed{\left[\hat{L}_j,\hat{L}_{j'}\right]=i\hbar\sum_{j''=1}^{3}\hat{L}_{j''}\varepsilon_{jj'j''},\quad \left[\hat{L}^2,\hat{L}_j\right]=0\,,} \tag{5.176}$$

$$\boxed{\hat{L}_z\left|l,m_l\right\rangle=\hbar m_l\left|l,m_l\right\rangle,\quad \hat{L}^2\left|l,m_l\right\rangle=\hbar^2 l(l+1)\left|l,m_l\right\rangle,\quad 0\le l,\quad -l\le m_l\le+l.} \tag{5.177}$$

In order to understand which eigenstates participating in Eqs. (169), (175), and (177) are compatible with each other, it is straightforward to use Eq. (172), together with Eqs. (168), (173), (174), and (176) to get the following relations:

$$\left[\hat{J}^2,\hat{L}^2\right]=0,\quad \left[\hat{J}^2,\hat{S}^2\right]=0, \tag{5.178}$$

$$\left[\hat{J}^2,\hat{L}_z\right]\neq 0,\quad \left[\hat{J}^2,\hat{S}_z\right]\neq 0. \tag{5.179}$$

This result is represented schematically on the Venn diagram shown in Fig. 12, in which the crossed arrows indicate the only *non*-commuting pairs of operators. The color lines in this figure encircle two operator groups that commute with each other and hence may share their eigenstates. The first group (encircled red), consists of all operators but $\hat{J}^2$; their shared eigenstates correspond to definite values of the corresponding quantum numbers: l, m_l, s, m_s, and m_j. Actually, only four of these numbers are independent, because due to Eq. (171) for these compatible operators, for each eigenstate of this group, their "magnetic" quantum numbers m have to satisfy the following relation:

[49] Let me hope that the difference between the quantum number j, and the indices j, j', j'' numbering the Cartesian components in relations like Eqs. (168) or (174), is absolutely clear from the context.

$$m_j = m_l + m_s. \tag{5.180}$$

Hence the common eigenstates of the operators of this group are fully defined by just four quantum numbers, for example, l, m_l, s, and m_s. For some calculations, especially those for the systems whose Hamiltonians include only the operators of this group, it is convenient to use this set of eigenstates as the basis; frequently this approach is called the *uncoupled representation*. The most important example of such a situation is a non-relativistic particle moving in a spherically-symmetric potential (3.155), whose Hamiltonian does not depend on its spin. As we have seen in the previous section, its stationary states correspond to definite l and m_l.

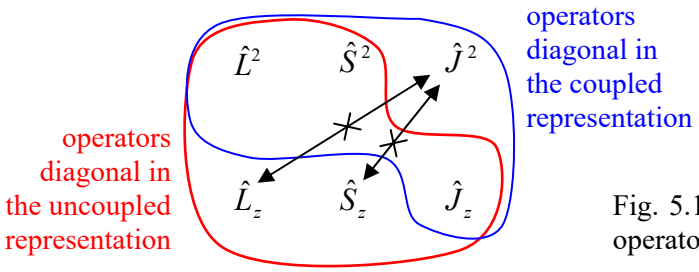


Fig. 5.12. The Venn diagram of angular momentum operators, and their mutually-commuting groups.

However, in some situations, interactions between the orbital and spin degrees of freedom (in the common jargon, the *spin-orbit coupling*) cannot be ignored; this interaction leads in particular to splitting (called the *fine structure*) of the atomic energy levels even in the absence of external magnetic field. I will discuss these effects in detail in the next chapter and now will only note that they may be described by a term proportional to the product $\hat{\mathbf{L}}\cdot\hat{\mathbf{S}}$ in the particle's Hamiltonian. If this term is substantial, the uncoupled representation becomes inconvenient. Indeed, writing

$$\hat{J}^2 = (\hat{\mathbf{L}}+\hat{\mathbf{S}})^2 = \hat{L}^2 + \hat{S}^2 + 2\hat{\mathbf{L}}\cdot\hat{\mathbf{S}}, \qquad \text{so that } 2\hat{\mathbf{L}}\cdot\hat{\mathbf{S}} = \hat{J}^2 - \hat{L}^2 - \hat{S}^2, \tag{5.181}$$

and looking at Fig. 12 again, we see that the operator $\hat{\mathbf{L}}\cdot\hat{\mathbf{S}}$ describing the spin-orbit coupling does not commute with operators $\hat{L}_z$ and $\hat{S}_z$. This means that stationary states of the system with such a term in the Hamiltonian do not belong to the uncoupled representation's basis. On the other hand, Eq. (181) shows that the operator $\hat{\mathbf{L}}\cdot\hat{\mathbf{S}}$ does commute with all four operators of another group, encircled blue in Fig. 12. According to Eqs. (178), (179), and (181), all operators of that group also commute with each other, so they have a group of common eigenstates, described by the quantum numbers l, s, j, and m_j. This group is the basis for the so-called *coupled representation* of particle states.

Excluding, for the notation briefness, the quantum numbers l and s that are common for both groups, it is convenient to denote the common ket-vectors of each group as, respectively,

$$\begin{aligned} &|m_l, m_s\rangle, && \text{for the uncolpled representation's basis,} \\ &|j, m_j\rangle, && \text{for the coupled representation's basis.} \end{aligned} \tag{5.182}$$

Coupled and uncoupled bases

As we will see in the next chapter, for the solution of some important problems (e.g., the fine structure of atomic spectra and the Zeeman effect), we will need the relation between the kets $|j, m_j\rangle$ and the kets $|m_l, m_s\rangle$. This relation may be represented as the usual linear superposition,

Clebsch-Jordan coefficients: definition

$$\left|j,m_j\right\rangle = \sum_{m_l,m_s} \left|m_l,m_s\right\rangle \left\langle m_l,m_s\left|j,m_j\right.\right\rangle . \tag{5.183}$$

The short brackets in this relation, essentially the elements of the unitary matrix of the transformation between two eigenstate bases (182), are called the *Clebsch-Gordan coefficients*.

The best (though imperfect) classical interpretation of Eq. (183) I can offer is as follows. If the lengths of the vectors **L** and **S** (in quantum mechanics associated with the numbers l and s, respectively), and also their scalar product **L**·**S**, are all fixed, then so is the length of the vector **J** = **L** + **S** – whose length in quantum mechanics is described by the number j. Hence, the classical image of a specific eigenket $|j, m_j\rangle$, in which l, s, j, and m_j are all fixed, is a state in which L^2, S^2, J^2, and J_z are fixed. However, this fixation still allows for an arbitrary rotation of the pair of vectors **L** and **S** (with a fixed angle between them, and hence fixed **L**·**S** and J^2) about the direction of the vector **J** – see Fig. 13.

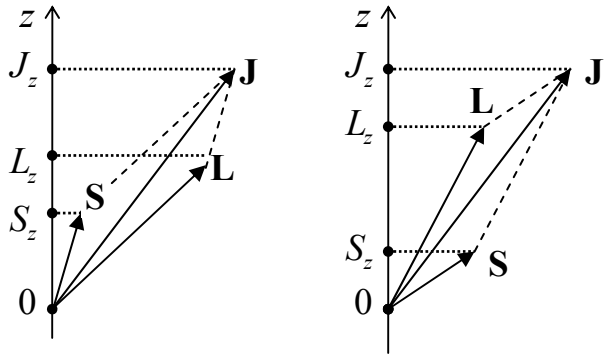


Fig. 5.13. A classical image of two different quantum states with the same quantum numbers l, s, j, and m_j, but different m_l and m_s.

Hence the components L_z and S_z in these conditions are *not* fixed, and in classical mechanics may take a continuum of values, two of which (with the largest and the smallest possible values of S_z) are shown in Fig. 13. In quantum mechanics, these components are quantized, with their states represented by eigenkets $|m_l, m_s\rangle$, so a linear combination of such kets is necessary to represent every ket $|j, m_j\rangle$. This is exactly what Eq. (183) does.

Some properties of the Clebsch-Gordan coefficients $\langle m_l, m_s|\, j, m_j\rangle$ may be readily established. For example, the coefficients do not vanish only if the involved magnetic quantum numbers satisfy Eq. (180). In our current case, this relation is not an elementary corollary of Eq. (171), because the Clebsch-Gordan coefficients, with the quantum numbers m_l, m_s in one state vector, and m_j in the other state vector, characterize the relationship between different groups of the basis states, so we need to prove this fact; let us do that. All matrix elements of the null-operator

$$\hat{J}_z - (\hat{L}_z + \hat{S}_z) = \hat{0} \tag{5.184}$$

should equal zero in any basis; in particular

$$\left\langle j,m_j\left|\hat{J}_z - (\hat{L}_z + \hat{S}_z)\right|m_l,m_s\right\rangle = 0. \tag{5.185}$$

Acting by the operator $\hat{J}_z$ upon the bra-vector, and by the sum $(\hat{L}_z + \hat{S}_z)$ upon the ket-vector, we get

$$\left[m_j - (m_l + m_s)\right]\left\langle j,m_j\left|m_l,m_s\right.\right\rangle = 0, \tag{5.186}$$

thus proving that

$$\langle m_l, m_s | j, m_s \rangle \equiv \langle j, m_s | m_l, m_s \rangle^* = 0, \qquad \text{if } m_j \neq m_l + m_s . \tag{5.187}$$

As we will see in a minute, this property will enable us, in particular, to establish the range of possible values of the quantum number j, at fixed l and s.

For the most important case of spin-½ particles (with $s = ½$, and hence $m_s = \pm ½$), whose uncoupled representation basis includes $2\times(2l + 1)$ states, the restriction (187) enables the representation of all non-zero Clebsch-Gordan coefficients on the simple "rectangular" diagram shown in Fig. 14. Indeed, each coupled-representation eigenket $|j, m_j\rangle$, with $m_j = m_l + m_s = m_l \pm ½$, may be related by non-zero Clebsch-Gordan coefficients to at most two uncoupled-representation eigenstates $|m_l, m_s\rangle$. Since m_l may only take integer values from $-l$ to $+l$, m_j may only take semi-integer values on the interval $[-l - ½, l + ½]$. Hence, by the definition of j as $(m_j)_{\max}$, its maximum value has to be $l + ½$, and for $m_j = l + ½$, this is the only possible value with this j. This means that the uncoupled state with $m_l = l$ and $m_s = ½$ should be identical to the coupled-representation state with $j = l + ½$ and $m_j = l + ½$:

$$\left| j = l + ½, m_j = l + ½ \right\rangle = \left| m_l = m_j - ½, m_s = +½ \right\rangle . \tag{5.188}$$

In Fig. 14, these two identical states are represented by the top-rightmost point (the uncoupled representation) and the sloped line passing through it (the coupled representation).

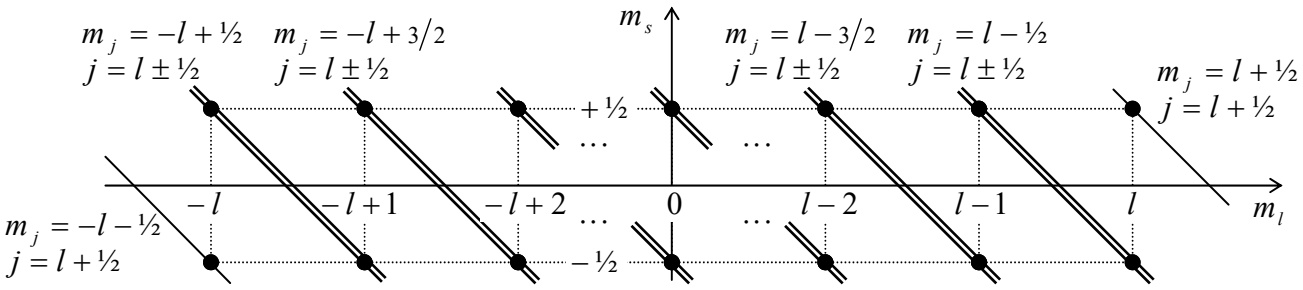


Fig. 5.14. A graphical representation of possible basis states of a spin-½ particle with a fixed l. Each dot corresponds to an uncoupled-representation ket-vector $|m_l, m_s\rangle$, while each sloped line corresponds to one coupled-representation ket-vector $|j, m_j\rangle$, related by Eq. (183) to the kets $|m_l, m_s\rangle$ whose dots it connects.

However, already the next value of this quantum number, $m_j = l - ½$, is compatible with two values of j, so each $|m_l, m_s\rangle$ ket has to be related to two $|j, m_j\rangle$ kets by two Clebsch-Gordan coefficients. Since j changes in unit steps, these values of j have to be $l \pm ½$. This choice,

$$j = l \pm ½, \tag{5.189}$$

where the alternating sign is independent of the sign of m_s, evidently satisfies all lower values of m_j as well – see Fig. 14.[50] (Again, only one value, $j = l + ½$, is necessary to represent the state with the lowest $m_j = -l - ½$ – see the bottom-leftmost point of that diagram.)

[50] Eq. (189) may be readily generalized to the case of arbitrary spin s: j may only take values that differ by 1, within the interval $[|l - s|, l + s]$. This important result (whose proof is left for the reader's exercise) allows a semi-quantitative classical interpretation in terms of the vector diagrams shown in Fig. 13: in them, the largest value of j corresponds to the parallel alignment of the vectors **L** and **S**, while its smallest value, to their antiparallel alignment.

Note that the total number of the coupled-representation states is $1 + 2\times 2l + 1 \equiv 2(2l + 1)$, i.e. is the same as those in the uncoupled representation. So, for spin-½ systems, each sum (183), for fixed j and m_j (plus the fixed common parameter l, plus the common $s = ½$), has at most two terms, i.e. involves at most two Clebsch-Gordan coefficients.

These coefficients may be calculated in a few steps, all but the last one rather simple even for arbitrary spin s. First, the similarity of the vector operators $\hat{\mathbf{J}}$ and $\hat{\mathbf{S}}$ to the operator $\hat{\mathbf{L}}$, expressed by Eqs. (169), (175), and (177), may be used to argue that the matrix elements of the operators $\hat{S}_\pm$ and $\hat{J}_\pm$, defined similarly to $\hat{L}_\pm$, have the matrix elements similar to those given by Eq. (164). Next, acting by the operator $\hat{J}_\pm = \hat{L}_\pm + \hat{S}_\pm$ upon both parts of Eq. (183), and then inner-multiplying the result by the bra vector $\langle m_l, m_s|$ and using the above matrix elements, we may get recurrence relations for the Clebsch-Gordan coefficients with adjacent values of m_l, m_s, and m_j. Finally, these relations may be sequentially applied to the adjacent states in both representations, starting from any of the two states common for them – for example, from the state with the ket-vector (188), corresponding to the top-rightmost point in Fig. 14.

Let me leave these straightforward but a bit tedious calculations for the reader's exercise, and just quote the final result of this procedure for $s = ½$:[51]

Clebsch – Gordan coefficients for $s = ½$

$$\left\langle m_l = m_j - ½, m_s = +½ \middle| j = l \pm ½, m_j \right\rangle = \pm\left(\frac{l \pm m_j + ½}{2l+1}\right)^{1/2},$$
$$\left\langle m_l = m_j + ½, m_s = -½ \middle| j = l \pm ½, m_j \right\rangle = +\left(\frac{l \mp m_j + ½}{2l+1}\right)^{1/2}. \tag{5.190}$$

As a simple example, let an electron be in the p-state ($l = 1$) with definite $j = ½$ and $m_j = ½$, and we want to know the probability of its spin being directed down ($m_s = -½$). Since in this case, $j = l - ½$, the above formulas should be used with the upper signs, giving

$$\left\langle m_l = 0, m_s = +½ \middle| j = ½, m_j = ½ \right\rangle = -\left(\frac{1 - ½ + ½}{2\cdot 1 + 1}\right)^{1/2} \equiv -\left(\frac{1}{3}\right)^{1/2},$$
$$\left\langle m_l = 1, m_s = -½ \middle| j = ½, m_j = ½ \right\rangle = +\left(\frac{1 + ½ + ½}{2l+1}\right)^{1/2} \equiv +\left(\frac{2}{3}\right)^{1/2}, \tag{5.191}$$

so the general Eq. (183) takes the form

$$\left| j = ½, m_j = ½ \right\rangle = -\left(\frac{1}{3}\right)^{1/2} \left| m_l = 0, m_s = +½ \right\rangle + \left(\frac{2}{3}\right)^{1/2} \left| m_l = 1, m_s = -½ \right\rangle, \tag{5.192}$$

and the probability of the spin-down state with $m_s = -½$ is $W_\downarrow = 2/3$.

In this course, Eqs. (190) will be used mostly in Sec. 6.4 for an analysis of the anomalous Zeeman effect. Also, the angular momentum addition rules described above are also valid for the addition of angular momenta of multiparticle system components, so we will revisit them in Chapter 8.

[51] For arbitrary spin s, the calculations and even the final expressions for the Clebsch-Gordan coefficients are rather bulky. They may be found, typically in a table form, mostly in special monographs – see, e.g., A. Edmonds, *Angular Momentum in Quantum Mechanics,* Princeton U. Press, 1957.

To conclude this section, I have to note that the Clebsch-Gordan coefficients (for arbitrary s) participate also in the so-called *Wigner-Eckart theorem* that expresses the matrix elements of *spherical tensor operators*, in the coupled-representation basis $|j, m_j\rangle$, via a reduced set of matrix elements. This theorem may be useful, for example, for the calculation of the rate of quantum transitions to/from high-n states in spherically symmetric potentials. Unfortunately, a discussion of this theorem and its applications would require a higher mathematical background than I can expect from my readers and more time/space than I can afford.[52]

5.8. Exercise problems

5.1. Use the discussion in Sec. 1 to find an alternative solution of Problem 4.18.

5.2. A spin-½ with a gyromagnetic ratio γ is placed into an external magnetic field, with a time-independent orientation, its magnitude $\mathscr{B}(t)$ being an arbitrary function of time. Find explicit expressions for the Heisenberg operators and the expectation values of all three Cartesian components of the spin as functions of time, in a coordinate system of your choice.

5.3. A two-level system is in the quantum state α described by the ket-vector $|\alpha\rangle = \alpha_\uparrow|\uparrow\rangle + \alpha_\downarrow|\downarrow\rangle$, with given (generally, complex) c-number coefficients $\alpha_{\uparrow\downarrow}$. Prove that we can always select such a geometric c-number vector $\mathbf{c} = \{c_x, c_y, c_z\}$ that α would be an eigenstate of $\mathbf{c}\cdot\hat{\boldsymbol{\sigma}}$, where $\hat{\boldsymbol{\sigma}}$ is the Pauli vector operator. Find all possible values of $\mathbf{c}$ satisfying this condition, and the second eigenstate (orthogonal to α) of the operator $\mathbf{c}\cdot\hat{\boldsymbol{\sigma}}$. Give a Bloch-sphere interpretation of your result.

5.4. Rewrite the key formulas of the solutions of Problems 4.27-4.29 in terms of the Bloch sphere angles, and verify at least one of them using the general relations of Sec. 5.1 of the lecture notes.

5.5. A spin-½ with a gyromagnetic ratio $\gamma > 0$ was placed into a time-independent magnetic field $\mathscr{B}_0 = \mathscr{B}_0\mathbf{n}_z$ and let relax into the lowest-energy state. At $t = 0$, an additional field $\mathscr{B}_1(t)$ is turned on; its vector has a constant magnitude but rotates within the $[x, y]$-plane with an angular velocity ω. Calculate the expectation values of all Cartesian components of the spin at $t \geq 0$, and discuss the representation of its dynamics on the Bloch sphere.

5.6.* Analyze statistics of the spacing $S \equiv E_+ - E_-$ between energy levels of a two-level system, assuming that all elements $H_{jj'}$ of its Hamiltonian matrix (2) are independent random numbers, with equal and constant probability densities within the energy interval of interest. Compare the result with that for a purely diagonal Hamiltonian matrix, with a similar probability distribution of its random diagonal elements.

5.7. For a periodic motion of a single particle in a confining potential $U(\mathbf{r})$, the *virial theorem* of non-relativistic classical mechanics[53] is reduced to the following equality:

[52] For the interested reader, I can recommend either Sec. 17.7 in E. Merzbacher, *Quantum Mechanics*, 3rd ed., Wiley, 1998, or Sec. 3.10 in J. Sakurai, *Modern Quantum Mechanics*, Addison-Wesley, 1994.
[53] See, e.g., CM Problem 1.12.

$$\overline{T} = \frac{1}{2}\overline{\mathbf{r}\cdot\nabla U},$$

where T is the particle's kinetic energy, and the top bar means averaging over the time period of motion. Prove the following quantum-mechanical version of the theorem for an arbitrary stationary state, in the absence of spin effects:

$$\langle T\rangle = \frac{1}{2}\langle \mathbf{r}\cdot\nabla U\rangle,$$

where the angular brackets denote (as usual in this course) the expectation values of the observables.

Hint: Mimicking the proof of the classical virial theorem, consider the time evolution of the following operator: $\hat{G} \equiv \hat{\mathbf{r}}\cdot\hat{\mathbf{p}}$.

5.8. A non-relativistic 1D particle moves in the spherically symmetric potential $U(r) = C\ln(r/R)$. Prove that for:

(i) $\langle v^2\rangle$ is the same in each eigenstate, and
(ii) the spacing between the energy levels is independent of the particle's mass.

5.9. Calculate, in the WKB approximation, the transparency $\mathscr{T}$ of the following saddle-shaped potential barrier:

$$U(x,y) = U_0\left(1 + \frac{xy}{a^2}\right),$$

where $U_0 > 0$ and a are real constants, for tunneling of a 2D particle with energy $E < U_0$.

5.10. In the WKB approximation, calculate the so-called *Gamow factor*[54] for the alpha decay of atomic nuclei, i.e. the exponential factor in the transparency of the potential barrier resulting from the following simple model for the alpha-particle's potential energy as a function of its distance from the nuclear center:

$$U(r) = \begin{cases} U_0 < 0, & \text{for } r < R, \\ \dfrac{ZZ'e^2}{4\pi\varepsilon_0 r}, & \text{for } R < r, \end{cases}$$

where $Ze = 2e > 0$ is the charge of the particle, $Z'e > 0$ is that of the nucleus after the decay, and R is the nucleus' radius.

5.11. Use the WKB approximation to calculate the average time of ionization of a hydrogen atom, initially in its ground state, made metastable by the application of an additional weak, uniform, time-independent electric field $\mathscr{E}$. Formulate the conditions of validity of your result.

5.12. For a 1D harmonic oscillator with mass m and frequency ω_0, calculate:

(i) all matrix elements $\langle n|\hat{x}^3|n'\rangle$, and

(ii) the diagonal matrix elements $\langle n|\hat{x}^4|n\rangle$,

where n and n' are arbitrary Fock states.

[54] Named after G. Gamow, who made this calculation as early as in 1928.

5.13. Calculate the sum (over all $n > 0$) of the so-called *oscillator strengths*,

$$f_n \equiv \frac{2m}{\hbar^2}(E_n - E_0)\left|\langle n|\hat{x}|0\rangle\right|^2,$$

(i) for a 1D harmonic oscillator, and
(ii) for a 1D particle confined in an arbitrary stationary potential.[55]

5.14. Prove the so-called *Bethe sum rule*,

$$\sum_{n'}(E_{n'} - E_n)\left|\langle n|e^{ik\hat{x}}|n'\rangle\right|^2 = \frac{\hbar^2 k^2}{2m}$$

(where k is any *c*-number constant), valid for a 1D particle moving in an arbitrary time-independent potential $U(x)$, and discuss its relation with the Thomas-Reiche-Kuhn sum rule whose derivation was the subject of the previous problem.

Hint: Calculate the expectation value, in a stationary state n, of the double commutator

$$\hat{D} \equiv \left[\left[\hat{H}, e^{ik\hat{x}}\right], e^{-ik\hat{x}}\right]$$

in two ways: first, just by spelling out both commutators, and, second, by using the commutation relations between operators $\hat{p}_x$ and $e^{ik\hat{x}}$, and compare the results.

5.15. Spell out the commutator $\left[\hat{a}, \exp\left\{\lambda\hat{a}^\dagger\right\}\right]$, where $\hat{a}^\dagger$ and $\hat{a}$ are the creation-annihilation operators (5.65), and λ is a *c*-number.

5.16. Given Eq. (116), prove Eq. (117) by using the hint given in the accompanying note.

5.17. Use Eqs. (116)-(117) to simplify the following operators:

(i) $\exp\{+ia\hat{x}\}\,\hat{p}_x \exp\{-ia\hat{x}\}$, and
(ii) $\exp\{+ia\hat{p}_x\}\,\hat{x} \exp\{-ia\hat{p}_x\}$,

where a is a *c*-number.

5.18.* Derive the commutation relation between the number operator (5.73) and a reasonably defined quantum-mechanical operator describing the harmonic oscillator's phase φ. Obtain the uncertainty relation for the corresponding observables, and explore its limit at $N >> 1$.

5.19. At $t = 0$, a 1D harmonic oscillator was in a state described by the ket-vector

$$|\alpha\rangle = \frac{1}{\sqrt{2}}\left(|31\rangle + |32\rangle\right),$$

where $|n\rangle$ are the ket-vectors of the stationary (Fock) states of the oscillator. Calculate:

(i) the expectation value of the oscillator's energy, and

[55] This *Thomas-Reiche-Kuhn sum rule* is important for applications because the coefficients f_n describe, in particular, the intensity of dipole quantum transitions between the n^{th} energy level and the ground state – see, e.g., Sec. 9.2 and also EM Sec. 7.2.

(ii) the time evolution of the expectation values of its coordinate and momentum.

5.20.* Re-derive the London dispersion force's potential of the interaction of two isotropic 3D harmonic oscillators (already calculated in Problem 3.20), using the language of mutually-induced polarization.

5.21. An external force pulse $F(t)$, of a finite time duration $\mathcal{T}$, is exerted on a 1D harmonic oscillator, initially in its ground state. Use the Heisenberg-picture equations of motion to calculate:

(i) the expectation values of the oscillator's coordinate and momentum and their uncertainties, at an arbitrary moment,

(ii) its total energy after the end of the pulse.

5.22. Use Eqs. (144)-(145) to calculate the uncertainties δx and δp for a harmonic oscillator in its squeezed ground state, and in particular, to prove Eqs. (143) for the case $\theta = 0$.

5.23. Calculate the energy of a harmonic oscillator in the squeezed ground state ζ.

5.24.* Prove that the squeezed ground state described by Eqs. (142) and (144)-(145) may be sustained by a sinusoidal modulation of a harmonic oscillator's parameter, and calculate the squeezing factor r as a function of the parameter modulation depth, assuming that the depth is small and the oscillator's damping is negligible.

5.25. Use Eqs. (148) to prove that at negligible spin effects, the operators $\hat{L}_j$ and $\hat{L}^2$ commute with the Hamiltonian of a particle placed in any central potential field.

5.26. Use Eqs. (149)-(150) and (153) to prove Eqs. (155).

5.27. Derive Eq. (164) by using any of the prior formulas.

5.28. Derive the expression $\langle L^2 \rangle = \hbar^2 l(l+1)$ from basic statistics, by assuming that all $(2l+1)$ values $L_z = \hbar m$ of a system with a fixed integer number l have equal probability, and that the system is isotropic. Explain why this statistical picture cannot be used for proof of Eq. (5.163).

5.29. In the basis of common eigenstates of the operators $\hat{L}_z$ and $\hat{L}^2$, described by kets $|l, m\rangle$:

(i) calculate the matrix elements $\left\langle l, m_1 \middle| \hat{L}_x \middle| l, m_2 \right\rangle$ and $\left\langle l, m_1 \middle| \hat{L}_x^2 \middle| l, m_2 \right\rangle$,

(ii) spell out your results for diagonal matrix elements (with $m_1 = m_2$) and their y-axis counterparts, and

(iii) calculate the diagonal matrix elements $\left\langle l, m \middle| \hat{L}_x \hat{L}_y \middle| l, m \right\rangle$ and $\left\langle l, m \middle| \hat{L}_y \hat{L}_x \middle| l, m \right\rangle$.

5.30. For the state described by the common eigenket $|l, m\rangle$ of the operators $\hat{L}_z$ and $\hat{L}^2$ in a reference frame $\{x, y, z\}$, calculate the expectation values $\langle L_{z'} \rangle$ and $\langle L_{z'}{}^2 \rangle$ in the reference frame whose z'-axis forms angle θ with the z-axis.

5.31. Write down the matrices of the following angular momentum operators: $\hat{L}_x, \hat{L}_y, \hat{L}_z$, and $\hat{L}_\pm$, in the z-basis of the $\{l, m\}$ states with $l = 1$.

5.32. Calculate the angular factor of the orbital wavefunction of a particle with a definite value of L^2, equal to $6\hbar^2$, and the largest possible value of L_x. What is this value?

5.33. For the state with the wavefunction $\psi = Cxye^{-\lambda r}$, with a real positive λ, calculate:

(i) the expectation values of the observables L_x, L_y, L_z, and L^2, and
(ii) the normalization constant C.

5.34. An angular state of a spinless particle is described by the following ket-vector:

$$|\alpha\rangle = \frac{1}{\sqrt{2}}\Big(|l = 3, m = 0\rangle + |l = 3, m = 1\rangle\Big).$$

Calculate the expectation values of the x- and y-components of its angular momentum. Is the result sensitive to a possible phase shift between the component eigenkets?

5.35. A particle is in a quantum state α with the orbital wavefunction proportional to the spherical harmonic $Y_1^1(\theta,\varphi)$. Find the angular dependence of the wavefunctions corresponding to the following ket-vectors:

(i) $\hat{L}_x|\alpha\rangle$, (ii) $\hat{L}_y|\alpha\rangle$, (iii) $\hat{L}_z|\alpha\rangle$, (iv) $\hat{L}_+\hat{L}_-|\alpha\rangle$, and (v) $\hat{L}^2|\alpha\rangle$.

5.36. A charged, spinless 2D particle of mass m is trapped in the potential well $U(x, y) = m\omega_0^2(x^2 + y^2)/2$. Calculate its energy spectrum in the presence of a uniform magnetic field $\mathscr{B}$ normal to the $[x, y]$-plane of the particle's motion.

5.37. Solve the previous problem for a spinless 3D particle, placed (in addition to a uniform magnetic field $\mathscr{B}$) into a spherically-symmetric potential well $U(\mathbf{r}) = m\omega_0^2 r^2/2$.

5.38. Calculate the spectrum of rotational energies of an axially symmetric rigid macroscopic body.

5.39. Simplify the double commutator $\Big[\hat{r}_j, \Big[\hat{L}^2, \hat{r}_{j'}\Big]\Big]$.

5.40. Prove the following commutation relation:

$$\Big[\hat{L}^2, \Big[\hat{L}^2, \hat{r}_j\Big]\Big] = 2\hbar^2\Big(\hat{r}_j\hat{L}^2 + \hat{L}^2\hat{r}_j\Big).$$

5.41. Use the commutation relation proved in the previous problem and Eq. (148) to prove the orbital electric-dipole transition selection rules mentioned in Sec. 6.

5.42. Express the commutators listed in Eq. (179), $\Big[\hat{J}^2, \hat{L}_z\Big]$ and $\Big[\hat{J}^2, \hat{S}_z\Big]$, via $\hat{L}_j$ and $\hat{S}_j$.

5.43. Find the operator $\hat{\mathcal{T}}_\phi$ describing a quantum state's rotation by angle ϕ about a certain axis, by using the similarity of this operation with the shift of a Cartesian coordinate, discussed in Sec. 5. Then use this operator to calculate the probabilities of measurements of spin-½ components of particles with *z*-polarized spin, by a Stern-Gerlach instrument turned by angle θ within the [*z*, *x*] plane, where *y* is the axis of particle propagation – see Fig. 4.1.[56]

5.44. The rotation operator $\hat{\mathcal{T}}_\phi$ analyzed in the previous problem and the linear translation operator $\hat{\mathcal{T}}_X$ discussed in Sec. 5 have a similar structure:

$$\hat{\mathcal{T}}_\lambda = \exp\left\{-i\hat{C}\lambda/\hbar\right\},$$

where λ is a real *c*-number scaling the shift and $\hat{C}$ is a Hermitian operator that does not explicitly depend on time.

(i) Prove that such operators are unitary.

(ii) Prove that if the shift by λ, induced by the operator $\hat{\mathcal{T}}_\lambda$, leaves the Hamiltonian of some system unchanged for any λ, then $\langle C\rangle$ is a constant of motion for any initial state of the system.

(iii) Discuss what the last conclusion means for the particular operators $\hat{\mathcal{T}}_X$ and $\hat{\mathcal{T}}_\phi$.

5.45. A particle with spin *s* is in a state with definite quantum numbers *l* and *j*. Prove that the observable $\mathbf{L}\cdot\mathbf{S}$ also has a definite value and calculate it.

5.46. For a spin-½ particle in a state with definite quantum numbers *l*, m_l, and m_s, calculate the expectation value of the observable J^2 and the probabilities of all its possible values. Interpret your results in terms of the Clebsch-Gordan coefficients (190).

5.47. Derive general recurrence relations for the Clebsch-Gordan coefficients for a particle with spin *s*.

Hint: By using the similarity of the commutation relations discussed in Sec. 7, write the relations similar to Eqs. (164) for other components of the angular momentum, and then apply them to Eq. (170).

5.48. Use the recurrence relations derived in the previous problem to prove Eqs. (190) for the spin-½ Clebsch-Gordan coefficients.

5.49. A spin-½ particle is in a state with definite values of L^2, J^2, and J_z. Find all possible values of the observables S^2, S_z, and L_z, the probability of each listed value, and the expectation value for each of these observables.

5.50. Re-solve the Landau-level problem discussed in Sec. 3.2, now for a spin-½ particle. Discuss the result for the particular case of an electron.

[56] Note that the last task is just a particular case of Problem 4.18 (see also Problem 1).

5.51. In the Heisenberg picture of quantum dynamics, find an explicit relation between the operators of velocity $\hat{\mathbf{v}} \equiv d\hat{\mathbf{r}}/dt$ and acceleration $\hat{\mathbf{a}} \equiv d\hat{\mathbf{v}}/dt$ of a nonrelativistic particle with an electric charge q, moving in an arbitrary external electromagnetic field. Compare the result with the corresponding classical expression.

Hint: For the orbital motion's description, you may use Eq. (3.26).

5.52. One byproduct of the solution of Problem 47 was the following relation for the spin operators (valid for any spin s):

$$\langle m_s \pm 1 | \hat{S}_\pm | m_s \rangle = \hbar \left[(s \pm m_s + 1)(s \mp m_s) \right]^{1/2}.$$

Use this result to spell out the matrices S_x, S_y, S_z, and S^2 of a particle with $s = 1$, in the z-basis – defined as the basis in which the matrix S_z is diagonal.

5.53.* For a particle with an arbitrary spin s, find the ranges of the quantum numbers m_j and j that are necessary to describe, in the coupled-representation basis:

(i) all states with a definite quantum number l, and
(ii) a state with definite values of not only l but also m_l and m_s.

Give an interpretation of your results in terms of the classical vector diagram – see, e.g., Fig. 13.

5.54. For a particle with spin s, find the range of the quantum numbers j necessary to describe, in the coupled-representation basis, all states with definite quantum numbers l and m_l.

5.55. A particle of mass m, with electric charge q and spin s, free to move along a planar circle of a radius R, is placed into a constant uniform magnetic field $\mathscr{B}$ directed normally to the circle's plane. Calculate the energy spectrum of the system. Explore and interpret the particular form the result takes when the particle is an electron with the g-factor $g_e \approx 2$.

Chapter 6. Perturbative Approaches

This chapter discusses several perturbative approaches to problems of quantum mechanics, and their simplest but important applications starting with the fine structure of atomic energy levels, and the effects of external dc and ac electric and magnetic fields on these levels. It continues with a discussion of quantum transitions to continuous spectrum and the Golden Rule of quantum mechanics, which naturally brings us to the issue of open quantum systems – to be discussed in the next chapter.

6.1. Time-independent perturbations

Unfortunately, only a few problems of quantum mechanics may be solved exactly in an analytical form. Actually, in the previous chapters we have solved a substantial part of such problems for a single particle, while for multiparticle systems, the exactly solvable cases are even more rare. However, most practical problems of physics feature a certain small parameter, and this smallness may be exploited by various approximate analytical methods giving *asymptotically correct* results – i.e. the results whose error tends to zero at the reduction of the small parameter(s). Earlier in the course, we explored one of them, the WKB approximation, which is adequate for a particle moving through a soft potential profile. In this chapter, we will discuss other techniques that are more suitable for other cases. The historical name for these techniques is *the perturbation theory*, but it is fairer to speak about *perturbative approaches* because they are substantially different for different situations.

The simplest version of the perturbation theory addresses the problem of stationary states and energy levels of systems described by time-independent Hamiltonians of the type

$$\hat{H} = \hat{H}^{(0)} + \hat{H}^{(1)}, \tag{6.1}$$

where the operator $\hat{H}^{(1)}$, describing the system's "perturbation", is relatively small – in the sense that its addition to the unperturbed operator $\hat{H}^{(0)}$ results in a relatively small change of the eigenenergies E_n and the corresponding eigenstates of the system. A typical problem of this type is the 1D *weakly anharmonic oscillator* (Fig. 1), described by the Hamiltonian (1) with

Weakly anharmonic oscillator

$$\hat{H}^{(0)} = \frac{\hat{p}^2}{2m} + \frac{m\omega_0^2 \hat{x}^2}{2}, \qquad \hat{H}^{(1)} = \alpha \hat{x}^3 + \beta \hat{x}^4 + \ldots \tag{6.2}$$

with sufficiently small coefficients α, β, ….

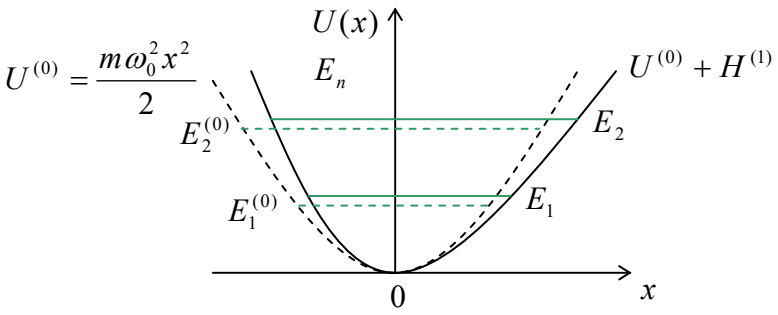


Fig. 6.1. The simplest application of the perturbation theory: a weakly anharmonic 1D oscillator. (Dashed lines characterize the unperturbed harmonic oscillator.)

I will use this system as our first example, but let me start by describing the perturbative approach to the general time-independent Hamiltonian (1). In the bra-ket formalism, the eigenproblem (4.68) for the perturbed Hamiltonian, i.e. the stationary Schrödinger equation of the system, is

$$\left(\hat{H}^{(0)} + \hat{H}^{(1)}\right)\left|n\right\rangle = E_n\left|n\right\rangle. \tag{6.3}$$

Let the eigenstates and eigenvalues of the unperturbed Hamiltonian, which satisfy the equation

$$\hat{H}^{(0)}\left|n^{(0)}\right\rangle = E_n^{(0)}\left|n^{(0)}\right\rangle, \tag{6.4}$$

be considered as known. In this case, the solution of problem (3) means finding, first, its perturbed eigenvalues E_n and, second, the coefficients $\langle n'^{(0)}|n\rangle$ of the expansion of the perturbed state's vectors $|n\rangle$ in the following series over the unperturbed ones, $|n'^{(0)}\rangle$:

$$\left|n\right\rangle = \sum_{n'}\left|n'^{(0)}\right\rangle\left\langle n'^{(0)}\right|n\rangle. \tag{6.5}$$

Let us plug Eq. (5), with the summation index n' replaced with n'' (just to have a more compact notation in our forthcoming result), into both sides of Eq. (3):

$$\sum_{n''}\left\langle n''^{(0)}\right|n\rangle\hat{H}^{(0)}\left|n''^{(0)}\right\rangle + \sum_{n''}\left\langle n''^{(0)}\right|n\rangle\hat{H}^{(1)}\left|n''^{(0)}\right\rangle = \sum_{n''}\left\langle n''^{(0)}\right|n\rangle E_n\left|n''^{(0)}\right\rangle. \tag{6.6}$$

and then inner-multiply all terms by an arbitrary unperturbed bra-vector $\langle n'^{(0)}|$ of the system. Assuming that the unperturbed eigenstates are orthonormal, $\langle n'^{(0)}|n''^{(0)}\rangle = \delta_{n'n''}$, and using Eq. (4) in the first term on the left-hand side, we get the following system of linear equations

$$\sum_{n''}\left\langle n''^{(0)}\right|n\rangle H^{(1)}_{n'n''} = \left\langle n'^{(0)}\right|n\rangle\left(E_n - E^{(0)}_{n'}\right), \tag{6.7}$$

where the matrix elements of the perturbation are calculated, by definition, in the *unperturbed* brackets:

$$H^{(1)}_{n'n''} \equiv \left\langle n'^{(0)}\right|\hat{H}^{(1)}\left|n''^{(0)}\right\rangle. \tag{6.8}$$

Perturbation's matrix elements

The linear equation system (7) is still exact,[1] and is frequently used for numerical calculations. (Since the matrix coefficients (8) typically decrease when n' and/or n'' become sufficiently large, the sum on the left-hand side of Eq. (7) may usually be truncated, still giving an acceptable accuracy of the solution.) To get analytical results, we need to make approximations. In the simple perturbation theory we are discussing now, this is achieved by the expansion of both the eigenenergies and the expansion coefficients into the Taylor series in a certain small parameter μ of the problem:

$$E_n = E_n^{(0)} + E_n^{(1)} + E_n^{(2)}..., \tag{6.9}$$

$$\left\langle n'^{(0)}\right|n\rangle = \left\langle n'^{(0)}\right|n\rangle^{(0)} + \left\langle n'^{(0)}\right|n\rangle^{(1)} + \left\langle n''^{(0)}\right|n\rangle^{(2)}..., \tag{6.10}$$

where

$$E_n^{(k)} \propto \left\langle n'^{(0)}\right|n\rangle^{(k)} \propto \mu^k. \tag{6.11}$$

[1] Please note the similarity of Eq. (7) to Eq. (2.215) of the 1D band theory. Indeed, the latter equation is just a particular form of Eq. (7) for the 1D wave mechanics, with a specific (periodic) potential $U(x)$ considered as the perturbation Hamiltonian. Moreover, the whole approximate treatment of the weak-potential limit in Sec. 2.7 was essentially a particular case of the perturbation theory we are discussing now (in its 1st order).

In order to explore the 1st-order approximation, which ignores all terms $O(\mu^2)$ and higher, let us plug only the two first terms of the expansions (9) and (10) into the basic equation (7):

$$\sum_{n''} H^{(1)}_{n'n''}\left(\delta_{n''n} + \left\langle n''^{(0)}\middle|n\right\rangle^{(1)}\right) = \left(\delta_{n'n} + \left\langle n'^{(0)}\middle|n\right\rangle^{(1)}\right)\left(E^{(0)}_n + E^{(1)}_n - E^{(0)}_{n'}\right). \tag{6.12}$$

Now let us open the parentheses, and disregard all the remaining terms $O(\mu^2)$. The result is

$$H^{(1)}_{n'n} = \delta_{n'n}E^{(1)}_n + \left\langle n'^{(0)}\middle|n\right\rangle^{(1)}(E^{(0)}_n - E^{(0)}_{n'}), \tag{6.13}$$

This relation is valid for any choice of the indices n and n'; let us start from the case $n = n'$, immediately getting a very simple (and practically, the most important!) result:

Energy: 1st-order correction

$$E^{(1)}_n = H^{(1)}_{nn} \equiv \left\langle n^{(0)}\middle|\hat{H}^{(1)}\middle|n^{(0)}\right\rangle. \tag{6.14}$$

For example, let us see what this result gives for two first perturbation terms in the weakly anharmonic oscillator (2):

$$E^{(1)}_n = \alpha\left\langle n^{(0)}\middle|\hat{x}^3\middle|n^{(0)}\right\rangle + \beta\left\langle n^{(0)}\middle|\hat{x}^4\middle|n^{(0)}\right\rangle. \tag{6.15}$$

As the reader knows (or should know :-) from the solution of Problem 5.12, the first bracket equals zero, while the second one yields

$$E^{(1)}_n = \frac{3}{4}\beta x_0^4\left(2n^2 + 2n + 1\right). \tag{6.16}$$

Naturally, there should be some non-vanishing contribution to the energies from the (typically, larger) perturbation proportional to α, so for its calculation, we need to explore the 2nd order of the theory. However, before doing that, let us complete our discussion of its 1st order.

For $n' \neq n$, Eq. (13) may be used to calculate the eigenstates rather than the eigenvalues:

$$\left\langle n'^{(0)}\middle|n\right\rangle^{(1)} = \frac{H^{(1)}_{n'n}}{E^{(0)}_n - E^{(0)}_{n'}}, \qquad \text{for } n' \neq n. \tag{6.17}$$

This means that the eigenket's expansion (5), in the 1st order, may be represented as

States: 1st-order result

$$\left|n^{(1)}\right\rangle = C\left|n^{(0)}\right\rangle + \sum_{n'\neq n}\frac{H^{(1)}_{n'n}}{E^{(0)}_n - E^{(0)}_{n'}}\left|n'^{(0)}\right\rangle. \tag{6.18}$$

The coefficient $C \equiv \langle n^{(0)}|n^{(1)}\rangle$ cannot be found from Eq. (17); however, requiring the final state n to be normalized, we see that other terms may provide only corrections $O(\mu^2)$, so in the 1st order we should take $C = 1$. The most important feature of Eq. (18) is its denominators: the closer the unperturbed eigenenergies of two states, the larger their mutual "interaction" due to the perturbation.

This feature also affects the 1st-order approximation's validity condition, which may be quantified using Eq. (17): the magnitudes of the brackets it describes have to be much less than the unperturbed bracket $\langle n|n\rangle^{(0)} = 1$, so all elements of the perturbation matrix have to be much less than the difference between the corresponding unperturbed energies. For the anharmonic oscillator's energy corrections (16), this requirement is reduced to $E_n^{(1)} << \hbar\omega_0$.

Now we are ready to go after the 2nd-order approximation to Eq. (7). Let us focus on the case $n' = n$, because as we already know, only this term will give us a correction to the eigenenergies. Moreover, since the left-hand side of Eq. (7) already has a small factor $H^{(1)}_{n'n''} \propto \mu$, the bracket coefficients in that part may be taken from the 1st-order result (17). As a result, we get

$$E_n^{(2)} = \sum_{n''} \left\langle n''^{(0)} \middle| n \right\rangle^{(1)} H^{(1)}_{nn''} = \sum_{n'' \neq n} \frac{H^{(1)}_{n''n} H^{(1)}_{nn''}}{E_n^{(0)} - E_{n''}^{(0)}}. \tag{6.19}$$

Since $\hat{H}^{(1)}$ has to be Hermitian, we may rewrite this expression as

$$E_n^{(2)} = \sum_{n' \neq n} \frac{\left| H^{(1)}_{n'n} \right|^2}{E_n^{(0)} - E_{n'}^{(0)}} \equiv \sum_{n' \neq n} \frac{\left| \left\langle n'^{(0)} \middle| \hat{H}^{(1)} \middle| n^{(0)} \right\rangle \right|^2}{E_n^{(0)} - E_{n'}^{(0)}}. \tag{6.20}$$

Energy: 2nd-order correction

This is the much-celebrated 2nd-order perturbation result, which frequently (in sufficiently symmetric problems) is the *first non-vanishing* correction to the state energy – for example, from the cubic term (proportional to α) in our weakly anharmonic oscillator problem (2). To calculate the corresponding correction, we may use another result of the solution of Problem 5.12:

$$\left\langle n' \middle| \hat{x}^3 \middle| n \right\rangle = \left(\frac{x_0}{\sqrt{2}} \right)^3 \times \left\{ \left[n(n-1)(n-2) \right]^{1/2} \delta_{n',n-3} + 3n^{3/2} \delta_{n',n-1} + 3(n+1)^{3/2} \delta_{n',n+1} + \left[(n+1)(n+2)(n+3) \right]^{1/2} \delta_{n',n+3} \right\}. \tag{6.21}$$

So, according to Eq. (20), we need to calculate

$$E_n^{(2)} = \alpha^2 \left(\frac{x_0}{\sqrt{2}} \right)^6 \times \sum_{n' \neq n} \frac{\left\{ \left[n(n-1)(n-2) \right]^{1/2} \delta_{n',n-3} + 3n^{3/2} \delta_{n',n-1} + 3(n+1)^{3/2} \delta_{n',n+1} + \left[(n+1)(n+2)(n+3) \right]^{1/2} \delta_{n',n+3} \right\}^2}{\hbar \omega_0 (n - n')}. \tag{6.22}$$

The summation is not as cumbersome as may look because, at the curly bracket's squaring, all mixed products are proportional to the products of different Kronecker deltas and hence vanish, so we need to sum up only the squares of each term, finally getting

$$E_n^{(2)} = -\frac{15}{4} \frac{\alpha^2 x_0^6}{\hbar \omega_0} \left(n^2 + n + \frac{11}{30} \right). \tag{6.23}$$

This formula shows that all 2nd-order energy level corrections are negative, regardless of the sign of α.[2] On the contrary, the 1st-order correction $E_n^{(1)}$ given by Eq. (16), does depend on the sign of β, so the net correction, $E_n^{(1)} + E_n^{(2)}$, may be of any sign.

The results (18) and (20) are clearly inapplicable to the degenerate case where, in the absence of perturbation, several states correspond to the same energy level, because of the divergence of their denominators.[3] This divergence hints that in this case, the largest effect of the perturbation is the

[2] Note that this is correct for the ground-state energy correction $E_g^{(2)}$ of *any* system, because for this state, the denominators of all terms of the sum (20) are negative, while their numerators are always non-negative.

[3] This is exactly the reason why this simple perturbation approach runs into serious problems for systems with a continuous spectrum, and other techniques (such as the WKB approximation) are often necessary.

degeneracy lifting, e.g., some splitting of the initially degenerate energy level $E^{(0)}$ (Fig. 2), and that for the analysis of this case, we can, in the first approximation, ignore the effect of all other energy levels. (A more detailed analysis shows that this is indeed the case until the level splitting becomes comparable with the distance to other energy levels.)

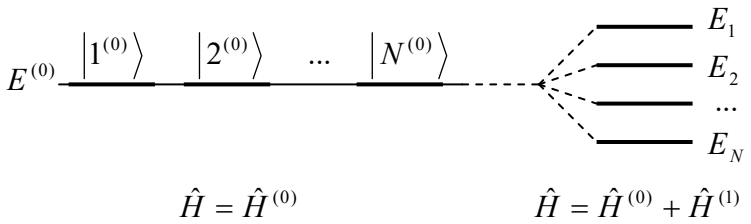


Fig. 5.2. Lifting the energy level degeneracy by a perturbation (schematically).

Limiting the summation in Eq. (7) to a group of N degenerate states with equal $E_{n'}^{(0)} \equiv E^{(0)}$, we reduce it to

$$\sum_{n''=1}^{N} \langle n''^{(0)} | n \rangle H^{(1)}_{n'n''} = \langle n'^{(0)} | n \rangle \left(E_n - E^{(0)} \right), \tag{6.24}$$

where now the indices n' and n'' number the N states of the group.[4] For $n = n'$, Eq. (24) may be rewritten as

$$\sum_{n''=1}^{N} \left(H^{(1)}_{n'n''} - E^{(1)}_{n''} \delta_{n'n''} \right) \langle n''^{(0)} | n' \rangle = 0, \quad \text{where } E^{(1)}_n \equiv E_n - E^{(0)}. \tag{6.25}$$

For each $n' = 1, 2, \ldots N$, this is a system of N linear, homogenous equations (with N terms each) for N unknown coefficients $\langle n''^{(0)} | n' \rangle$. In this problem, we may readily recognize the problem of diagonalization of the perturbation matrix $\mathrm{H}^{(1)}$ – cf. Sec. 4.4 and in particular Eq. (4.101). As in the general case, the condition of self-consistency of the system is:

Initially degenerate system: energy levels

$$\begin{vmatrix} H^{(1)}_{11} - E^{(1)}_n & H^{(1)}_{12} & \ldots \\ H^{(1)}_{21} & H^{(1)}_{22} - E^{(1)}_n & \ldots \\ \ldots & \ldots & \ldots \end{vmatrix} = 0, \tag{6.26}$$

where now the index n numbers the N roots of this equation, in arbitrary order. According to the definition (25) of $E_n^{(1)}$, the resulting N energy levels E_n may be found as $E^{(0)} + E_n^{(1)}$. If the perturbation matrix is diagonal in the chosen basis $n^{(0)}$, the result is extremely simple,

$$E_n - E^{(0)} \equiv E^{(1)}_n = H^{(1)}_{nn}, \tag{6.27}$$

and formally coincides with Eq. (14) for the non-degenerate case, but now it may give a different result for each of N previously degenerate states n.

[4] Note that here the choice of the basis is to some extent arbitrary because due to the linearity of equations of quantum mechanics, any linear combination of the states $n''^{(0)}$ is also an eigenstate of the unperturbed Hamiltonian. However, for using Eq. (25), these combinations have to be orthonormal, as was supposed in the derivation of Eq. (7).

Now let us see what this general theory gives for several important examples. First of all, let us consider a system with just two degenerate states with energy sufficiently far from all other levels. Then, in the basis of these two degenerate states, the most general perturbation matrix is

$$\mathrm{H}^{(1)} = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} \tag{6.28}$$

This matrix coincides with the general matrix (5.2) of a two-level system. Hence, we come to the very important conclusion: for a weak perturbation, all properties of any double-degenerate system are identical to those of the genuine two-level systems, which were the subject of numerous discussions in Chapter 4 and again in Sec. 5.1. In particular, its eigenenergies are given by Eq. (5.6), and may be described by the level-anticrossing diagram shown in Fig. 5.1.

6.2. The linear Stark effect

As a more involved example of the level degeneracy lifting by a perturbation, let us discuss the *Stark effect*[5] – the atomic level splitting by an external electric field. Let us study this effect, in the linear approximation, for a hydrogen-like atom/ion.[6] Taking the direction of the external electric field $\mathscr{E}$ (which is practically always uniform on the atomic scale) for the z-axis, the perturbation may be represented by the following Hamiltonian:

$$\hat{H}^{(1)} = -F\hat{z} = -q\mathscr{E}\hat{z} = -q\mathscr{E}r\cos\theta\,. \tag{6.29}$$

(In the last form, the operator sign is dropped, because we will work in the coordinate representation.)

As you (should :-) remember, energy levels of a hydrogen-like atom/ion depend only on the principal quantum number n – see Eq. (3.201); hence all the states, besides the ground 1s state with $n = 1$ and $l = m = 0$, have some orbital degeneracy, which grows rapidly with n. Let us consider the lowest degenerate level with $n = 2$. Since, according to Eq. (3.203), $0 \le l \le n - 1$, at this level the orbital quantum number l may equal either 0 (one 2s state, with $m = 0$) or 1 (three 2p states, with $m = 0, \pm 1$). Due to this 4-fold degeneracy, $\mathrm{H}^{(1)}$ is a 4×4 matrix with 16 elements:

$$\mathrm{H}^{(1)} = \begin{matrix} \overbrace{m=0}^{l=0} & \overbrace{m=0\ m=+1\ m=-1}^{l=1} & & \\ \begin{pmatrix} H_{11} & H_{12} & H_{13} & H_{14} \\ H_{21} & H_{22} & H_{23} & H_{24} \\ H_{31} & H_{32} & H_{33} & H_{34} \\ H_{41} & H_{42} & H_{43} & H_{44} \end{pmatrix} & \begin{matrix} m=0, \\ m=0, \\ m=+1, \\ m=-1, \end{matrix} & \left.\begin{matrix} \\ \\ \\ \end{matrix}\right\} & \begin{matrix} l=0, \\ \\ l=1. \end{matrix} \end{matrix} \tag{6.30}$$

[5] This effect was discovered experimentally in 1913 by Johannes Stark and independently by Antonio Lo Surdo, so it is sometimes (and more fairly) called the "Stark – Lo Surdo effect". Sometimes this name is used with the qualifier "dc" to distinguish it from the *ac Stark* effect – the energy level shift under the effect of an ac field – see Sec. 5 below.

[6] An analysis of the *quadratic Stark effect* for the ground-state energy in the same system, changing with the field only as $\mathscr{E}^2$, is left for the reader's exercise.

However, there is no need to be scared. First, due to the Hermitian nature of the operator, only 10 of these 16 matrix elements (4 diagonal and 6 off-diagonal ones) may be substantially different from each other. Moreover, due to the high symmetry of the problem, there are a lot of zeros even among these elements. Indeed, let us have a look at the angular components Y_l^m of the corresponding wavefunctions, with $l = 0$ and $l = 1$, described by Eqs. (3.174)-(3.175). For the states with $m = \pm 1$, the azimuthal parts of wavefunctions are proportional to $\exp\{\pm i\varphi\}$; hence the off-diagonal elements H_{34} and H_{43} of the matrix (30), relating these functions, are proportional to

$$\oint d\Omega\, Y_1^{\pm *} \hat{H}^{(1)} Y_1^{\mp} \propto \int_0^{2\pi} d\varphi \left(e^{\pm i\varphi}\right)^* \left(e^{\mp i\varphi}\right) = 0. \tag{6.31}$$

The azimuthal-angle symmetry also kills the off-diagonal elements H_{13}, H_{14}, H_{23}, H_{24} (and hence their complex conjugates H_{31}, H_{41}, H_{32}, and H_{42}), because they relate states with $m = 0$ and $m = \pm 1$, and hence are proportional to

$$\oint d\Omega\, Y_1^{0*} \hat{H}^{(1)} Y_1^{\pm 1} \propto \int_0^{2\pi} d\varphi\, e^{\pm i\varphi} = 0. \tag{6.32}$$

For the diagonal matrix elements H_{33} and H_{44}, corresponding to $l = 1$ and $m = \pm 1$, the azimuthal-angle integrals do not vanish, but since the corresponding spherical harmonics depend on the polar angle as $\sin\theta$, these elements are proportional to

$$\oint d\Omega\; Y_1^{\pm 1 *} \hat{H}^{(1)} Y_1^{\pm 1} \propto \int_0^{\pi} \sin\theta d\theta\, \sin\theta\cos\theta\sin\theta = \int_{-1}^{+1} \cos\theta\left(1-\cos^2\theta\right) d(\cos\theta), \tag{6.33}$$

and hence are equal to zero – as any limit-symmetric integral of an odd function. Finally, for the states $2s$ and $2p$ with $m = 0$, the diagonal elements H_{11} and H_{22} are also killed by the polar-angle integration:

$$\oint d\Omega\, Y_0^{0*} \hat{H}^{(1)} Y_0^0 \propto \int_0^{\pi} \sin\theta d\theta \cos\theta = \int_{-1}^{1} \cos\theta\; d(\cos\theta) = 0\,, \tag{6.34}$$

$$\oint d\Omega\, Y_0^{1*} \hat{H}^{(1)} Y_0^1 \propto \int_0^{\pi} \sin\theta d\theta \cos^3\theta = \int_{-1}^{+1} \cos^3\theta\; d(\cos\theta) = 0. \tag{6.35}$$

Hence, the only non-zero elements of the matrix (30) are two off-diagonal elements H_{12} and H_{21}, which relate two states with the same $m = 0$, but different $l = \{0, 1\}$, because they are proportional to

$$\oint d\Omega\, Y_0^{0*} \cos\theta\, Y_1^0 = \frac{\sqrt{3}}{4\pi} \int_0^{2\pi} d\varphi \int_0^{\pi} \sin\theta d\theta\, \cos^2\theta = \frac{1}{\sqrt{3}} \neq 0. \tag{6.36}$$

What remains is to use Eqs. (3.209) for the radial parts of these functions to complete the calculation of those two matrix elements:

$$H_{12} = H_{21} = -\frac{q\mathscr{E}}{\sqrt{3}} \int_0^{\infty} r^2 dr\, \mathcal{R}_{2,0}(r) r \mathcal{R}_{2,1}(r)\,. \tag{6.37}$$

Due to the additive structure of the function $\mathcal{R}_{2,0}(r)$, the integral falls into a sum of two table integrals, both of the type MA Eq. (6.7d), finally giving

$$H_{12} = H_{21} = 3q\mathscr{E}r_0, \tag{6.38}$$

where r_0 is the spatial scale (3.192); for the hydrogen atom, it is just the Bohr radius r_B – see Eq. (1.10).

Thus, the perturbation matrix (30) is reduced to

$$\mathrm{H}^{(1)} = \begin{pmatrix} 0 & 3q\mathscr{E}r_0 & 0 & 0 \\ 3q\mathscr{E}r_0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{6.39}$$

so the condition (26) of self-consistency of the system (25),

$$\begin{vmatrix} -E_2^{(1)} & 3q\mathscr{E}r_0 & 0 & 0 \\ 3q\mathscr{E}r_0 & -E_2^{(1)} & 0 & 0 \\ 0 & 0 & -E_2^{(1)} & 0 \\ 0 & 0 & 0 & -E_2^{(1)} \end{vmatrix} = 0, \tag{6.40}$$

gives a very simple characteristic equation

$$\left(E_2^{(1)}\right)^2\left[\left(E_2^{(1)}\right)^2 - \left(3q\mathscr{E}r_0\right)^2\right] = 0\,. \tag{6.41}$$

with four roots:

$$\boxed{\left(E_2^{(1)}\right)_{1,2} = 0, \qquad \left(E_2^{(1)}\right)_{3,4} = \pm 3q\mathscr{E}r_0\,.} \tag{6.42}$$

Linear Stark effect for $n = 2$

so the degeneracy is only partly lifted – see the levels in Fig. 3.

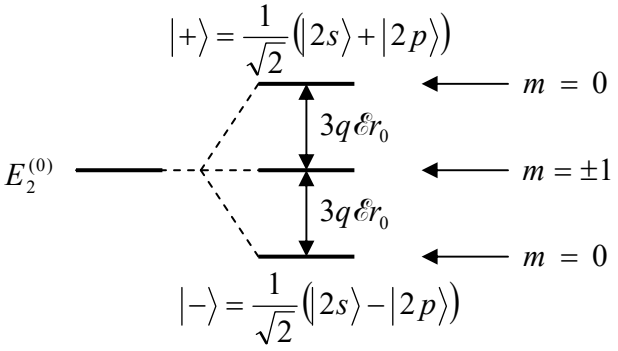


Fig. 6.3. The linear Stark effect for the level $n = 2$ of a hydrogen-like atom.

Generally, in order to understand the nature of states corresponding to these levels, we should return to Eq. (25) with each calculated value of $E_2^{(1)}$, and find the corresponding expansion coefficients $\langle n''^{(0)}|n'\rangle$ that describe the perturbed states. However, in our simple case, the outcome of this procedure is clear in advance. Indeed, since the states with $\{l = 1, m = \pm 1\}$ are not affected by the perturbation at all (in the linear approximation in the electric field), their degeneracy is not lifted, and energy is not affected – see the middle line in Fig. 3. On the other hand, the partial perturbation matrix connecting the states $2s$ and $2p$, i.e. the top left 2×2 part of the full matrix (39), is proportional to the Pauli matrix σ_x, and we already know the result of its diagonalization – see Eqs. (4.113)-(4.114). This means that the upper and lower split levels correspond to very simple linear combinations of the previously degenerate states with $m = 0$,

$$|\pm\rangle = \frac{1}{\sqrt{2}}\Big(|2s\rangle \pm |2p\rangle\Big). \tag{6.43}$$

Finally, let us estimate the magnitude of the linear Stark effect for a hydrogen atom. For a very high dc electric field of $\mathscr{E} = 3\times10^6$ V/m,[7] $|q| = e \approx 1.6\times10^{-19}$ C, and $r_0 = r_B \approx 0.5\times10^{-10}$ m, we get a level splitting of $3q\mathscr{E}r_0 \approx 0.8\times10^{-22}$ J ≈ 0.5 meV. This number is much lower than the unperturbed energy of the level, $E_2 = -E_H/(2\times2^2) \approx -3.4$ eV, so the perturbative result is quite applicable. On the other hand, the calculated splitting is much larger than the resolution limit imposed by the line's natural width ($\sim10^{-7}\,E_2$, see Chapter 9), so the effect is quite observable even in substantially lower electric fields. Note, however, that our simple results are quantitatively correct only when the Stark splitting (42) is much larger than the fine-structure splitting of the same level in the absence of the field– see the next section.

6.3. Fine structure of atomic levels

Now let us use the same perturbation theory to analyze, also for the simplest case of a hydrogen-like atom/ion, the so-called *fine structure* of atomic levels – their degeneracy lifting even in the absence of external fields. Since the effective speed v of the electron motion in atoms is much smaller than the speed of light c, the fine structure may be analyzed as a sum of two independent relativistic effects. To analyze the first of them, let us expand the well-known classical relativistic expression[8] for the kinetic energy $T = E - mc^2$ of a free particle with the rest mass m,[9]

$$T = \left(m^2c^4 + p^2c^2\right)^{1/2} - mc^2 \equiv mc^2\left[\left(1+\frac{p^2}{m^2c^2}\right)^{1/2} - 1\right], \tag{6.44}$$

into the Taylor series in the small ratio $(p/mc)^2 \approx (v/c)^2$:

$$T = mc^2\left[1+\frac{1}{2}\left(\frac{p}{mc}\right)^2 - \frac{1}{8}\left(\frac{p}{mc}\right)^4 + \ldots - 1\right] \equiv \frac{p^2}{2m} - \frac{p^4}{8m^3c^2} + \ldots, \tag{6.45}$$

and drop all the terms besides the two spelled-out ones. Of them, the first term is non-relativistic, while the second one represents the main relativistic correction to T.

Following the correspondence principle, the quantum-mechanical problem in this approximation may be described by Eq. (1) with the unperturbed Hamiltonian

$$\hat{H}^{(0)} = \frac{\hat{p}^2}{2m} + \hat{U}(r), \qquad \hat{U}(r) = -\frac{C}{r}, \tag{6.46}$$

(whose eigenstates and eigenenergies were discussed in Sec. 3.5) and the kinetic-relativistic perturbation

Kinetic-relativistic perturbation

$$\hat{H}^{(1)} = -\frac{\hat{p}^4}{8m^3c^2} \equiv -\frac{1}{2mc^2}\left(\frac{\hat{p}^2}{2m}\right)^2. \tag{6.47}$$

Using Eq. (46), we may rewrite the last formula as

[7] This value approximately corresponds to the threshold of electric breakdown in the air at ambient conditions, due to the impact ionization. As a result, experiments with higher dc fields are rather difficult.

[8] See, e.g., EM Eq. (9.78).

[9] This fancy font is used, as in Secs. 3.5-3.8, to distinguish the mass m from the magnetic quantum number m.

$$\hat{H}^{(1)} = -\frac{1}{2mc^2}\left(\hat{H}^{(0)} - \hat{U}(r)\right)^2, \tag{6.48}$$

so its matrix elements participating in the characteristic equation (25) for a given degenerate energy level (3.201), i.e. a given principal quantum number n, are

$$\langle nlm|\hat{H}^{(1)}|nl'm'\rangle = -\frac{1}{2mc^2}\langle nlm|\left(\hat{H}^{(0)} - \hat{U}(r)\right)\left(\hat{H}^{(0)} - \hat{U}(r)\right)|nl'm'\rangle, \tag{6.49}$$

where the bra- and ket-vectors describe the unperturbed eigenstates, whose eigenfunctions (in the coordinate representation) are given by Eq. (3.200): $\psi_{n,l,m} = \mathcal{R}_{n,l}(r)Y_l^m(\theta,\varphi)$.

It is straightforward (and hence left for the reader's exercise) to prove that all off-diagonal elements of the set (49) are equal to 0. Thus we may use Eq. (27) for each set of the quantum numbers $\{n, l, m\}$:

$$\begin{aligned} E^{(1)}_{n,l,m} \equiv E_{n,l,m} - E^{(0)}_n &= \langle nlm|\hat{H}^{(1)}|nlm\rangle = -\frac{1}{2mc^2}\left\langle\left(\hat{H}^{(0)} - \hat{U}(r)\right)^2\right\rangle_{n,l,m} \\ &= -\frac{1}{2mc^2}\left(E_n^2 - 2E_n\left\langle\hat{U}\right\rangle_{n,l} + \left\langle\hat{U}^2\right\rangle_{n,l}\right) = -\frac{1}{2mc^2}\left(\frac{E_0^2}{4n^4} - \frac{E_0}{n^2}C\left\langle\frac{1}{r}\right\rangle_{n,l} + C^2\left\langle\frac{1}{r^2}\right\rangle_{n,l}\right), \end{aligned} \tag{6.50}$$

where the index m has been dropped, because the radial wavefunctions $\mathcal{R}_{n,l}(r)$, which affect these expectation values, do not depend on that quantum number. Now using Eqs. (3.191), (3.201) and the first two of Eqs. (3.211), we finally get

$$E^{(1)}_{n,l} = -\frac{mC^2}{2\hbar^2c^2n^4}\left(\frac{n}{l+½} - \frac{3}{4}\right) \equiv -\frac{2E_n^2}{mc^2}\left(\frac{n}{l+½} - \frac{3}{4}\right). \tag{6.51}$$

Kinetic-relativistic energy correction

Let us discuss this result. First of all, its last form confirms that the correction (51) is indeed much smaller than the unperturbed energy E_n (and hence the perturbation theory is valid) if the latter is much smaller than the relativistic rest energy mc^2 of the particle – as it is for the hydrogen atom. Next, since in the Bohr problem's solution, $n \geq l + 1$, the first fraction in the parentheses of Eq. (51) is always larger than 1, and hence than ¾, so the kinetic relativistic correction to energy is negative for all n and l. (Actually, this fact could be predicted already from Eq. (47), which shows that the perturbation's Hamiltonian is a negatively defined form.) Finally, for a fixed principal number n, the negative correction's magnitude decreases with the growth of l. This fact may be interpreted using the second of Eqs. (3.211): the larger is l (at fixed n), the larger is the particle's effective distance from the center, and hence the smaller is its effective velocity, i.e. the smaller is the magnitude of the quantum-mechanical average of the negative relativistic correction (47) to the kinetic energy.

The result (51) is valid for the Coulomb interaction $U(r) = -C/r$ of any physical nature. However, if we speak specifically about hydrogen-like atoms/ions, there is also another relativistic correction to energy, due to the so-called *spin-orbit interaction* (alternatively called the "spin-orbit coupling"). Its physics may be understood from the following semi-quantitative classical reasoning: from the "the point of view" of an electron rotating about the nucleus at distance r with velocity $\mathbf{v}$, it is the nucleus, of the electric charge Ze, that rotates about the electron with the velocity $(-\mathbf{v})$ and hence the time period $\mathcal{T} = 2\pi r/v$. From the point of view of magnetostatics, such circular motion of the electric charge $Q = Ze$, is

equivalent to a circular dc electric current $I = Q/\mathcal{T} = (Ze)(v/2\pi r)$. At the electron's location, i.e. in the center of the current loop, it creates the magnetic field with the following magnitude:[10]

$$\mathscr{B}_a = \frac{\mu_0}{2r} I = \frac{\mu_0}{2r}\frac{Zev}{2\pi r} \equiv \frac{\mu_0 Zev}{4\pi r^2}. \tag{6.52}$$

The field's direction **n** is perpendicular to the apparent plane of the nucleus' rotation (i.e. that of the real rotation of the electron), and hence its vector may be readily expressed via the similarly directed vector $\mathbf{L} = m_e v r \mathbf{n}$ of the electron's angular (orbital) momentum:

$$\boldsymbol{\mathscr{B}}_a = \frac{\mu_0 Zev}{4\pi r^2}\mathbf{n} \equiv \frac{\mu_0 Ze}{4\pi r^3 m_e} m_e v r \mathbf{n} \equiv \frac{\mu_0 Ze}{4\pi r^3 m_e}\mathbf{L} \equiv \frac{Ze}{4\pi\varepsilon_0 r^3 m_e c^2}\mathbf{L}, \tag{6.53}$$

where the last step used the basic relation between the SI-unit constants: $\mu_0 \equiv 1/c^2\varepsilon_0$.

A more careful (but still classical) analysis of the problem[11] brings both good and bad news. The bad news is that the result (53) is wrong by the so-called *Thomas factor* of two even for the circular motion, because the electron moves with acceleration, and the reference frame bound to it cannot be inertial (as was implied in the above reasoning), so the effective magnetic field felt by the electron is actually

$$\boldsymbol{\mathscr{B}} = \frac{Ze}{8\pi\varepsilon_0 r^3 m_e c^2}\mathbf{L}. \tag{6.54}$$

The good news is that this result is valid not only for circular but an arbitrary orbital motion in the Coulomb field $U(r)$. Hence from the discussion in Sec. 4.1 and Sec. 4.4 we may expect that the quantum-mechanical description of the interaction between this effective magnetic field and the electron's spin moment (4.115) is given by the following perturbation Hamiltonian[12]

$$\hat{H}^{(1)} = -\hat{\mathbf{m}}\cdot\hat{\boldsymbol{\mathscr{B}}} = -\gamma_e\hat{\mathbf{S}}\cdot\left(\frac{Ze}{8\pi\varepsilon_0 r^3 m_e c^2}\hat{\mathbf{L}}\right) \equiv \frac{1}{2m_e^2c^2}\frac{Ze^2}{4\pi\varepsilon_0}\frac{1}{r^3}\hat{\mathbf{S}}\cdot\hat{\mathbf{L}}, \tag{6.55}$$

where at spelling out the electron's gyromagnetic ratio $\gamma_e \equiv -g_e e/2m_e$, the small correction to the value $g_e = 2$ of the electron's g-factor (see Sec. 4.4) is ignored, because Eq. (55) is already a small correction. This expectation is confirmed by the fully-relativistic Dirac theory, to be discussed in Sec. 9.7 below: it yields, for an arbitrary central potential $U(r)$, the following spin-orbit coupling Hamiltonian:

Spin-orbit coupling

$$\hat{H}^{(1)} = \frac{1}{2m_e^2c^2}\frac{1}{r}\frac{dU(r)}{dr}\hat{\mathbf{S}}\cdot\hat{\mathbf{L}}. \tag{6.56}$$

For the Coulomb potential $U(r) = -Ze^2/4\pi\varepsilon_0 r$, this formula is reduced to Eq. (55).

[10] See, e.g., EM Sec. 5.1, in particular, Eq. (5.24). Note that such an effective magnetic field is induced by any motion of electrons, in particular that in solids, leading to a variety of spin-orbit effects there – see, e.g., a concise review by R. Winkler *et al.*, in B. Kramer (ed.), *Advances in Solid State Physics* **41**, 211 (2001).

[11] It was carried out first by Llewellyn Thomas in 1926; for a simple review see, e.g., R. Harr and L. Curtis, *Am. J. Phys*. **55**, 1044 (1987).

[12] In the Gaussian units, Eq. (55) is valid without the factor $4\pi\varepsilon_0$ in the denominator; while Eq. (56), "as is".

As we already know from the discussion in Sec. 5.7, the angular factor of this Hamiltonian commutes with all the operators of the coupled-representation group (inside the blue line in Fig. 5.12): $\hat{L}^2$, $\hat{S}^2$, $\hat{J}^2$, and $\hat{J}_z$, and hence is diagonal in the coupled-representation basis with definite quantum numbers l, j, and m_j (and of course $s = ½$). Hence, using Eq. (5.181) to rewrite Eq. (56) as

$$\hat{H}^{(1)} = \frac{1}{2m_e^2c^2}\frac{Ze^2}{4\pi\varepsilon_0}\frac{1}{r^3}\frac{1}{2}\left(\hat{J}^2 - \hat{L}^2 - \hat{S}^2\right), \tag{6.57}$$

we may again use Eq. (27) for each set $\{s, l, j, m_j\}$, with common n:

$$E^{(1)}_{n,j,l} = \frac{1}{2m_e^2c^2}\frac{Ze^2}{4\pi\varepsilon_0}\left\langle\frac{1}{r^3}\right\rangle_{n,l}\frac{1}{2}\left\langle\hat{J}^2 - \hat{L}^2 - \hat{S}^2\right\rangle_{j,s}, \tag{6.58}$$

where the indices irrelevant for each particular factor have been dropped. Now using the last of Eqs. (3.211), and similar expressions (5.169), (5.175), and (5.177) for eigenvalues of the involved operators, we get an explicit expression for the spin-orbit corrections[13]

$$E^{(1)}_{n,j,l} = \frac{1}{2m_e^2c^2}\frac{Ze^2}{4\pi\varepsilon_0}\frac{\hbar^2}{2r_0^3}\frac{j(j+1)-l(l+1)-¾}{n^3 l(l+½)(l+1)} \equiv \frac{E_n^2}{m_ec^2}n\frac{j(j+1)-l(l+1)-¾}{l(l+½)(l+1)}, \tag{6.59}$$

Spin-orbit energy correction

with l and j related by Eq. (5.189): $j = l \pm ½$.

The last form of its result shows clearly that this correction has the same magnitude scale as the kinetic correction (51).[14] In the 1st order of the perturbation theory, they may be just added (with $m = m_e$), giving a surprisingly simple formula for the net fine structure of the n^{th} energy level:

$$E^{(1)}_{\text{fine}} = \frac{E_n^2}{2m_ec^2}\left(3 - \frac{4n}{j+½}\right). \tag{6.60}$$

Fine structure of atomic levels

This simplicity, as well as the independence of the result of the orbital quantum number l, will become less surprising when (in Sec. 9.7) we see that this formula follows in one shot from the Dirac theory, in which the Bohr atom's energy spectrum is numbered only with n and j, but not l. Let us recall that for an electron ($s = ½$), according to Eq. (5.189) with $0 \le l \le n - 1$, the quantum number j may take n positive half-integer values, from ½ to $n - ½$. Hence, Eq. (60) shows that the fine structure of the n^{th} Bohr's energy level has n sub-levels – see Fig. 4.

Please note that according to Eq. (5.175), each of these sub-levels is still $(2j + 1)$-times degenerate in the quantum number m_j. This degeneracy is very natural, because all m-numbers describe the state orientation in a certain direction, while in the absence of an external field, the system is still isotropic. Moreover, on each fine-structure level (besides the highest one with $j = n - ½$), each of the m_j-states is doubly degenerate in the orbital quantum number $l = j \mp ½$ – see the labels of l in Fig. 4. (According to Eq. (5.190), each of these states, with fixed j and m_j, may be represented as a linear

[13] The factor l in the denominator does not give a divergence at $l = 0$, because in this case $j = s = ½$, so $j(j + 1) = ¾$, and the numerator turns into 0 as well. A careful analysis of this case (see, e.g., G. Woolgate, *Elementary Atomic Structure*, 2nd ed., Oxford, 1983), including the so-called *Darwin term* not described by Eqs. (51) and (59), shows that the final Eq. (60), which does not include l, is valid even in this case.

[14] This is natural because the magnetic interaction of charged particles is essentially a relativistic effect, of the same order ($\sim v^2/c^2$) as the kinetic correction (47) – see, e.g., EM Sec. 5.1, in particular Eq. (5.3).

combination of two states with adjacent values of l, and hence different electron spin orientations, $m_s = \pm ½$, weighed with the Clebsch-Gordan coefficients.)

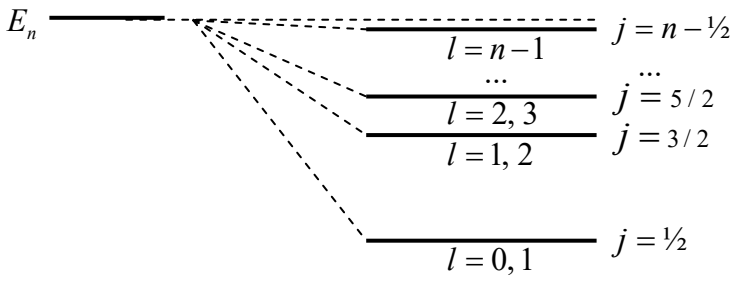


Fig. 6.4. The fine structure of a hydrogen-like atom/ion's level.

These details aside, one may crudely say that the relativistic corrections combined make the total eigenenergy grow with l, contributing to the effect already mentioned in the discussion of the periodic table of elements in Sec. 3.7. The relative scale of this increase may be quantified by the largest deviation from the unperturbed energy E_n, reached for the s-states (with $l = 0$):

$$\frac{\left|E^{(1)}_{\max}\right|}{E_n} = \frac{E_n}{2m_e c^2}(4n-3) \equiv \left(\frac{Ze^2}{4\pi\varepsilon_0\hbar c}\right)^2\left(\frac{1}{n}-\frac{3}{4n^2}\right) \equiv Z^2\alpha^2\left(\frac{1}{n}-\frac{3}{4n^2}\right). \tag{6.61}$$

where α is the *fine-structure* ("Sommerfeld's") *constant*,

$$\alpha \equiv \frac{e^2}{4\pi\varepsilon_0\hbar c} \approx \frac{1}{137}, \tag{6.62}$$

(which was already mentioned in Sec. 4.4), which characterizes the relative strength (or rather weakness) of the electromagnetic effects in quantum mechanics – which in particular makes perturbative quantum electrodynamics possible.[15] These expressions show that the fine-structure splitting is a very small effect ($\sim\alpha^2 \sim 10^{-6}$) for the hydrogen atom, but it rapidly grows (as Z^2) with the nuclear charge (i.e. the atomic number) Z, and becomes rather substantial for the heaviest stable atoms with $Z \sim 10^2$.

6.4. The Zeeman effect

Now, we are ready to review the *Zeeman effect* – the atomic level splitting by an external magnetic field.[16] Using Eq. (3.26), with $q = -e$, for the description of the electron's orbital motion in the field, and the Pauli Hamiltonian (4.163) with $\gamma = -e/m_e$, for the electron spin's interaction with the field, we see that even for a hydrogen-like (i.e. single-electron) atom/ion, neglecting the relativistic effects, the full Hamiltonian is rather involved:

$$\hat{H} = \frac{1}{2m_e}\left(\hat{\mathbf{p}} + e\hat{\mathbf{A}}\right)^2 - \frac{Ze^2}{4\pi\varepsilon_0 r} + \frac{e}{m_e}\mathscr{B}\cdot\hat{\mathbf{S}}. \tag{6.63}$$

[15] The expression $\alpha^2 = E_H/m_e c^2$, where E_H is the Hartree energy (1.13), i.e. the scale of the basic energies E_n, is also very revealing.

[16] It was discovered experimentally in 1896 by Pieter Zeeman who, amazingly, was fired from the University of Leiden for unauthorized use of lab equipment for this work – just to receive a Nobel Prize for it in a few years!

There are several simplifications we may make. First, let us assume that the external field is spatial-uniform on the atomic scale (which is a very good approximation for most cases), so we can take its vector potential in an axially symmetric gauge – cf. Eq. (3.132):

$$\mathbf{A} = \frac{1}{2}\mathscr{B} \times \mathbf{r}. \tag{6.64}$$

Second, let us neglect the terms proportional to $\mathscr{B}^2$, which are small in practical magnetic fields of the order of a few teslas.[17] The remaining term in the effective kinetic energy, describing the interaction with the magnetic field, is linear in the momentum operator, so we may repeat the standard classical calculation[18] to reduce it to the product of $\mathscr{B}$ by the orbital magnetic moment's component $m_z = -eL_z/2m_e$ – besides that both m_z and L_z should be understood as operators now. As a result, the Hamiltonian (63) reduces to Eq. (1), $\hat{H}^{(0)} + \hat{H}^{(1)}$, where $\hat{H}^{(0)}$ is that of the atom at $\mathscr{B} = 0$, and

$$\hat{H}^{(1)} = \frac{e\mathscr{B}}{2m_e}\left(\hat{L}_z + 2\hat{S}_z\right) \tag{6.65}$$

Zeeman effect's perturbation

This expression immediately reveals the major complication with the Zeeman effect's analysis. Namely, in comparison with the equal orbital and spin contributions to the total angular momentum (5.170) of the electron, its spin produces a twice larger contribution to the magnetic moment, so the right-hand side of Eq. (65) is *not* proportional to $\hat{J}_z = \hat{L}_z + \hat{S}_z$. As a result, the effect's description is quite simple only in two limits.

If the magnetic field is so *high* that its effects are much stronger than the relativistic (fine-structure) effects discussed in the previous section, we may treat the two terms in Eq. (65) as independent perturbations of different (orbital and spin) degrees of freedom. Since each of the perturbation matrices is diagonal in its own z-basis, we can again use Eq. (27) to write

$$E - E^{(0)} = \frac{e\mathscr{B}}{2m_e}\left(\langle n,l,m_l|\hat{L}_z|n,l,m_l\rangle + 2\langle m_s|\hat{S}_z|m_s\rangle\right) = \frac{e\mathscr{B}}{2m_e}\left(\hbar m_l + 2\hbar m_s\right) = \mu_B\mathscr{B}(m_l \pm 1). \tag{6.66}$$

Paschen-Back effect

This result describes the splitting of each $2\times(2l + 1)$-degenerate energy level, with certain n and l, into $(2l + 3)$ levels (Fig. 5), with the adjacent level distance of $\mu_B\mathscr{B}$, of the order of 10^{-4} eV per tesla.

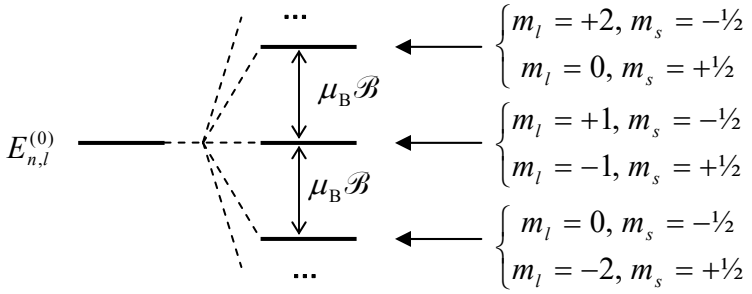


Fig. 6.5. The Paschen-Back effect.

[17] Despite its smallness, the quadratic term is necessary for a description of the negative contribution of the orbital motion to the magnetic susceptibility χ_m (the so-called *orbital diamagnetism*, see EM Sec. 5.5), whose analysis, using Eq. (63), is left for the reader's exercise.

[18] See, e.g., EM Sec. 5.4, in particular, Eqs. (5.95) and (5.100).

Note that all these levels, besides the top and bottom ones, remain doubly degenerate. This limit of the Zeeman effect is sometimes called the *Paschen-Back effect* – whose simplicity was recognized only in the 1920s, due to the need in very high magnetic fields for its observation.

In the opposite limit of relatively *low* magnetic fields, the Zeeman effect takes place on the background of the much larger fine-structure splitting. As was discussed in Sec. 3, at $\mathscr{B} = 0$ each split sub-level has a $2\times(2j + 1)$-fold degeneracy corresponding to $(2j + 1)$ different values of the half-integer quantum number m_j, ranging from $-j$ to $+j$, and two values of the integer $l = j \mp ½$ – see Fig. 4.[19] The magnetic field lifts this degeneracy. Indeed, in the coupled representation discussed in Sec. 5.7, the perturbation (65) is described by the matrix with elements

$$\begin{aligned} H^{(1)} &= \frac{e\mathscr{B}}{2m_e}\left\langle j, m_j \left| \hat{L}_z + 2\hat{S}_z \right| j', m_{j'} \right\rangle \equiv \frac{e\mathscr{B}}{2m_e}\left\langle j, m_j \left| \hat{J}_z + \hat{S}_z \right| j', m_{j'} \right\rangle \\ &= \frac{e\mathscr{B}}{2m_e}\left(\hbar m_j \delta_{m_j m_{j'}} + \left\langle j, m_j \left| \hat{S}_z \right| j', m_{j'} \right\rangle \right). \end{aligned} \tag{6.67}$$

To spell out the second term, let us use the general expansion (5.183) for the particular case $s = ½$, when (as was discussed at the end of Sec. 5.7) it has at most two non-vanishing terms, with the Clebsh-Gordan coefficients (5.190):

$$\begin{aligned} &\left| j = l \pm ½, m_j \right\rangle \\ &= \pm\left(\frac{l \pm m_j + ½}{2l+1} \right)^{1/2} \left| m_l = m_j - ½, m_s = +½ \right\rangle + \left(\frac{l \mp m_j + ½}{2l+1} \right)^{1/2} \left| m_l = m_j + ½, m_s = -½ \right\rangle. \end{aligned} \tag{6.68}$$

Taking into account that the operator $\hat{S}_z$ gives non-zero brackets only for $m_s = m_{s'}$, the 2×2 matrix of elements $\left\langle m_l = m_j \pm ½, m_s = \mp ½ \left| \hat{S}_z \right| m_l = m_j \pm ½, m_s = \mp ½ \right\rangle$ is diagonal, so we may use Eq. (27) to get

Anomalous Zeeman effect for $s = 1/2$

$$\begin{aligned} E - E^{(0)} &= \frac{e\mathscr{B}}{2m_e}\left[\hbar m_j + \frac{\hbar}{2}\frac{\left(l \pm m_j + ½\right)}{2l+1} - \frac{\hbar}{2}\frac{\left(l \mp m_j + ½\right)}{2l+1} \right] \\ &\equiv \frac{e\mathscr{B}}{2m_e}\hbar m_j \left(1 \pm \frac{1}{2l+1} \right) \equiv \mu_B \mathscr{B} m_j \left(1 \pm \frac{1}{2l+1} \right), \quad \text{for} - j \le m_j \le +j, \end{aligned} \tag{6.69}$$

where the two signs correspond to the two possible values of $l = j \mp ½$ – see Fig. 6.

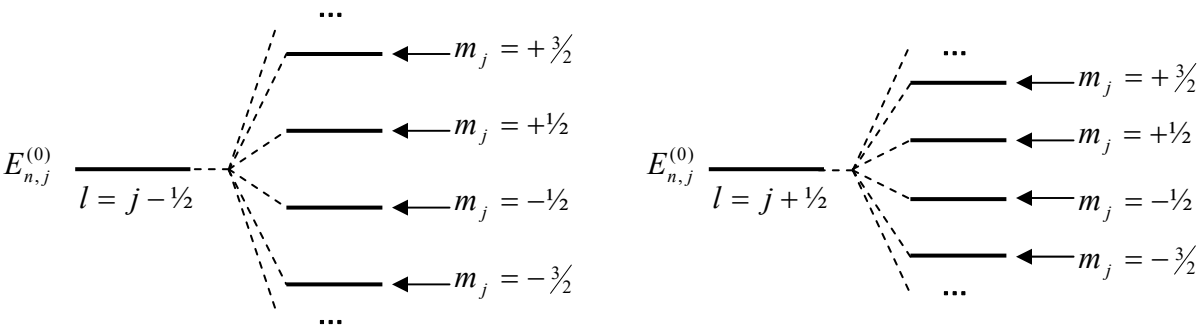


Fig. 6.6. The anomalous Zeeman effect in a hydrogen-like atom/ion.

[19] In the almost-hydrogen-like, but more complex atoms (such as those of alkali metals), the degeneracy in l may be lifted by electron-electron Coulomb interaction even in the absence of an external magnetic field.

We see that the magnetic field splits each sub-level of the fine structure, with a given l, into $2j + 1$ equidistant levels, with the distance between the levels depending on l. In the late 1890s when this effect was first observed (by T. Preston), there was no notion of spin at all, so this puzzling result was called the *anomalous Zeeman effect*.[20]

The strict quantum-mechanical analysis of the anomalous Zeeman effect for arbitrary s (which is important for applications to multi-electron atoms) is conceptually not too complex but requires explicit expressions for the corresponding Clebsch-Gordan coefficients, which are rather bulky. Let me just cite the unexpectedly simple result of this analysis:

$$\Delta E = \mu_B \mathcal{B} m_j g, \tag{6.70a}$$

Anomalous Zeeman effect for arbitrary s

where g is the so-called *Lande factor*:[21]

$$g = 1 + \frac{j(j+1) + s(s+1) - l(l+1)}{2j(j+1)}. \tag{6.70b}$$

For $s = ½$ (and hence $j = l \pm ½$), this factor is reduced to the parentheses in the last forms of Eq. (69).

It is remarkable that Eqs. (70) may be readily derived using very plausible classical arguments, similar to those used in Sec. 5.7 – see Fig. 5.13 and its discussion. As was discussed in Sec. 5.6, in the absence of spin, the quantization of the observable L_z is a modification of the classical picture of the torque-induced precession of the vector $\mathbf{L}$ about the magnetic field's direction, so the interaction energy, proportional to $\mathcal{B}L_z = \boldsymbol{\mathcal{B}} \cdot \mathbf{L}$, remains constant – see Fig. 7a. On the other hand, at the spin-orbit interaction without an external magnetic field, the Hamiltonian function of the system includes the product $\mathbf{S}\cdot\mathbf{L}$, so in the stationary state it has to be constant, together with J^2, L^2, and S^2. Hence, this system's classical image is a joint precession of the vectors $\mathbf{S}$ and $\mathbf{L}$ about the direction of the vector $\mathbf{J} = \mathbf{L} + \mathbf{S}$, in such a manner that the spin-orbit interaction energy, proportional to the product $\mathbf{L}\cdot\mathbf{S}$, remains constant (Fig. 7b). On this backdrop, the anomalous Zeeman effect in a relatively weak magnetic field $\boldsymbol{\mathcal{B}} = \mathcal{B}\mathbf{n}_z$ corresponds to a much slower additional precession of the vector $\mathbf{J}$ about the z-axis, "dragging" with it the vectors $\mathbf{L}$ and $\mathbf{S}$, rapidly rotating around it.

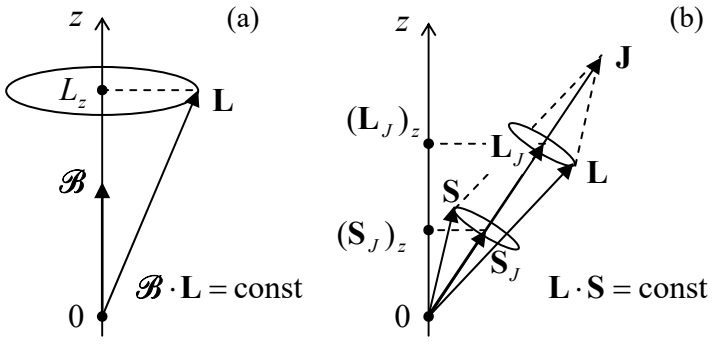


Fig. 6.7. Classical images of (a) the orbital angular momentum's quantization in a magnetic field, and (b) the fine-structure level splitting.

[20] In this terminology, the *normal Zeeman effect*, observed in atoms with zero net spin, is the one with no spin splitting, i.e. without the second terms in the parentheses of Eqs. (66), (67), and (69).

[21] This formula is frequently used with capital letters J, S, and L, which denote the quantum numbers of the atom as a whole.

This physical picture allows us to conjecture that what is important for the slow precession rate are only the vectors **L** and **S** averaged over the period of their much faster precession around vector **J** – in other words, only their components $\mathbf{L}_J$ and $\mathbf{S}_J$ along the vector **J**. Classically, these components may be calculated as

$$\mathbf{L}_J = \frac{\mathbf{L}\cdot\mathbf{J}}{J^2}\mathbf{J}, \qquad \text{and} \quad \mathbf{S}_J = \frac{\mathbf{S}\cdot\mathbf{J}}{J^2}\mathbf{J}. \tag{6.71}$$

The scalar products participating in these expressions may be readily expressed via the squared lengths of the vectors, using the following geometric formulas:

$$S^2 = (\mathbf{J}-\mathbf{L})^2 \equiv J^2 + L^2 - 2\mathbf{L}\cdot\mathbf{J}, \qquad L^2 = (\mathbf{J}-\mathbf{S})^2 \equiv J^2 + S^2 - 2\mathbf{J}\cdot\mathbf{S}. \tag{6.72}$$

As a result, we get the following time average:

$$\begin{aligned} \overline{L_z + 2S_z} &= \left(\mathbf{L}_J + 2\mathbf{S}_J\right)_z = \left(\frac{\mathbf{L}\cdot\mathbf{J}}{J^2}\mathbf{J} + 2\frac{\mathbf{S}\cdot\mathbf{J}}{J^2}\mathbf{J}\right)_z = \frac{J_z}{J^2}\left(\mathbf{L}\cdot\mathbf{J} + 2\mathbf{S}\cdot\mathbf{J}\right) \\ &= J_z\frac{(J^2+L^2-S^2)+2(J^2+S^2-L^2)}{2J^2} \equiv J_z\left(1+\frac{J^2+S^2-L^2}{2J^2}\right). \end{aligned} \tag{6.73}$$

The last move is to smuggle in some quantum mechanics by using, instead of the vector lengths squared and the *z*-component of J_z, their eigenvalues given by Eqs. (5.169), (5.175), and (5.177). As a result, we immediately arrive at the exact Eqs. (70). This coincidence encourages thinking about quantum mechanics of angular momenta in the classical terms of torque-induced precession, which turns out to be very fruitful in some more complex problems of atomic and molecular physics.

The high-field limit and low-field limits of the Zeeman effect, described respectively by Eqs. (66) and (69), are separated by a medium field range, in which the Zeeman splitting is of the order of the fine-structure splitting analyzed in Sec. 3. There is no time in this course for a quantitative analysis of this (conceptually simple) crossover, which involves rather cumbersome algebra.[22]

6.5. Time-dependent perturbations

Now let us proceed to the case when the perturbation $\hat{H}^{(1)}$ in Eq. (1) is a function of time, while $\hat{H}^{(0)}$ is time-independent. The adequate perturbative approach to this problem, and its results, depend critically on the relation between the characteristic frequency ω of the perturbation and the distance between the initial system's energy levels:

$$\hbar\omega \leftrightarrow \left|E_n - E_{n'}\right|. \tag{6.74}$$

In the case when all essential frequencies of a perturbation are very small in the sense of Eq. (74), we are dealing with the so-called *adiabatic* change of parameters, that may be treated essentially as a time-independent perturbation – see the previous sections of this chapter). The most interesting observation here is that the adiabatic perturbation does not allow any significant transfer of the system's

[22] For a more complete discussion of the Stark, Zeeman, and fine-structure effects in atoms, I can recommend, for example, either the monograph by G. Woolgate cited above, or the one by I. Sobelman, *Theory of Atomic Spectra*, Alpha Science, 2006.

probability from one eigenstate to another. For example, in the WKB limit of the orbital motion, the Bohr quantization rule and its Wilson-Sommerfeld modification (2.110) guarantee that the integral

$$\oint_C \mathbf{p}\cdot d\mathbf{r}\,, \tag{6.75}$$

taken along the particle's classical trajectory, is an *adiabatic invariant*, i.e. does not change at a slow change of system's parameters. (It is curious that classical mechanics also guarantees the invariance of the integral (75), but its proof there[23] is much harder than the quantum-mechanical derivation of this fact, carried out in Sec. 2.4.) This is why even if the perturbation becomes large with time (while changing sufficiently slowly), we can expect the classification of eigenstates and eigenvalues to persist.

Let us proceed to the harder case when both sides of Eq. (74) are comparable, using for this discussion the Schrödinger picture of quantum dynamics, given by Eq. (4.158). Combining it with Eq. (1), we get the Schrödinger equation in the form

$$i\hbar\frac{\partial}{\partial t}\left|\alpha(t)\right\rangle = \left(\hat{H}^{(0)} + H^{(1)}(t)\right)\left|\alpha(t)\right\rangle. \tag{6.76}$$

Very much in the spirit of our treatment of the time-independent case in Sec. 1, let us represent the time-dependent ket-vector of the system with its expansion,

$$\left|\alpha(t)\right\rangle = \sum_n \left|n\right\rangle\left\langle n\middle|\alpha(t)\right\rangle, \tag{6.77}$$

over the full and orthonormal set of the unperturbed, stationary ket-vectors defined by the equation

$$\hat{H}^{(0)}\left|n\right\rangle = E_n\left|n\right\rangle. \tag{6.78}$$

(Note that these kets $|n\rangle$ are exactly what was called $|n^{(0)}\rangle$ in Sec. 1; we may afford a less bulky notation in this section because only the lowest orders of the perturbation theory will be discussed.) Plugging the expansion (77), with n replaced with n', into both sides of Eq. (76), and then inner-multiplying both its sides by the bra-vector $\langle n|$ of another unperturbed (and hence time-independent) state of the system, we get the following set of linear, ordinary differential equations for the expansion coefficients:

$$i\hbar\frac{d}{dt}\left\langle n\middle|\alpha(t)\right\rangle = E_n\left\langle n\middle|\alpha(t)\right\rangle + \sum_{n'} H^{(1)}_{nn'}(t)\left\langle n'\middle|\alpha(t)\right\rangle, \tag{6.79}$$

where the matrix elements of the perturbation, in the *unperturbed* state basis, defined similarly to Eq. (8), are now functions of time:

$$H^{(1)}_{nn'}(t) \equiv \left\langle n\middle|\hat{H}^{(1)}(t)\middle|n'\right\rangle. \tag{6.80}$$

The set of differential equations (79), which are still exact, may be useful for numerical calculations.[24] However, it has a certain technical inconvenience, which becomes clear if we consider its (evident) solution in the absence of perturbation:[25]

[23] See, e.g., CM Sec. 10.2.

[24] Even if the problem under analysis may be described by the wave-mechanics Schrödinger equation (1.25), direct numerical integration of that *partial* differential equation is typically less convenient than that of the *ordinary* differential equations (79).

[25] This is of course just a more general form of Eq. (1.62) of the wave mechanics of time-independent systems.

$$\langle n|\alpha(t)\rangle = \langle n|\alpha(0)\rangle \exp\left\{-i\frac{E_n}{\hbar}t\right\}. \tag{6.81}$$

We see that these solutions oscillate very fast, and their numerical modeling may represent a challenge for even the fastest computers. These spurious oscillations (whose frequency, in particular, depends on the energy reference level) may be partly tamed by looking for the general solution of Eqs. (79) in a form inspired by Eq. (81):

$$\langle n|\alpha(t)\rangle \equiv a_n(t)\exp\left\{-i\frac{E_n}{\hbar}t\right\}. \tag{6.82}$$

Here $a_n(t)$ are new functions of time (essentially, the stationary states' probability amplitudes), which may be used, in particular, to calculate the time-dependent *level occupancies*, i.e. the probabilities W_n to find the perturbed system on the corresponding energy levels of the unperturbed system:

$$W_n(t) = \left|\langle n|\alpha(t)\rangle\right|^2 = \left|a_n(t)\right|^2. \tag{6.83}$$

Plugging Eq. (82) into Eq. (79), for these functions, we readily get a slightly modified system of equations:

Probability amplitudes: evolution

$$i\hbar\dot{a}_n = \sum_{n'} a_{n'} H^{(1)}_{nn'}(t) e^{i\omega_{nn'}t}, \tag{6.84}$$

where the factors $\omega_{nn'}$, defined by the relation

Quantum transition frequencies

$$\hbar\omega_{nn'} \equiv E_n - E_{n'}, \tag{6.85}$$

have the physical sense of frequencies of *potential* quantum transitions between the n^{th} and n'^{th} energy levels of the unperturbed system. (The conditions when such transitions indeed take place will be clear soon.) The advantages of Eq. (84) over Eq. (79), for both analytical and numerical calculations, are their independence of the energy reference, and lower frequencies of oscillations of the right-hand side terms, especially when the energy levels of interest are close to each other.[26]

In order to continue our analytical treatment, let us focus on a particular but very important problem of a sinusoidal perturbation *turned on* at some moment – which may be taken for $t = 0$:

Turning on sinusoidal perturbation

$$\hat{H}^{(1)}(t) = \begin{cases} 0, & \text{for } t < 0, \\ \hat{A}e^{-i\omega t} + \hat{A}^\dagger e^{+i\omega t}, & \text{for } t \geq 0, \end{cases} \tag{6.86}$$

where the *perturbation amplitude operators* $\hat{A}$ and $\hat{A}^\dagger$,[27] and hence their matrix elements,

[26] Note that the relation of Eq. (84) to the initial Eq. (79) is very close to the relation of the interaction picture of quantum dynamics, discussed at the end of Sec. 4.6, to its Schrödinger picture, with the perturbation Hamiltonian playing the role of the interaction one – compare Eqs. (1) and Eq. (4.206). Indeed, Eq. (84) could be readily obtained from the interaction picture, and I did not do this just to avoid using this heavy bra-ket artillery for our current (relatively) simple problem, and hence to keep its physics more transparent.

[27] The notation of the amplitude operators in Eq. (86) is justified by the fact that the perturbation Hamiltonian has to be self-adjoint (Hermitian), and hence each term on the right-hand side of that relation has to be a Hermitian conjugate of its counterpart, which is evidently true only if the amplitude operators are also the Hermitian conjugates of each other. Note, however, that each of these amplitude operators is generally *not* Hermitian.

$$\langle n|\hat{A}|n'\rangle \equiv A_{nn'}, \qquad \langle n|\hat{A}^{\dagger}|n'\rangle = A^{*}_{n'n}, \tag{6.87}$$

are time-independent after the turn-on moment. In this case, Eq. (84) yields

$$i\hbar \dot{a}_n = \sum_{n'} a_{n'} \left[A_{nn'} e^{i(\omega_{nn'} - \omega)t} + A^{*}_{n'n} e^{i(\omega_{nn'} + \omega)t} \right], \qquad \text{for } t > 0 . \tag{6.88}$$

This is, generally, still a nontrivial system of coupled differential equations; however, it allows simple and explicit solutions in two very important limits. First, let us assume that our system initially was definitely in one eigenstate *n'* (usually, though not necessarily, in the ground state), and that the occupancies W_n of all other levels stay very low all the time. (We will find the condition when the second assumption is valid *a posteriori* – from the solution.) With these assumptions,

$$a_{n'} = 1; \qquad |a_n| << 1, \quad \text{for } n \neq n', \tag{6.89}$$

Eq. (88) may be readily integrated, giving

$$a_n = -\frac{A_{nn'}}{\hbar(\omega_{nn'} - \omega)} \left[e^{i(\omega_{nn'} - \omega)t} - 1 \right] - \frac{A^{*}_{n'n}}{\hbar(\omega_{nn'} + \omega)} \left[e^{i(\omega_{nn'} + \omega)t} - 1 \right], \qquad \text{for } n \neq n'. \tag{6.90}$$

This expression describes what is colloquially called the *ac excitation of* (other) *energy levels*. Qualitatively, it shows that the probability W_n (83) of finding the system in each state ("on each energy level") of the system does not tend to any constant value but rather oscillates in time. It also shows that the ac-field-induced transfer of the system from one state to the other one has a clearly resonant character: the maximum occupancy W_n of a level with number $n \neq n'$ grows infinitely when the corresponding *detuning*[28]

$$\Delta_{nn'} \equiv \omega - \omega_{nn'}, \tag{6.91}$$

tends to zero. This conclusion is clearly unrealistic, and is an artifact of our initial assumption (89); according to Eq. (90), it is satisfied only if[29]

$$|A_{nn'}| << \hbar |\omega \pm \omega_{nn'}|, \tag{6.92}$$

and hence which does not allow a deeper analysis of the resonant excitation.

In order to overcome this limitation, we may perform the following trick – very similar to the one we used for the transfer to the degenerate case in Sec. 1. Let us assume that for a certain level *n*,

$$|\Delta_{nn'}| << \omega, |\omega \pm \omega_{n''n}|, |\omega \pm \omega_{n''n'}|, \qquad \text{for all } n'' \neq n, n' \tag{6.93}$$

– the condition illustrated in Fig. 8. Then, according to Eq. (90), we may ignore the occupancy of all but two levels, *n* and *n'*, and also the second, non-resonant term with frequency $\omega_{nn'} + \omega \approx 2\omega >> |\Delta_{nn'}|$ in Eqs. (88),[30] now written for two probability amplitudes, a_n and $a_{n'}$.

[28] The notion of detuning is also very useful in the classical theory of oscillations (see, e.g., CM Chapter 5), where the role of $\omega_{nn'}$ is played by the own frequency ω_0 of the oscillator.

[29] Strictly speaking, one more condition is that the number of "resonance" levels is also not too high – see Sec. 6.

[30] The second assumption, i.e. the omission of non-resonant terms in the equations for amplitudes is called the *Rotating Wave Approximation* (RWA); the same idea in the classical theory of oscillations is the basis of what is usually called the *van der Pol method*, and its result, the *reduced equations* – see, e.g., CM Secs. 5.3-5.5.

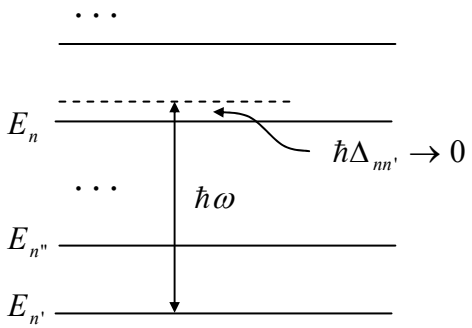


Fig. 6.8. The resonant excitation of an energy level.

The result is the following system of two linear equations:

$$i\hbar\dot{a}_n = a_{n'}Ae^{-i\Delta t}, \qquad i\hbar\dot{a}_{n'} = a_n A^* e^{i\Delta t}, \tag{6.94}$$

which uses the shorthand notation $A \equiv A_{nn'}$ and $\Delta \equiv \Delta_{nn'}$. (I will use this simplified notation for a while – until other energy levels become involved, at the beginning of the next section). This system may be readily reduced to a form without explicit time dependence of the right-hand parts – for example, by introducing the following new probability amplitudes, with the same moduli:

$$b_n \equiv a_n e^{i\Delta t/2}, \qquad b_{n'} \equiv a_{n'}e^{-i\Delta t/2}, \tag{6.95}$$

so

$$a_n = b_n e^{-i\Delta t/2}, \quad a_{n'} = b_{n'}e^{i\Delta t/2}. \tag{6.96}$$

Plugging these relations into Eq. (94), we get two usual linear first-order differential equations:

$$i\hbar\dot{b}_n = -\frac{\hbar\Delta}{2}b_n + Ab_{n'}, \qquad i\hbar\dot{b}_{n'} = A^* b_n + \frac{\hbar\Delta}{2}b_{n'}. \tag{6.97}$$

As the reader knows very well by now, the general solution of such a system is a linear combination of two exponential functions, $\exp\{\lambda_\pm t\}$, with the exponents $\lambda_\pm$ that may be found by plugging any of these functions into Eq. (97), and requiring the consistency of the two resulting linear algebraic equations. In our case, the consistency condition (i.e. the characteristic equation of the system) is

$$\begin{vmatrix} -\hbar\Delta/2 - i\hbar\lambda_\pm & A \\ A^* & \hbar\Delta/2 - i\hbar\lambda_\pm \end{vmatrix} = 0, \tag{6.98}$$

and has two solutions $\lambda_\pm = \pm i\Omega$, where

Rabi oscillations: frequency

$$\Omega \equiv \left(\frac{\Delta^2}{4} + \frac{|A|^2}{\hbar^2}\right)^{1/2}, \qquad \text{i.e. } 2\Omega = \left(\Delta^2 + 4\frac{|A|^2}{\hbar^2}\right)^{1/2}. \tag{6.99}$$

The coefficients at the exponents are determined by initial conditions. If, as was assumed before, the system was on the level *n'* initially (at $t = 0$), i.e. if $a_{n'}(0) = 1$, $a_n(0) = 0$, so $b_{n'}(0) = 1$, $b_n(0) = 0$ as well, then Eqs. (97) yield, in particular:

$$b_n(t) = -i\frac{A}{\hbar\Omega}\sin\Omega t, \tag{6.100}$$

so the n^{th} level occupancy is

$$W_n = |b_n|^2 = \frac{|A|^2}{\hbar^2\Omega^2}\sin^2\Omega t \equiv \frac{|A|^2}{|A|^2 + (\hbar\Delta/2)^2}\sin^2\Omega t\,. \qquad (6.101)$$

Rabi formula

This is the famous *Rabi oscillation formula*.[31] It shows that if the detuning is large in comparison with $|A|/\hbar$, though still small in the sense of Eq. (93), the frequency 2Ω of the Rabi oscillations is completely determined by the detuning, and their amplitude is small:

$$W_n(t) = 4\frac{|A|^2}{\hbar^2\Delta^2}\sin^2\frac{\Delta t}{2} << 1, \qquad \text{for } |A|^2 << (\hbar\Delta)^2, \qquad (6.102)$$

– the result which could be obtained directly from Eq. (90), just neglecting the second term on its right-hand side. However, now we may also analyze the results of an increase of the perturbation amplitude: it leads not only to an increase of the amplitude of the probability oscillations but also of their frequency – see Fig. 9. Ultimately, at $|A| >> \hbar|\Delta|$ (for example, at the exact resonance, $\Delta = 0$., i.e. $\omega_{nn'} = \omega$, so $E_n = E_{n'} + \hbar\omega$), Eqs. (101)-(102) give $\Omega = |A|/\hbar$ and $(W_n)_{max} = 1$, i.e. describe a periodic, full "repumping" of the system from one level to another and back, with a frequency proportional to the perturbation amplitude.[32]

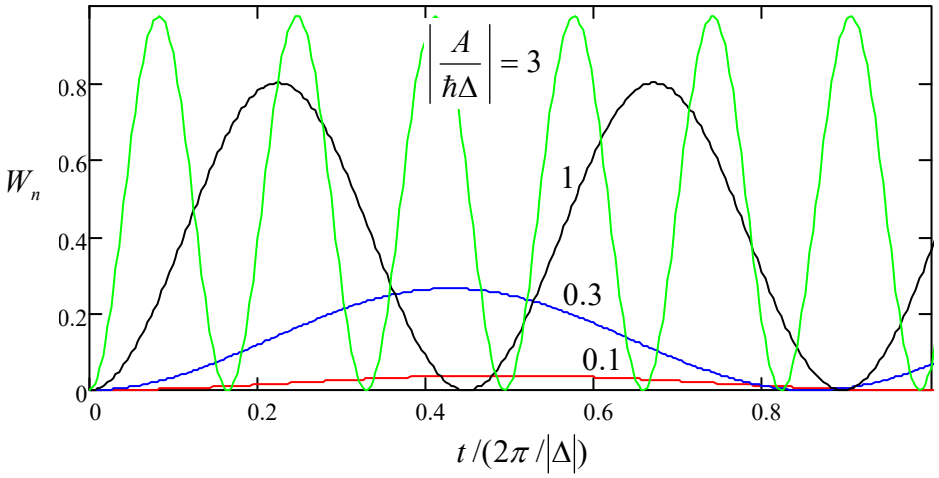


Fig. 6.9. The Rabi oscillations for several values of the normalized amplitude of ac perturbation.

This effect is a close analog of the quantum oscillations in two-level systems with time-independent Hamiltonians, which were discussed in Secs. 2.6 and 5.1. Indeed, let us revisit, for a moment, their discussion started at the end of Sec.1 of this chapter, now paying more attention to the time evolution of the system under a perturbation. As was argued in that section, the most general perturbation Hamiltonian lifting the two-fold degeneracy of an energy level, in an arbitrary basis, has the matrix (28). Let us describe the system's dynamics using, again, the Schrödinger picture, representing the ket-vector of an arbitrary state of the system in the form (5.1), where ↑ and ↓ are the

[31] It was derived in 1952 by Isaac Rabi, in the context of his group's pioneering experiments with the ac (practically, microwave) excitation of quantum states, using molecular beams in vacuum.

[32] As Eqs. (82), (96), and (99) show, the lowest frequency in the system is $\omega_1 = \omega_{n'} - \Delta/2 + \Omega$, so at $A \to 0$, $\hbar\omega_1 \approx \hbar\omega_{n'} + 2|A|^2/\hbar\Delta$. This effective shift of the lowest energy level (which may be measured by another "probe" field of a different frequency) is a particular case of the *ac Stark effect*, which was already mentioned in Sec. 2.

time-independent states of the basis in that Eq. (28) is written (now without any obligation to associate these states with the *z*-basis of any spin-½.) Then, the Schrödinger equation (4.158) yields

$$i\hbar\begin{pmatrix}\dot{\alpha}_\uparrow\\ \dot{\alpha}_\downarrow\end{pmatrix} = \mathrm{H}^{(1)}\begin{pmatrix}\alpha_\uparrow\\ \alpha_\downarrow\end{pmatrix} \equiv \begin{pmatrix}H_{11} & H_{12}\\ H_{21} & H_{22}\end{pmatrix}\begin{pmatrix}\alpha_\uparrow\\ \alpha_\downarrow\end{pmatrix} \equiv \begin{pmatrix}H_{11}\alpha_\uparrow + H_{12}\alpha_\downarrow\\ H_{21}\alpha_\uparrow + H_{22}\alpha_\downarrow\end{pmatrix}. \tag{6.103}$$

As we know (for example, from the discussion in Sec. 5.1), the average of the diagonal elements of the matrix gives just a common shift of the system's energy; for the purpose of the analysis, it may be absorbed into the energy reference level. Also, the Hamiltonian operator has to be Hermitian, so the off-diagonal elements of its matrix have to be complex-conjugate. With this, Eqs. (103) are reduced to the form,

$$i\hbar\dot{\alpha}_\uparrow = -\frac{\xi}{2}\alpha_\uparrow + H_{12}\alpha_\downarrow, \qquad i\hbar\dot{\alpha}_\downarrow = H_{12}^{*}\alpha_\uparrow + \frac{\xi}{2}\alpha_\downarrow, \qquad \text{with } \hbar\xi \equiv H_{22} - H_{11}, \tag{6.104}$$

which is absolutely similar to Eqs. (97). In particular, these equations describe the quantum oscillations of the probabilities $W_\uparrow = |\alpha_\uparrow|^2$ and $W_\downarrow = |\alpha_\downarrow|^2$ with the frequency[33]

$$2\Omega = \left(\xi^2 + 4\frac{\left|H_{12}\right|^2}{\hbar^2}\right)^{1/2}. \tag{6.105}$$

The similarity of Eqs. (97) and (104), and hence of Eqs. (99) and (105), shows that the "usual" quantum oscillations and the Rabi oscillations have essentially the same physical nature, besides that in the latter case the external ac signal quantum $\hbar\omega$ bridges the separated energy levels, effectively reducing their difference $(E_n - E_{n'})$ to a much smaller difference $-\Delta \equiv (E_n - E_{n'}) - \hbar\omega$. Also, since the Hamiltonian (28) is similar to that given by Eq. (5.2), the dynamics of such a system with two ac-coupled energy levels, within the limits (93) of the perturbation theory, is completely similar to that of a time-independent two-level system. In particular, its state may be similarly represented by a point on the Bloch sphere shown in Fig. 5.3, with its dynamics described, in the Heisenberg picture, by Eq. (5.19). This fact is very convenient for the experimental implementation of quantum information processing systems (to be discussed in more detail in Sec. 8.5), because it enables qubit manipulations in a broad variety of physical systems with well-separated energy levels, using external ac (usually either microwave or optical) sources.

Note, however, that according to Eq. (90), if a system has energy levels other than *n* and *n'*, they also become occupied to some extent. Since the sum of all occupancies equals 1, this means that $(W_n)_{\max}$ may approach 1 only if the other excitation amplitude is very small, and hence the state manipulation time scale $\mathcal{T} = 2\pi/\Omega = 2\pi\hbar/|A|$ is very long. The ultimate limit in this sense is provided by the harmonic oscillator where all energy levels are equidistant, and the probability repumping between all of them occurs at comparable rates. In particular, in this system, the implementation of the full Rabi oscillations is impossible even at the exact resonance.[34]

---

[33] By the way, Eq. (105) gives a natural generalization of the relations obtained for the frequency of such oscillations in Sec. 2.6, where the coupled potential wells were assumed to be exactly similar, so $\xi = 0$. Moreover, Eqs. (104) gives a long-promised proof of Eqs. (2.201), and hence a better justification of Eqs. (2.203).

[34] From Sec. 5.5, we already know what happens to the ground state of an oscillator at its external sinusoidal (or any other) excitation: it turns into a Glauber state, i.e. a superposition of *all* Fock states – see Eq. (5.134).

However, I would not like these quantitative details to obscure from the reader the most important qualitative (OK, maybe semi-quantitative :-) conclusion of this section's analysis: a resonant increase of the interlevel transition intensity at $\omega \to \omega_{nn'}$. As will be shown later in the course, in a quantum system coupled to its environment at least slightly (hence in reality, in *any* quantum system), such increase is accompanied by a sharp increase of the external field's *absorption*, which may be measured. This increase is used in numerous applications, notably including the magnetic resonance techniques already mentioned in Sec. 5.1.

6.6. Quantum-mechanical Golden Rule

One of the results of the past section, Eq. (102), may be used to derive one of the most important and nontrivial results of quantum mechanics. For that, let us consider the case when the perturbation causes quantum transitions from a *discrete* energy level $E_{n'}$ into a group of eigenstates with a very dense (essentially *continuous*) spectrum E_n – see Fig. 10a.

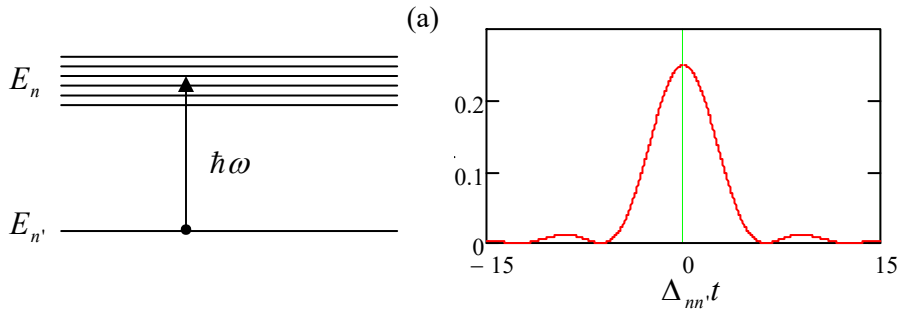


Fig. 6.10. Deriving the Golden Rule: (a) the energy level scheme, and (b) the function under the integral in Eq. (108).

If, for all states n of the group, the following conditions are satisfied

$$\left|A_{nn'}\right|^2 << \left(\hbar\Delta_{nn'}\right)^2 << \left(\hbar\omega_{nn'}\right)^2, \tag{6.106}$$

then Eq. (102) coincides with the result that would follow from Eq. (90). This means that we may apply Eq. (102), with the indices n and n' duly restored, to any level n of our tight group. As a result, the total probability of having our system transferred from the initial level n' to that group is

$$W_\Sigma(t) = \sum_n W_n(t) = \frac{4}{\hbar^2}\sum_n \frac{\left|A_{nn'}\right|^2}{\Delta_{nn'}^2}\sin^2\frac{\Delta_{nn'}t}{2}. \tag{6.107}$$

Now comes the main, absolutely beautiful trick: let us assume that the summation over n is limited to a tight group of very similar states whose matrix elements $A_{nn'}$ are virtually similar (we will check the validity of this assumption later on), so we can take $|A_{nn'}|^2$ out of the sum in Eq. (107) and then replace the sum with the corresponding integral:

$$W_\Sigma(t) = \frac{4\left|A_{nn'}\right|^2}{\hbar^2}\int \frac{1}{\Delta_{nn'}^2}\sin^2\frac{\Delta_{nn'}t}{2}\,dn \equiv \frac{4\left|A_{nn'}\right|^2\rho_n t}{\hbar}\int \frac{1}{\left(\Delta_{nn'}t\right)^2}\sin^2\frac{\Delta_{nn'}t}{2}\,d(-\Delta_{nn'}t), \tag{6.108}$$

where ρ_n is the density of the states n on the energy axis:

$$\rho_n \equiv \frac{dn}{dE_n}. \tag{6.109}$$

Density of states

This density and the matrix element $A_{nn'}$ have to be evaluated at $\Delta_{nn'} = 0$, i.e. at energy $E_n = E_{n'} + \hbar\omega$, and are assumed to be constant within the final state group. At fixed $E_{n'}$, the function under integral (108) is even and decreases fast at $|\Delta_{nn'}t| >> 1$ – see Fig. 10b. Hence we may introduce a dimensionless integration variable $\xi \equiv \Delta_{nn'}t$, and extend the integration over it formally from $-\infty$ to $+\infty$. Then the integral in Eq. (108) is reduced to a table one,[35] and yields

$$W_\Sigma(t) = \frac{4|A_{nn'}|^2 \rho_n t}{\hbar} \int_{-\infty}^{+\infty} \frac{1}{\xi^2} \sin^2 \frac{\xi}{2} d\xi = \frac{4|A_{nn'}|^2 \rho_n t}{\hbar} \frac{\pi}{2} \equiv \Gamma t , \tag{6.110}$$

where the constant

Golden Rule

$$\boxed{\Gamma = \frac{2\pi}{\hbar}|A_{nn'}|^2 \rho_n} \tag{6.111}$$

is called the *transition rate*.[36] This is one of the most famous and useful results of quantum mechanics, its *Golden Rule*[37], which deserves much discussion.

First of all, let us reproduce the reasoning already used in Sec. 2.5 to show that the meaning of the rate Γ is much deeper than Eq. (110) seems to imply. Indeed, due to the conservation of the total probability, $W_{n'} + W_\Sigma = 1$, we can rewrite that equation as

$$\dot{W}_{n'}\big|_{t=0} = -\Gamma. \tag{6.112}$$

Evidently, this result cannot be true for all times, otherwise the probability $W_{n'}$ would become negative. The reason for this apparent contradiction is that Eq. (110) was obtained in the assumption that initially, the system was completely on level n': $W_{n'}(0) = 1$. Now, if at the initial moment the value of $W_{n'}$ is different, the result (110) has to be multiplied by that number, due to the linear relation (88) between da_n/dt and $a_{n'}$. Hence, instead of Eq. (112), we get a differential equation similar to Eq. (2.159),

$$\dot{W}_{n'}\big|_{t\geq 0} = -\Gamma W_{n'}, \tag{6.113}$$

which, for a time-independent Γ, has the evident solution,

Initial occupancy's decay

$$\boxed{W_{n'}(t) = W_{n'}(0)e^{-\Gamma t}}, \tag{6.114}$$

describing the exponential decay of the initial state's occupancy, with the time constant $\tau = 1/\Gamma$.

I am inviting the reader to review this fascinating result again: by the summation of *periodic oscillations* (102) over many levels n, we have got an *exponential decay* (114) of the probability. This trick becomes possible because the effective range ΔE_n of the state energies E_n giving substantial

[35] See, e.g., MA Eq. (6.12).

[36] In some texts, the density of states in Eq. (111) is replaced with a formal expression $\Sigma_n \delta(E_n - E_{n'} - \hbar\omega)$. Indeed, applied to a finite energy interval ΔE_n with $\Delta n >> 1$ levels, it gives the same result: $\Delta n \equiv (dn/dE_n)\Delta E_n \equiv \rho_n \Delta E_n$. Such replacement may be technically useful in some cases, but is incorrect for $\Delta n \sim 1$, and hence should be used with the utmost care, so for most applications, the more explicit form (111) is preferable.

[37] Sometimes Eq. (111) is called "Fermi's Golden Rule". This is rather unfair, because this result had been developed mostly by the same P. A. M. Dirac in 1927, and Enrico Fermi's role was not much more than advertising it, under the name of "Golden Rule No. 2", in his influential lecture notes on nuclear physics that were published much later, in 1950. (To be fair to Fermi, he has never tried to pose as the Golden Rule's author.)

contributions to the integral (108), shrinks with time: $\Delta E_n \sim \hbar/t$.[38] However, since most of the decay (114) takes place within the time interval of the order of $\tau \equiv 1/\Gamma$, the range of the participating final energies may be estimated as

$$\Delta E_n \sim \frac{\hbar}{\tau} \equiv \hbar\Gamma . \tag{6.115}$$

This estimate is very instrumental for the formulation of conditions of the Golden Rule's validity. First, we have assumed that the matrix elements of the perturbation and the density of states are independent of the energy within the interval (115). This gives the following requirement

$$\Delta E_n \sim \hbar\Gamma << E_n - E_{n'} \sim \hbar\omega , \tag{6.116}$$

Second, for the transfer from the sum (107) to the integral (108), we need the number of states within that energy interval, $\Delta N_n = \rho_n \Delta E_n$, to be much larger than 1. Merging Eq. (116) with Eq. (92) for all the energy levels $n'' \neq n, n'$ not participating in the resonant transition, we may summarize all conditions of the Golden Rule validity as

$$\rho_n^{-1} << \hbar\Gamma << \hbar\left|\omega \pm \omega_{n'n''}\right| . \tag{6.117}$$

Golden Rule's validity

(The reader may ask whether I have neglected the condition expressed by the first of Eqs. (106). However, for $\Delta_{nn'} \sim \Delta E_n/\hbar \sim \Gamma$, this condition is just $|A_{nn'}|^2 << (\hbar\Gamma)^2$, so plugging it into Eq. (111),

$$\Gamma << \frac{2\pi}{\hbar}(\hbar\Gamma)^2 \rho_n , \tag{6.118}$$

and canceling one Γ and one $\hbar$, we see that it coincides with the first relation in Eq. (117) above.)

Let us have a look at whether these conditions may be satisfied in practice, at least in some cases. For example, let us consider the optical ionization of an atom, with the released electron confined in a volume of the order of 1 $cm^3 \equiv 10^{-6}$ m^3. According to Eq. (1.90), with E of the order of the atomic ionization energy $E_n - E_{n'} = \hbar\omega \sim 1$ eV, the density of electron states in that volume is of the order of 10^{21} 1/eV, while the right-hand side of Eq. (117) is of the order of $E_n \sim 1$ eV. Thus the conditions (117) provide an approximately 20-orders-of-magnitude range for acceptable values of $\hbar\Gamma$. This illustration should give the reader a taste of why the Golden Rule is applicable to so many situations.

The physical picture of the initial state's decay is also very important. According to Eq. (114), the external excitation transfers the system into the continuous spectrum of levels n, and it never comes back to the initial level n'. However, it was derived from the quantum mechanics of Hamiltonian systems, whose equations are invariant with respect to time reversal.[39] This paradox is a result of our generalization (113) of the exact result (112) This trick, breaking the time-reversal symmetry, is absolutely adequate for the physics under study. Indeed, some gut feeling of the physical sense of the resulting irreversibility may be obtained from the following observation. As Eq. (1.86) illustrates, the distance between the adjacent orbital energy levels tends to zero only if the system's size goes to infinity. This means that our assumption of the continuous energy spectrum of the finial states n

[38] This is one more appearance of the "energy-time uncertainty relation", which was discussed in Sec. 2.5.

[39] This situation is similar to the irreversible increase of entropy of macroscopic systems, despite the fact that their microscopic components obey reversible laws of motion, which is postulated in thermodynamics and explained in statistical physics – see, e.g., SM Secs. 1.2 and 2.2.

essentially requires these states to be broadly extended in space – being either fully free or virtually free de Broglie waves. Thus the Golden Rule corresponds to the (physically justified) assumption that in an infinitely large system, the traveling de Broglie waves excited by a local source and propagating outward from it, would never come back, and even if they did, unpredictable phase shifts introduced by minor uncontrollable perturbations on their way would never allow them to sum up in the coherent way necessary to bring the system back into the initial state *n'*. (This is essentially the same situation that was discussed, for a particular 1D wave-mechanical system, in Sec. 2.5.)

To get a feeling of the Golden Rule at work, let us apply it to the following simple problem – which is a toy model of the photoelectric effect, briefly discussed in Sec. 1.1(ii). A 1D particle is initially trapped in the ground state of a narrow potential well described by Eq. (2.158):

$$U(x) = -\mathcal{W}\delta(x), \qquad \text{with } \mathcal{W} > 0. \tag{6.119}$$

Let us calculate the rate Γ of the particle's "ionization" (i.e. its excitation into a group of extended, delocalized states) by a weak classical sinusoidal force of amplitude F_0 and frequency ω, suddenly turned on at some instant, say $t = 0$.

As a reminder, the initial localized state (in our current notation, *n'*) of such a particle was already found in Sec. 2.6:

$$\psi_{n'}(x) = \kappa^{1/2}\exp\{-\kappa|x|\}, \qquad \text{with } \kappa \equiv \frac{m\mathcal{W}}{\hbar^2}, \quad E_{n'} \equiv -\frac{\hbar^2\kappa^2}{2m} = -\frac{m\mathcal{W}^2}{2\hbar^2}. \tag{6.120}$$

The final, extended states *n*, with a continuous spectrum, for this problem exist only at energies $E_n > 0$, so the excitation rate is different from zero only for frequencies

$$\omega > \omega_{\min} \equiv \frac{|E_{n'}|}{\hbar} = \frac{m\mathcal{W}^2}{2\hbar^3}. \tag{6.121}$$

The weak sinusoidal force may be described by the following perturbation Hamiltonian,

$$\hat{H}^{(1)} = -F(t)\hat{x} = -F_0\hat{x}\cos\omega t \equiv -\frac{F_0}{2}\hat{x}\left(e^{i\omega t} + e^{-i\omega t}\right), \qquad \text{for } t > 0, \tag{6.122}$$

so according to Eq. (86), which serves as the amplitude operator's definition, in this case

$$\hat{A} = \hat{A}^{\dagger} = -\frac{F_0}{2}\hat{x}. \tag{6.123}$$

The matrix elements $A_{nn'}$ that participate in Eq. (111) may be readily calculated in the coordinate representation:

$$A_{nn'} = \int_{-\infty}^{+\infty}\psi_n^*(x)\hat{A}(x)\psi_{n'}(x)dx = -\frac{F_0}{2}\int_{-\infty}^{+\infty}\psi_n^*(x)x\psi_{n'}(x)dx. \tag{6.124}$$

Since, according to Eq. (120), the initial $\psi_{n'}$ is a *symmetric* function of *x*, any non-vanishing contributions to this integral are given only by *antisymmetric* functions $\psi_n(x)$, proportional to $\sin k_n x$, with the wave number k_n related to the final energy by the well-familiar equality (1.89):

$$\frac{\hbar^2 k_n^2}{2m} = E_n. \tag{6.125}$$

As we know from Sec. 2.6 (see in particular Eq. (2.167) and its discussion), such antisymmetric functions, with $\psi_n(0) = 0$, are not affected by the zero-centered delta-functional potential (119), so their density ρ_n is the same as that in completely free space, and we could use Eq. (1.93). However, since that relation was derived for traveling waves, it is more prudent to repeat its derivation for standing waves, confining them to an artificial segment $[-l/2, +l/2]$ – long in the sense

$$k_n l, \kappa l >> 1, \tag{6.126}$$

so it does not affect the initial localized state and the excitation process. Then the confinement requirement $\psi_n(\pm l/2) = 0$ immediately yields the condition $k_n l/2 = n\pi$, so Eq. (1.93) is indeed valid, but only for positive values of k_n, because $\sin k_n x$ with $k_n \to -k_n$ does not describe an independent standing-wave eigenstate. Hence the final state density is

$$\rho_n \equiv \frac{dn}{dE_n} \equiv \frac{dn}{dk_n} \Big/ \frac{dE_n}{dk_n} = \frac{l}{2\pi} \Big/ \frac{\hbar^2 k_n}{m} \equiv \frac{lm}{2\pi\hbar^2 k_n}. \tag{6.127}$$

It may look troubling that the density of states depends on the artificial segment's length l, but the same l also participates in the final wavefunctions' normalization factor,[40]

$$\psi_n = \left(\frac{2}{l}\right)^{1/2} \sin k_n x, \tag{6.128}$$

and hence in the matrix element (124):

$$A_{nn'} = -\frac{F_0}{2}\left(\frac{2\kappa}{l}\right)^{1/2} \int_{-l/2}^{+l/2} \sin k_n x \, e^{-\kappa|x|} x dx. = -\frac{F_0}{2i}\left(\frac{2\kappa}{l}\right)^{1/2} \left(\int_0^{l/2} e^{(ik_n - \kappa)x} x dx - \int_0^{l/2} e^{-(ik_n + \kappa)x} x dx \right). \tag{6.129}$$

These two integrals may be readily worked out by parts. Taking into account that due to the condition (126), their upper limits may be extended to ∞, the result is

$$A_{nn'} = -\left(\frac{2\kappa}{l}\right)^{1/2} F_0 \frac{2k_n\kappa}{\left(k_n^2 + \kappa^2\right)^2}. \tag{6.130}$$

Note that the matrix element is a smooth function of k_n (and hence of E_n), so an important condition of the Golden Rule, the virtual constancy of $A_{nn'}$ on the interval $\Delta E_n \sim \hbar\Gamma << E_n$, is satisfied. So, the general Eq. (111) is reduced, for our problem, to the following expression:

$$\Gamma = \frac{2\pi}{\hbar}\left[\left(\frac{2\kappa}{l}\right)^{1/2} F_0 \frac{2k_n\kappa}{(k_n^2 + \kappa^2)^2}\right]^2 \frac{lm}{2\pi\hbar^2 k_n} \equiv \frac{8F_0^2 m k_n \kappa^3}{\hbar^3 (k_n^2 + \kappa^2)^4}, \tag{6.131}$$

which is independent of the artificially introduced l – thus justifying its use.

Note that due to the above definitions of k_n and κ, the expression in the parentheses in the denominator of the last expression does not depend on the potential well's "weight" $\mathcal{W}$, and is a function of only the excitation frequency ω (and the particle's mass):

[40] The normalization to infinite volume, by using Eq. (4.263), is also possible, but physically less transparent.

$$\frac{\hbar^2\left(k_n^2+\kappa^2\right)}{2m}=E_n-E_{n'}=\hbar\omega. \tag{6.132}$$

As a result, Eq. (131) may be recast simply as

$$\hbar\Gamma=\frac{F_0^2\mathcal{W}^3k_n}{2(\hbar\omega)^4}. \tag{6.133}$$

What is hidden here is that k_n, defined by Eq. (125) with $E_n = E_{n'} + \hbar\omega$, is a function of the external force's frequency, changing as $\omega^{1/2}$ at $\omega >> \omega_{min}$ (so Γ drops as $\omega^{-7/2}$ at $\omega \to \infty$), and as $(\omega - \omega_{min})^{1/2}$ when ω approaches the "red boundary" (121) of the ionization effect, so $\Gamma \propto (\omega - \omega_{min})^{1/2} \to 0$ in that limit as well.

A conceptually very similar but a bit more involved analysis of this effect in a more realistic 3D case, namely the hydrogen atom's ionization by an optical wave, is left for the reader's exercise.

6.7. Golden Rule for step-like perturbations

Now let us reuse some of our results for a perturbation being turned on at $t = 0$, but after that time-independent:

Step-like perturbation

$$\hat{H}^{(1)}(t)=\begin{cases}0, & \text{for } t<0,\\ \hat{H}=\text{const}, & \text{for } t\ge 0.\end{cases} \tag{6.134}$$

A superficial comparison of this equality and the former Eq. (86) seems to indicate that we may use all our previous results, taking $\omega = 0$ and replacing $\hat{A}+\hat{A}^\dagger$ with $\hat{H}^{(1)}$. However, that conclusion (which would give us a wrong factor of 2 in the result) does not take into account the fact that when analyzing both the two-level approximation in Sec. 5 and the Golden Rule in Sec. 6, we have dropped the second (non-resonant) term in Eq. (90). In our current case (134), with $\omega = 0$, there is no such difference between these terms. This is why it is more prudent to use the general Eq. (84),

$$i\hbar\dot{a}_n=\sum_{n'}a_{n'}H_{nn'}e^{i\omega_{nn'}t}, \tag{6.135}$$

in which the matrix element of the perturbation is now time-independent at $t > 0$. We see that it is formally equivalent to Eq. (88) with only the first (resonant) term kept, provided that we make the following replacements:

$$\hat{A}\to\hat{H},\quad \Delta_{nn'}\equiv\omega-\omega_{nn'}\to-\omega_{nn'}. \tag{6.136}$$

Let us use this equivalency to consider the results of coupling between a discrete-energy state n', to which the particle is initially placed, and a dense group of states with a quasi-continuum spectrum, in the same energy range. Figure 11a shows an example of such a system: a particle is initially (say, at $t = 0$) placed into a potential well separated by a penetrable potential barrier from a formally infinite region with a continuous energy spectrum. Let me hope that the physical discussion in the last section makes the outcome of such an experiment evident: the particle will gradually and irreversibly tunnel out of the well, so the probability $W_{n'}(t)$ of its still residing in the well will decay in accordance with Eq. (114). The rate of this decay may be found by making the replacements (136) in Eq. (111):

$$\Gamma = \frac{2\pi}{\hbar}\left|H_{nn'}\right|^2 \rho_n, \tag{6.137}$$

where the states n and n' now have virtually the same energy.[41]

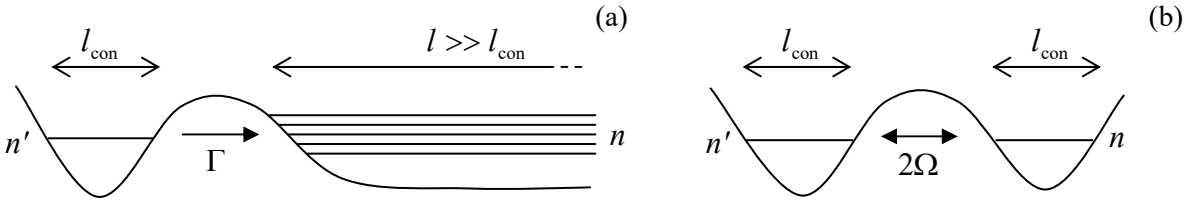


Fig. 6.11. Tunneling from a discrete-energy state n': (a) to a state continuum, and (b) to another discrete-energy state n.

It is very informative to compare this result, semi-quantitatively, with Eq. (105) for a symmetric ($E_n = E_{n'}$) system of two potential wells separated by a similar potential barrier – see Fig. 11b. For the symmetric case, i.e. $\xi = 0$, Eq. (105) is reduced to simply

$$\Omega = \frac{1}{\hbar}\left|H_{nn'}\right|_{con}. \tag{6.138}$$

Here I have used the index "con" (from "confinement") to emphasize that this matrix element is somewhat different from the one participating in Eq. (137), even if the potential barriers are similar. Indeed, in the latter case, the matrix element,

$$H_{nn'} \equiv \langle n|\hat{H}|n'\rangle = \int \psi_{n'}^* \hat{H} \psi_n dx, \tag{6.139}$$

has to be calculated for two wavefunctions ψ_n and $\psi_{n'}$ confined to spatial intervals of the same scale l_{con}, while in Eq. (137), the wavefunctions ψ_n are extended over a much larger distance $l >> l_{con}$ – see Fig. 11. As Eq. (128) tells us, in the 1D model this means an additional small factor of the order of $(l_{con}/l)^{1/2}$. Now using Eq. (128) as a crude but suitable model for the final-state wavefunctions, we arrive at the following estimate, which is independent of the artificially introduced length l:

$$\hbar\Gamma \sim 2\pi\left|H_{nn'}\right|_{con}^2 \frac{l_{con}}{l}\rho_n \sim 2\pi\left|H_{nn'}\right|_{con}^2 \frac{l_{con}}{l}\frac{lm}{2\pi\hbar^2 k_n} \sim \frac{\left|H_{nn'}\right|_{con}^2}{\Delta E_{n'}} \equiv \frac{(\hbar\Omega)^2}{\Delta E_{n'}}, \tag{6.140}$$

where $\Delta E_{n'} \sim \hbar^2/ml^2_{con}$ is the scale of the distances between the adjacent eigenenergies of the particle in an unperturbed potential well. Since the condition of validity of Eq. (138) is $\hbar\Omega << \Delta E_{n'}$, we see that

$$\hbar\Gamma \sim \frac{\hbar\Omega}{\Delta E_n}\hbar\Omega << \hbar\Omega.. \tag{6.141}$$

This (sufficiently general[42]) perturbative result confirms the conclusion of a more particular analysis carried out at the end of Sec. 2.6: the rate of the (irreversible) quantum tunneling into a state continuum is always *much lower* than the frequency of (reversible) quantum oscillations between

[41] The condition of validity of Eq. (137) is again given by Eq. (117), just with $\omega = 0$ in the upper limit for Γ.

[42] It is straightforward to verify that the estimate (141) is valid for similar problems of any spatial dimensionality, not just for the 1D case we have analyzed.

discrete states separated with the same potential barrier – at least for the case when both are much lower than $\Delta E_n/\hbar$, so the perturbation theory is valid. A very handwaving interpretation of this result is that the particle performs the quantum oscillations between the confined state in the well and the space-extended states behind the barrier many times before finally "deciding to perform" an irreversible transition into the unconfined continuum. This qualitative picture is consistent with experimentally observable effects of dispersive electromagnetic environments on electron tunneling.[43]

Let me conclude this section (and this chapter) with the application of Eq. (137) to a very important case that will provide a smooth transition to the next chapter's topic. Consider a *composite* system consisting of two *component* systems, *a* and *b*, with the energy spectra sketched in Fig. 12.

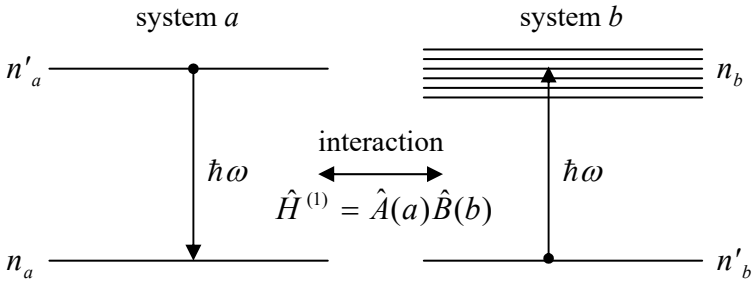


Fig. 6.12. Energy relaxation in system *a* due to its weak coupling to system *b* (which serves as the environment of *a*).

Let the systems be completely independent initially. The independence means that in the absence of their coupling, the total Hamiltonian of the system may be represented as a sum of two operators:

$$\hat{H}^{(0)} = \hat{H}_a(a) + \hat{H}_b(b), \tag{6.142}$$

where arguments *a* and *b* symbolize the non-overlapping sets of the degrees of freedom of the two systems. Such operators, belonging to their individual, different Hilbert spaces, naturally commute. Similarly, the eigenkets of the system may be naturally factored as

Direct product

$$|n\rangle = |n_a\rangle \otimes |n_b\rangle . \tag{6.143}$$

The *direct product* sign ⊗ is used here (and below) to denote the formation of a joint ket-vector from the kets of the independent systems, belonging to different Hilbert spaces. Evidently, the order of operands in such a product may be changed at will. As a result, its eigenenergies separate into a sum, just as the Hamiltonian (142) does:

$$\hat{H}^{(0)}|n\rangle = \left(\hat{H}_a + \hat{H}_b\right)|n_a\rangle \otimes |n_b\rangle \equiv \left(\hat{H}_a|n_a\rangle\right)\otimes|n_b\rangle + \left(\hat{H}_b|n_b\rangle\right)\otimes|n_a\rangle = \left(E_{na} + E_{nb}\right)|n\rangle. \tag{6.144}$$

In such composite systems, the relatively weak interaction of its components may be usually represented as a product of two Hermitian operators, each depending only on the degrees of freedom of one component system:

$$\hat{H}^{(1)} = \hat{A}(a)\hat{B}(b) . \tag{6.145}$$

A very common example of such an interaction is the electric-dipole interaction between an atomic-scale system (with a linear size of the order of the Bohr radius $r_B \sim 10^{-10}$ m) and the electromagnetic field at optical frequencies $\omega \sim 10^{16}$ s^{-1}, with the wavelength $\lambda = 2\pi c/\omega \sim 10^{-6}$ m >> r_B:

[43] See, e.g., P. Delsing *et al.*, *Phys. Rev. Lett.* **63**, 1180 (1989).

$$\hat{H}^{(1)} = -\hat{\mathbf{d}} \cdot \hat{\mathscr{E}}, \qquad \text{with } \hat{\mathbf{d}} \equiv \sum_k q_k \hat{\mathbf{r}}_k \,, \tag{6.146}$$

where the dipole electric moment $\mathbf{d}$ depends only on the positions $\mathbf{r}_k$ of the charged particles (numbered with index k) of the atomic system, while the electric field $\mathscr{E}$ is a function of only the electromagnetic field's degrees of freedom – to be discussed in Chapter 9 below.

Returning to the general situation shown in Fig. 12, if the component system a was initially in an excited state n'_a, the interaction (145), turned on at some moment of time, may bring it into another discrete state n_a of lower energy – for example, the ground state. In the process of this transition, the released energy, in the form of an energy quantum

$$\hbar\omega \equiv E_{n'a} - E_{na}, \tag{6.147}$$

is picked up by system b:

$$E_{nb} = E_{n'b} + \hbar\omega \equiv E_{n'b} + \left(E_{n'a} - E_{na}\right), \tag{6.148}$$

so the total energy $E = E_a + E_b$ of the system does not change. (If the states n_a and n'_b are the ground states of the component systems, as they are in most applications of this analysis, and we take the ground state energy $E_g = E_{na} + E_{n'b}$ of the composite system for the reference, then Eq. (148) gives merely $E_{nb} = E_{n'a}$.) If the final state n_b of system b is inside a state group with a quasi-continuous energy spectrum (Fig. 12), the process has the exponential character (114)[44] and may be interpreted as the effect of *energy relaxation* of system a, with the released energy quantum $\hbar\omega$ *absorbed* by system b.

If the relaxation rate Γ is sufficiently low, it may be described by the Golden Rule (137). Since the perturbation (145) does not depend on time explicitly, and the total energy E does not change, this relation, with the account of Eqs. (143) and (145), takes the form

$$\Gamma = \frac{2\pi}{\hbar}\left|A_{nn'}\right|^2 \left|B_{nn'}\right|^2 \rho_n, \quad \text{where } A_{nn'} \equiv \left\langle n_a \middle| \hat{A} \middle| n'_a \right\rangle, \quad \text{and } B_{nn'} = \left\langle n_b \middle| \hat{B} \middle| n'_b \right\rangle, \tag{6.149}$$

Golden Rule for coupled systems

where ρ_n is the density of the final states of system b at the relevant energy (147).[45] In particular, Eq. (149), with the dipole Hamiltonian (146), will enable us to readily calculate, in Chapter 9, the natural linewidth of atomic electric-dipole transitions.

Instead, I will now proceed to a general discussion of the effects of quantum systems' interaction with their environment, toward which the situation shown in Fig. 12 provides a clear conceptual path. Indeed, in this case the transition from the Hamiltonian (and hence reversible) quantum dynamics of the whole composite system $a + b$ to the Golden-Rule-governed (and hence irreversible) dynamics of system a has been achieved essentially by following this component system alone, i.e. ignoring the details of the exact state of system b. (As was argued in the previous section, the quasi-continuous spectrum of the latter system essentially requires it to have a large spatial size, so it may be legitimately called the *environment* of the "open" system a.) This is exactly the approach that will be pursued in the next chapter.

[44] This process is *spontaneous*: it starts as soon as either the interaction (145) has been turned on or (if it had been already on) as soon as the system a is placed into the excited state n'_a.

[45] Note that these partial matrix elements may be calculated in the Heisenberg picture as well, because due to the general Eq. (4.149) and the energy balance (147), the additional time dependences of these elements would be proportional to $\exp\{\pm i\omega t\}$, and cancel at their multiplication.

6.8. Exercise problems

6.1. Use Eq. (6.14) of the lecture notes to prove the following general form of the Hellmann-Feynman theorem:[46]

$$\frac{\partial E_n}{\partial \lambda} = \left\langle n \middle| \frac{\partial \hat{H}}{\partial \lambda} \middle| n \right\rangle ,$$

where λ is an arbitrary *c*-number parameter.

6.2. Establish a relation between Eq. (16) and the result of the classical theory of weakly anharmonic ("nonlinear") oscillations at negligible damping.

Hint: You may like to use N. Bohr's reasoning that was discussed in Problem 1.1.

6.3. An additional weak time-independent force F is exerted on a 1D particle that had been placed into a hard-wall potential well

$$U(x) = \begin{cases} 0, & \text{for } 0 < x < a, \\ +\infty, & \text{otherwise.} \end{cases}$$

Calculate, sketch, and discuss the 1st-order perturbation of its ground-state wavefunction.

6.4. A time-independent force $\mathbf{F} = \mu(\mathbf{n}_x y + \mathbf{n}_y x)$, where μ is a small constant, is applied to a 3D isotropic harmonic oscillator of mass m and frequency ω_0, located at the origin. Calculate, in the first order of the perturbation theory, the effect of the force upon the ground-state energy of the oscillator and its lowest excited energy level. How small should the constant μ be for your results to be quantitatively correct?

6.5. A 1D particle of mass m is localized at a narrow potential well that may be approximated with a delta function:

$$U(x) = -\mathcal{W}\delta(x), \qquad \text{with } \mathcal{W} > 0.$$

Calculate the change of its ground state energy by an additional weak time-independent force F, in the first non-vanishing approximation of the perturbation theory. Discuss the limits of validity of this result, taking into account that at $F \neq 0$, the localized state of the particle is metastable.

6.6. Use Eq. (16) to calculate the eigenvalues of the operator $\hat{L}^2$, in the limit $|m| \approx l >> 1$, by purely wave-mechanical means.

Hint: Try the following substitution: $\Theta(\theta) = f(\theta)/\sin^{1/2}\theta$.

6.7. In the lowest non-vanishing order of the perturbation theory, calculate the shift of the ground-state energy of an electrically charged spherical rotor (i.e. a particle of mass $\mathcal{m}$, free to move over a spherical surface of radius R) due to a weak uniform time-independent electric field $\mathcal{E}$.

6.8. Use the perturbation theory to evaluate the effect of a time-independent uniform electric field $\mathcal{E}$ on the ground state energy E_g of a hydrogen atom. In particular:

[46] As a reminder, proof of its wave-mechanics form was the task of Problem 1.7.

(i) calculate the 2nd-order shift of E_g, neglecting the extended unperturbed states with $E > 0$, and bring the result to the simplest analytical form you can,

(ii) find the lower and the upper bounds on the shift, and

(iii) discuss the simplest experimental manifestation of this *quadratic Stark effect*.

6.9. A particle of mass m, with electric charge q, is in its ground s-state with a given energy $E_g < 0$, being localized by a very-short-range, spherically symmetric potential well. Calculate its static electric polarizability.

6.10. In some atoms, the effect of nuclear charge screening by electrons on the motion of each of them may be reasonably well approximated by the replacement of the Coulomb potential (3.190), $U = -C/r$, with the so-called *Hulthén potential*

$$U = -\frac{C/a}{\exp\{r/a\}-1} \to -C\times\begin{cases} 1/r, & \text{for } r << a, \\ \exp\{-r/a\}/a, & \text{for } a << r. \end{cases}$$

Assuming that the effective screening radius a is much larger than $r_0 \equiv \hbar^2/mC$, use the perturbation theory to calculate the energy spectrum of a single particle of mass m, moving in this potential, in the lowest order needed to lift the l-degeneracy of the energy levels.

6.11. In the lowest non-vanishing order of the perturbation theory, calculate the correction to energies of the ground state and all lowest excited states of a hydrogen-like atom/ion, due to the electron's penetration into the nucleus, by modeling it the latter a spinless, uniformly charged sphere of radius $R << r_B/Z$.

6.12. A particle of mass m is placed inside a hard-wall ellipsoid whose surface is described by the equation

$$\frac{x^2+y^2}{a^2}+\frac{z^2}{b^2}=1, \qquad \text{with } b=(1+\varepsilon)a, \quad |\varepsilon| << 1.$$

Calculate its ground-state energy in the 1st order in the small parameter ε, and interpret the result.

6.13. Prove that the relativistic correction operator (48) indeed has only diagonal matrix elements in the basis of unperturbed Bohr atom states (3.200).

6.14. Calculate the lowest-order relativistic correction to the ground-state energy of a 1D harmonic oscillator.

6.15. Use the perturbation theory to calculate the contribution to the magnetic susceptibility χ_m of a dilute gas, that is due to the orbital motion of a single electron inside each gas particle. Spell out your result for a spherically symmetric ground state of the electron, and give an estimate of the magnitude of this *orbital susceptibility*.

6.16. A certain energy level degeneracy is not lifted in the 1st order of the stationary perturbation theory. Calculate its lifting in the 2nd order of the theory. Apply the result to a planar rotor of mass m and radius R, with electric charge q, placed into a weak, uniform, time-independent electric field $\mathcal{E}$.

6.17.* The Hamiltonian of a quantum system is slowly changed over time.

(i) Develop a theory of quantum transitions in the system, and spell out its result in the 1st approximation in the speed of the change.

(ii) Use this approximation to calculate the probability that a finite-time pulse of a slowly changing force $F(t)$ drives a 1D harmonic oscillator, initially in its ground state, into an excited state.

(iii) Compare the last result with the exact one.

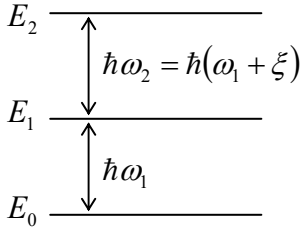
6.18. Use the single-particle model to calculate the complex electric permittivity $\varepsilon(\omega)$ of a dilute gas of similar atoms, due to their induced electric polarization by a weak external ac field, for a field frequency ω very close to one of the quantum transition frequencies $\omega_{nn'}$. Based on the result, calculate and estimate the absorption cross-section of each atom.

Hint: In the single-particle model, the atom's properties are determined by Z similar, non-interacting electrons, each moving in a similar static attracting potential, generally different from the Coulomb one, because it is contributed not only by the nucleus but also by other electrons.

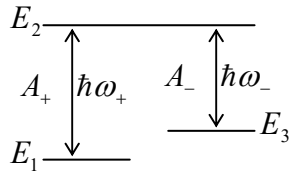
6.19. Use the solution of the previous problem to generalize the expression for the London dispersion force between two atoms (whose calculation in the harmonic oscillator model was the subject of Problems 3.20 and 5.20) to the single-particle model with an arbitrary energy spectrum.

6.20. Use the solution of the previous problem to calculate the potential energy of the interaction of two hydrogen atoms, both in their ground state, separated by distance $r >> r_B$.

6.21. In a certain quantum system, distances between the three lowest energy levels are slightly different – see the figure on the right ($|\xi| << \omega_{1,2}$). Assuming that the involved matrix elements of the perturbation Hamiltonian are known and are all proportional to the external ac field's amplitude, find the time necessary to populate the first excited level almost completely (with a given precision $\varepsilon << 1$), by using the Rabi oscillation effect, if at $t = 0$, the system is in its ground state. Spell out your result for a weakly anharmonic 1D oscillator.



6.22.* Analyze the possibility of a slow transfer of a system from one of its energy levels to another one (in the figure on the right, from level 1 to level 3), by using the scheme shown in that figure, in which the monochromatic external excitation amplitudes A_+ and A_- may be slowly changed at will.



6.23. A weak external force pulse $F(t)$, of a finite time duration, is applied to the particle in a system with a discrete energy spectrum, which initially was in its ground state.

(i) Derive, in the lowest non-vanishing order of the perturbation theory, a formula for the probability that the pulse drives the particle into its lowest excited state.

(ii) Specify this formula for a 1D harmonic oscillator and compare the result with the exact solution of the problem.

(iii) Spell out the perturbative result for the Gaussian-shaped waveform $F(t) = F_0 \exp\{-t^2/\tau^2\}$ and analyze its dependence on the scale τ of the pulse duration.

6.24. A spatially uniform but time-dependent external electric field $\mathscr{E}(t)$ is applied, starting from $t = 0$, to a charged planar rotor, initially in its ground state.

(i) Calculate, in the lowest non-vanishing order in the field's strength, the probability that by a certain time $t > 0$, the rotor is in its m^{th} excited state.

(ii) Spell out and analyze your results for a constant-magnitude field rotating, with a constant angular velocity ω, within the rotor's plane.

(iii) Do the same for a monochromatic field of frequency ω, with a fixed direction.

6.25. A heavy relativistic particle, with electric charge $q = Ze$, flies by a hydrogen atom, initially in its ground state, with an impact parameter b within the range $r_B << b << r_B/\alpha$, where $\alpha \approx 1/137$ is the fine structure constant. Calculate the total probability of the atom's transition to one of its lowest excited states.

6.26. Develop a general theory of quantum excitations of the higher levels of a discrete-spectrum system, initially in the ground state, by a weak time-dependent perturbation, up to the 2nd order. Spell out and discuss the result for the case of monochromatic excitation, with a nearly perfect tuning of its frequency ω to *the half* of a certain quantum transition frequency $\omega_{n0} \equiv (E_n - E_0)/\hbar$.

6.27. A particle of mass m is initially in a localized ground state, with energy $E_g < 0$, of a very-short-range, spherically symmetric potential well. Calculate the rate of its delocalization by an applied classical force $\mathbf{F}(t) = \mathbf{n}F_0\cos\omega t$ with a time-independent direction $\mathbf{n}$.

6.28.* Calculate the rate of ionization of a hydrogen atom, initially in its ground state, by a classical, linearly polarized electromagnetic wave with an electric field's amplitude $\mathscr{E}_0$, and a frequency ω within the range

$$\frac{\hbar}{m_e r_B^2} << \omega << \frac{c}{r_B}.$$

Recast your result in terms of the cross-section of electromagnetic wave absorption. Discuss briefly what changes of the theory would be necessary if either of the above conditions had been violated.

6.29.* Use the quantum-mechanical Golden Rule to derive the general expression for the electric current I through a weak tunnel junction between two conductors, biased with dc voltage V, treating the conductors as degenerate Fermi gases of electrons with negligible direct interaction. Simplify the result in the low-voltage limit.

Hint: The electric current flowing through a weak tunnel junction is so low that it does not substantially perturb the electron states inside each conductor.

6.30.* Generalize the result of the previous problem to the case when a weak tunnel junction is biased with voltage $V(t) = V_0 + A\cos\omega t$, with $\hbar\omega$ generally comparable with eV_0 and eA.

6.31.* Use the quantum-mechanical Golden Rule to derive the Landau-Zener formula (2.257).

Chapter 7. Open Quantum Systems

This chapter discusses the effects of a weak interaction of a quantum system with its environment. Some part of this material is on the fine line between quantum mechanics and (quantum) statistical physics. Here I will only cover those aspects of the latter field[1] *that are of key importance for the major goals of this course, including the discussion of quantum measurements in Chapter 10.*

7.1. Open systems, and the density matrix

All the way until the last part of the previous chapter, we have discussed quantum systems isolated from their environment. Indeed, from the very beginning, we have assumed that we are dealing with the statistical ensembles of systems as similar to each other as only allowed by the laws of quantum mechanics. Each member of such an ensemble, called *pure* (or *coherent*), may be described by the same state vector $|\alpha\rangle$ – in the wave mechanics case, by the same wavefunction Ψ_α. Even if we deal with a composite system, say, a two-components system like those discussed at the end of the previous chapter, we still can consider its coherent states of the type

$$|\alpha\rangle = \sum_n \alpha_n |n\rangle = \sum_n \alpha_n |n_a\rangle \otimes |n_b\rangle, \tag{7.1}$$

with a *unique* correspondence between the pure states n_a and n_b of the two subsystems.

However, in many important cases, our knowledge of a quantum system's state is less complete.[2] These cases fall into two categories. The first case is when a relatively simple quantum system *s* of our interest (say, an electron or an atom) is in substantial contact with its environment *e* – here understood in the most general sense, say, as all the whole Universe less the system *s* – see Fig. 1. Then there is virtually no chance of making two or more experiments with exactly the same composite system because that would imply a repeated preparation of the whole environment (including the experimenter :-) in a certain quantum state – a rather challenging task, to put it mildly. Then it makes much more sense to consider a statistical ensemble of another kind – a *mixed ensemble*, with *random* states of the environment, though possibly with its macroscopic parameters (e.g., temperature, pressure, etc.) known with high precision. Such ensembles will be the focus of the analysis in this chapter.

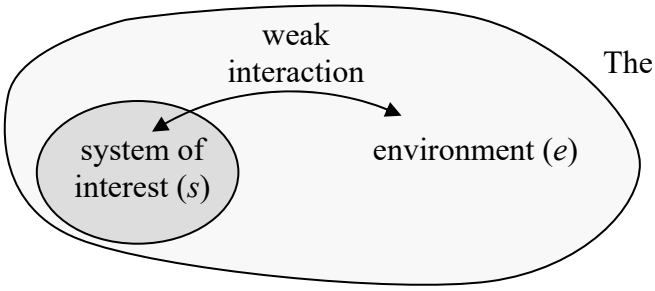


Fig. 7.1. A quantum system and its environment (VERY schematically :-).

[1] A broader discussion of statistical mechanics and physical kinetics, including those of quantum systems, may be found in the SM part of this series.

[2] Actually, no system, possibly apart from our Universe as a whole (see below), is ever *exactly* coherent, though in many cases, deviations from the coherence may be ignored with acceptable accuracy.

© K. Likharev

Much of this analysis will also pertain to another category of cases – when the system of our interest is, at present, isolated from its environment with acceptable precision, but our knowledge of its state is still incomplete for some reason. Most typically, the system could be in contact with its environment at earlier times. So, this second category of cases may be considered as a particular case of the first one, and may be described by the results of its analysis, with certain simplifications – which will be spelled out in appropriate places of my narrative.

In classical physics, the analysis of mixed statistical ensembles is based on the notion of the probability W of each detailed ("microscopic") state of the system of interest.[3] Let us see how such an ensemble may be described in quantum mechanics. In the case when the coupling between the system of our interest and its environment is so weak that they may be clearly separated, we can still use state vectors of their states, defined in completely different Hilbert spaces. Then the most general quantum state of the whole Universe, still assumed to be pure,[4] may be described as the following linear superposition:

$$|\alpha\rangle = \sum_{j,k} \alpha_{jk} |s_j\rangle \otimes |e_k\rangle . \qquad (7.2)$$

Universe: quantum state

The "only" difference of such a state from the superposition described by Eq. (1), is that there is no one-to-one correspondence between the states of our system and its environment. In other words, a certain quantum state s_j of the system of interest may coexist with different states e_k of its environment. This is the quantum-mechanical description of a *mixed state* of system s.[5]

Of course, the huge size of the Hilbert space of the environment, i.e. of the number of the $|e_k\rangle$ factors in the superposition (2), strips us of any practical opportunity to make direct calculations using that sum. For example, according to the basic Eq. (4.125), in order to find the expectation value of an arbitrary observable A in the state (2), we would need to calculate all long brackets in the sum

$$\langle A \rangle = \langle \alpha | A | \alpha \rangle \equiv \sum_{j,j';k,k'} \alpha^*_{jk} \alpha_{j'k'} \langle e_k | \otimes \langle s_j | \hat{A} | s_{j'} \rangle \otimes | e_{k'} \rangle . \qquad (7.3)$$

Even if we assume that each of the sets $\{s\}$ and $\{e\}$ is full and orthonormal, Eq. (3) still includes a double sum over the enormous basis state set of the environment!

However, let us consider a limited, but the most important subset of all operators – those of *intrinsic* observables, which depend only on the degrees of freedom of the system of our interest (s). These operators do not act upon the environment's degrees of freedom, and hence in Eq. (3), we may move the environment's bra-vectors $\langle e_k|$ over all the way to the ket-vectors $|e_{k'}\rangle$. Assuming, again, that the set of environmental eigenstates is full and orthonormal, Eq. (3) is now reduced to

[3] In systems with a continuum of states, we have to discuss the *probability density* w instead – see below.

[4] Whether this assumption is true is an interesting issue, still being debated (more by philosophers than by physicists), but it is widely believed that its solution is not critical for the validity of the results of this approach to all systems available for our experimentation.

[5] Note also that in this definition, the notion of a mixed state includes pure states as particular cases, while in some texts, this term is limited to states that are *not exactly pure*. Due to the already discussed prevalence of the mixed states (in any definition!) in the world, with the exactly pure states serving only as (sometimes useful and acceptable) abstractions, this is not much of a difference.

$$\langle A\rangle = \sum_{j,j';k,k'} \alpha^*_{jk}\alpha_{j'k'}\langle s_j|\hat{A}|s_{j'}\rangle\langle e_k|e_{k'}\rangle = \sum_{jj'} A_{jj'}\sum_k \alpha^*_{jk}\alpha_{j'k}. \tag{7.4}$$

This is already a big relief because we have "only" a single sum over k, but the main trick is still ahead. After the summation over k, the second sum in the last form of Eq. (4) is some function w of the indices j and j', so according to Eq. (4.96), this relation may be represented as

Intrinsic observable: expectation value

$$\boxed{\langle A\rangle = \sum_{jj'} A_{jj'}w_{j'j} \equiv \mathrm{Tr}(\mathrm{Aw}),} \tag{7.5}$$

where the matrix w, with the elements

Density matrix: definition

$$\boxed{w_{j'j} \equiv \sum_k \alpha^*_{jk}\alpha_{j'k}, \quad \text{i.e. } w_{jj'} \equiv \sum_k \alpha_{jk}\alpha^*_{j'k},} \tag{7.6}$$

is called the *density matrix* of the system.[6] Most importantly, Eq. (5) shows that the knowledge of this matrix allows the calculation of the expectation value of *any* intrinsic observable A (and, according to the general Eqs. (1.33)-(1.34), its r.m.s. fluctuation as well, if needed), even for the very general state (2). This is why let us have a good look at the density matrix.

First of all, we know from the general discussion in Chapter 4, fully applicable to the pure state (2), the expansion coefficients in superpositions of this type may be always expressed as short brackets of the type (4.40); in our current case, we may write

$$\alpha_{jk} = \left(\langle e_k| \otimes \langle s_j|\right)|\alpha\rangle. \tag{7.7}$$

Plugging this expression into Eq. (6), we get

$$w_{jj'} \equiv \sum_k \alpha_{jk}\alpha^*_{j'k} = \langle s_j| \otimes \left(\sum_k \langle e_k|\alpha\rangle\langle\alpha|e_k\rangle\right) \otimes |s_{j'}\rangle = \langle s_j|\hat{w}|s_{j'}\rangle. \tag{7.8}$$

We see that from the point of our system (i.e. in its Hilbert space whose basis states may be numbered by the index j only), the density matrix is indeed just the matrix of some construct,[7]

Density operator: definition

$$\boxed{\hat{w} \equiv \sum_k \langle e_k|\alpha\rangle\langle\alpha|e_k\rangle,} \tag{7.9}$$

which is called the *density* (or "statistical") *operator*. As it follows from the definition (9), in contrast to the density matrix, this operator does not depend on the choice of a particular basis s_j – just as all linear operators considered earlier in this course. However, in contrast to them, the density operator *does* depend on the composite system's state α, including the state of the system s as well. Still, in the j-space, it is mathematically just an operator whose matrix elements obey all relations of the bra-ket formalism.

In particular, due to its definition (6), the density operator is Hermitian:

$$w^*_{jj'} = \sum_k \alpha^*_{jk}\alpha_{j'k} = \sum_k \alpha_{j'k}\alpha^*_{jk} = w_{j'j}, \tag{7.10}$$

[6] This notion was introduced in 1927 by John von Neumann.

[7] Note that the "short brackets" in this expression are not c-numbers, because the state α is defined in a larger Hilbert space (of the environment plus the system of interest) than the basis states e_k (of the environment only).

so according to the general analysis of Sec. 4.3, in the Hilbert space of the system s, there should be a certain basis $\{w\}$ in that the matrix of this operator is diagonal:

$$w_{jj'}\big|_{\text{in } w} = w_j \delta_{jj'}. \tag{7.11}$$

Since any operator, in any basis, may be represented in the form (4.59), in the basis $\{w\}$ we may write

$$\hat{w} = \sum_j \left|w_j\right\rangle w_j \left\langle w_j\right|. \tag{7.12}$$

Density operator in w-basis

This expression reminds but is not equivalent to Eq. (4.44) for the identity operator, that has been used so many times in this course, and in the basis w_j has the form

$$\hat{I} = \sum_j \left|w_j\right\rangle\left\langle w_j\right|. \tag{7.13}$$

In order to comprehend the meaning of the coefficients w_j participating in Eq. (12), let us use Eq. (5) to calculate the expectation value of any observable A whose eigenstates coincide with those of the special basis $\{w\}$, and whose matrix is, therefore, diagonal in this basis:

$$\left\langle A\right\rangle = \mathrm{Tr}\,(\mathrm{Aw}) = \sum_{jj'} A_{jj'} w_j \delta_{jj'} = \sum_j A_j w_j\,, \tag{7.14}$$

Expectation value of w_j-compatible variable

where A_j is just the expectation value of the observable A in the state w_j. Hence, to comply with the general Eq. (1.37), the real c-number w_j must have the physical sense of the probability W_j of finding the system in the state j. As the result, we may rewrite Eq. (12) in the form[8]

$$\hat{w} = \sum_j \left|w_j\right\rangle W_j \left\langle w_j\right|. \tag{7.15}$$

In the ultimate case when only one of the probabilities (say, $W_{j''}$) is different from zero,

$$W_j = \delta_{jj''}, \tag{7.16}$$

the system is in a pure (coherent) state $w_{j''}$. Indeed, it is fully described by one ket-vector $|w_{j''}\rangle$, and we can use the general rule (4.86) to represent it in another (arbitrary) basis $\{s\}$ as a coherent superposition

$$\left|w_{j''}\right\rangle = \sum_{j'} U^{\dagger}_{j''j'} \left|s_{j'}\right\rangle = \sum_{j'} U^{*}_{j'j''} \left|s_{j'}\right\rangle, \tag{7.17}$$

where U is the unitary matrix of transform from the basis $\{w\}$ to the basis $\{s\}$. According to Eqs. (11) and (16), in such a pure state, the density matrix is diagonal in the $\{w\}$ basis,

$$w_{jj'}\big|_{\text{in } w} = \delta_{j,j''}\delta_{j',j''}, \tag{7.18a}$$

but not in an arbitrary basis. Indeed, using the general rule (4.92), we get

$$w_{jj'}\big|_{\text{in } s} = \sum_{l,l'} U^{\dagger}_{jl} w_{ll'}\big|_{\text{in } w} U_{lj'} = U^{\dagger}_{jj''} U_{j''j'} = U^{*}_{j''j} U_{j''j'}. \tag{7.18b}$$

To make this result more transparent, let us denote the matrix elements $U_{j''j} \equiv \langle w_{j''}|s_j\rangle$ (which, for a fixed j'', depend on just one index j) as α_j; then

[8] In some textbooks, this relation is taken for the theory's starting point, leaving the physical sense of the density operator obscure.

Density matrix: pure state

$$w_{jj'}\big|_{\text{in } s} = \alpha_j^* \alpha_{j'}, \tag{7.19}$$

so N^2 elements of the whole $N\times N$ matrix are determined by just *one* string of N c-numbers α_j. For example, for a two-level system ($N = 2$),

$$\mathrm{w}\big|_{\text{in } s} = \begin{pmatrix} \alpha_1\alpha_1^* & \alpha_2\alpha_1^* \\ \alpha_1\alpha_2^* & \alpha_2\alpha_2^* \end{pmatrix}. \tag{7.20}$$

We see that the off-diagonal terms are, colloquially, "as large as the diagonal ones", in the following sense:

$$w_{12}w_{21} = w_{11}w_{22}. \tag{7.21}$$

Since the diagonal terms have the sense of the probabilities $W_{1,2}$ to find the system in the corresponding state, we may represent Eq. (20) as

$$\mathrm{w}\big|_{\text{pure state}} = \begin{pmatrix} W_1 & (W_1W_2)^{1/2}e^{i\varphi} \\ (W_1W_2)^{1/2}e^{-i\varphi} & W_2 \end{pmatrix}. \tag{7.22}$$

The physical sense of the (real) constant φ is the phase shift between the coefficients in the linear superposition (17), which represents the pure state $w_{j''}$ in the basis $\{s_{1,2}\}$.

Now let us consider a different statistical ensemble of two-level systems, that includes the member states identical in all aspects (including similar probabilities $W_{1,2}$ in the same basis $s_{1,2}$), besides that the phase shifts φ are random, with the phase's probability uniformly distributed over the trigonometric circle. Then the ensemble averaging is equivalent to the averaging over φ from 0 to 2π,[9] which kills the off-diagonal terms of the density matrix (22), so the matrix becomes diagonal:

$$\mathrm{w}\big|_{\text{classical mixture}} = \begin{pmatrix} W_1 & 0 \\ 0 & W_2 \end{pmatrix}. \tag{7.23}$$

The mixed statistical ensemble with the density matrix diagonal in the stationary state basis is called the *classical mixture* and represents the limit opposite to the pure state.

After this example, the reader should not be much shocked by the main claim[10] of statistical mechanics that any large ensemble of similar systems in *thermodynamic* (or "thermal") *equilibrium* is exactly such a classical mixture. Moreover, for systems in thermal equilibrium with a much larger environment of a fixed temperature T (such an environment is frequently called either a "heat bath" or a "thermostat"), statistical physics gives a very simple expression, called the *Gibbs distribution*, for the probabilities W_n:

Gibbs distribution

$$W_n = \frac{1}{Z}\exp\left\{-\frac{E_n}{k_{\mathrm{B}}T}\right\}, \qquad \text{with } Z \equiv \sum_n \exp\left\{-\frac{E_n}{k_{\mathrm{B}}T}\right\}. \tag{7.24}$$

[9] For a system with a time-independent Hamiltonian, such averaging is especially plausible in the basis of the stationary states n of the system, in which the phase φ is just the difference of integration constants in Eq. (4.158), and its randomness may be naturally produced by minor fluctuations of the energy difference $E_1 - E_2$.

[10] This fact follows from the basic postulate of statistical physics, called the *microcanonical distribution* – see, e.g., SM Sec. 2.2.

where E_n is the eigenenergy of the corresponding stationary state, and the normalization coefficient Z is called the *statistical sum*.[11]

A detailed analysis of classical and quantum ensembles in thermodynamic equilibrium is a major focus of statistical physics courses (such as the SM part of this series) rather than this course of quantum mechanics. However, I would still like to attract the reader's attention to the key fact that, in contrast with the similarly-looking Boltzmann distribution for single particles,[12] the Gibbs distribution is *general*, *not* limited to classical statistics. In particular, for a quantum gas of indistinguishable particles, it is absolutely compatible with the quantum statistics (such as the Bose-Einstein or Fermi-Dirac distributions) of the component particles. For example, if we use Eq. (24) to calculate the average energy of a 1D harmonic oscillator of frequency ω_0 in thermal equilibrium, we easily get[13]

$$W_n = \exp\left\{-n\frac{\hbar\omega_0}{k_B T}\right\}\left(1-\exp\left\{-\frac{\hbar\omega_0}{k_B T}\right\}\right), \qquad Z = \exp\left\{-\frac{\hbar\omega_0}{2k_B T}\right\}\Bigg/\left(1-\exp\left\{-\frac{\hbar\omega_0}{k_B T}\right\}\right), \quad (7.25)$$

$$\langle E\rangle \equiv \sum_{n=0}^{\infty} W_n E_n = \frac{\hbar\omega_0}{2}\coth\frac{\hbar\omega_0}{2k_B T} \equiv \frac{\hbar\omega_0}{2} + \frac{\hbar\omega_0}{\exp\{\hbar\omega_0/k_B T\}-1}. \quad (7.26a)$$

The final form of the last result,

$$\langle E\rangle = \frac{\hbar\omega_0}{2} + \hbar\omega_0\langle n\rangle, \quad \text{with } \langle n\rangle = \frac{1}{\exp\{\hbar\omega_0/k_B T\}-1} \to \begin{cases} 0, & \text{for } k_B T \ll \hbar\omega_0, \\ k_B T/\hbar\omega_0, & \text{for } \hbar\omega_0 \ll k_B T, \end{cases} \quad (7.26b)$$

may be interpreted as an addition, to the ground-state energy $\hbar\omega_0/2$, of the average number $\langle n\rangle$ of thermally-induced excitations, with the energy $\hbar\omega_0$ each. In the harmonic oscillator, whose energy levels are equidistant, such a language is completely appropriate, because the transfer of the system from any level to the one just above it adds the same amount of energy, $\hbar\omega_0$. Note that the above expression for $\langle n\rangle$ is actually the Bose-Einstein distribution (for the particular case of zero chemical potential); we see that it does not contradict the Gibbs distribution (24) of the total energy of the system, but rather immediately follows from it.

Because of the fundamental importance of Eq. (26) for virtually all fields of physics, let me draw the reader's attention to its two main properties. At low temperatures, $k_B T \ll \hbar\omega_0$, there are virtually no excitations, $\langle n\rangle \to 0$, and the average energy of the oscillator is dominated by that of its ground state. In the opposite limit of high temperatures, $\langle n\rangle \to k_B T/\hbar\omega_0 \gg 1$, and $\langle E\rangle$ approaches the classical value $k_B T$.

7.2. Coordinate representation and the Wigner function

For many applications of the density operator, its coordinate representation is convenient. (I will only discuss it for the 1D case; the generalization to multi-dimensional cases is straightforward.)

[11] See. e.g., SM Sec. 2.4. The Boltzmann constant k_B is only needed if the temperature is measured in non-energy units – say in kelvins.

[12] See, e.g., SM Sec. 2.8.

[13] See, e.g., SM Sec. 2.5 – but mind a different energy reference level, $E_0 = \hbar\omega_0/2$, used for example in SM Eqs. (2.68)-(2.69), affecting the expression for Z. Actually, the calculation, using Eqs. (24) and (5.86), is so straightforward that it is highly recommended to the reader as a simple exercise.

Following Eq. (4.47), it is natural to define the following function of two arguments – traditionally, but a bit misleadingly, also called the *density matrix*:

Density matrix: coordinate representation

$$w(x,x') \equiv \langle x|\hat{w}|x'\rangle . \tag{7.27}$$

Inserting, into the right-hand side of this definition, two closure conditions (4.44) for an arbitrary (but full and orthonormal) basis $\{s\}$, and then using Eq. (4.233),[14] we get

$$w(x,x') = \sum_{j,j'} \langle x|s_j\rangle \langle s_j|\hat{w}|s_{j'}\rangle \langle s_{j'}|x'\rangle = \sum_{j,j'} \psi_j(x) w_{jj'}|_{\text{in } s} \psi_{j'}^*(x') . \tag{7.28}$$

In the special basis $\{w\}$, in which the density matrix is diagonal, this expression is reduced to

$$w(x,x') = \sum_j \psi_j(x) W_j \psi_j^*(x') . \tag{7.29}$$

Let us discuss the properties of this function. At coinciding arguments, $x' = x$, this is just the probability density:[15]

$$w(x,x) = \sum_j \psi_j(x) W_j \psi_j^*(x) = \sum_j w_j(x) W_j = w(x) . \tag{7.30}$$

However, the density matrix gives more information about the system than just the probability density. As the simplest example, let us consider a pure quantum state, with $W_j = \delta_{j,j'}$, so $\psi(x) = \psi_{j'}(x)$, and

$$w(x,x') = \psi_{j'}(x)\psi_{j'}^*(x') \equiv \psi(x)\psi^*(x') , \tag{7.31}$$

so

$$|w(x,x')|^2 \equiv w(x,x')w^*(x,x') = \psi(x)\psi^*(x)\psi(x')\psi^*(x') = w(x)w(x') . \tag{7.32}$$

For example, for a simple wave packet with a small spatial extent δx, $w(x, x')$ has an appreciable magnitude only if both points are not farther than $\sim\delta x$ from the packet center, and hence from each other. Note that the density matrix carries information not only about the modulus but also the phase of the wavefunction. However, in the ultimate limit of a pure state, the density-matrix description is redundant because all this information is contained in the wavefunction itself.

The density matrix becomes really invaluable when we deal with an incoherent mixture of several wavefunctions, for example, with the classical mixture describing the thermodynamic equilibrium. In this case, we can use Eq. (24) to rewrite Eq. (29) as follows:

$$w(x,x') = \sum_n \psi_n(x) W_n \psi_n^*(x') = \frac{1}{Z}\sum_n \psi_n(x) \exp\left\{-\frac{E_n}{k_B T}\right\} \psi_n^*(x') . \tag{7.33}$$

As the simplest example, let us see what is the density matrix of a *free* particle in thermal equilibrium. As we know very well by now, in this case, the set of energies $E_p = p^2/2m$ of stationary states (monochromatic waves) forms a continuum, so we need to replace the sum (33) with an integral, using, for example, the "delta-normalized" traveling-wave eigenfunctions (4.264):

[14] For now, I will focus on a fixed time instant (say, $t = 0$), and hence write $\psi(x)$ instead of $\Psi(x, t)$.

[15] This fact is the origin of the density matrix's name.

$$w(x,x') = \frac{1}{2\pi\hbar Z}\int_{-\infty}^{+\infty} \exp\left\{-\frac{ipx}{\hbar}\right\}\exp\left\{-\frac{p^2}{2mk_BT}\right\}\exp\left\{\frac{ipx'}{\hbar}\right\}dp\,. \tag{7.34}$$

This is a usual Gaussian integral and may be worked out, as we have done repeatedly in Chapter 2 and beyond, by complementing the exponent to the full square of the momentum p plus a constant. The statistical sum Z may be also readily calculated, [16]

$$Z = (2\pi mk_BT)^{1/2}, \tag{7.35}$$

However, for what follows it is more useful to write the result for the product wZ (the so-called *un-normalized density matrix*):

$$w(x,x')Z = \left(\frac{mk_BT}{2\pi\hbar^2}\right)^{1/2}\exp\left\{-\frac{mk_BT(x-x')^2}{2\hbar^2}\right\}. \tag{7.36}$$

Free particle: thermal equilibrium

This is a very interesting result: the density matrix depends only on the difference of its arguments, dropping to zero fast as the distance between the points x and x' exceeds the following characteristic scale (sometimes called the *correlation length* or *correlation distance*[17])

$$x_c \equiv \left\langle (x-x')^2\right\rangle^{1/2} = \frac{\hbar}{(mk_BT)^{1/2}}. \tag{7.37}$$

Correlation length

Some gut feeling of this length may be obtained from the following observation. It is straightforward to use Eq. (24) to verify that the average energy $\langle E\rangle = \langle p^2/2m\rangle$ of a free particle in thermal equilibrium, i.e. in the classical mixture (33), equals $k_BT/2$. (This value agrees with the *equipartition theorem* of classical statistics.[18]) Hence the average magnitude of the particle's momentum may be estimated as

$$p_c \equiv \left\langle p^2\right\rangle^{1/2} = \left(2m\langle E\rangle\right)^{1/2} = (mk_BT)^{1/2}, \tag{7.38}$$

so x_c is of the order of the minimal length allowed by the Heisenberg-like "uncertainty relation":

$$x_c = \frac{\hbar}{p_c}. \tag{7.39}$$

With the growth of temperature, the correlation length (37) goes to zero, and the density matrix (36) tends to a delta function:

$$w(x,x')Z\Big|_{T\to\infty} \to \delta(x-x')\,. \tag{7.40}$$

Since in this limit the average kinetic energy of the particle is not smaller than its potential energy in *any* fixed potential profile, Eq. (40) is the general property of the density matrix (33).

---

[16] Due to the delta-normalization of the eigenfunction, the density matrix (34) for the free particle (and any system with a continuous eigenvalue spectrum) is normalized as

$$\int_{-\infty}^{+\infty} w(x,x')Z dx' = \int_{-\infty}^{+\infty} w(x,x')Z dx = 1.$$

[17] Note that in some other fields of physics, the same term is used for a differently defined notion – see, e.g., SM Eqs. (4.30)-(4.31).

[18] See, e.g., SM Sec. 2.2.

Moreover, as will be shown in a minute (see Eq. (61) below), Eq. (36) may be used to calculate not only the r.m.s. magnitude p_c of the free particle's momentum[19] but also the whole distribution of the momentum's probability density. Thus for mixed states, the density function $w(x, x')$ is indeed much more informative than the probability distribution $w(x)$ – which, in our current case, is flat.

Now note the following important feature of Eq. (36): if we replace $k_B T$ with $\hbar/i(t - t_0)$, and x' with x_0, the un-normalized density matrix wZ for a free particle turns into the particle's propagator – cf. Eq. (2.49). This is not just an occasional coincidence. Indeed, in Chapter 2 we saw that the propagator of a system with an arbitrary stationary Hamiltonian may be expressed via the stationary eigenfunctions as

$$G(x,t;x_0,t_0)=\sum_n \psi_n(x)\exp\left\{-i\frac{E_n}{\hbar}(t-t_0)\right\}\psi_n^*(x_0). \tag{7.41}$$

Comparing this expression with Eq. (33), we see that the replacement

$$\frac{i(t-t_0)}{\hbar}\to\frac{1}{k_B T}, \tag{7.42}$$

plus the notation replacement $x' \to x_0$, turn the pure-state propagator G into the un-normalized density matrix wZ of the same system in thermodynamic equilibrium. This important fact, rooted in the mathematical similarity of the Gibbs distribution (24) with the Schrödinger equation's solution (1.69), enables a theoretical technique of the so-called *thermodynamic Green's functions*, which is especially productive in condensed matter physics.[20]

For our current purposes, we can employ Eq. (42) to reuse some of the wave mechanics results, in particular, the following formula for the harmonic oscillator's propagator

$$G(x,t;x_0,t_0)=\left(\frac{m\omega_0}{2\pi i\hbar\sin[\omega_0(t-t_0)]}\right)^{1/2}\exp\left\{-\frac{m\omega_0\left[\left(x^2+x_0^2\right)\cos[\omega_0(t-t_0)]-2xx_0\right]}{2i\hbar\sin[\omega_0(t-t_0)]}\right\}. \tag{7.43}$$

which may be readily proved to satisfy the Schrödinger equation for the Hamiltonian (5.62), with the appropriate initial condition: $G(x, t_0; x_0, t_0) = \delta(x - x_0)$. Making the substitution (42), we immediately get

Harmonic oscillator in thermal equilibrium

$$w(x,x')Z=\left[\frac{m\omega_0}{2\pi\hbar\sinh\left(\hbar\omega_0/k_B T\right)}\right]^{1/2}\exp\left\{-\frac{m\omega_0\left[\left(x^2+x'^2\right)\cosh\left(\hbar\omega_0/k_B T\right)-2xx'\right]}{2\hbar\sinh\left(\hbar\omega_0/k_B T\right)}\right\}. \tag{7.44}$$

As a sanity check, at very low temperatures, $k_B T << \hbar\omega_0$, both hyperbolic functions participating in this expression are very large and nearly equal, and it yields

$$w(x,x')Z\Big|_{T\to 0}\to\left[\left(\frac{m\omega_0}{\pi\hbar}\right)^{1/4}\exp\left\{-\frac{m\omega_0 x^2}{\hbar}\right\}\right]\times\exp\left\{-\frac{\hbar\omega_0}{2k_B T}\right\}\times\left[\left(\frac{m\omega_0}{\pi\hbar}\right)^{1/4}\exp\left\{-\frac{m\omega_0 x'^2}{\hbar}\right\}\right]. \tag{7.45}$$

[19] This is what could be expected from the basic equation (5): the density matrix enables the calculation of the statistical average of any variable pertaining to the system – including its momentum.

[20] I will have no time to discuss this technique and have to refer the interested reader to special literature. Probably, the most famous text of that field is A. Abrikosov, L. Gor'kov, and I. Dzyaloshinski, *Methods of Quantum Field Theory in Statistical Physics*, Prentice-Hall, 1963. (Later reprintings are available from Dover.)

In each of the expressions in square brackets, we can readily recognize the ground state's wavefunction (2.275) of the oscillator, while the middle exponent is just the statistical sum (24) in the low-temperature limit when it is dominated by the ground-level contribution:

$$Z\Big|_{T\to 0} \to \exp\left\{-\frac{\hbar\omega_0}{2k_B T}\right\}. \tag{7.46}$$

As a result, Z in both parts of Eq. (45) may be canceled, and the density matrix in this limit is described by Eq. (31), with the ground state as the only state of the oscillator. This is natural when the temperature is too low for the thermal excitation of any other state.

Returning to arbitrary temperatures, Eq. (44) in coinciding arguments gives the following expression for the probability density:[21]

$$w(x,x)Z \equiv w(x)Z = \left[\frac{m\omega_0}{2\pi\hbar\sinh\left(\hbar\omega_0/k_B T\right)}\right]^{1/2} \exp\left\{-\frac{m\omega_0 x^2}{\hbar}\tanh\frac{\hbar\omega_0}{2k_B T}\right\}. \tag{7.47}$$

This is just a Gaussian function of x, with the following variance:

$$\left\langle x^2\right\rangle = \frac{\hbar}{2m\omega_0}\coth\frac{\hbar\omega_0}{2k_B T}. \tag{7.48}$$

To compare this result with our earlier ones, it is useful to recast it as

$$\left\langle U\right\rangle = \frac{m\omega_0^2}{2}\left\langle x^2\right\rangle = \frac{\hbar\omega_0}{4}\coth\frac{\hbar\omega_0}{2k_B T}. \tag{7.49}$$

Comparing this expression with Eq. (26), we see that the average value of potential energy is exactly one-half of the total energy – the other half being the average kinetic energy. This is what we could expect, because according to Eqs. (5.96)-(5.97), such relation holds for each Fock state and hence should also hold for their classical mixture.

Unfortunately, besides the trivial case (30) of coinciding arguments, it is hard to give a straightforward interpretation of the density function in terms of the system's measurements. This is a fundamental difficulty, which has been well explored in terms of the *Wigner function* (sometimes called the "Wigner-Ville distribution")[22] defined as

$$W(X,P) \equiv \frac{1}{2\pi\hbar}\int w\left(X+\frac{\widetilde{X}}{2}, X-\frac{\widetilde{X}}{2}\right)\exp\left\{-\frac{iP\widetilde{X}}{\hbar}\right\}d\widetilde{X}. \tag{7.50}$$

Wigner function: definition

From the mathematical standpoint, this is just the Fourier transform of the density matrix in one of two new coordinates defined by the following relations (see Fig. 2):

[21] I have to confess that this notation is imperfect, because strictly speaking, $w(x, x')$ and $w(x)$ are different functions, and so are the functions $w(p, p')$ and $w(p)$ used below. In a perfect world, I would use different letters for them all, but I desperately want to stay with "w" for all the probability densities, and there are not so many good fonts for this letter. Let me hope that the difference between these functions is clear from their arguments and the context.

[22] It was introduced in 1932 by Eugene Wigner on the basis of a general (*Weyl-Wigner*) transform suggested by Hermann Weyl in 1927 and re-derived in 1948 by Jean Ville on a different mathematical basis.

$$X \equiv \frac{x+x'}{2}, \quad \widetilde{X} \equiv x - x', \qquad \text{so that } x \equiv X + \frac{\widetilde{X}}{2}, \quad x' \equiv X - \frac{\widetilde{X}}{2}. \tag{7.51}$$

Physically, the new argument X may be interpreted as the average position of the particle during the time interval $(t - t')$, while $\widetilde{X}$, as the distance passed by it during that time interval, so P characterizes the momentum of the particle during that motion. As a result, the Wigner function may be understood as a mathematical construct intended to characterize the system's probability distribution simultaneously in the coordinate and the momentum space – for 1D systems, on the phase plane $[X, P]$ that was discussed earlier – see Fig. 5.8. Let us see how fruitful this intention is.

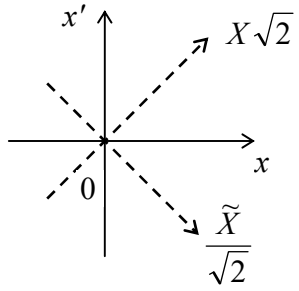


Fig. 7.2. The coordinates X and $\widetilde{X}$ employed in the Weyl-Wigner transform (50). They differ from the coordinates obtained by the rotation of the reference frame by the angle $\pi/4$ only by factors $\sqrt{2}$ and $1/\sqrt{2}$, describing scale stretch.

First of all, we may write the Fourier transform reciprocal to Eq. (50):

$$w\left(X + \frac{\widetilde{X}}{2}, X - \frac{\widetilde{X}}{2}\right) = \int W(X,P) \exp\left\{\frac{iP\widetilde{X}}{\hbar}\right\} dP\,. \tag{7.52}$$

For the particular case $\widetilde{X} = 0$, this relation yields

$$w(X) \equiv w(X,X) = \int W(X,P) dP\,. \tag{7.53}$$

Hence the integral of the Wigner function over the momentum P gives the probability density to find the system at point X – just as it does for a classical distribution function $w_{cl}(X, P)$.[23]

Next, the Wigner function has a similar property for its integration over X. To prove this fact, we may first introduce the momentum representation of the density matrix, in full analogy with its coordinate representation (27):

$$w(p,p') \equiv \langle p | \hat{w} | p' \rangle. \tag{7.54}$$

Inserting, as usual, two identity operators, in the form given by Eq. (4.252), into the right-hand side of this equality, we get the following relation between the momentum and coordinate representations:

$$w(p,p') = \int\int dx dx' \langle p|x\rangle\langle x|\hat{w}|x'\rangle\langle x'|p'\rangle = \frac{1}{2\pi\hbar}\int\int dx dx' \exp\left\{-\frac{ipx}{\hbar}\right\} w(x,x') \exp\left\{+\frac{ip'x'}{\hbar}\right\}. \tag{7.55}$$

This is of course nothing else than the unitary transform of an operator from the x-basis to the p-basis, similar to the first form of Eq. (4.272). For coinciding arguments, $p = p'$, Eq. (55) is reduced to

[23] Such function, used to express the probability dW to find the system in a small area of the phase plane as $dW = w_{cl}(X, P)dXdP$, is a major notion of the (1D) classical statistics – see, e.g., SM Sec. 2.1.

$$w(p) \equiv w(p,p) = \frac{1}{2\pi\hbar}\int\int dx dx'\, w(x,x') \exp\left\{-\frac{ip(x-x')}{\hbar}\right\}. \tag{7.56}$$

Now using Eq. (29) and then Eq. (4.265), this function may be represented as

$$w(p) = \frac{1}{2\pi\hbar}\sum_j W_j \int\int dx dx'\, \psi_j(x)\psi_j^*(x) \exp\left\{-\frac{ip(x-x')}{\hbar}\right\} = \sum_j W_j \varphi_j(p)\varphi_j^*(p), \tag{7.57}$$

and hence interpreted as the probability density of the particle's momentum at value p. Now, in the variables (51), Eq. (56) has the form

$$w(p) = \frac{1}{2\pi\hbar}\int\int w\left(X + \frac{\widetilde{X}}{2}, X - \frac{\widetilde{X}}{2}\right)\exp\left\{-\frac{ip\widetilde{X}}{\hbar}\right\} d\widetilde{X} dX. \tag{7.58}$$

Comparing this equality with the definition (50) of the Wigner function, we see that

$$w(P) = \int W(X,P) dX\,. \tag{7.59}$$

Thus, according to Eqs. (53) and (59), the integrals of the Wigner function over either the coordinate or momentum give the probability densities to find the system at a certain value of the counterpart variable. This is of course the main requirement for any quantum-mechanical candidate for the best analog of the classical probability density, $w_{cl}(X, P)$.

Let us see how the Wigner function looks for the simplest systems at thermodynamic equilibrium. For a free 1D particle, we can use Eq. (34), ignoring for simplicity the normalization issues:

$$W(X,P) \propto \int_{-\infty}^{+\infty} \exp\left\{-\frac{mk_B T\widetilde{X}^2}{2\hbar^2}\right\}\exp\left\{-\frac{iP\widetilde{X}}{\hbar}\right\} d\widetilde{X}\,. \tag{7.60}$$

The usual Gaussian integration yields

$$W(X,P) = \text{const} \times \exp\left\{-\frac{P^2}{2mk_B T}\right\}. \tag{7.61}$$

We see that the function is independent of X (as it should be for this translational-invariant system), and coincides with the Gibbs distribution (24).[24] We could get the same result directly from classical statistics. This is natural because as we know from Sec. 2.2, the free motion is essentially not quantized – at least in terms of its energy and momentum.

Now let us consider a "more quantum" system, the harmonic oscillator. Plugging Eq. (44) into Eq. (50), for that system in thermal equilibrium it is easy (and hence is left for the reader's exercise) to show that the Wigner function is also Gaussian, now in both its arguments:

$$W(X,P) = \text{const} \times \exp\left\{-C\left[\frac{m\omega_0^2 X^2}{2} + \frac{P^2}{2m}\right]\right\}, \tag{7.62}$$

though the coefficient C is now different from $1/k_B T$, and tends to that limit only at high temperatures, $k_B T >> \hbar\omega_0$. Moreover, for a Glauber state, the Wigner function also gives a very vivid result – a

[24] Note also that the width of this Gaussian distribution of momentum is indeed given by Eq. (38).

Gaussian distribution similar to Eq. (62), but properly shifted from the origin to the central point of the state – see Sec. 5.5.

However, for some other possible states of the harmonic oscillator, e.g., any pure Fock state with $n > 0$, the Wigner function takes negative values in some regions of the $[X, P]$ plane – see Fig. 3.[25] (Such plots were the basis of my, admittedly very imperfect, classical images of the Fock states in Fig. 5.8.)

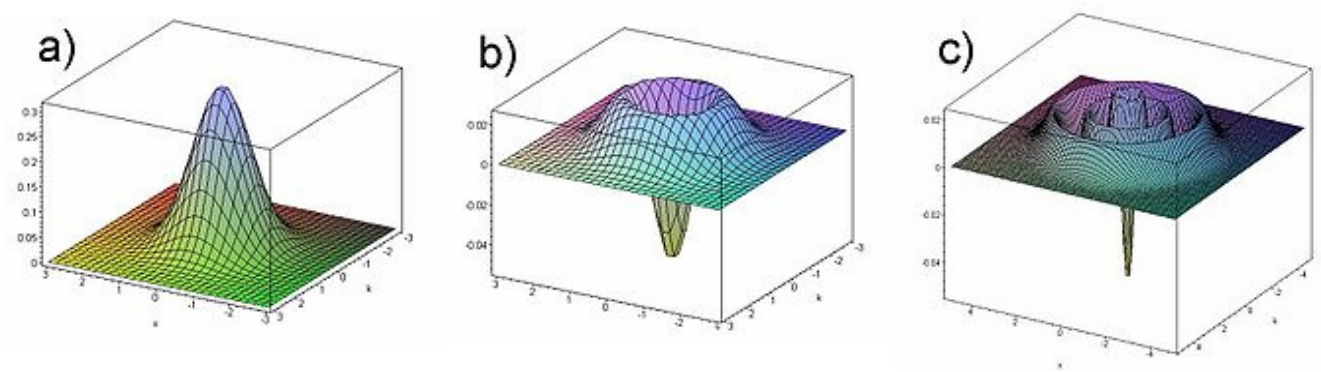


Fig. 7.3. The Wigner functions $W(X, P)$ of a harmonic oscillator, in a few of its stationary (Fock) states n: (a) $n = 0$, (b) $n = 1$; (c) $n = 5$. Graphics by J. S. Lundeen; adapted from http://en.wikipedia.org/wiki/Wigner_function as a public-domain material.

The same is true for most other quantum systems and their states. Indeed, this fact could be predicted just by looking at the definition (50) applied to a pure quantum state, in which the density function may be factored – see Eq. (31):

$$W(X,P) = \frac{1}{2\pi\hbar}\int \psi\left(X + \frac{\tilde{X}}{2}\right)\psi^*\left(X - \frac{\tilde{X}}{2}\right)\exp\left\{-\frac{iP\tilde{X}}{\hbar}\right\}d\tilde{X}\,. \tag{7.63}$$

By changing the argument P (say, at fixed X), we are essentially changing the spatial "frequency" (wave number) of the wavefunction product's Fourier component we are calculating, and we know that their Fourier images typically change sign as the frequency is changed. Hence the wavefunctions should have some high-symmetry properties to avoid this effect. Indeed, the Gaussian functions (describing, for example, the Glauber states, the squeezed states, and in their particular case, the ground state of the harmonic oscillator) have such symmetry, but many other functions do not.

Hence the Wigner function cannot serve as a direct quantum-mechanical analog of the classical probability density $w_{cl}(X, P)$. However, the function is useful for semi-quantitative interpretation of complicated mixed states of quantum systems.

7.3. Open system dynamics: Dephasing

So far we have discussed the density operator as something *given* at a particular time instant. Now let us discuss how is it *formed*, i.e. its evolution in time, starting from the simplest case when the probabilities W_j participating in Eq. (15) are *time-independent* – by this or that reason, to be discussed in a moment. In this case, in the Schrödinger picture, we may rewrite Eq. (15) as

[25] Spectacular experimental measurements of this function (for $n = 0$ and $n = 1$) were carried out recently by E. Bimbard *et al.*, *Phys. Rev. Lett.* **112**, 033601 (2014).

$$\hat{w}(t) = \sum_j \left|w_j(t)\right\rangle W_j \left\langle w_j(t)\right|. \tag{7.64}$$

Taking a time derivative of both sides of this equation, multiplying them by $i\hbar$, and applying Eq. (4.158) to the basis states w_j, with the account of the fact that the Hamiltonian operator is Hermitian, we get

$$\begin{aligned} i\hbar\dot{\hat{w}} &= i\hbar\sum_j \left(\left|\dot{w}_j(t)\right\rangle W_j \left\langle w_j(t)\right| + \left|w_j(t)\right\rangle W_j \left\langle \dot{w}_j(t)\right|\right) \\ &= \sum_j \left(\hat{H}\left|w_j(t)\right\rangle W_j \left\langle w_j(t)\right| - \left|w_j(t)\right\rangle W_j \left\langle w_j(t)\right|\hat{H}\right) \\ &\equiv \hat{H}\sum_j \left|w_j(t)\right\rangle W_j \left\langle w_j(t)\right| - \sum_j \left|w_j(t)\right\rangle W_j \left\langle w_j(t)\right|\hat{H}. \end{aligned} \tag{7.65}$$

Now using Eq. (64) again (twice), we get the so-called *von Neumann equation*[26]

$$i\hbar\dot{\hat{w}} = \left[\hat{H}, \hat{w}\right]. \tag{7.66}$$

von Neumann equation

Note that this equation is similar in structure to Eq. (4.199) describing the time evolution of time-independent operators in the Heisenberg picture operators:

$$i\hbar\dot{\hat{A}} = \left[\hat{A}, \hat{H}\right], \tag{7.67}$$

besides the opposite order of the operators in the commutator – equivalent to the change of sign on the right-hand side. This should not be too surprising, because Eq. (66) belongs to the Schrödinger picture of quantum dynamics, while Eq. (67), belongs to its Heisenberg picture.

Unfortunately, in most cases when the system is open, i.e. interacts with its environment, Eq. (66) is *not* valid, because the probabilities W_j may change in time.[27] However, the von Neumann equation serves an important role even in this case, provided that the interaction with the environment is so weak that its effect on the system's evolution may be considered as a weak perturbation. The analysis of this situation may be based on the following Hamiltonian:

$$\hat{H} = \hat{H}_s + \hat{H}_e\{\lambda\} + \hat{H}_{\text{int}}, \tag{7.68}$$

Interaction with environment

describing the system of our interest (s), its environment as such (e), and their interaction.[28] Here $\{\lambda\}$ denotes the (typically, huge) set of degrees of freedom of the environment. If the energy scale of the interaction Hamiltonian is relatively small, the time evolution of the probabilities W_j is relatively slow, and Eq. (66) may be used, in the 0^{th} approximations of the perturbation theory, for the density matrices of both the system s (with the Hamiltonian $\hat{H}_s$) and the environment (with the Hamiltonian $\hat{H}_e$).

[26] In some texts, it is called the "Liouville equation", due to its conceptual proximity to the classical Liouville theorem for the classical distribution function $w_{\text{cl}}(X, P)$ – see, e.g., SM Sec. 6.1 and in particular Eq. (6.5).

[27] Very unfortunately, this fact is not explained in some textbooks, which quote the von Neumann equation without proper qualifications.

[28] Note that by writing Eq. (68), we are treating the whole system, including the environment, as a Hamiltonian one. This can always be done if the accounted part of the environment is large enough so the processes in the system s of our interest do not depend on the type of boundary between this part of the environment and its "external" (even larger) part; in particular, we may assume the total system to be closed, i.e. Hamiltonian.

This approach, which will considered later in this chapter, turns out to be very useful if the elementary act of interaction of the system of interest with the environment is in some sense small. The classical example is the *Brownian particle* interacting with the molecules of the surrounding gas or fluid.[29] (In this example, a single hit by a molecule changes the particle's momentum by a minor fraction.) On the other hand, the model (68) is *not* very productive for a particle interacting with the environment consisting of *similar* particles, when a single collision may change its momentum (or other variables) dramatically. In such cases, the methods discussed in the next chapter are more adequate.

If the interaction with the environment is stronger, the perturbative approach to Eq. (68) also becomes invalid, but in this case, the von Neumann equation may still be used for a discussion of one major effect of the environment, namely *dephasing* (also called "decoherence"). Let us analyze a simple model of an open two-level quantum system, with its intrinsic Hamiltonian having the form

$$\hat{H}_s = c_z \hat{\sigma}_z , \tag{7.69}$$

similar to the Pauli Hamiltonian (4.163),[30] and a factorable interaction with environment – cf. Eq. (6.145) and its discussion:

$$\hat{H}_{\text{int}} = \hat{f}\{\lambda\}\hat{\sigma}_z , \tag{7.70}$$

where $\hat{f}$ is a Hermitian operator depending only on the set $\{\lambda\}$ of environmental degrees of freedom ("coordinates") defined in their Hilbert space – different from that of the two-level system. As a result, the operators $\hat{f}\{\lambda\}$ and $\hat{H}_e\{\lambda\}$ commute with $\hat{\sigma}_z$ – and with any other intrinsic operator of the two-level system. Of course, any realistic $\hat{H}_e\{\lambda\}$ is extremely complex, so how much we will be able to achieve without specifying it, may become a pleasant surprise for the reader.

Before we proceed to the analysis, let us recognize two examples of two-level systems that may be described by this model. The first example is a spin-½ in an external magnetic field of a fixed direction (taken for the z-axis), which includes both an average component $\overline{\mathscr{B}}$ and a random (fluctuating) component $\widetilde{\mathscr{B}}_z(t)$ induced by the environment. As it follows from Eq. (4.163b), it may be described by the Hamiltonian (68)-(70) with

$$c_z = -\frac{\hbar\gamma}{2}\overline{\mathscr{B}_z} \quad \text{and} \quad \hat{f} = -\frac{\hbar\gamma}{2}\widehat{\widetilde{\mathscr{B}}}_z(t) . \tag{7.71}$$

Another example is a particle in a symmetric double-well potential U_s (Fig. 4), with a barrier between them sufficiently high to be practically impenetrable, and an additional force $F(t)$, exerted by the environment, so the total potential energy is $U(x, t) = U_s(x) - F(t)x$. If the force, including its static

[29] The theory of the Brownian motion, the effect first observed experimentally by biologist Robert Brown in the 1820s, was pioneered by Albert Einstein in 1905 and developed in detail by Marian Smoluchowski in 1906-1907 and Adriaan Fokker in 1913. Due to this historical background, in some older texts, the approach described in the balance of this chapter is called the "quantum theory of the Brownian motion". Let me, however, emphasize that due to the later progress of experimental techniques, quantum-mechanical behaviors, including the environmental effects in them, have been observed in a rapidly growing number of various quasi-macroscopic systems, for which this approach is quite applicable. In particular, this is true for most systems being explored as possible qubits of prospective quantum computing and encryption systems – see Sec. 8.5 below.

[30] As we know from Secs. 4.6 and 5.1, such Hamiltonian is sufficient to lift the energy level degeneracy.

part $\overline{F}$ and fluctuations $\widetilde{F}(t)$, is sufficiently weak, we can neglect its effects on the shape of potential wells and hence on the localized wavefunctions $\psi_{L,R}$, so the force's effect is reduced to the variation of the difference $E_L - E_R = F(t)\Delta x$ between the eigenenergies. As a result, the system may be described by Eqs. (68)-(70) with

$$c_z = -\overline{F}\Delta x/2 \quad \text{and} \quad \hat{f} = -\hat{\widetilde{F}}(t)\Delta x/2\,. \tag{7.72}$$

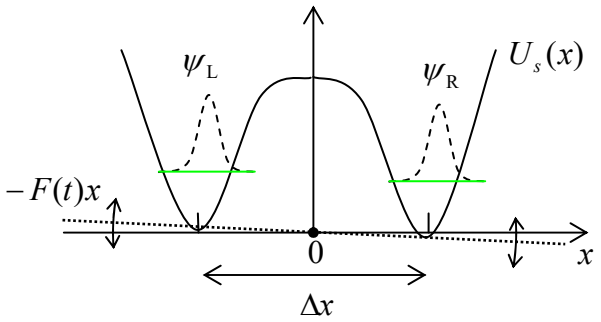


Fig. 7.4. Dephasing in a double-well system.

Now let us start our general analysis of the model described by Eqs. (68)-(70) by writing the equation of motion for the Heisenberg-picture operator $\hat{\sigma}_z(t)$:

$$i\hbar\dot{\hat{\sigma}}_z = \left[\hat{\sigma}_z, \hat{H}\right] = (c_z + \hat{f})\left[\hat{\sigma}_z, \hat{\sigma}_z\right] = 0, \tag{7.73}$$

showing that in our simple model (68)-(70), the operator $\hat{\sigma}_z$ does not evolve in time. What does this mean for the observables? For an arbitrary density matrix of any two-level system,

$$w = \begin{pmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{pmatrix}, \tag{7.74}$$

we can readily calculate the trace of the operator $\hat{\sigma}_z\hat{w}$. Indeed, since the operator traces are basis-independent, we may do this in the usual z-basis:

$$\mathrm{Tr}(\hat{\sigma}_z\hat{w}) = \mathrm{Tr}(\sigma_z \mathrm{w}) = \mathrm{Tr}\left[\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{pmatrix}\right] = w_{11} - w_{22} = W_1 - W_2\,. \tag{7.75}$$

Since, according to Eq. (5), $\hat{\sigma}_z$ may be considered the operator for the difference of the number of particles in the basis states 1 and 2, in the case (73), the difference $W_1 - W_2$ does not depend on time, and since the sum of these probabilities is also fixed, $W_1 + W_2 = 1$, both of them are constant. The physics of this simple result is especially clear for the model shown in Fig. 4: since the potential barrier separating the potential wells is so high that tunneling through it is negligible, the interaction with the environment cannot move the system from one well into another one.

It may look like nothing interesting may happen in such a simple situation, but in a minute we will see that this is not true. Due to the time independence of W_1 and W_2, we may use the von Neumann equation (66) to describe the density matrix evolution. In the z-basis:

$$i\hbar\dot{\mathrm{w}} \equiv i\hbar\begin{pmatrix} \dot{w}_{11} & \dot{w}_{12} \\ \dot{w}_{21} & \dot{w}_{22} \end{pmatrix} = [\mathrm{H}, \mathrm{w}] \equiv \left(c_z + \hat{f}\right)[\sigma_z, \mathrm{w}]$$

$$\equiv \left(c_z + \hat{f}\right)\left[\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{pmatrix}\right] = \left(c_z + \hat{f}\right)\begin{pmatrix} 0 & 2w_{12} \\ -2w_{21} & 0 \end{pmatrix}. \tag{7.76}$$

This result means that while the diagonal elements, i.e., the probabilities of the states, do not evolve in time (as we already know), the off-diagonal elements do change; for example,

$$i\hbar \dot{w}_{12} = 2(c_z + \hat{f})w_{12}, \tag{7.77}$$

with a similar but complex conjugate equation for w_{21}. The solution of this linear differential equation is straightforward, and yields

$$w_{12}(t) = w_{12}(0)\exp\left\{-i\frac{2c_z}{\hbar}t\right\}\exp\left\{-i\frac{2}{\hbar}\int_0^t \hat{f}(t')dt'\right\}. \tag{7.78}$$

The first exponent is a deterministic *c*-number factor, while in the second one $\hat{f}(t) \equiv \hat{f}\{\lambda(t)\}$ is still an operator in the Hilbert space of the environment, but from the point of view of the two-level system of our interest, it is just a random function of time. The time-average part of this function may be included in c_z, so in what follows, we will assume that it equals zero.

Let us start from the limit when the environment behaves classically.[31] In this case, the operator in Eq. (78) may be considered a *classical* random function $f(t)$, provided that we average its effects over a statistical ensemble of many such functions describing many (macroscopically similar) experiments. For a small time interval $t = dt \to 0$, we can use the Taylor expansion of the exponent, truncating it after the quadratic term:

$$\left\langle \exp\left\{-i\frac{2}{\hbar}\int_0^{dt} f(t')dt'\right\}\right\rangle \approx 1 + \left\langle -i\frac{2}{\hbar}\int_0^{dt} f(t')dt'\right\rangle + \left\langle \frac{1}{2}\left(-i\frac{2}{\hbar}\int_0^{dt} f(t')dt'\right)\left(-i\frac{2}{\hbar}\int_0^{dt} f(t'')dt''\right)\right\rangle$$
$$\equiv 1 - i\frac{2}{\hbar}\int_0^{dt}\langle f(t')\rangle dt' - \frac{2}{\hbar^2}\int_0^{dt}dt'\int_0^{dt}dt''\langle f(t')f(t'')\rangle \equiv 1 - \frac{2}{\hbar^2}\int_0^{dt}dt'\int_0^{dt}dt''K_f(t'-t''). \tag{7.79}$$

Here we have used the facts that the statistical average of $f(t)$ is equal to zero, while the second average, called the *correlation function*, in a statistically- (i.e. macroscopically-) stationary state of the environment may only depend on the time difference $\tau \equiv t' - t''$:

Correlation function

$$\boxed{\langle f(t')f(t'')\rangle = K_f(t'-t'') \equiv K_f(\tau).} \tag{7.80}$$

If this difference is much larger than some time scale τ_c, called the *correlation time* of the environment, the values $f(t')$ and $f(t'')$ are completely independent (*uncorrelated*), as illustrated in Fig. 5a, so at $\tau \to \infty$, the correlation function has to tend to zero. On the other hand, at $\tau = 0$, i.e. $t' = t''$, the correlation function is just the variance of f:

$$K_f(0) = \langle f^2 \rangle, \tag{7.81}$$

and has to be positive. As a result, the function looks approximately as shown in Fig. 5b. (On the way to zero at $\tau \to \pm\infty$, it may or may not change sign.)

[31] This assumption is not in contradiction with the quantum treatment of the two-level system *s*, because a typical environment is large, and hence has a very dense energy spectrum, with small adjacent level distances that may be readily bridged by thermal excitations of minor energy, often making it essentially classical.

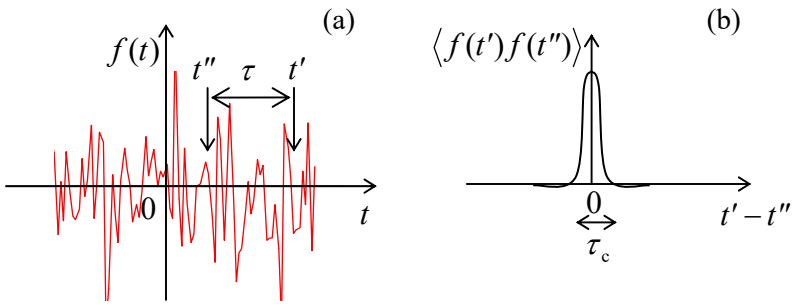


Fig. 7.5. (a) A typical random process and (b) its correlation function – schematically.

Hence, if we are only interested in time differences τ much longer than τ_c, which is typically very short, we may approximate $K_f(\tau)$ well with a delta function of the time difference. Let us take it in the following form, convenient for later discussion:

$$K_f(\tau) \approx \hbar^2 D_\varphi \delta(\tau), \tag{7.82}$$

Phase diffusion coefficient

where D_φ is a positive constant called the *phase diffusion coefficient*. The origin of this term stems from the very similar effect of classical diffusion of Brownian particles in a highly viscous medium. Indeed, the particle's velocity in such a medium is approximately proportional to the external force. Hence, if the random hits of a particle by the medium's molecules may be described by a force that obeys a law similar to Eq. (82), the velocity (along any Cartesian coordinate) is also *delta-correlated*:

$$\langle v(t) \rangle = 0, \quad \langle v(t')v(t'') \rangle = 2D\delta(t' - t''). \tag{7.83}$$

Now we can integrate the kinematic relation $\dot{x} = v$, to calculate the particle's displacement from its initial position during a time interval [0, t] and its variance:

$$x(t) - x(0) = \int_0^t v(t')dt', \tag{7.84}$$

$$\left\langle \left(x(t) - x(0)\right)^2 \right\rangle = \left\langle \int_0^t v(t')dt' \int_0^t v(t'')dt'' \right\rangle = \int_0^t dt' \int_0^t dt'' \langle v(t')v(t'') \rangle = \int_0^t dt' \int_0^t dt'' 2D\delta(t' - t'') = 2Dt. \tag{7.85}$$

This is the famous law of diffusion, showing that the r.m.s. deviation of the particle from the initial point grows with time as $(2Dt)^{1/2}$, where the constant D is called the *diffusion coefficient*.

Returning to the diffusion of the quantum-mechanical phase, with Eq. (82) the last double integral in Eq. (79) yields $\hbar^2 D_\varphi dt$, so the statistical average of Eq. (78) is

$$\langle w_{12}(dt) \rangle = w_{12}(0) \exp\left\{ -i \frac{2c_z}{\hbar} dt \right\} \left(1 - 2D_\varphi dt\right). \tag{7.86}$$

Applying this formula to sequential time intervals,

$$\langle w_{12}(2dt) \rangle = \langle w_{12}(dt) \rangle \exp\left\{ -i \frac{2c_z}{\hbar} dt \right\} \left(1 - 2D_\varphi dt\right) = w_{12}(0) \exp\left\{ -i \frac{2c_z}{\hbar} 2dt \right\} \left(1 - 2D_\varphi dt\right)^2, \tag{7.87}$$

etc., for a finite time $t = Ndt$, in the limit $N \to \infty$ and $dt \to 0$ (at fixed t) we get

$$\langle w_{12}(t) \rangle = w_{12}(0) \exp\left\{ -i \frac{2c_z}{\hbar} t \right\} \times \lim_{N \to \infty} \left(1 - 2D_\varphi t \frac{1}{N} \right)^N. \tag{7.88}$$

By the definition of the natural logarithm base e,[32] this limit is just $\exp\{-2D_\varphi t\}$, so, finally:

Two-level system: dephasing

$$\langle w_{12}(t)\rangle = w_{12}(0)\exp\left\{-i\frac{2a}{\hbar}t\right\}\exp\{-2D_\varphi t\} \equiv w_{12}(0)\exp\left\{-i\frac{2a}{\hbar}t\right\}\exp\left\{-\frac{t}{T_2}\right\}. \quad (7.89)$$

So, due to coupling to the environment, the off-diagonal elements of the density matrix decay with some *dephasing time* $T_2 = 1/2D_\varphi$,[33] providing a natural evolution from the density matrix (22) of a pure state to the diagonal matrix (23), with the same probabilities $W_{1,2}$, describing a fully dephased (incoherent) classical mixture.[34]

This simple model offers a very clear look at the nature of the decoherence: the random "force" $f(t)$ exerted by the environment, "shakes" the energy difference between two eigenstates of the system and hence the instantaneous velocity $2(c_z + f)/\hbar$ of their mutual phase shift $\varphi(t)$ – cf. Eq. (22). Due to the randomness of the force, $\varphi(t)$ performs a random walk around the trigonometric circle, so the average of its trigonometric functions $\exp\{\pm i\varphi\}$ over time gradually tends to zero, killing the off-diagonal elements of the density matrix. Our analysis, however, has left open two important issues:

(i) Is this approach valid for a *quantum* description of a typical environment?

(ii) If yes, what is physically the D_φ that was formally defined by Eq. (82)?

7.4. Fluctuation-dissipation theorem

Similar questions may be asked about a more general situation, when the Hamiltonian $\hat{H}_s$ of the system of interest (s), in the composite Hamiltonian (68), is not specified at all, but the interaction between that system and its environment still has the form similar to Eqs. (70) and (6.130):

$$\hat{H}_{\text{int}} = -\hat{F}\{\lambda\}\,\hat{x}, \quad (7.90)$$

where x is some observable of our system s – say, one of its generalized coordinates. It may look incredible that in this very general situation, one still can make a very simple and powerful statement about the statistical properties of the generalized force F, under only two (interrelated) conditions – which are satisfied in a huge number of cases of interest:

(i) the coupling of system s of interest to its environment e is not too strong – in the sense that the perturbation theory (see Chapter 6) is applicable, and

(ii) the environment may be considered as staying in thermodynamic equilibrium, with a certain temperature T, regardless of the process in the system of interest.[35]

---

[32] See, e.g., MA Eq. (1.2a) with the substitution $n = -N/2D_\varphi t$.

[33] In context of the spin magnetic resonance (see below), T_2 is frequently called the "spin-spin relaxation time".

[34] Note that this result is valid only if the approximation (82) may be applied at time interval dt which, in turn, should be much smaller than the T_2 in Eq. (88), i.e. if the dephasing time is much longer than the environment's correlation time τ_c. This requirement may be always satisfied by making the coupling to the environment sufficiently weak. In addition, in typical environments, τ_c is very short. For example, in the original Brownian motion experiments with a-few-μm pollen grains in water, it is of the order of the duration of each molecular impact (~10^{-12} s) because the average interval between sequential impacts is even much shorter.

[35] The most frequent example of the violation of this condition is the environment's overheating by the energy flow from system s. Let me leave it to the reader to estimate the overheating of a standard physical laboratory

This famous statement is called the *fluctuation-dissipation theorem* (FDT).[36] Due to the importance of this fundamental result, let me derive it.[37] Since by writing Eq. (68), we treat the whole system ($s + e$) as a Hamiltonian one, we may use the Heisenberg equation (4.199) to write

$$i\hbar\dot{\hat{F}} = \left[\hat{F}, \hat{H}\right] \equiv \left[\hat{F}, \hat{H}_e\right]. \tag{7.91}$$

(The second step uses the fact, discussed in the last section, that $\hat{F}\{\lambda\}$ commutes with both $\hat{H}_s$ and $\hat{x}$.) Generally, very little may be done with this equation, because the time evolution of the environment's Hamiltonian depends, in turn, on that of the force. This is where the perturbation theory becomes indispensable. Let us decompose the force operator into the following sum:

$$\hat{F}\{\lambda\} = \left\langle \hat{F} \right\rangle + \hat{\tilde{F}}(t), \quad \text{with } \left\langle \hat{\tilde{F}}(t) \right\rangle = 0, \tag{7.92}$$

where (here and on, until further notice) the sign $\langle \ldots \rangle$ means the statistical averaging *over the environment alone*, i.e. over an ensemble with absolutely similar evolutions of the system s, but random states of its environment.[38] From the point of view of the system s, the first term of the sum (still an operator!) describes the average response of the environment to the system's dynamics (possibly, including such irreversible effects as friction), and has to be calculated with a proper account of their interaction – as we will do later in this section. On the other hand, the last term in Eq. (92) represents random *fluctuations* of the environment, which exist even in the absence of the system s. Hence, in the first nonvanishing approximation in the interaction strength, the fluctuation part may be calculated ignoring the interaction, i.e. treating the environment as being in thermodynamic equilibrium:

$$i\hbar\dot{\hat{\tilde{F}}} = \left[\hat{\tilde{F}}, \hat{H}_e\Big|_{\text{eq}}\right]. \tag{7.93}$$

Since, in this approximation, the environment's Hamiltonian does not have an explicit dependence on time, the solution of this equation may be written by combining Eqs. (4.190) and (4.175):

$$\hat{F}(t) = \exp\left\{+\frac{i}{\hbar}\hat{H}_e\Big|_{\text{eq}} t\right\} \hat{F}(0) \exp\left\{-\frac{i}{\hbar}\hat{H}_e\Big|_{\text{eq}} t\right\}. \tag{7.94}$$

Let us use this relation to calculate the correlation function of the fluctuations $F(t)$, defined similarly to Eq. (80), but taking care of the order of the time arguments (very soon we will see why):

---

room by a typical dissipative quantum process – the emission of an optical photon by an atom. (*Hint*: it is extremely small.)

[36] The FDT was first derived by Herbert Callen and Theodore Allen Welton in 1951, on the background of an earlier derivation of its classical limit by Harry Nyquist in 1928 (for a particular case of electric circuits).

[37] The FDT may be proved in several ways that are shorter than the one given below – see, e.g., either the proof in SM Secs. 5.5 and 5.6 (based on H. Nyquist's arguments), or the original paper by H. Callen and T. Welton, *Phys. Rev.* **83**, 34 (1951) – wonderful in its clarity. The longer approach I will describe here, besides giving the important *Green-Kubo formula* (109) as a byproduct, is a very useful exercise in operator manipulation and the perturbation theory in its integral form – different from the differential forms used in Chapter 6. The reader not interested in this exercise may skip the derivation and jump straight to the result expressed by Eq. (134), which uses the notions defined by Eqs. (114) and (123).

[38] For usual ("ergodic") environments, without intrinsic long-term memories, this statistical averaging over an ensemble of environments is equivalent to averaging over intermediate times – much longer than the correlation time τ_c of the environment, but still much shorter than the characteristic time of evolution of the system under analysis, such as the dephasing time T_2 and the energy relaxation time T_1 – both still to be calculated.

$$\left\langle \tilde{F}(t)\tilde{F}(t')\right\rangle = \left\langle \exp\left\{+\frac{i}{\hbar}\hat{H}_e t\right\}\hat{F}(0)\exp\left\{-\frac{i}{\hbar}\hat{H}_e t\right\}\exp\left\{+\frac{i}{\hbar}\hat{H}_e t'\right\}\hat{F}(0)\exp\left\{-\frac{i}{\hbar}\hat{H}_e t'\right\}\right\rangle. \quad (7.95)$$

(Here and in Eq. (96), the thermal equilibrium of the environment is also meant but just implied, for the notation brevity.) We may calculate this expectation value in any basis, and the best choice for it is evident: in the environment's stationary-state basis, the density operator of the environment, its Hamiltonian and hence the exponents in Eq. (95) are all represented by diagonal matrices. Using Eq. (5), the correlation function becomes

$$\begin{aligned}\left\langle \tilde{F}(t)\tilde{F}(t')\right\rangle &= \mathrm{Tr}\left[\hat{w}\exp\left\{+\frac{i}{\hbar}\hat{H}_e t\right\}\hat{F}(0)\exp\left\{-\frac{i}{\hbar}\hat{H}_e t\right\}\exp\left\{+\frac{i}{\hbar}\hat{H}_e t'\right\}\hat{F}(0)\exp\left\{-\frac{i}{\hbar}\hat{H}_e t'\right\}\right] \\ &\equiv \sum_n \left[\hat{w}\exp\left\{+\frac{i}{\hbar}\hat{H}_e t\right\}\hat{F}(0)\exp\left\{-\frac{i}{\hbar}\hat{H}_e t\right\}\exp\left\{+\frac{i}{\hbar}\hat{H}_e t'\right\}\hat{F}(0)\exp\left\{-\frac{i}{\hbar}\hat{H}_e t'\right\}\right]_{nn} \\ &= \sum_{n,n'} W_n \exp\left\{+\frac{i}{\hbar}E_n t\right\}\hat{F}_{nn'}\exp\left\{-\frac{i}{\hbar}E_{n'} t\right\}\exp\left\{+\frac{i}{\hbar}E_{n'} t'\right\}\hat{F}_{n'n}\exp\left\{-\frac{i}{\hbar}E_n t'\right\} \quad (7.96) \\ &\equiv \sum_{n,n'} W_n \left|F_{nn'}\right|^2 \exp\left\{+\frac{i}{\hbar}(E_n - E_{n'})(t-t')\right\}.\end{aligned}$$

Here W_n are the Gibbs-distribution probabilities given by Eq. (24), with the environment's temperature T, and $F_{nn'} \equiv F_{nn'}(0)$ are the Schrödinger-picture matrix elements of the interaction force operator.

We see that though the correlator (96) is a function of the difference $\tau \equiv t - t'$ only (as it should be for fluctuations in a macroscopically stationary system), it may depend on the order of its arguments. This is why let us mark this particular correlation function with the upper index "+",

$$K_F^+(\tau) \equiv \left\langle \tilde{F}(t)\tilde{F}(t')\right\rangle = \sum_{n,n'} W_n \left|F_{nn'}\right|^2 \exp\left\{+\frac{i\tilde{E}\tau}{\hbar}\right\}, \quad \text{where } \tilde{E} \equiv E_n - E_{n'}, \quad (7.97)$$

while its counterpart, with swapped times t and t', with the upper index "–":

$$K_F^-(\tau) \equiv K_F^+(-\tau) = \left\langle \tilde{F}(t')\tilde{F}(t)\right\rangle = \sum_{n,n'} W_n \left|F_{nn'}\right|^2 \exp\left\{-\frac{i\tilde{E}\tau}{\hbar}\right\}. \quad (7.98)$$

So, in contrast with classical processes, in quantum mechanics the correlation function of fluctuations $\tilde{F}$ is not necessarily time-symmetric:

$$K_F^+(\tau) - K_F^-(\tau) \equiv K_F^+(\tau) - K_F^+(-\tau) = \left\langle \tilde{F}(t)\tilde{F}(t') - \tilde{F}(t')\tilde{F}(t)\right\rangle = 2i\sum_{n,n'} W_n \left|F_{nn'}\right|^2 \sin\frac{\tilde{E}\tau}{\hbar} \neq 0, \quad (7.99)$$

so $\hat{F}(t)$ gives one more example of a Heisenberg-picture operator whose "values" at different moments generally do not commute – see Footnote 47 in Chapter 4. (A good sanity check here is that at $\tau = 0$, i.e. at $t = t'$, the difference (99) between K_F^+ and K_F^- vanishes.)

Now let us return to the force operator's decomposition (92), and calculate its first (average) component. To do that, let us write the formal solution of Eq. (91) as follows:

$$\hat{F}(t) = \frac{1}{i\hbar}\int_{-\infty}^{t}\left[\hat{F}(t'), \hat{H}_e(t')\right]dt'. \quad (7.100)$$

On the right-hand side of this relation, we still cannot treat the Hamiltonian of the environment as an unperturbed (equilibrium) one, even if the effect of our system (s) on the environment is very weak, because this would give us zero statistical average of the force $F(t)$. Hence, we should make one more step in our perturbative treatment, taking into account the effect of the force on the environment. To do this, let us use Eqs. (68) and (90) to write the (so far, exact) Heisenberg equation of motion for the environment's Hamiltonian,

$$i\hbar\dot{\hat{H}}_e = \left[\hat{H}_e, \hat{H}\right] \equiv -\hat{x}\left[\hat{H}_e, \hat{F}\right], \tag{7.101}$$

and its formal solution, similar to Eq. (100), but for time t' rather than t:

$$\hat{H}_e(t') = -\frac{1}{i\hbar}\int_{-\infty}^{t'} \hat{x}(t'')\left[\hat{H}_e(t''), \hat{F}(t'')\right]dt''. \tag{7.102}$$

Plugging this equality into the right-hand side of Eq. (100), and averaging the result (again, over the environment only!), we get

$$\left\langle \hat{F}(t)\right\rangle = \frac{1}{\hbar^2}\int_{-\infty}^{t} dt' \int_{-\infty}^{t'} dt''\, \hat{x}(t'')\left\langle \left[\hat{F}(t'), \left[\hat{H}_e(t''), \hat{F}(t'')\right]\right]\right\rangle. \tag{7.103}$$

This is still an exact result, but now it is ready for an approximate treatment, implemented by averaging in its right-hand side over the unperturbed (thermal-equilibrium) state of the environment.[39] This may be done absolutely similarly to that in Eq. (96), at the last step using Eq. (94):

$$\begin{aligned}
&\left\langle \left[\hat{F}(t'), \left[\hat{H}_e(t''), \hat{F}(t'')\right]\right]\right\rangle = \mathrm{Tr}\left\{\mathrm{w}\left[\mathrm{F}(t'), \left[\mathrm{H}_e\mathrm{F}(t'')\right]\right]\right\} \\
&\equiv \mathrm{Tr}\left\{\mathrm{w}\left[\mathrm{F}(t')\mathrm{H}_e\mathrm{F}(t'') - \mathrm{F}(t')\mathrm{F}(t'')\mathrm{H}_e - \mathrm{H}_e\mathrm{F}(t'')\mathrm{F}(t') + \mathrm{F}(t'')\mathrm{H}_e\mathrm{F}(t')\right]\right\} \\
&= \sum_{n,n'} W_n\left[F_{nn'}(t')E_{n'}F_{n'n}(t'') - F_{nn'}(t')F_{n'n}(t'')E_n - E_nF_{nn'}(t'')F_{n'n}(t') + F_{nn'}(t'')E_{n'}F_{n'n}(t'')\right] \\
&\equiv -\sum_{n,n'} W_n\widetilde{E}\left|F_{nn'}\right|^2\left[\exp\left\{\frac{i\widetilde{E}(t'-t'')}{\hbar}\right\} + \mathrm{c.c.}\right].
\end{aligned} \tag{7.104}$$

Now, if we try to integrate each term of this sum, as Eq. (103) seems to require, we will see that the lower-limit substitution (at t', $t'' \to -\infty$) is uncertain because the exponents oscillate without decay. This mathematical difficulty may be overcome by the following physical reasoning. As illustrated by the example considered in the previous section, coupling to a disordered environment makes the "memory horizon" of the system of our interest (s) finite: its current state does not depend on its history beyond a certain time scale.[40] As a result, the function under the integrals of Eq. (103), i.e. the sum (104), should self-average at a certain finite time. A simplistic technique for expressing this fact mathematically is just dropping the lower-limit substitution; this would give the correct result for Eq. (103). However, a better

[39] This is exactly the moment when, in this approach, the reversible quantum dynamics of the formally-Hamiltonian full system (s + e) is replaced with irreversible dynamics of the system s. At the conditions formulated at the beginning of this section, this replacement is physically justified – see also the discussions in Secs. 2.5, 6.6, and 6.7.

[40] Indeed, this is true for virtually any real physical system – in contrast to idealized models such as a dissipation-free oscillator that swings for ever and ever with the same amplitude and phase, thus "remembering" the initial conditions.

(mathematically more acceptable) trick is to first multiply the functions under the integrals by, respectively, $\exp\{\varepsilon(t-t')\}$ and $\exp\{\varepsilon(t'-t'')\}$, where ε is a very small positive constant, then carry out the integration, and after that follow the limit $\varepsilon \to 0$. The physical justification of this procedure may be provided by saying that the system's behavior should not be affected if its interaction with the environment was not kept constant but rather turned on gradually and very slowly – say, exponentially with an infinitesimal rate ε. With this modification, Eq. (103) becomes

$$\left\langle \hat{F}(t) \right\rangle = -\frac{1}{\hbar^2} \sum_{n,n'} W_n \widetilde{E} \left| F_{nn'} \right|^2 \lim_{\varepsilon \to 0} \int_{-\infty}^{t} dt' \int_{-\infty}^{t'} dt'' \, \hat{x}(t'') \left[\exp\left\{ \frac{i\widetilde{E}(t'-t'')}{\hbar} + \varepsilon(t''-t) \right\} + \text{c.c.} \right]. \quad (7.105)$$

This double integration is over the area shaded in Fig. 6, which makes it obvious that the order of integration may be changed to the opposite one as

$$\int_{-\infty}^{t} dt' \int_{-\infty}^{t'} dt'' \ldots = \int_{-\infty}^{t} dt'' \int_{t''}^{t} dt' \ldots = \int_{-\infty}^{t} dt'' \int_{t''-t}^{0} d(t'-t) \ldots \equiv \int_{-\infty}^{t} dt'' \int_{0}^{\tau} d\tau' \ldots, \quad (7.106)$$

where $\tau' \equiv t - t'$, and $\tau \equiv t - t''$.

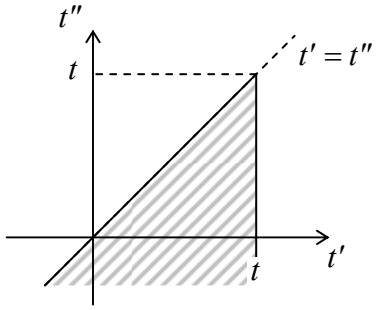


Fig. 7.6. The 2D integration area in Eqs. (105) and (106).

As a result, Eq. (105) may be rewritten as a single integral,

$$\left\langle \hat{F}(t) \right\rangle = \int_{-\infty}^{t} G(t-t'') \, \hat{x}(t'') dt'' \equiv \int_{0}^{\infty} G(\tau) \, \hat{x}(t-\tau) d\tau, \quad (7.107)$$

whose kernel,

Environment's response function

$$\begin{aligned} G(\tau > 0) &\equiv -\frac{1}{\hbar^2} \sum_{n,n'} W_n \widetilde{E} \left| F_{nn'} \right|^2 \lim_{\varepsilon \to 0} \int_{0}^{\tau} \left[\exp\left\{ \frac{i\widetilde{E}(\tau - \tau')}{\hbar} - \varepsilon\tau \right\} + \text{c.c.} \right] d\tau' \\ &= \lim_{\varepsilon \to 0} \frac{2}{\hbar} \sum_{n,n'} W_n \left| F_{nn'} \right|^2 \sin\frac{\widetilde{E}\tau}{\hbar} e^{-\varepsilon\tau} \equiv \frac{2}{\hbar} \sum_{n,n'} W_n \left| F_{nn'} \right|^2 \sin\frac{\widetilde{E}\tau}{\hbar}, \end{aligned} \quad (7.108)$$

does not depend on the particular law of evolution of the system (s) under study, i.e. provides a general characterization of its coupling to the environment.

In Eq. (107) we may readily recognize the most general form of the linear response of a system (in our case, the environment), taking into account the causality principle, where $G(\tau)$ is the *response function* (also called the "temporal Green's function") of the environment. Now comparing Eq. (108) with Eq. (99), we get a wonderfully simple universal relation,

Green-Kubo formula

$$\left\langle \left[\hat{\tilde{F}}(\tau), \hat{\tilde{F}}(0) \right] \right\rangle = i\hbar G(\tau). \quad (7.109)$$

that emphasizes once again the quantum nature of the correlation function's time asymmetry. (This relation, called the *Green-Kubo* (or just "Kubo") *formula* after the works by Melville Green in 1954 and Ryogo Kubo in 1957, did not come up in the easier derivations of the FDT, mentioned in the beginning of this section.)

However, for us the relation between the function $G(\tau)$ and the force's *anti*-commutator,

$$\left\langle \left\{ \hat{\tilde{F}}(t+\tau), \hat{\tilde{F}}(t) \right\} \right\rangle \equiv \left\langle \hat{\tilde{F}}(t+\tau)\hat{\tilde{F}}(t) + \hat{\tilde{F}}(t)\hat{\tilde{F}}(t+\tau) \right\rangle \equiv K_F^+(\tau) + K_F^-(\tau), \tag{7.110}$$

is much more important, because of the following reason. Eqs. (97)-(98) show that the so-called *symmetrized correlation function*,

$$K_F(\tau) \equiv \frac{K_F^+(\tau) + K_F^-(\tau)}{2} = \frac{1}{2}\left\langle \left\{ \hat{\tilde{F}}(\tau), \hat{\tilde{F}}(0) \right\} \right\rangle = \lim_{\varepsilon \to 0} \sum_{n,n'} W_n \left| F_{nn'} \right|^2 \cos\frac{\tilde{E}\tau}{\hbar} e^{-2\varepsilon|\tau|}$$
$$\equiv \sum_{n,n'} W_n \left| F_{nn'} \right|^2 \cos\frac{\tilde{E}\tau}{\hbar}, \tag{7.111}$$

Symmetrized correlation function

which is an even function of the time difference τ, looks very similar to the response function (108), "only" with another trigonometric function under the sum, and a different constant front factor.[41] This similarity may be used to obtain a direct algebraic relation between the Fourier images of these two functions of τ. Indeed, the function (111) may be represented as the Fourier integral[42]

$$K_F(\tau) = \int_{-\infty}^{+\infty} S_F(\omega) e^{-i\omega\tau} d\omega = 2\int_0^{+\infty} S_F(\omega) \cos\omega\tau \, d\omega, \tag{7.112}$$

with the reciprocal transform

$$S_F(\omega) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} K_F(\tau) e^{i\omega\tau} d\tau = \frac{1}{\pi}\int_0^{+\infty} K_F(\tau) \cos\omega\tau \, d\tau, \tag{7.113}$$

of the *symmetrized spectral density* of the variable F, defined as

$$S_F(\omega)\delta(\omega - \omega') \equiv \frac{1}{2}\left\langle \hat{F}_\omega \hat{F}_{-\omega'} + \hat{F}_{-\omega'}\hat{F}_\omega \right\rangle \equiv \frac{1}{2}\left\langle \left\{ \hat{F}_\omega, \hat{F}_{-\omega'} \right\} \right\rangle, \tag{7.114}$$

Symmetrized spectral density

where the function $\hat{F}_\omega$ (also a Heisenberg-picture operator rather than a c-number!) is defined as

$$\hat{F}_\omega \equiv \frac{1}{2\pi}\int_{-\infty}^{+\infty} \hat{\tilde{F}}(t) e^{i\omega t} dt, \qquad \text{so that } \hat{\tilde{F}}(t) = \int_{-\infty}^{+\infty} \hat{F}_\omega e^{-i\omega t} d\omega. \tag{7.115}$$

The physical meaning of the function $S_F(\omega)$ becomes clear if we write Eq. (112) for the particular case $\tau = 0$:

$$K_F(0) \equiv \left\langle \hat{\tilde{F}}^2 \right\rangle = \int_{-\infty}^{+\infty} S_F(\omega) d\omega = 2\int_0^{+\infty} S_F(\omega) d\omega. \tag{7.116}$$

[41] For the heroic reader who has suffered through the calculations up to this point: our conceptual work is done! What remains is just some simple math to bring the relation between Eqs. (108) and (111) to an explicit form.

[42] Due to their practical importance, and certain mathematical issues of their justification for random functions, Eqs. (112)-(113) have their own grand name, the *Wiener-Khinchin theorem*, though the math rigor aside, they are just a straightforward corollary of the standard Fourier integral transform (115).

This formula infers that if we pass the function $F(t)$ through a linear filter cutting from its frequency spectrum a narrow band $d\omega$ of physical (positive) frequencies, then the variance $\langle F_{\rm f}^2\rangle$ of the filtered signal $F_{\rm f}(t)$ would be equal to $2S_F(\omega)d\omega$ – hence the name "spectral density".[43]

Let us use Eqs. (111) and (113) to calculate the spectral density of fluctuations $\widetilde{F}(t)$ in our model, using the same ε-trick as at the deviation of Eq. (108), to quench the upper-limit substitution:

$$\begin{aligned} S_F(\omega) &= \sum_{n,n'} W_n \left|F_{nn'}\right|^2 \frac{1}{2\pi} \lim_{\varepsilon\to 0} \int_{-\infty}^{+\infty} \cos\frac{\widetilde{E}\tau}{\hbar} e^{-\varepsilon|\tau|} e^{i\omega\tau} d\tau \\ &\equiv \frac{1}{2\pi} \sum_{n,n'} W_n \left|F_{nn'}\right|^2 \lim_{\varepsilon\to 0} \int_0^{+\infty} \left[\exp\left\{\frac{i\widetilde{E}\tau}{\hbar}\right\} + \text{c.c.}\right] e^{-\varepsilon\tau} e^{i\omega\tau} d\tau \\ &= \frac{1}{2\pi} \sum_{n,n'} W_n \left|F_{nn'}\right|^2 \lim_{\varepsilon\to 0} \left[\frac{1}{i\left(\widetilde{E}/\hbar+\omega\right)-\varepsilon} + \frac{1}{i\left(-\widetilde{E}/\hbar+\omega\right)-\varepsilon}\right]. \end{aligned} \tag{7.117}$$

Now it is a convenient time to recall that each of the two summations here is over the eigenenergies of the environment, whose spectrum is virtually continuous because of its large size, so we may transform each sum into an integral – just as this was done in Sec. 6.6:

$$\sum_n \ldots \to \int \ldots dn = \int \ldots \rho(E_n)\, dE_n, \tag{7.118}$$

where $\rho(E) \equiv dn/dE$ is the environment's density of states at a given energy. This transformation yields

$$S_F(\omega) = \frac{1}{2\pi} \lim_{\varepsilon\to 0} \int dE_n W(E_n) \rho(E_n) \int dE_{n'} \rho(E_{n'}) \left|F_{nn'}\right|^2 \left[\frac{1}{i\left(\widetilde{E}/\hbar-\omega\right)-\varepsilon} + \frac{1}{i\left(-\widetilde{E}/\hbar-\omega\right)-\varepsilon}\right]. \tag{7.119}$$

Since the expression inside the square bracket depends only on a specific linear combination of two energies, namely on $\widetilde{E} \equiv E_n - E_{n'}$, it is convenient to introduce another, linearly independent combination of the energies, for example, the average energy $\overline{E} \equiv (E_n + E_{n'})/2$, so the state energies may be represented as

$$E_n = \overline{E} + \frac{\widetilde{E}}{2}, \quad E_{n'} = \overline{E} - \frac{\widetilde{E}}{2}. \tag{7.120}$$

With this notation, Eq. (119) becomes

$$\begin{aligned} S_F(\omega) = -\frac{\hbar}{2\pi} \lim_{\varepsilon\to 0} \int d\overline{E} \Bigg[&\int d\widetilde{E}\, W\left(\overline{E}+\frac{\widetilde{E}}{2}\right) \rho\left(\overline{E}+\frac{\widetilde{E}}{2}\right) \rho\left(\overline{E}-\frac{\widetilde{E}}{2}\right) \left|F_{nn'}\right|^2 \frac{1}{i\left(\widetilde{E}-\hbar\omega\right)-\hbar\varepsilon} \\ &+ \int d\widetilde{E}\, W\left(\overline{E}+\frac{\widetilde{E}}{2}\right) \rho\left(\overline{E}+\frac{\widetilde{E}}{2}\right) \rho\left(\overline{E}-\frac{\widetilde{E}}{2}\right) \left|F_{nn'}\right|^2 \frac{1}{i\left(-\widetilde{E}-\hbar\omega\right)-\hbar\varepsilon} \Bigg]. \end{aligned} \tag{7.121}$$

Due to the smallness of the parameter $\hbar\varepsilon$ (which should be much smaller than all genuine energies of the problem, including $k_{\rm B}T$, $\hbar\omega$, E_n, and $E_{n'}$), each of the internal integrals in Eq. (121) is dominated by

[43] An alternative popular measure of the spectral density of a process $F(t)$ is $\mathcal{S}_F(\nu) \equiv \langle F_{\rm f}^2\rangle/d\nu = 4\pi S_F(\omega)$, where $\nu = \omega/2\pi$ is the "cyclic" frequency (measured in Hz).

an infinitesimal vicinity of one point, $\widetilde{E}_{\pm} = \pm\hbar\omega$. In these vicinities, the state densities, the matrix elements, and the Gibbs probabilities do not change considerably and may be taken out of the integral, which may be then worked out explicitly:[44]

$$S_F(\omega) = -\frac{\hbar}{2\pi}\lim_{\varepsilon\to 0}\int d\overline{E}\,\rho_+\rho_-\left[W_+\left|F_+\right|^2\int_{-\infty}^{+\infty}\frac{d\widetilde{E}}{i\left(\widetilde{E}-\hbar\omega\right)-\hbar\varepsilon}+W_-\left|F_-\right|^2\int_{-\infty}^{+\infty}\frac{d\widetilde{E}}{i\left(-\widetilde{E}-\hbar\omega\right)-\hbar\varepsilon}\right]$$

$$\equiv -\frac{\hbar}{2\pi}\lim_{\varepsilon\to 0}\int d\overline{E}\,\rho_+\rho_-\left[W_+\left|F_+\right|^2\int_{-\infty}^{+\infty}\frac{-i\left(\widetilde{E}-\hbar\omega\right)-\hbar\varepsilon}{\left(\widetilde{E}-\hbar\omega\right)^2+\left(\hbar\varepsilon\right)^2}d\widetilde{E}+W_-\left|F_-\right|^2\int_{-\infty}^{+\infty}\frac{i\left(\widetilde{E}+\hbar\omega\right)-\hbar\varepsilon}{\left(\widetilde{E}+\hbar\omega\right)^2+\left(\hbar\varepsilon\right)^2}d\widetilde{E}\right]$$

$$= \frac{\hbar}{2}\int\rho_+\rho_-\left[W_+\left|F_+\right|^2+W_-\left|F_-\right|^2\right]d\overline{E}\,, \tag{7.122}$$

where the indices ± mark the functions' values at the special points $\widetilde{E}_{\pm} = \pm\hbar\omega$, i.e. $E_n = E_{n'} \pm \hbar\omega$. The physics of these points becomes simple if we interpret the state n, for which the equilibrium Gibbs distribution function equals W_n, as the initial state of the environment, and n' as its final state. Then the top-sign point corresponds to $E_{n'} = E_n - \hbar\omega$, i.e. to the result of emission of one energy quantum $\hbar\omega$ of the "observation" frequency ω by the environment to the system s of our interest, while the bottom-sign point $E_{n'} = E_n + \hbar\omega$, corresponds to the absorption of such quantum by the environment. As Eq. (122) shows, both processes give similar, positive contributions to the force fluctuations.

The situation is different for the Fourier image of the response function $G(\tau)$,[45]

$$\chi(\omega) \equiv \int_0^{+\infty} G(\tau)e^{i\omega\tau}d\tau\,, \tag{7.123}$$

Generalized susceptibility

that is usually called *generalized susceptibility* – in our case, of the environment. Its physical meaning is that according to Eq. (107), the complex function $\chi(\omega) = \chi'(\omega) + i\chi''(\omega)$ relates the Fourier amplitudes of the generalized coordinate and the generalized force: [46]

$$\left\langle\hat{F}_\omega\right\rangle = \chi(\omega)\hat{x}_\omega\,. \tag{7.124}$$

The physics of its imaginary part $\chi''(\omega)$ is especially clear. Indeed, if x_ω represents a sinusoidal *classical* oscillation waveform, say

$$x(t) = x_0\cos\omega t \equiv \frac{x_0}{2}e^{-i\omega t} + \frac{x_0}{2}e^{+i\omega t}, \quad \text{i.e. } x_\omega = x_{-\omega} = \frac{x_0}{2}, \tag{7.125}$$

[44] Using, e.g., MA Eq. (6.5a). (The imaginary parts of the integrals vanish, because the integration in infinite limits may be always re-centered to the finite points $\pm\hbar\omega$.) A math-enlightened reader may have noticed that the integrals might be taken without the introduction of small ε, by using the Cauchy theorem – see MA Eq. (15.1).

[45] The integration in Eq. (123) may be extended to the whole time axis, $-\infty < \tau < +\infty$, if we complement the definition (107) of the function $G(\tau)$ for $\tau > 0$ with its definition as $G(\tau) = 0$ for $\tau < 0$, in correspondence with the causality principle.

[46] To prove this relation, it is sufficient to plug the expression $\hat{x}_s = \hat{x}_\omega e^{-i\omega t}$, or any sum of such exponents, into Eqs. (107) and then use the definition (123). This (simple) exercise is highly recommended to the reader.

then, in accordance with the correspondence principle, Eq. (124) should hold for the *c*-number complex amplitudes F_ω and x_ω, enabling us to calculate the time dependence of the force as

$$F(t) = F_\omega e^{-i\omega t} + F_{-\omega} e^{+i\omega t} = \chi(\omega) x_\omega e^{-i\omega t} + \chi(-\omega) x_{-\omega} e^{+i\omega t} = \frac{x_0}{2}\left[\chi(\omega) e^{-i\omega t} + \chi^*(\omega) e^{+i\omega t}\right]$$
$$= \frac{x_0}{2}\left[(\chi' + i\chi'') e^{-i\omega t} + (\chi' - i\chi'') e^{+i\omega t}\right] \equiv x_0\left[\chi'(\omega)\cos\omega t + \chi''(\omega)\sin\omega t\right]. \quad (7.126)$$

We see that $\chi''(\omega)$ weighs the force's part (frequently called *quadrature*) that is $\pi/2$-shifted from the coordinate x, i.e. is in phase with its velocity and hence characterizes the time-average power flow from the system into its environment, i.e. the *energy dissipation* rate:[47]

$$\overline{\mathscr{P}} = \overline{-F(t)\dot{x}(t)} = \overline{-x_0\left[\chi'(\omega)\cos\omega t + \chi''(\omega)\sin\omega t\right](-\omega x_0 \sin\omega t)} = \frac{x_0^2}{2}\omega\chi''(\omega). \quad (7.127)$$

Let us calculate this function from Eqs. (108) and (123), just as we have done for the spectral density of fluctuations:

$$\chi''(\omega) = \mathrm{Im}\int_0^{+\infty} G(\tau) e^{i\omega\tau} d\tau = \frac{2}{\hbar}\sum_{n,n'} W_n \left|F_{nn'}\right|^2 \lim_{\varepsilon\to 0} \mathrm{Im}\int_0^{+\infty} \frac{1}{2i}\left(\exp\left\{i\frac{\widetilde{E}\tau}{\hbar}\right\} - \text{c.c.}\right) e^{i\omega\tau} e^{-\varepsilon\tau} d\tau$$
$$= \sum_{n,n'} W_n \left|F_{nn'}\right|^2 \lim_{\varepsilon\to 0} \mathrm{Im}\left(\frac{1}{-\widetilde{E} - \hbar\omega - i\hbar\varepsilon} - \frac{1}{\widetilde{E} - \hbar\omega - i\hbar\varepsilon}\right) \quad (7.128)$$
$$\equiv \sum_{n,n'} W_n \left|F_{nn'}\right|^2 \lim_{\varepsilon\to 0}\left(\frac{\hbar\varepsilon}{\left(\widetilde{E} + \hbar\omega\right)^2 + (\hbar\varepsilon)^2} - \frac{\hbar\varepsilon}{\left(\widetilde{E} - \hbar\omega\right)^2 + (\hbar\varepsilon)^2}\right).$$

Making the transfer (118) from the double sum to the double integral, and then the integration variable transfer (120), we get

$$\chi''(\omega) = \lim_{\varepsilon\to 0}\int d\overline{E}\left[\int_{-\infty}^{+\infty} W\left(\overline{E} + \frac{\widetilde{E}}{2}\right)\rho\left(\overline{E} + \frac{\widetilde{E}}{2}\right)\rho\left(\overline{E} - \frac{\widetilde{E}}{2}\right)\left|F_{nn'}\right|^2 \frac{\hbar\varepsilon}{\left(\widetilde{E} + \hbar\omega\right)^2 + (\hbar\varepsilon)^2} d\widetilde{E}\right.$$
$$\left. - \int_{-\infty}^{+\infty} W\left(\overline{E} + \frac{\widetilde{E}}{2}\right)\rho\left(\overline{E} + \frac{\widetilde{E}}{2}\right)\rho\left(\overline{E} - \frac{\widetilde{E}}{2}\right)\left|F_{nn'}\right|^2 \frac{\hbar\varepsilon}{\left(\widetilde{E} - \hbar\omega\right)^2 + (\hbar\varepsilon)^2} d\widetilde{E}\right]. \quad (7.129)$$

Now using the same argument about the smallness of parameter ε as above, we may take the spectral densities, the matrix elements of force, and the Gibbs probabilities out of the integrals, and work out the remaining integrals, getting a result very similar to Eq. (122):

$$\chi''(\omega) = \pi\int \rho_+\rho_-\left[W_-\left|F_-\right|^2 - W_+\left|F_+\right|^2\right] d\overline{E}. \quad (7.130)$$

To relate these results, it is sufficient to notice that according to Eq. (24), the Gibbs probabilities $W_\pm$ are related by a coefficient depending on only the temperature T and observation frequency ω:

[47] The sign minus in Eq. (127) is due to the fact that according to Eq. (90), F is the force exerted *on* our system (s) by the environment, so the force exerted *by* our system on the environment is $-F$. With this sign clarification, the expression $\mathscr{P} = -F\dot{x} = -Fv$ for the instant power flow is evident if x is the usual Cartesian coordinate of a 1D particle. However, according to analytical mechanics (see, e.g., CM Chapters 2 and 10), it is also valid for any {generalized coordinate, generalized force} pair that forms the interaction Hamiltonian (90).

$$W_{\pm} \equiv W\left(\overline{E} + \frac{\widetilde{E}_{\pm}}{2}\right) \equiv W\left(\overline{E} \pm \frac{\hbar\omega}{2}\right) = \frac{1}{Z}\exp\left\{-\frac{\overline{E} \pm \hbar\omega/2}{k_B T}\right\} = W\left(\overline{E}\right)\exp\left\{\mp\frac{\hbar\omega}{2k_B T}\right\}, \tag{7.131}$$

so both the spectral density (122) and the dissipative part (130) of the generalized susceptibility may be expressed via the same integral over the environment energies:

$$S_F(\omega) = \hbar\cosh\left(\frac{\hbar\omega}{2k_B T}\right)\int \rho_+\rho_- W\left(\overline{E}\right)\left[\left|F_+\right|^2 + \left|F_-\right|^2\right]d\overline{E}, \tag{7.132}$$

$$\chi''(\omega) = 2\pi\sinh\left(\frac{\hbar\omega}{2k_B T}\right)\int \rho_+\rho_- W\left(\overline{E}\right)\left[\left|F_+\right|^2 + \left|F_-\right|^2\right]d\overline{E}, \tag{7.133}$$

and hence are universally related as

$$S_F(\omega) = \frac{\hbar}{2\pi}\chi''(\omega)\coth\frac{\hbar\omega}{2k_B T}. \tag{7.134}$$

Fluctuation-dissipation theorem

This is, finally, the much-celebrated Callen-Welton's fluctuation-dissipation theorem (FDT). It reveals a fundamental, intimate relationship between these two effects of the environment ("no dissipation without fluctuation") – hence the name. A curious feature of the FDT is that Eq. (134) includes the same function of temperature as the average energy (26) of a quantum oscillator of frequency ω, though, as the reader could witness, the notion of the oscillator was by no means used in its derivation. As we will see in the next section, this fact leads to rather interesting consequences and even conceptual opportunities.

In the classical limit, $\hbar\omega << k_B T$, the FDT is reduced to

$$S_F(\omega) = \frac{\hbar}{2\pi}\chi''(\omega)\frac{2k_B T}{\hbar\omega} = \frac{k_B T}{\pi}\frac{\mathrm{Im}\,\chi(\omega)}{\omega}. \tag{7.135}$$

In most systems of interest, the last fraction is close to a finite (positive) constant within a substantial range of relatively low frequencies. Indeed, expanding the right-hand side of Eq. (123) into the Taylor series in small ω, we get

$$\chi(\omega) = \chi(0) + i\omega\eta + \ldots, \quad \text{with } \chi(0) = \int_0^\infty G(\tau)d\tau, \quad \text{and } \eta \equiv \int_0^\infty G(\tau)\tau d\tau. \tag{7.136}$$

Since the temporal Green's function G is real by definition, the Taylor expansion of $\chi''(\omega) \equiv \mathrm{Im}\chi(\omega)$ at $\omega = 0$ starts with the linear term $\omega\eta$, where η is a certain real coefficient, and unless $\eta = 0$, is dominated by this term at small ω. The physical sense of the constant η becomes clear if we consider an environment that provides a force described by a simple, well-known kinematic friction law

$$\left\langle\hat{F}\right\rangle = -\eta\,\dot{\hat{x}}, \qquad \text{with } \eta \geq 0, \tag{7.137}$$

where η is usually called the *drag coefficient*. For the Fourier images of coordinate and force, this gives the relation $F_\omega = i\omega\eta x_\omega$, so according to Eq. (124),

$$\chi(\omega) = i\omega\eta, \qquad \text{i.e. } \frac{\chi''(\omega)}{\omega} \equiv \frac{\mathrm{Im}\,\chi(\omega)}{\omega} = \eta \geq 0. \tag{7.138}$$

Ohmic dissipation

With this approximation, and in the classical limit, Eq. (135) is reduced to the well-known *Nyquist formula*:[48]

Nyquist formula

$$S_F(\omega) = \frac{k_B T}{\pi}\eta, \qquad \text{i.e. } \left\langle F_f^2 \right\rangle = 4k_B T \eta d\nu . \tag{7.139}$$

According to Eq. (112), if such a constant spectral density[49] persisted at all frequencies, it would correspond to a delta-correlated process $F(t)$, with

$$K_F(\tau) = 2\pi S_F(0)\delta(\tau) = 2k_B T\eta\delta(\tau) \tag{7.140}$$

– cf. Eqs. (82) and (83). Since in the classical limit, the right-hand side of Eq. (109) is negligible, and the correlation function may be considered an even function of time, the symmetrized function under the integral in Eq. (113) may be rewritten just as $\langle F(\tau)F(0)\rangle$. In the limit of relatively low observation frequencies (in the sense that ω is much smaller than not only the quantum frontier $k_B T/\hbar$ but also the frequency scale of the function $\chi''(\omega)/\omega$), Eq. (138) may be used to recast Eq. (135) in the form[50]

$$\eta \equiv \lim_{\omega\to 0}\frac{\chi''(\omega)}{\omega} = \frac{1}{k_B T}\int_0^\infty \langle F(\tau)F(0)\rangle d\tau . \tag{7.141}$$

To conclude this section, let me return for a minute to the questions formulated in our earlier discussion of dephasing in the two-level model. In that problem, the dephasing time scale is $T_2 = 1/2D_\varphi$. Hence the classical approach to the dephasing, used in Sec. 3, is adequate if $\hbar D_\varphi << k_B T$. Next, we may identify the operators $\hat{f}$ and $\hat{\sigma}_z$ participating in Eq. (70) with, respectively, $\left(-\hat{F}\right)$ and $\hat{x}$ participating in the general Eq. (90). Then the comparison of Eqs. (82), (89), and (140) yields

Classical dephasing time

$$\frac{1}{T_2} \equiv 2D_\varphi = \frac{4k_B T}{\hbar^2}\eta, \tag{7.142}$$

so, for the model described by Eq. (137) with a temperature-independent drag coefficient η, the rate of dephasing by a classical environment is proportional to its temperature.

---

[48] Actually, the 1928 work by H. Nyquist was about the electronic noise in resistors, just discovered experimentally by his Bell Labs colleague John Bertrand Johnson. For an Ohmic resistor, as the dissipative "environment" of the electric circuit it is connected with, Eq. (137) is just the Ohm's law, and may be recast as either $\langle V\rangle = -R(dQ/dt) = RI$, or $\langle I\rangle = -G(d\Phi/dt) = GV$. Thus for the voltage V across an open circuit, η corresponds to its resistance R, while for current I in a short circuit, to its conductance $G = 1/R$. In this case, the fluctuations described by Eq. (139) are referred to as the *Johnson-Nyquist noise*. (Because of this important application, any model leading to Eq. (138) is commonly referred to as the *Ohmic dissipation*, even if the physical nature of the variables x and F is quite different from voltage and current.)

[49] A random process whose spectral density may be reasonably approximated by a constant is frequently called *white noise*, because it is a random mixture of all possible sinusoidal components with equal weights, reminding the spectral composition of natural white light.

[50] Note that in some fields (especially in physical kinetics and chemical physics), this particular limit of the Nyquist formula is called the Green-Kubo (or just the "Kubo") formula. However, in view of the FDT development history (described above), it is much more reasonable to associate these names with Eq. (109) – as it is done in most fields of physics.

7.5. The Heisenberg-Langevin approach

The fluctuation-dissipation theorem offers a very simple and efficient though limited approach to the analysis of the system of interest (*s* in Fig. 1). It is to write its Heisenberg equations (4.199) of motion of the relevant operators, which would now include the environmental force operator, and explore these equations using the Fourier transform and the Wiener-Khinchin theorem (112)-(113). This approach to classical equations of motion is commonly associated with the name Langevin,[51] so its extension to dynamics of Heisenberg-picture operators is frequently referred to as the *Heisenberg-Langevin* (or "quantum Langevin", or "Langevin-Lax"[52]) *approach* to open system analysis.

Perhaps the best way to describe this method is to demonstrate how it works for the very important case of a 1D harmonic oscillator, so the generalized coordinate x of Sec. 4 is just the oscillator's coordinate. For the sake of simplicity, let us assume that the environment provides the simple Ohmic dissipation described by Eq. (137) – which is a very good approximation in many cases. As we already know from Chapter 5, the Heisenberg equations of motion for operators of coordinate and momentum of the oscillator, in the presence of an external force $F(t)$, are

$$\dot{\hat{x}} = \frac{\hat{p}}{m}, \qquad \dot{\hat{p}} = -m\omega_0^2 \hat{x} + \hat{F}, \tag{7.143}$$

so using Eqs. (92) and (137), we get

$$\dot{\hat{x}} = \frac{\hat{p}}{m}, \qquad \dot{\hat{p}} = -m\omega_0^2 \hat{x} - \eta\dot{\hat{x}} + \hat{\tilde{F}}(t). \tag{7.144}$$

Combining both Eqs. (144), we may write their system as a single differential equation

$$m\ddot{\hat{x}} + \eta\dot{\hat{x}} + m\omega_0^2 \hat{x} = \hat{\tilde{F}}(t), \tag{7.145}$$

which is similar to the well-known classical equation of motion of a damped oscillator under the effect of an external force. In the view of Eqs. (5.29) and (5.35), whose corollary the Ehrenfest theorem (5.36) is, this may look not surprising, but please note again that the approach discussed in the previous section justifies such quantitative description of the drag force in quantum mechanics – necessarily in parallel with the accompanying fluctuation force.

For the Fourier images of the operators, defined similarly to Eq. (115), Eq. (145) gives the following relation,

$$\hat{x}_\omega = \frac{\hat{F}_\omega}{m\left(\omega_0^2 - \omega^2\right) - i\eta\omega}, \tag{7.146}$$

which should be also well-known to the reader from the classical theory of forced oscillations.[53] However, since these Fourier components are still Heisenberg-picture operators and their "values" for

[51] A 1908 work by Paul Langevin was the first systematic development of Einstein's ideas (1905) on the Brownian motion, using the random force language, as an alternative to M. Smoluchowski's approach using the probability density language – see Sec. 6 below.

[52] Indeed, perhaps the largest credit for the extension of the Langevin approach to quantum systems belongs to Melvin J. Lax, whose work in the early 1960s was motivated mostly by quantum electronics applications – see, e.g., his monograph M. Lax, *Fluctuation and Coherent Phenomena in Classical and Quantum Physics*, Gordon and Breach, 1968, and references therein.

[53] If necessary, see CM Sec. 5.1.

different ω generally do not commute, we have to tread carefully. The best way to proceed is to write a copy of Eq. (146) for frequency $(-\omega')$, and then combine these equations to form a symmetrical combination similar to that used in Eq. (114). The result is

$$\frac{1}{2}\left\langle \hat{x}_{\omega}\hat{x}_{-\omega'} + \hat{x}_{-\omega'}\hat{x}_{\omega} \right\rangle = \frac{1}{\left| m\left(\omega_0^2 - \omega^2\right) - i\eta\omega \right|^2} \frac{1}{2}\left\langle \hat{F}_{\omega}\hat{F}_{-\omega'} + \hat{F}_{-\omega'}\hat{F}_{\omega} \right\rangle. \tag{7.147}$$

Since the spectral density definition similar to Eq. (114) is valid for any observable, in particular for x, Eq. (147) allows us to relate the symmetrized spectral densities of coordinate and force:

$$S_x(\omega) = \frac{S_F(\omega)}{\left| m\left(\omega_0^2 - \omega^2\right) - i\eta\omega \right|^2} \equiv \frac{S_F(\omega)}{m^2\left(\omega_0^2 - \omega^2\right)^2 + (\eta\omega)^2}. \tag{7.148}$$

Now using an analog of Eq. (116) for x, we can calculate the coordinate's variance:

$$\left\langle x^2 \right\rangle = K_x(0) = \int_{-\infty}^{+\infty} S_x(\omega)\,d\omega = 2\int_0^{+\infty} \frac{S_F(\omega)\,d\omega}{m^2\left(\omega_0^2 - \omega^2\right)^2 + (\eta\omega)^2}, \tag{7.149}$$

where now, in contrast to the notation used in Sec. 4, the sign $\langle \ldots \rangle$ means averaging over the usual statistical ensemble of many systems of interest – in our current case, of many harmonic oscillators.

If the coupling to the environment is so weak that the drag coefficient η is small (in the sense that the oscillator's dimensionless *Q-factor* is large, $Q \equiv m\omega_0/\eta >> 1$), this integral is dominated by the resonance peak in a narrow vicinity, $|\omega - \omega_0| \equiv |\xi| << \omega_0$, of its resonance frequency, and we can take the relatively smooth function $S_F(\omega)$ out of the integral, thus reducing it to a table form:[54]

$$\begin{aligned} \left\langle x^2 \right\rangle &\approx 2S_F(\omega_0)\int_0^{+\infty} \frac{d\omega}{m^2\left(\omega_0^2 - \omega^2\right)^2 + (\eta\omega)^2} \approx 2S_F(\omega_0)\int_{-\infty}^{+\infty} \frac{d\xi}{\left(2m\omega_0\xi\right)^2 + \left(\eta\omega_0\right)^2} \\ &\equiv 2S_F(\omega_0)\frac{1}{(\eta\omega_0)^2}\int_{-\infty}^{+\infty} \frac{d\xi}{(2m\xi/\eta)^2 + 1} = 2S_F(\omega_0)\frac{1}{(\eta\omega_0)^2}\frac{\pi\eta}{2m} = \frac{\pi S_F(\omega_0)}{\eta m\omega_0^2}. \end{aligned} \tag{7.150}$$

With the account of the FDT (134) and of Eq. (138), this gives[55]

$$\left\langle x^2 \right\rangle = \frac{\pi}{\eta m\omega_0^2}\frac{\hbar}{2\pi}\eta\omega_0 \coth\frac{\hbar\omega_0}{2k_BT} \equiv \frac{\hbar}{2m\omega_0}\coth\frac{\hbar\omega_0}{2k_BT}. \tag{7.151}$$

But this is exactly Eq. (48), which was derived in Sec. 2 from the Gibbs distribution, without any explicit account of the environment – though implying it by using the notion of the thermally-equilibrium ensemble.[56]

---

[54] See, e.g., MA Eq. (6.5a).

[55] Note that this calculation remains correct even if the dissipation's dispersion law deviates from the Ohmic model (138), provided that the drag coefficient η is replaced with its effective value $\mathrm{Im}\chi(\omega_0)/\omega_0$.

[56] By the way, the simplest way to calculate $S_F(\omega)$, i.e. to derive the FDT, is to *require* that Eqs. (48) and (150) give the same result for an oscillator with any eigenfrequency ω. This was exactly the approach used by H. Nyquist (for the classical case) – see also SM Sec. 5.5.

Notice that in the final form of Eq. (151), the coefficient η, which characterizes the oscillator-to-environment interaction strength, has canceled! Does this mean that in Sec. 4 we toiled in vain? By no means. First of all, the result (150), augmented by the FDT (134), has an important conceptual value. For example, let us consider the low-temperature limit $k_B T << \hbar\omega_0$ where Eq. (151) is reduced to

$$\left\langle x^2 \right\rangle = \frac{\hbar}{2m\omega_0} \equiv \frac{x_0^2}{2}. \tag{7.152}$$

Let us ask a naïve question: what exactly is the origin of this coordinate's uncertainty? From the point of view of the usual quantum mechanics of absolutely closed (Hamiltonian) systems, there is no doubt: this non-vanishing variance of the coordinate is the result of the final spatial extension of the ground-state wavefunction (2.275), reflecting Heisenberg's uncertainty relation – which in turn results from the fact that the operators of coordinate and momentum do not commute. However, from the point of view of the Heisenberg-Langevin equation (145), the variance (152) is an inalienable part of the oscillator's response to the fluctuation force $\widetilde{F}(t)$ *exerted by the environment* at frequencies $\omega \approx \omega_0$. Though it is impossible to refute the former, absolutely legitimate point of view, in many applications it is easier to subscribe to the latter standpoint and treat the coordinate's uncertainty as the result of the so-called *quantum noise* of the environment, which, in equilibrium, obeys the FTD (134). This notion has received numerous confirmations in experiments that did not include *any* oscillators with their own frequencies ω_0 close to the noise measurement frequency ω.[57]

The second advantage of the Heisenberg-Langevin approach is that it is possible to use Eq. (148) to calculate the (experimentally measurable!) distribution $S_x(\omega)$, i.e. decompose the fluctuations into their spectral components. This procedure is not restricted to the limit of small η (i.e. of large Q); for any damping, we may just plug the FDT (134) into Eq. (148). For example, let us have a look at the so-called *quantum diffusion*. A free 1D particle, moving in a viscous medium providing it with the Ohmic damping (137), may be considered as a particular case of a 1D harmonic oscillator (145), but with $\omega_0 = 0$, so combining Eqs. (134) and (149), we get

$$\left\langle x^2 \right\rangle = 2\int_0^{+\infty} \frac{S_F(\omega)d\omega}{(m\omega^2)^2 + (\eta\omega)^2} = 2\eta\int_0^{+\infty} \frac{1}{(m\omega^2)^2 + (\eta\omega)^2} \frac{\hbar\omega}{2\pi} \coth\frac{\hbar\omega}{2k_B T} d\omega. \tag{7.153}$$

This integral has two divergences. The first one, of the type $\int d\omega/\omega^2$ at the lower limit, is just a classical effect: according to Eq. (85), the particle's displacement variance grows with time, so it cannot have a finite time-independent value that Eq. (153) tries to calculate. However, we still can use that result to single out the quantum effects on diffusion – say, by comparing it with a similar but purely classical case. These effects are prominent at high frequencies, especially if the quantum noise overcomes the thermal noise before the dynamic cut-off, i.e. if

$$\frac{k_B T}{\hbar} << \frac{\eta}{m}. \tag{7.154}$$

In this case, there is a broad range of frequencies where the quantum noise gives a substantial contribution to the integral:

[57] See, for example, R. Koch *et al.*, *Phys. Lev. B* **26**, 74 (1982).

Quantum diffusion

$$\left\langle x^2 \right\rangle_{\rm Q} \approx 2\eta \int_{k_{\rm B}T/\hbar}^{\eta/m} \frac{1}{(\eta\omega)^2}\frac{\hbar\omega}{2\pi}d\omega \equiv \frac{\hbar}{\pi\eta}\int_{k_{\rm B}T/\hbar}^{\eta/m}\frac{d\omega}{\omega} = \frac{\hbar}{\pi\eta}\ln\frac{\hbar\eta}{mk_{\rm B}T} \sim \frac{\hbar}{\eta}. \qquad (7.155)$$

Formally, this contribution diverges at either $m \to 0$ or $T \to 0$, but this logarithmic (i.e. extremely weak) divergence is readily quenched by almost any change of the environment model at very high frequencies, where the "Ohmic" approximation (136) becomes unrealistic.

The Heisenberg-Langevin approach is very powerful because its straightforward generalizations enable analyses of fluctuations in virtually arbitrary linear systems, i.e. the systems described by linear differential (or integro-differential) equations of motion, including those with many degrees of freedom, and distributed systems (*continua*), and such systems dominate many fields of physics. However, this approach also has a major limitation: if the equations of motion of the Heisenberg operators are *not* linear, then there is no linear relation, such as Eq. (146), between the Fourier images of the generalized forces and the generalized coordinates, and as the result, there is no simple relation, such as Eq. (148), between their spectral densities. In other words, if the Heisenberg equations of motion are nonlinear, there is no regular simple way to use them to calculate the statistical properties of the observables.

For example, let us return for a second to the dephasing problem described by Eqs. (68)-(70), and assume that the deterministic and fluctuating parts of the effective force $-f$ exerted by the environment, are characterized by relations similar, respectively, to Eqs. (124) and (134). Now writing the Heisenberg equations of motion for the two remaining spin operators, and using the commutation relations between them, we get

$$\dot{\hat{\sigma}}_x = \frac{1}{i\hbar}\left[\hat{\sigma}_x, \hat{H}\right] = \frac{1}{i\hbar}\left[\hat{\sigma}_x, \left(c_z + \hat{f}\right)\hat{\sigma}_z\right] = -\frac{2}{\hbar}\hat{\sigma}_y\left(c_z + \hat{f}\right) = -\frac{2}{\hbar}\hat{\sigma}_y\left(c_z + \eta\hat{\sigma}_z + \hat{\tilde{f}}\right), \qquad (7.156)$$

and a similar equation for $\dot{\hat{\sigma}}_y$. Such nonlinear equations cannot be used to calculate the statistical properties of the Pauli operators in this system exactly – at least analytically.

For some calculations, this problem may be circumvented by *linearization*: if we are only interested in small fluctuations of the observables, their nonlinear Heisenberg equations of motion, such as Eq. (156), may be linearized with respect to small deviations of the operators from their (generally, deterministic "values", and then the resulting linear equations for the operator variations may be solved either as has been demonstrated above, or (if the deterministic "values" evolve in time) using their Fourier expansions. Sometimes this approach gives relatively simple and important results,[58] but for many other problems, this approach is insufficient, leaving a lot of space for alternative methods.

7.6. Density matrix approach

The main alternative approach to the dynamics of open quantum systems, which is essentially a generalization of the one discussed in Sec. 2, is to extract the final results of interest from the dynamics of the density operator of our system *s*. Let us discuss this approach in detail.[59]

[58] For example, the formula used for processing the experimental results by R. Koch *et al*. (mentioned above), had been derived in this way. (This derivation will be suggested to the reader as an exercise.)

[59] As in Sec. 4, the reader not interested in the derivation of the basic equation (181) of the density matrix evolution may immediately jump to the discussion of this equation and its applications.

We already know that the density matrix allows the calculation of the expectation value of any observable of the system *s* – see Eq. (5). However, our initial recipe (6) for the density matrix element calculation, which requires the knowledge of the exact state (2) of the whole Universe, is not too practicable, while the von Neumann equation (66) for the density matrix evolution is limited to cases in which probabilities W_j of the system states are fixed – thus excluding such important effects as the energy relaxation. However, such effects may be analyzed using a different assumption – that the system of interest interacts only with a *local environment* that is very close to its thermally-equilibrium state described, in the stationary-state basis, by a diagonal density matrix with the elements (24).

This calculation is facilitated by the following general observation. Let us number the basis states of the full local system (the system of our interest plus its local environment) by *l*, and use Eq. (5) to write

$$\langle A \rangle = \mathrm{Tr}\left(\hat{A}\hat{w}_l\right) \equiv \sum_{l,l'} A_{ll'} w_{l'l} = \sum_{l,l'} \langle l|\hat{A}|l'\rangle\langle l'|\hat{w}_l|l\rangle, \tag{7.157}$$

where $\hat{w}_l$ is the full density operator of this local system. At a weak interaction between the system *s* and the local environment *e*, their states reside in different Hilbert spaces, so we can write[60]

$$|l\rangle = |s_j\rangle \otimes |e_k\rangle, \tag{7.158}$$

and if the observable *A* depends only on the coordinates of the system *s* of our interest, we may reduce Eq. (157) to the form similar to Eq. (5):

$$\begin{aligned} \langle A \rangle &= \sum_{j,j';k,k'} \langle e_k| \otimes \langle s_j|\hat{A}|s_{j'}\rangle \otimes |e_{k'}\rangle\langle e_{k'}| \otimes \langle s_{j'}|\hat{w}_l|s_j\rangle \otimes |e_k\rangle \\ &= \sum_{i,i'} A_{jj'} \langle s_{j'}| \otimes \left(\sum_k \langle e_k|\hat{w}_l|e_k\rangle \right) \otimes |s_j\rangle = \mathrm{Tr}_j(\hat{A}\hat{w}), \end{aligned} \tag{7.159}$$

where

$$\hat{w} \equiv \sum_k \langle e_k|\hat{w}_l|e_k\rangle = \mathrm{Tr}_k \hat{w}_l, \tag{7.160}$$

showing how the density operator $\hat{w}$ of the system *s* (sometimes called the *reduced density operator*) may be calculated from the full operator $\hat{w}_l$.

Now comes the key physical assumption of this approach: since we may select the local environment *e* to be much larger than the system *s* of our interest, we may consider the composite system *l* as a Hamiltonian one, with time-independent probabilities of its stationary states, so for the description of the evolution in time of *its* full density operator $\hat{w}_l$ (again, in contrast to that, $\hat{w}$, of the system of our interest) we *may* use the von Neumann equation (66). Partitioning its right-hand side in accordance with Eq. (68), we get:

$$i\hbar\dot{\hat{w}}_l = \left[\hat{H}_s, \hat{w}_l\right] + \left[\hat{H}_e, \hat{w}_l\right] + \left[\hat{H}_{\mathrm{int}}, \hat{w}_l\right]. \tag{7.161}$$

The next step is to use the perturbation theory to solve this equation in the lowest order in $\hat{H}_{\mathrm{int}}$, that would yield a non-vanishing contribution due to the interaction. For that, Eq. (161) is not very

[60] Let me emphasize that this simple representation is valid only for the *basis states* of our local system but, generally, not for its *quantum state*! Indeed, the calculation we are performing is valid (and is most relevant) even when the local system is not in *any* definite quantum state, and may be only described by the density matrix $w_{ll'}$.

convenient, because its right-hand side contains two other terms, of a much larger scale than the interaction Hamiltonian. To mitigate this technical difficulty, the interaction picture that was discussed at the end of Sec. 4.6, is very natural. (It is not necessary though, and I will use this picture mostly as an exercise of its application – unfortunately, the only example I can afford to give in this course.)

As a reminder, in that picture (whose entities will be marked with index "I", with the unmarked operators assumed to be in the Schrödinger picture), both the operators *and* the state vectors (and hence the density operator) depend on time. However, the time evolution of the operator of any observable A is described by an equation similar to Eq. (67), but with the *unperturbed* part of the Hamiltonian only – see Eq. (4.214). In the model (68), this means

$$i\hbar\dot{\hat{A}}_{\rm I} = \left[\hat{A}_{\rm I}, \hat{H}_0\right]. \tag{7.162}$$

where the unperturbed Hamiltonian consists of two parts defined in different Hilbert spaces:

$$\hat{H}_0 \equiv \hat{H}_s + \hat{H}_e. \tag{7.163}$$

On the other hand, the state vector's dynamics is governed by the interaction evolution operator $\hat{u}_{\rm I}$ that obeys Eqs. (4.215). Since this equation, using the interaction-picture Hamiltonian (4.216),

$$\hat{H}_{\rm I} \equiv \hat{u}_0^\dagger \hat{H}_{\rm int}\hat{u}_0, \tag{7.164}$$

is absolutely similar to the ordinary Schrödinger equation using the full Hamiltonian, we may repeat all arguments given at the beginning of Sec. 3 to prove that the dynamics of the density operator in the interaction picture of a Hamiltonian system is governed by the following analog of the von Neumann equation (66):

$$i\hbar\dot{\hat{w}}_{\rm I} = \left[\hat{H}_{\rm I}, \hat{w}_{\rm I}\right], \tag{7.165}$$

where the index l is dropped for the notation simplicity. Since this equation is similar in structure (with the opposite sign) to the Heisenberg equation (67), we may use the solution Eq. (4.190) of the latter equation to write its analog:

$$\hat{w}_{\rm I}(t) = \hat{u}_{\rm I}(t,0)\hat{w}_l(0)\hat{u}_{\rm I}^\dagger(t,0). \tag{7.166}$$

It is also straightforward (and hence is left for the reader) to verify that in this picture, the expectation value of any observable A may be found from an expression similar to the basic Eq. (5):

$$\langle A\rangle = {\rm Tr}\left(\hat{A}_{\rm I}\hat{w}_{\rm I}\right), \tag{7.167}$$

showing again that the interaction and Schrödinger pictures give the same final results.

In the case of the factorable interaction (90),[61] Eq. (162) is simplified for both operators participating in that product – for each one in its own way. In particular, for $\hat{A} = \hat{x}$, it yields

$$i\hbar\dot{\hat{x}}_{\rm I} = \left[\hat{x}_{\rm I}, \hat{H}_0\right] \equiv \left[\hat{x}_{\rm I}, \hat{H}_s\right] + \left[\hat{x}_{\rm I}, \hat{H}_e\right]. \tag{7.168}$$

[61] An analysis of a much more general case when the interaction with the environment is represented as a sum of several/many products of the type (90), may be found, for example, in the monograph by K. Blum, *Density Matrix Theory and Applications*, 3rd ed., Springer, 2012.

Since the coordinate operator is defined in the Hilbert space of our system *s*, it commutes with the Hamiltonian of the environment, so we get

$$i\hbar\dot{\hat{x}}_{\rm I} = \left[\hat{x}_{\rm I}, \hat{H}_s\right]. \tag{7.169}$$

On the other hand, if $\hat{A} = \hat{F}$, this operator is defined in the Hilbert space of the environment, and commutes with the Hamiltonian of the unperturbed system *s*. As a result, we get

$$i\hbar\dot{\hat{F}}_{\rm I} = \left[\hat{F}_{\rm I}, \hat{H}_e\right]. \tag{7.170}$$

This means that with our time-independent unperturbed Hamiltonians, $\hat{H}_s$ and $\hat{H}_e$, the time evolution of the interaction-picture operators is rather simple. In particular, the analogy between Eq. (170) and Eq. (93) allows us to immediately write the following analog of Eq. (94):

$$\hat{F}_{\rm I}(t) = \exp\left\{+\frac{i}{\hbar}\hat{H}_e t\right\}\hat{F}(0)\exp\left\{-\frac{i}{\hbar}\hat{H}_e t\right\}, \tag{7.171}$$

so in the stationary-state basis *n* of the environment,

$$\left(\hat{F}_{\rm I}\right)_{nn'}(t) = \exp\left\{+\frac{i}{\hbar}E_n t\right\}F_{nn'}(0)\exp\left\{-\frac{i}{\hbar}E_{n'}t\right\} \equiv F_{nn'}(0)\exp\left\{-i\frac{E_n - E_{n'}}{\hbar}t\right\}, \tag{7.172}$$

and similarly (but in the basis of the stationary states of system *s*) for the operator $\hat{x}$. As a result, the right-hand side of Eq. (164) may be also factored:

$$\begin{aligned}\hat{H}_{\rm I}(t) &\equiv \hat{u}_0^{\dagger}(t,0)\hat{H}_{\rm int}\hat{u}_0(t,0) = \exp\left\{\frac{i}{\hbar}\left(\hat{H}_s + \hat{H}_e\right)t\right\}\left(-\hat{x}\hat{F}\right)\exp\left\{-\frac{i}{\hbar}\left(\hat{H}_s + \hat{H}_e\right)t\right\} \\ &= -\left(\exp\left\{+\frac{i}{\hbar}\hat{H}_s t\right\}\hat{x}\exp\left\{-\frac{i}{\hbar}\hat{H}_s t\right\}\right)\left(\exp\left\{+\frac{i}{\hbar}\hat{H}_e t\right\}\hat{F}(0)\exp\left\{-\frac{i}{\hbar}\hat{H}_e t\right\}\right) \equiv -\hat{x}_I(t)\hat{F}_{\rm I}(t).\end{aligned} \tag{7.173}$$

So, the transfer to the interaction picture has taken some time, but now it enables a smooth ride.[62] Indeed, just as in Sec. 4, we may rewrite Eq. (165) in the integral form:

$$\hat{w}_{\rm I}(t) = \frac{1}{i\hbar}\int_{-\infty}^{t}\left[\hat{H}_{\rm I}(t'), \hat{w}_{\rm I}(t')\right]dt'; \tag{7.174}$$

plugging this result into the right-hand side of Eq. (165), we get

$$\dot{\hat{w}}_{\rm I}(t) = -\frac{1}{\hbar^2}\int_{-\infty}^{t}\left[\hat{H}_{\rm I}(t), \left[\hat{H}_{\rm I}(t'), \hat{w}_{\rm I}(t')\right]\right]dt' = -\frac{1}{\hbar^2}\int_{-\infty}^{t}\left[\hat{x}(t)\hat{F}(t), \left[\hat{x}(t')\hat{F}(t'), \hat{w}_{\rm I}(t')\right]\right]dt', \tag{7.175}$$

where, for the notation's brevity, from this point on I will strip the operators $\hat{x}$ and $\hat{F}$ of their index "I". (I hope their time dependence indicates the interaction picture clearly enough.)

So far, this equation is exact (and cannot be solved analytically), but this is a good time to notice that even if we approximate the density operator on its right-hand side by its unperturbed, factorable "value"

[62] If we used either the Schrödinger or the Heisenberg picture instead, the forthcoming Eq. (175) would pick up a rather annoying multitude of fast-oscillating exponents, of different time arguments, on its right-hand side.

$$\hat{w}_{\rm I}(t') \to \hat{w}(t')\hat{w}_e, \qquad \text{with } \langle e_n | \hat{w}_e | e_{n'} \rangle = W_n \delta_{nn'}, \tag{7.176}$$

corresponding to no interaction between the system *s* and its thermally-equilibrium environment *e*, where e_n are the stationary states of the environment and W_n are the Gibbs probabilities (24), Eq. (175) still describes a nontrivial time evolution of the density operator.[63] This is exactly the first non-vanishing approximation (in the weak interaction) we have been looking for. Now using Eq. (160), we find the equation of evolution of the density operator of the system of our interest: [64]

$$\dot{\hat{w}}(t) = -\frac{1}{\hbar^2}\int_{-\infty}^{t} {\rm Tr}_n\left[\hat{x}(t)\hat{F}(t), \left[\hat{x}(t')\hat{F}(t'), \hat{w}(t')\hat{w}_e\right]\right]dt', \tag{7.177}$$

where the trace is over the stationary states of the environment. To spell out the right-hand side of Eq. (177), note again that the coordinate and force operators commute with each other (but not with themselves at different time moments!) and hence may be swapped at will, so we may write

$$\begin{aligned}
{\rm Tr}_n[\ldots,[\ldots,\ldots]] &= \hat{x}(t)\hat{x}(t')\hat{w}(t'){\rm Tr}_n\left[\hat{F}(t)\hat{F}(t')\hat{w}_e\right] - \hat{x}(t)\hat{w}(t')\hat{x}(t'){\rm Tr}_n\left[\hat{F}(t)\hat{w}_e\hat{F}(t')\right] \\
&\quad - \hat{x}(t')\hat{w}(t')\hat{x}(t){\rm Tr}_n\left[\hat{F}(t')\hat{w}_e\hat{F}(t)\right] + \hat{w}(t')\hat{x}(t')\hat{x}(t){\rm Tr}_n\left[\hat{w}_e\hat{F}(t')\hat{F}(t)\right] \\
&= \hat{x}(t)\hat{x}(t')\hat{w}(t')\sum_{n,n'} F_{nn'}(t)F_{n'n}(t')W_n - \hat{x}(t)\hat{w}(t')\hat{x}(t')\sum_{n,n'} F_{nn'}(t)W_{n'}F_{n'n}(t') \\
&\quad - \hat{x}(t')\hat{w}(t')\hat{x}(t)\sum_{n,n'} F_{nn'}(t')W_{n'}F_{n'n}(t) + \hat{w}(t')\hat{x}(t')\hat{x}(t)\sum_{n,n'} W_n F_{nn'}(t')F_{n'n}(t).
\end{aligned} \tag{7.178}$$

Since the summation over both indices *n* and *n'* in this expression is over the same energy level set (of all stationary states of the environment), we may swap these indices in any of the sums. Doing this only in the terms including the factors $W_{n'}$, we turn them into W_n, so this factor becomes common:

$$\begin{aligned}
{\rm Tr}_n[\ldots,[\ldots,\ldots]] &= \sum_{n,n'} W_n\left[\hat{x}(t)\hat{x}(t')\hat{w}(t')F_{nn'}(t)F_{n'n}(t') - \hat{x}(t)\hat{w}(t')\hat{x}(t')F_{n'n}(t)F_{nn'}(t')\right. \\
&\quad \left. - \hat{x}(t')\hat{w}\hat{x}(t)F_{n'n}(t')F_{nn'}(t) + \hat{w}\hat{x}(t')\hat{x}(t)F_{nn'}(t')F_{n'n}(t)\right].
\end{aligned} \tag{7.179}$$

Now using Eq. (172), we get

$$\begin{aligned}
{\rm Tr}_n[\ldots,[\ldots,\ldots]] &= \sum_{n,n'} W_n |F_{nn'}|^2 \times \left[\begin{array}{l} \hat{x}(t)\hat{x}(t')\hat{w}(t')\exp\left\{+\dfrac{i\widetilde{E}(t-t')}{\hbar}\right\} - \hat{x}(t)\hat{w}(t')\hat{x}(t')\exp\left\{-\dfrac{i\widetilde{E}(t-t')}{\hbar}\right\} \\ - \hat{x}(t')\hat{w}(t')\hat{x}(t)\exp\left\{+\dfrac{i\widetilde{E}(t-t')}{\hbar}\right\} + \hat{w}(t')\hat{x}(t')\hat{x}(t)\exp\left\{-\dfrac{i\widetilde{E}(t-t')}{\hbar}\right\} \end{array}\right] \\
&\equiv \sum_{n,n'} W_n |F_{nn'}|^2 \cos\frac{\widetilde{E}(t-t')}{\hbar}\left[\hat{x}(t),\left[\hat{x}(t'),\hat{w}(t')\right]\right] + i\sum_{n,n'} W_n |F_{nn'}|^2 \sin\frac{\widetilde{E}(t-t')}{\hbar}\left[\hat{x}(t),\left\{\hat{x}(t'),\hat{w}(t')\right\}\right].
\end{aligned} \tag{7.180}$$

Comparing the two double sums participating in this expression with Eqs. (108) and (111), we see that they are nothing else than, respectively, the symmetrized correlation function and the temporal Green's

[63] This is exactly the moment of transition from the reversible quantum dynamics to irreversible one, in this approach.

[64] For the notation simplicity, the fact that here (and in all following formulas) the density operator $\hat{w}$ of the system *s* of our interest is taken in the interaction picture, is just implied.

function (multiplied by $\hbar/2$) of the time-difference argument $\tau = t - t' \geq 0$. As a result, Eq. (177) takes a compact form:

$$\dot{\hat{w}}(t) = -\frac{1}{\hbar^2}\int_{-\infty}^{t} K_F(t-t')\left[\hat{x}(t),\left[\hat{x}(t'),\hat{w}(t')\right]\right]dt' - \frac{i}{2\hbar}\int_{-\infty}^{t} G(t-t')\left[\hat{x}(t),\left\{\hat{x}(t'),\hat{w}(t')\right\}\right]dt' . \quad (7.181)$$

Density matrix: time evolution

Let me hope that the readers (especially the ones who have braved this derivation) enjoy this beautiful result as much as I do. It gives an equation for the time evolution of the density operator of the system of our interest (s), with the effects of its environment represented only by two real, c-number functions of τ: one (K_F) describing the fluctuation force exerted by the environment, and the other one (G) representing its ensemble-averaged environment's response to the system's evolution. And most spectacularly, these are exactly the same functions that participate in the alternative, Heisenberg-Langevin approach to the problem, and hence related to each other by the fluctuation-dissipation theorem (134).

After a short celebration, let us acknowledge that Eq. (181) is still an integro-differential equation, and needs to be solved together with Eq. (169) for the system coordinate's evolution. Such equations do not allow explicit analytical solutions, besides a few very simple (and not very interesting) cases. For most applications, further simplifications should be made. One of them is based on the fact (which was already discussed in Sec. 3) that both environmental functions participating in Eq. (181) tend to zero when their argument $\tau \equiv t - t'$ becomes much larger than the environment's correlation time τ_c, which is independent of the system-to-environment coupling strength. If the coupling is sufficiently weak, the time scales $T_{nn'}$ of the evolution of the density matrix elements, following from Eq. (181), are much longer than this correlation time, and also than the characteristic time scale of the coordinate operator's evolution. In this limit, all arguments t' of the density operator, giving substantial contributions to the right-hand side of Eq. (181), are so close to t that it does not matter whether its argument is t' or just t. This simplification, $w(t') \approx w(t)$, is known as the *Markov approximation*.[65]

However, this approximation alone is still insufficient for finding the general solution of Eq. (181). Substantial further progress is possible in two important cases. The most important of them is when the intrinsic Hamiltonian $\hat{H}_s$ of the system s of our of interest does not depend on time explicitly and has a *discrete* eigenenergy spectrum E_n,[66] with well-separated levels:

$$\left|E_n - E_{n'}\right| >> \frac{\hbar}{T_{nn'}} . \quad (7.182)$$

Let us see what this condition yields for Eq. (181) rewritten for the matrix elements in the stationary state basis, in the Markov approximation:

[65] Named after A. A. Markov (1856-1922; in older Western literature, "Markoff"), a mathematician famous for his general theory of the so-called *Markov processes*, whose future development is completely determined by its present state, but not its pre-history.

[66] Here, rather reluctantly, I will use this standard notation, E_n, for the eigenenergies of our system of interest (s), in the hope that the reader would not confuse these discrete energy levels with the quasi-continuous energy levels of its environment (e), participating in particular in Eqs. (108) and (111). As a reminder, by this stage of our calculations, the environment's levels have disappeared from our formulas, leaving behind their functionals $K_F(\tau)$ and $G(\tau)$.

$$\dot{w}_{nn'} = -\frac{1}{\hbar^2}\int_{-\infty}^{t} K_F(t-t')\left[\hat{x}(t),\left[\hat{x}(t'),\hat{w}\right]\right]_{nn'} dt' - \frac{i}{2\hbar}\int_{-\infty}^{t} G(t-t')\left[\hat{x}(t),\left\{\hat{x}(t'),\hat{w}\right\}\right]_{nn'} dt' . \quad (7.183)$$

After spelling out the commutators, the right-hand side of this expression includes four operator products, which differ "only" by the operator order. Let us first have a look at one of these products,

$$\left[\hat{x}(t)\hat{x}(t')\hat{w}\right]_{nn'} \equiv \sum_{m,m'} x_{nm}(t)x_{mm'}(t')w_{m'n'} , \quad (7.184)$$

where the indices m and m' run over the same set of stationary states of the system s of our interest as the indices n and n'. According to Eq. (169) with a time-independent H_s, the matrix elements $x_{nn'}$ (in the stationary state basis) oscillate in time as $\exp\{i\omega_{nn'}t\}$, so

$$\left[\hat{x}(t)\hat{x}(t')\hat{w}\right]_{nn'} = \sum_{m,m'} x_{nm}x_{mm'}\exp\left\{i\left(\omega_{nm}t + \omega_{mm'}t'\right)\right\}w_{m'n'} , \quad (7.185)$$

where on the right-hand side, the coordinate matrix elements are in the Schrödinger picture, and the usual notation (6.85) is used for the quantum transition frequencies:

$$\hbar\omega_{nn'} \equiv E_n - E_{n'} . \quad (7.186)$$

According to condition (182), frequencies $\omega_{nn'}$ with $n \neq n'$ are much higher than the speed of evolution of the density matrix elements (in the interaction picture!) – on both the left-hand and right-hand sides of Eq. (183). Hence, on the right-hand side of Eq. (183), we may keep only the terms that do not oscillate with these frequencies $\omega_{nn'}$, because rapidly-oscillating terms would give negligible contributions to the density matrix dynamics.[67] For that, in the double sum (185) we should save only the terms proportional to the difference $(t - t')$ because they will give (after the integration over t') a slowly changing contribution to the right-hand side.[68] These terms should have $\omega_{nm} + \omega_{mm'} = 0$, i.e. $(E_n - E_m) + (E_m - E_{m'}) \equiv E_n - E_{m'} = 0$. For a non-degenerate energy spectrum, this requirement means $m' = n$; as a result, the double sum is reduced to a single one:

$$\left[\hat{x}(t)\hat{x}(t')\hat{w}\right]_{nn'} \approx w_{nn'}\sum_{m} x_{nm}x_{mn}\exp\left\{i\omega_{nm}(t-t')\right\} \equiv w_{nn'}\sum_{m}\left|x_{nm}\right|^2\exp\left\{i\omega_{nm}(t-t')\right\} . \quad (7.187)$$

Another product, $\left[\hat{w}\hat{x}(t')\hat{x}(t)\right]_{nn'}$, which appears on the right-hand side of Eq. (183) may be simplified absolutely similarly, giving

$$\left[\hat{w}\hat{x}(t')\hat{x}(t)\right]_{nn'} \approx \sum_{m}\left|x_{n'm}\right|^2\exp\left\{i\omega_{n'm}(t'-t)\right\}w_{nn'} . \quad (7.188)$$

These expressions hold whether n and n' are equal or not. The situation is different for two other products on the right-hand side of Eq. (183), with w sandwiched between $x(t)$ and $x(t')$. For example,

$$\left[\hat{x}(t)\hat{w}\hat{x}(t')\right]_{nn'} = \sum_{m,m'} x_{nm}(t)w_{mm'}x_{m'n'}(t') = \sum_{m,m'} x_{nm}w_{mm'}x_{m'n'}\exp\left\{i\left(\omega_{nm}t + \omega_{m'n'}t'\right)\right\} . \quad (7.189)$$

[67] This is essentially the same rotating-wave approximation (RWA) as was used in Sec. 6.5.

[68] As was already discussed in Sec. 4, the lower-limit substitution ($t' = -\infty$) in the integrals participating in Eq. (183) gives zero, due to the finite-time "memory" of the system, expressed by the decay of the correlation and response functions at large values of the time delay $\tau = t - t'$.

For this term, the same requirement of keeping an oscillating function of $(t - t')$ only, yields a different condition: $\omega_{nm} + \omega_{m'n'} = 0$, i.e.

$$\left(E_n - E_m\right) + \left(E_{m'} - E_{n'}\right) = 0 . \tag{7.190}$$

Here the double sum's reduction is possible only if we make an additional assumption that all interlevel energy distances are unique, i.e. our system of interest has no equidistant levels (such as in the harmonic oscillator). For the diagonal elements ($n = n'$), the RWA requirement is reduced to $m = m'$, giving sums over all diagonal elements of the density matrix:

$$\left[\hat{x}(t)\hat{w}\hat{x}(t')\right]_{nn} = \sum_m \left|x_{nm}\right|^2 \exp\left\{i\omega_{nm}\left(t - t'\right)\right\} w_{mm} . \tag{7.191}$$

(Another similar term, $\left[\hat{x}(t')\hat{w}\hat{x}(t)\right]_{nn}$, is just a complex conjugate of this one.) However, for off-diagonal matrix elements ($n \neq n'$), the situation is different: Eq. (190) may be satisfied only if $m = n$ and $m' = n'$, so the double sum is reduced to just one, non-oscillating term:

$$\left[\hat{x}(t)\hat{w}\hat{x}(t')\right]_{nn'} = x_{nn} w_{nn'} x_{n'n'}, \quad \text{for } n \neq n' . \tag{7.192}$$

The second similar term, $\left[\hat{x}(t')\hat{w}\hat{x}(t)\right]_{nn}$, is exactly the same, so in the first of the integrals in Eq. (183), these terms add up, while in the second one, they cancel.

This is why the final equations of evolution look differently for diagonal and off-diagonal elements of the density matrix. For the former case ($n = n'$), Eq. (183) is reduced to the so-called *master equation*[69] relating diagonal elements w_{nn} of the density matrix, i.e. the energy level occupancies W_n: [70]

$$\dot{W}_n = \sum_{m \neq n} \left| x_{nm} \right|^2 \int_0^\infty \left[-\frac{1}{\hbar^2} K_F(\tau)\left(W_n - W_m\right)\left(\exp\{i\omega_{nm}\tau\} + \exp\{-i\omega_{nm}\tau\}\right) - \frac{i}{2\hbar} G(\tau)\left(-W_n - W_m\right)\left(\exp\{i\omega_{nm}\tau\} - \exp\{-i\omega_{nm}\tau\}\right)\right] d\tau, \tag{7.193}$$

where $\tau \equiv t - t'$. Changing the summation index notation from m to n', we may rewrite the master equation in its canonical form

$$\dot{W}_n = \sum_{n' \neq n} \left(\Gamma_{n' \to n} W_{n'} - \Gamma_{n \to n'} W_n\right), \tag{7.194}$$

Master equation

where the coefficients

$$\Gamma_{n' \to n} \equiv \left|x_{nn'}\right|^2 \int_0^\infty \left[\frac{2}{\hbar^2} K_F(\tau)\cos\omega_{nn'}\tau - \frac{1}{\hbar} G(\tau)\sin\omega_{nn'}\tau\right] dt', \tag{7.195}$$

Interlevel transition rates

are called the *interlevel transition rates*.[71] Formula (194) has a very clear physical meaning of the level occupancy dynamics (i.e. the balance of the probability flows ΓW) due to quantum transitions between

[69] The master equations, which were first introduced to quantum mechanics in 1928 by W. Pauli, are sometimes called the "Pauli master equations", or "kinetic equations", or "rate equations".

[70] As Eq. (193) shows, the term with $m = n$ would vanish and thus may be legitimately excluded from the sum.

[71] As Eq. (193) shows, the result for $\Gamma_{n \to n'}$ is described by Eq. (195) as well, provided that the indices n and n' are swapped in all components of its right-hand side, including the swap $\omega_{nn'} \to \omega_{n'n} = -\omega_{nn'}$.

the energy levels (see Fig. 7), in our current case caused by the interaction between the system of our interest and its environment.

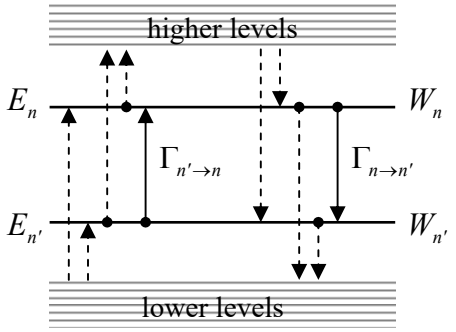


Fig. 7.7. Probability flows in a discrete-spectrum system. Solid arrows: the exchange between the two energy levels, n and n', described by one term in the master equation (194); dashed arrows: other transitions to/from these two levels.

The Fourier transforms (113) and (123) enable us to express the two integrals in Eq. (195) via, respectively, the symmetrized spectral density $S_F(\omega)$ of environment force fluctuations and the imaginary part $\chi''(\omega)$ of the generalized susceptibility, both at frequency $\omega = \omega_{nn'}$. After that we may use the fluctuation-dissipation theorem (134) to exclude the former function, getting finally[72]

Transition rates via $\chi''(\omega)$

$$\Gamma_{n'\to n} = \frac{1}{\hbar}\left|x_{nn'}\right|^2 \chi''(\omega_{nn'})\left(\coth\frac{\hbar\omega_{nn'}}{2k_{\rm B}T} - 1\right) \equiv \frac{2}{\hbar}\left|x_{nn'}\right|^2 \frac{\chi''(\omega_{nn'})}{\exp\{(E_n - E_{n'})/k_{\rm B}T\} - 1}. \qquad (7.196)$$

Note that since the imaginary part χ'' of the generalized susceptibility is an odd function of frequency, Eq. (196) is in compliance with the Gibbs distribution for arbitrary temperature. Indeed, according to this formula, the ratio of the "up" and "down" rates for each pair of levels equals

$$\frac{\Gamma_{n'\to n}}{\Gamma_{n\to n'}} = \frac{\chi''(\omega_{nn'})}{\exp\{(E_n - E_{n'})/k_{\rm B}T\} - 1} \Big/ \frac{\chi''(\omega_{n'n})}{\exp\{(E_{n'} - E_n)/k_{\rm B}T\} - 1} \equiv \exp\left\{-\frac{E_n - E_{n'}}{k_{\rm B}T}\right\}. \qquad (7.197)$$

On the other hand, according to the Gibbs distribution (24), in thermal equilibrium the level populations should be in the same proportion. Hence, Eq. (196) complies with the so-called *detailed balance equation*,

Detailed balance

$$W_n\Gamma_{n\to n'} = W_{n'}\Gamma_{n'\to n}, \qquad (7.198)$$

valid in the equilibrium for each pair $\{n, n'\}$, so all right-hand sides of all Eqs. (194), and hence the time derivatives of all W_n vanish – as they should. Thus, the stationary solution of the master equations indeed describes the thermal equilibrium correctly. As will be shown in the next section, the speed of reaching this equilibrium, within each pair of levels, is determined by the sum of these two rates:

[72] It is straightforward (and highly recommended to the reader as an exercise) to show that at low temperatures ($k_{\rm B}T << |E_{n'} - E_n|$), Eq. (196) gives the same result as the Golden Rate formula (6.111), with $A = x$. (The low-temperature condition ensures that the initial occupancy of the excited level n is negligible, as was assumed at the derivation of Eq. (6.111).)

$$\left(\Gamma_{\Sigma}\right)_{nn'} \equiv \Gamma_{n\to n'} + \Gamma_{n'\to n} = \Gamma_{nn'} \coth\frac{\hbar\omega_{nn'}}{2k_{\mathrm{B}}T} \to \Gamma_{nn'} \times \begin{cases} 1, & \text{for } k_{\mathrm{B}}T << \hbar\omega_{nn'}, \\ 2k_{\mathrm{B}}T/\hbar\omega_{nn'}, & \text{for } k_{\mathrm{B}}T >> \hbar\omega_{nn'}, \end{cases} \quad (7.199)$$

$$\text{with } \Gamma_{nn'} \equiv \frac{2}{\hbar}\left|x_{nn'}\right|^2 \chi''(\omega_{nn'}).$$

In situations when the level specification is obvious, the rate Γ_{Σ} is frequently denoted as the reciprocal *energy relaxation time* $1/T_1$.

The system of master equations (194), frequently complemented by additional right-hand-side terms describing interlevel transitions due to other factors (e.g., by an external ac force with a frequency close to one of $\omega_{nn'}$), is the key starting point for practical analyses of many applied quantum systems, notably including optical quantum amplifiers and generators (lasers). It is important to remember that they are strictly valid only in the rotating-wave approximation, i.e. if the condition (182) is satisfied for all n and n' of substance.

The relaxation times T_1, characterizing the dynamics of the *diagonal* elements of the density matrix, should not be confused with the characteristic times T_2 of the *off-diagonal* element decay, i.e. of the dephasing, which was preliminary discussed in Sec. 3. Now let us see what Eqs. (183) have to say about the dephasing rates. Taking into account our intermediate results (187)-(192), and merging the non-oscillating components (with $m = n$ and $m = n'$) of the sums Eq. (187) and (188) with the terms (192), which also do not oscillate in time, we get the following equation:[73]

$$\dot{w}_{nn'} = -\left\{\int_0^{\infty}\left[\frac{1}{\hbar^2}K_F(\tau)\left(\sum_{m\neq n}\left|x_{nm}\right|^2\exp\{i\omega_{nm}\tau\} + \sum_{m\neq n'}\left|x_{n'm}\right|^2\exp\{-i\omega_{n'm}\tau\} + \left(x_{nn}-x_{n'n'}\right)^2\right)\right.\right.$$
$$\left.\left. + \frac{i}{2\hbar}G(\tau)\left(\sum_{m\neq n}\left|x_{nm}\right|^2\exp\{i\omega_{nm}\tau\} - \sum_{m\neq n'}\left|x_{n'm}\right|^2\exp\{-i\omega_{n'm}\tau\}\right)\right]d\tau\right\}w_{nn'}, \text{ for } n\neq n'. \quad (7.200)$$

In contrast with Eq. (194), the right-hand side of this equation includes both a real and an imaginary part, and hence may be represented as

$$\dot{w}_{nn'} = -\left(1/T_{nn'} + i\Delta_{nn'}\right)w_{nn'}, \quad (7.201)$$

where both factors $1/T_{nn'}$ and $\Delta_{nn'}$ are real. As Eq. (201) shows, the second term in the right-hand side of this equation causes slow oscillations of the matrix elements $w_{nn'}$, which, after returning to the Schrödinger picture, add just small corrections[74] to the unperturbed frequencies (186) of their oscillations, and are not important for most applications. More important is the first term, proportional to

[73] Sometimes Eq. (200) (in any of its numerous alternative forms) is called the *Redfield equation*, after the 1965 work by A. Redfield. Note, however, that in the mid-1960s, several other authors, notably including (in the alphabetical order) H. Haken, W. Lamb, M. Lax, W. Louisell, and M. Scully, also made major contributions to the very fast development of the density-matrix approach to open quantum systems.

[74] Such corrections are sometimes called *Lamb shifts*, because they are the generic form of the particular effect more commonly called the Lamb shift: a minute (~1 GHz) difference between the frequencies of transitions $2s \leftrightarrow 1s$ and $2p \leftrightarrow 1s$ (with all states having $j = ½$) of hydrogen, due to the electric-dipole coupling to the free-space electromagnetic environment, first observed experimentally in 1947 by Willis Lamb and Robert Retherford. (These frequencies have to be equal not only in the non-relativistic theory described in Sec. 3.6 but also in the relativistic quantum theory (see Secs. 6.3, 9.7), if the electromagnetic environment is ignored.) The explanation of

$$\frac{1}{T_{nn'}} = \int_0^\infty \left[\frac{1}{\hbar^2} K_F(\tau) \left(\sum_{m \neq n} |x_{nm}|^2 \cos \omega_{nm} \tau + \sum_{m \neq n'} |x_{n'm}|^2 \cos \omega_{n'm} \tau + (x_{nn} - x_{n'n'})^2 \right) \right.$$
$$\left. - \frac{1}{2\hbar} G(\tau) \left(\sum_{m \neq n} |x_{nm}|^2 \sin \omega_{nm} \tau + \sum_{m \neq n'} |x_{n'm}|^2 \sin \omega_{n'm} \tau \right) \right] d\tau, \quad \text{for } n \neq n', \tag{7.202}$$

because it describes the effect completely absent without the environment coupling: exponential decay of the off-diagonal matrix elements, i.e. the dephasing. Comparing the first two terms of Eq. (202) with Eq. (195), we see that the dephasing rates may be described by a very simple formula:

$$\frac{1}{T_{nn'}} = \frac{1}{2} \left(\sum_{m \neq n} \Gamma_{n \to m} + \sum_{m \neq n'} \Gamma_{n' \to m} \right) + \frac{\pi}{\hbar^2} (x_{nn} - x_{n'n'})^2 S_F(0). \tag{7.203a}$$

Moreover, since at low frequencies, the dissipation provided by most real environments is Ohmic (138), this expression may be further simplified:

Dephasing rate

$$\frac{1}{T_{nn'}} \equiv \frac{1}{2} \left(\sum_{m \neq n} \Gamma_{n \to m} + \sum_{m \neq n'} \Gamma_{n' \to m} \right) + \frac{k_B T}{\hbar^2} \eta (x_{nn} - x_{n'n'})^2, \quad \text{for } n \neq n'. \tag{7.203b}$$

This result shows that two effects yield independent contributions to the dephasing. The first of them may be interpreted as a result of "virtual" transitions of the system from the levels n and n' of our interest to other energy levels m; according to Eq. (196), this contribution is proportional to the strength of coupling to the environment at relatively high frequencies ω_{nm} and $\omega_{n'm}$. On the contrary, the second contribution is due to low-frequency, essentially classical fluctuations of the environment, and hence to the low-frequency dissipative susceptibility. In the Ohmic dissipation case, when the ratio $\eta \equiv \chi''(\omega)/\omega$ is frequency-independent, both contributions are of the same order, but their exact relation depends on the matrix elements $x_{nn'}$ of a particular system. Note also that Eq. (203a), as well as the analysis carried out in Sec. 3, implies that low-frequency fluctuations of any other origin, not taken into account in our analysis (say, an unintentional noise from experimental equipment), may also contribute to dephasing. Such "technical fluctuations" are indeed a very serious challenge for the experimental implementation of coherent qubit systems – see Sec. 8.5 below. On the other hand, in optical systems, the low-frequency contribution to dephasing is usually negligible.

7.7. Application to two-level systems

Let us see what these results mean for the particular but very important case of a two-level system, with just two relevant states. As was discussed in Sec. 3, it may be described by the Hamiltonian (7.68), with the interaction term in the form (7.70) which coincides with Eq. (7.90) used for the calculations in the previous section, at the replacement $\hat{x} \to \hat{\sigma}_z$. However, instead of the simple intrinsic Hamiltonian (7.69) excluding interstate transitions, which was discussed in Sec. 3, the theory described above is valid for the most general Hamiltonian (5.3) in the two-function Hilbert space. As was discussed in Sec. 5.1, after dropping just the trivial term $b\mathrm{I}$, which may be always removed by selecting the proper energy reference, this Hamiltonian, in the usual z-representation, is described by the following matrix

---

the Lamb shift by H. Bethe, in the same 1947, essentially launched the whole field of quantum electrodynamics – to be briefly discussed in Chapter 9.

$$\mathrm{H}_s = \mathbf{c}\cdot\hat{\boldsymbol{\sigma}} \equiv \begin{pmatrix} c_z & c_x - ic_y \\ c_x + ic_y & -c_z \end{pmatrix} \equiv c\begin{pmatrix} \cos\theta & \sin\theta\, e^{-i\varphi} \\ \sin\theta\, e^{i\varphi} & -\cos\theta \end{pmatrix}, \tag{7.204}$$

where θ and φ are the angles describing the direction of the c-number "field vector" $\mathbf{c}$:

$$\mathbf{c} \equiv c_x\mathbf{n}_x + c_y\mathbf{n}_y + c_z\mathbf{n}_z \equiv c\left(\mathbf{n}_x \sin\theta\cos\varphi + \mathbf{n}_y \sin\theta\sin\varphi + \mathbf{n}_z\cos\theta\right). \tag{7.205}$$

As was discussed in Sec. 5.1, this Hamiltonian may describe a large variety of two-level quantum systems, from spin-½ in a magnetic field, where it is just the Pauli Hamiltonian (4.163), to a particle placed into a system of two coupled quantum wells – see Figs. 2.19 and 2.21.

First, let us consider the case when the field vector $\mathbf{c}$ is time-independent. The two stationary states of the system (let us call them + and –), or rather the coefficients $\langle\uparrow\downarrow|\pm\rangle$ of their expansion in the z-basis, and their energies $E_\pm$ may be easily found from the corresponding system of equations (4.102):[75]

$$\begin{aligned} \left(c\cos\theta - E_\pm\right)\left\langle\uparrow\middle|\pm\right\rangle + c\sin\theta\, e^{-i\varphi}\left\langle\downarrow\middle|\pm\right\rangle = 0, \\ c\sin\theta\, e^{i\varphi}\left\langle\uparrow\middle|\pm\right\rangle + \left(-c\cos\theta - E_\pm\right)\left\langle\downarrow\middle|\pm\right\rangle = 0. \end{aligned} \tag{7.206}$$

The results are $E_\pm = \pm c$ (so that $E_+ - E_- = 2c$), and (up to an arbitrary common phase multiplier):

$$\left\langle\uparrow\middle|+\right\rangle = \left\langle\downarrow\middle|-\right\rangle = \cos\frac{\theta}{2}, \qquad \left\langle\downarrow\middle|+\right\rangle = -\left\langle\uparrow\middle|-\right\rangle^* = \sin\frac{\theta}{2}e^{i\varphi}. \tag{7.207}$$

Now we may readily calculate the matrix elements that participate (after the replacement $\hat{x} \to \hat{\sigma}_z$) in Eqs. (196) and (203):

$$\begin{aligned} \left(\sigma_z\right)_{\pm\pm} &\equiv \left\langle\pm\middle|\hat{\sigma}_z\middle|\pm\right\rangle = \left(\left\langle\uparrow\middle|\pm\right\rangle^*, \left\langle\downarrow\middle|\pm\right\rangle^*\right)\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} \left\langle\uparrow\middle|\pm\right\rangle \\ \left\langle\downarrow\middle|\pm\right\rangle \end{pmatrix} = \pm\cos\theta, \\ \left(\sigma_z\right)_{\pm\mp} &\equiv \left\langle\pm\middle|\hat{\sigma}_z\middle|\mp\right\rangle = \left(\left\langle\uparrow\middle|\pm\right\rangle^*, \left\langle\downarrow\middle|\pm\right\rangle^*\right)\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} \left\langle\uparrow\middle|\mp\right\rangle \\ \left\langle\downarrow\middle|\mp\right\rangle \end{pmatrix} = -\sin\theta\, e^{\mp i\varphi}. \end{aligned} \tag{7.208}$$

With these substitutions, Eq. (199) with $\omega_{nn'} = \Omega \equiv 2c/\hbar$ gives

$$\frac{1}{T_1} \equiv \Gamma_\Sigma = \frac{2}{\hbar}\chi''(\Omega)\coth\left(\frac{\hbar\Omega}{k_\mathrm{B}T}\right)\sin^2\theta = \frac{4\Omega}{\hbar}\eta\coth\left(\frac{\hbar\Omega}{k_\mathrm{B}T}\right)\sin^2\theta, \tag{7.209}$$

where the last expression is valid only for the Ohmic dissipation (138), while Eq. (203) yields

$$\frac{1}{T_2} \equiv \frac{1}{T_{+-}} = \frac{1}{2T_1} + \frac{4k_\mathrm{B}T}{\hbar^2}\eta\cos^2\theta. \tag{7.210}$$

So, the relaxation (T_1) and dephasing (T_2) times generally depend on the angle θ between the z-axis (i.e., with our assumption (70), the Bloch-sphere direction of the system's dissipative coupling to the environment) and the dc "field" $\mathbf{c}$ that determines the position of the stationary states on the sphere –

[75] Since particular cases of this procedure were repeatedly performed in Chapter 4 (see also the solutions of Problems 4.27-4.29 and 5.2-5.4), I hope that more detailed explanations are not needed.

see Fig. 5.2.[76] In particular, at $\theta = 0$, i.e. in the absence of interlevel transitions, Eq. (210) reduces to the result (142) obtained under the same assumption, by using the Heisenberg-Langevin approach.

Now let us have a brief look at what the master equations (194), now taking the form

$$\dot{W}_+ = \Gamma_\uparrow W_- - \Gamma_\downarrow W_+, \qquad \dot{W}_- = \Gamma_\downarrow W_+ - \Gamma_\uparrow W_-, \tag{7.211}$$

say about the environment's effects on the two-level system's dynamics. Since the total probability $W_+ + W_-$ of finding the system on some energy level has to equal 1, we can eliminate one of these probabilities from these (compatible) equations, and readily integrate the remaining linear equation, for arbitrary initial conditions, getting

$$W_\pm(t) = W_\pm(0)\exp\left\{-\frac{t}{T_1}\right\} + W_\pm(\infty)\left(1 - \exp\left\{-\frac{t}{T_1}\right\}\right), \tag{7.212}$$

where

$$W_+(\infty) = \frac{\Gamma_\uparrow}{\Gamma_\Sigma} = \frac{1}{\exp\{\hbar\Omega / k_B T\} + 1}, \qquad W_-(\infty) = \frac{\Gamma_\downarrow}{\Gamma_\Sigma} = \frac{1}{\exp\{-\hbar\Omega / k_B T\} + 1}. \tag{7.213}$$

If the system was initially incoherent (i.e. was a classical mixture of two states), then this result, describing the exponential transient of both probabilities from their initial values $W_\pm(0)$ to the final (stationary) values, would be all we could say about its dynamics. However, if the initial state is pure, there is more to the story, and we also need to use the dephasing equation (201) for the off-diagonal matrix element w_{+-}. (According to Eq. (6), $w_{-+} = w_{+-}{}^*$.) Using the relation reciprocal to Eq. (166) to transfer from the interaction picture back to the Schrödinger one, and neglecting the small Lamb-shift correction due to the environment, we get a simple equation:

$$\dot{w}_{+-} = \left(-\frac{1}{T_2} + i\Omega\right) w_{+-}, \tag{7.214}$$

with the obvious solution

$$w_{+-}(t) = w_{+-}(0)\exp\left\{\left(-\frac{1}{T_2} + i\Omega\right)t\right\}, \quad \text{so } w_{-+}(t) = w_{-+}(0)\exp\left\{\left(-\frac{1}{T_2} - i\Omega\right)t\right\}. \tag{7.215}$$

For example, let the system be initially in a pure state $\uparrow$. (For a spin-½, this would mean that it is z-polarized, while for the two-well implementation shown in Fig. 4, this is the state with the particle is definitely in the right well.) Then, according to Eqs. (18b) or (20),

$$\begin{aligned} &W_+(0) \equiv w_{++}(0) = \left|\langle + | \uparrow \rangle\right|^2 = \cos^2\frac{\theta}{2}, \qquad W_-(0) \equiv w_{--}(0) = \left|\langle - | \uparrow \rangle\right|^2 = \sin^2\frac{\theta}{2}, \\ &w_{+-}(0) = \langle - | \uparrow \rangle^* \langle + | \uparrow \rangle = -\sin\frac{\theta}{2}\cos\frac{\theta}{2}e^{i\varphi}, \qquad w_{-+}(0) = w_{+-}^*(0) = -\sin\frac{\theta}{2}\cos\frac{\theta}{2}e^{-i\varphi}, \end{aligned} \tag{7.216}$$

where Eqs. (207) were used. Formally, Eqs. (212)-(216) give the full solution of the problem, but in order to comprehend its meaning, let us have a look at the probability $W_\uparrow(t)$ to find our system in the initial state $\uparrow$ at an arbitrary moment t. Since $t > 0$, due to dephasing, the system's state is no longer pure, in order to recalculate the density matrix elements (216) calculated in the stationary-state basis,

[76] Since in most optical systems, where the low-frequency contribution to dephasing is small, Eq. (210) gives a very simple (and frequently used) relation $T_2 \approx 2T_1$.

into $W_\uparrow(t)$, i.e. one of the elements of the density matrix in the z-basis, we need to use the general Eq. (4.93), which is valid for any operator, in particular for the density operator:

$$W_\uparrow(t) \equiv \langle\uparrow|+\rangle w_{++}(t)\langle\uparrow|+\rangle^* + \langle\uparrow|+\rangle w_{+-}(t)\langle\uparrow|-\rangle^* + \langle\uparrow|-\rangle w_{-+}(t)\langle\uparrow|+\rangle^* + \langle\uparrow|-\rangle w_{--}(t)\langle\uparrow|-\rangle^*. \quad (7.217)$$

Plugging in Eqs. (207), (212), (213), (215), and (216), we finally get

$$W_\uparrow(t) \equiv \left(\sin^4\frac{\theta}{2} + \cos^4\frac{\theta}{2}\right)\exp\left\{-\frac{t}{T_1}\right\} + W_\uparrow(\infty)\left(1 - \exp\left\{-\frac{t}{T_1}\right\}\right) + 2\sin^2\frac{\theta}{2}\cos^2\frac{\theta}{2}\cos\Omega t\,\exp\left\{-\frac{t}{T_2}\right\}, \quad (7.218)$$

where

$$W_\uparrow(\infty) = \frac{\Gamma_\uparrow}{\Gamma_\Sigma}\cos^2\frac{\theta}{2} + \frac{\Gamma_\downarrow}{\Gamma_\Sigma}\sin^2\frac{\theta}{2} \equiv \frac{1}{2}\left(1 - \frac{\Gamma_\downarrow - \Gamma_\uparrow}{\Gamma_\Sigma}\cos\theta\right) \equiv \frac{1}{2}\left[1 - \tanh\left(\frac{\hbar\Omega}{2k_B T}\right)\cos\theta\right]. \quad (7.219)$$

So, the probability of the initial state of the system not only relaxes with the time constant T_1 to its final, thermal-equilibrium values $W_\uparrow(\infty)$ and $W_\downarrow(\infty) = 1 - W_\uparrow(\infty)$, but also performs the quantum oscillations with frequency $\Omega \equiv 2c/\hbar = (E_+ - E_-)/\hbar$ between these two states, decaying with the time constant T_2. (For spin-½ implementations of the two-level system, this means that the spin's precession about the field's direction decays, with the expectation values $\langle S_{x,y}\rangle$ tending to zero at $t \sim T_2$, and $\langle S_z\rangle$ tending to its thermally-equilibrium value $\hbar[W_\uparrow(\infty) - ½]$ at $t \sim T_1$.) In many experimental situations when Eqs. (211) and (214) are valid but the constants T_1 and T_2 cannot be reliably calculated,[77] they may be measured by observation of these two relaxation effects.

More complex problems of this type, for example those described by time-dependent Hamiltonians, may evade such simple analytical solutions because the very notions of stationary state and energy levels cannot be used. However, if the time-dependent part of a Hamiltonian is small and may be considered a perturbation, the two-level system may be still described by Eqs. (211) and (214), with additional terms describing this perturbation. In some of these cases, the Bloch equation (5.22), also with additional terms on its right-hand side, may be very useful for analysis. In Sec. 10, one of such problems is offered to the reader as an exercise.

7.8. Damped harmonic oscillator

As was explained in Section 6, the performed calculations starting from Eq. (191) are not valid for systems with equidistant energy spectra – for example, the harmonic oscillator. For this particular but very important system, with its simple matrix elements $x_{nn'}$ given by Eqs. (5.92), it is longish but straightforward to perform similar calculations, starting from (183), to obtain an equation similar in structure to Eq. (200), but with two other terms, proportional to $w_{n\pm1,n'\pm1}$, on its right-hand side. Neglecting the minor Lamb-shift term, the equation reads

$$\dot{w}_{nn'} = -\delta\left\{\left[(n_e+1)(n+n') + n_e(n+n'+2)\right]w_{nn'} - 2(n_e+1)\left[(n+1)(n'+1)\right]^{1/2}w_{n+1,n'+1} - 2n_e(nn')^{1/2}w_{n-1,n'-1}\right\}. \quad (7.220)$$

[77] For example, in systems whose coupling to the environment cannot be expressed as a single product (46).

Here δ is the effective damping coefficient,[78]

$$\delta \equiv \frac{x_0^2}{2\hbar}\mathrm{Im}\,\chi(\omega_0) \equiv \frac{\mathrm{Im}\,\chi(\omega_0)}{2m\omega_0}, \tag{7.221}$$

equal to just $\eta/2m$ for the Ohmic dissipation, and n_e is the equilibrium number of oscillator's excitations, given by Eq. (26b), with the environment's temperature T. (I am using this new notation because in dynamics, the instant expectation value $\langle n\rangle$ may be time-dependent, and is generally different from its equilibrium value n_e.)

Alternatively, the derivation of Eq. (220) may be started at a bit earlier point, from the Markov approximation applied to Eq. (181), by expressing the coordinate operator via the creation-annihilation operators (5.65). This procedure gives the result in the operator (i.e. basis-independent) form:[79]

$$\dot{\hat{w}} = -\delta\left[(n_e+1)\left(\left\{\hat{a}^\dagger\hat{a},\hat{w}\right\} - 2\hat{a}\hat{w}\hat{a}^\dagger\right) + n_e\left(\left\{\hat{a}\hat{a}^\dagger,\hat{w}\right\} - 2\hat{a}^\dagger\hat{w}\hat{a}\right)\right]. \tag{7.222}$$

In the Fock state basis, this equation immediately reduces to Eq. (220); however, Eq. (222) may be more convenient for some applications.

Returning to Eq. (220), we see that it inter-relates only the elements $w_{nn'}$ located at the same distance $(n - n')$ from the principal diagonal of the density matrix. This means, in particular, that the dynamics of the diagonal elements w_{nn} of the matrix, i.e. the Fock state probabilities W_n, is independent of the off-diagonal elements, and may be represented in the form (194) truncated to the transitions between the adjacent energy levels only ($n' = n \pm 1$):

$$\dot{W}_n = \left(\Gamma_{n+1\to n}W_{n+1} - \Gamma_{n\to n+1}W_n\right) + \left(\Gamma_{n-1\to n}W_{n-1} - \Gamma_{n\to n-1}W_n\right), \tag{7.223}$$

with the following rates:

$$\begin{aligned} &\Gamma_{n+1\to n} = 2\delta(n+1)(n_e+1), \qquad &&\Gamma_{n\to n+1} = 2\delta(n+1)n_e, \\ &\Gamma_{n-1\to n} = 2\delta n n_e, &&\Gamma_{n\to n-1} = 2\delta n(n_e+1). \end{aligned} \tag{7.224}$$

According to the definition of n_e, given by Eq. (26b),

$$n_e = \frac{1}{\exp\{\hbar\omega_0/k_BT\}-1}, \qquad \text{so } n_e+1 = \frac{1}{1-\exp\{-\hbar\omega_0/k_BT\}}, \tag{7.225}$$

---

[78] This coefficient participates prominently in the classical theory of damped oscillations (see, e.g., CM Sec. 5.1), in particular defining the oscillator's Q-factor as $Q \equiv \omega_0/2\delta$, and the decay time of the amplitude A and the energy E of free oscillations: $A(t) = A(0)\exp\{-\delta t\}$, $E(t) = E(0)\exp\{-2\delta t\}$.

[79] Sometimes Eq. (222) is called the *Lindblad equation*, but I believe this terminology is inappropriate. It is true that its structure falls into a general category of equations suggested by G. Lindblad in 1976 for the density operators in the Markov approximation, whose diagonalized form in the interaction picture is

$$\dot{\hat{w}} = \sum_j \gamma_j\left(2\hat{L}_j\hat{w}\hat{L}_j^\dagger - \left\{\hat{L}_j\hat{L}_j^\dagger,\hat{w}\right\}\right).$$

However, Eq. (222) was derived much earlier (by L. Landau in 1927 for zero temperature, and by M. Lax in 1960 for an arbitrary temperature), and in contrast to the general Lindblad equation, spells out the participating operators $\hat{L}_j$ and coefficients γ_j for a particular physical system – the harmonic oscillator.

so taking into account Eqs. (5.92), (186), and (221), and the asymmetry of the function $\chi''(\omega)$, we see that these rates are again described by Eq. (196), even though the last formula was derived for non-equidistant energy spectra.

Hence the only substantial new feature of the master equation for the harmonic oscillator is that the decay of the off-diagonal elements of its density matrix is scaled by the same parameter (2δ) as that of the decay of its diagonal elements, i.e. there is no radical difference between the dephasing and energy-relaxation times T_2 and T_1. This fact may be interpreted as the result of the independence of the energy level distances, $\hbar\omega_0$, of the fluctuations $F(t)$ exerted on the oscillator by the environment, so their low-frequency density, $S_F(0)$, does not contribute to dephasing. (This fact formally follows also from Eq. (203) as well, taking into account that for the oscillator, $x_{nn} = x_{n'n'} = 0$.)

The simple equidistant structure of the oscillator's spectrum makes it possible to readily solve the system of Eqs. (223), with $n = 0, 1, 2, \ldots$, for some important cases. In particular, if the initial state of the oscillator is a classical mixture, with no off-diagonal elements, its further relaxation proceeds as such a mixture: $w_{nn'}(t) = 0$ for all $n' \neq n$.[80] In particular, it is straightforward to use Eq. (208) to verify that if the initial classical mixture obeys the Gibbs distribution (25), but with a temperature T_i different from that of the environment (T_e), then the relaxation process is reduced to a simple exponential transient of the effective temperature from T_i to T_e:

$$W_n(t) = \exp\left\{-n\frac{\hbar\omega_0}{k_B T_{ef}(t)}\right\}\left(1-\exp\left\{-\frac{\hbar\omega_0}{k_B T_{ef}(t)}\right\}\right), \qquad \text{with } T_{ef}(t) = T_i e^{-2\delta t} + T_e\left(1-e^{-2\delta t}\right), \quad (7.226)$$

with the corresponding evolution of the expectation value of the full energy E – cf. Eq. (26b):

$$\langle E\rangle(t) = \frac{\hbar\omega_0}{2} + \hbar\omega_0\langle n\rangle(t), \qquad \langle n\rangle(t) = \frac{1}{\exp\{\hbar\omega_0 / k_B T_{ef}(t)\}-1} \to_{t\to\infty} n_e\,. \quad (7.227)$$

However, if the initial state of the oscillator is different (say, corresponds to some excited Fock state), the relaxation process described by Eqs. (223)-(224), is more complex – see, e.g., Fig. 8.

At low temperatures (Fig. 8a), it may be interpreted as a gradual "roll" of the probability distribution down the energy staircase, with a gradually decreasing velocity $d\langle n\rangle/dt \propto \langle n\rangle$. However, at substantial temperatures, with $k_B T \sim \hbar\omega_0$, i.e. $n_e \sim 1$ (Fig. 8b), this "roll-down" is saturated when the level occupancies $W_n(t)$ approach their equilibrium values (25).[81]

The analysis of this process may be simplified in the case when $W(n, t) \equiv W_n(t)$ is a smooth function of the energy level number n, limited to high levels: $n >> 1$. In this limit, we may use the Taylor expansion of this function (written for the points $\Delta n = \pm 1$), truncated to three leading terms:

$$W_{n\pm1}(t) \equiv W(n\pm1, t) \approx W(n,t) \pm \frac{\partial W(n,t)}{\partial n} + \frac{1}{2}\frac{\partial^2 W(n,t)}{\partial n^2}\,. \quad (7.228)$$

[80] Note, however, that this is not true for many applications, in which a damped oscillator is also under the effect of an external time-dependent field, which may be described by additional, typically off-diagonal terms on the right-hand side of Eqs. (220).

[81] The reader may like to have a look at the results of very nice measurements of such functions $W_n(t)$ in microwave oscillators, performed using their coupling with Josephson-junction circuits: H. Wang *et al.*, *Phys. Rev. Lett.* **101**, 240401 (2008), and with Rydberg atoms: M. Brune *et al.*, *Phys. Rev. Lett.* **101**, 240402 (2008).

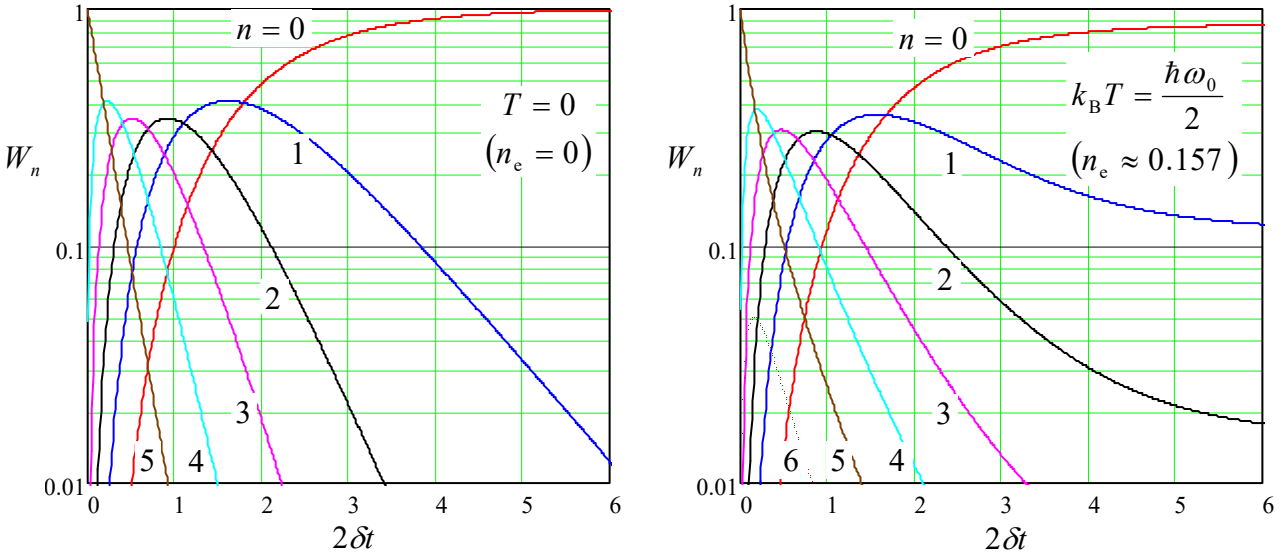


Fig. 7.8. Relaxation of a harmonic oscillator, initially in its 5th Fock state, at: (a) $T = 0$, and (b) $T > 0$. Note that in the latter case, even the energy levels with $n > 5$ get populated, due to their thermal excitation.

Plugging this expression into Eqs. (223)-(224), we get for the function $W(n, t)$ a partial differential equation, which may be recast in the following form:

$$\frac{\partial W}{\partial t} = -\frac{\partial}{\partial n}\left[f(n)W\right] + \frac{\partial^2}{\partial n^2}\left[d(n)W\right], \qquad \text{with } f(n) \equiv 2\delta(n_e - n), \quad d(n) \equiv 2\delta(n_e + \tfrac{1}{2})n. \quad (7.229)$$

Since the energy E of an oscillator with $n >> 1$ is close to $\hbar\omega_0 n$, this *energy diffusion equation* essentially describes the time evolution of the continuous probability density $w(E, t)$ that in this case, may be defined as $w(E, t) \equiv W(E/\hbar\omega_0, t)/\hbar\omega_0$. In the classical limit $n_e >> 1$, Eq. (229) is analytically solvable for arbitrary initial conditions.[82] Note, however, that the most important properties of the damped harmonic oscillator (including its relaxation dynamics) may be analyzed much simpler by using the Heisenberg-Langevin approach which was discussed in Section 5.

7.9. Continuous-spectrum systems

The continuous approximation explored at the end of the last section naturally reminds us of the need to discuss dissipative systems with *continuous* spectra. Unfortunately, for such systems the few (relatively :-) simple results that may be obtained from the basic Eq. (181), are essentially classical in nature and are discussed in detail in the SM part of this series. Here, I will give only a simple illustration.

Let us consider a 1D particle that interacts weakly with a thermally-equilibrium environment, but otherwise is free to move along the x-axis. As we know from Chapters 2 and 5, in this case, the most convenient basis is that of the momentum eigenstates p. In the momentum representation, the density matrix is just the c-number function $w(p, p')$ defined by Eq. (54), which was already discussed in brief

[82] See, e.g., the paper by B. Zeldovich *et al.*, *Sov. Phys. JETP* **28**, 308 (1969), which also gives some more intricate solutions of Eqs. (223)-(224).

in Sec. 2. On the other hand, the coordinate operator, which participates on the right-hand side of Eq. (181), has the form given by the first of Eqs. (4.269),

$$\hat{x} = i\hbar \frac{\partial}{\partial p}, \tag{7.230}$$

dual to the coordinate-representation formula (4.268). As we already know, such operators are local – see, e.g., Eq. (4.244). Due to this locality, the whole right-hand side of Eq. (181) is local as well, and hence (within the framework of our perturbative treatment) the interaction with the environment affects only the diagonal values $w(p, p)$ of the density matrix, i.e. the momentum's probability density $w(p)$.

Let us find the equation governing the evolution of this function in time in the Markov approximation, when the time scale of the density matrix evolution is much longer than the correlation time τ_c of the environment, i.e. the time scale of the functions $K_F(\tau)$ and $G(\tau)$. In this approximation, we may take the matrix elements out of the first integral of Eq. (181),

$$\begin{aligned} -\frac{1}{\hbar^2}\int_{-\infty}^{t} K_F(t-t')dt' \left[\hat{x}(t),\left[\hat{x}(t'),\hat{w}(t')\right]\right] &\approx -\frac{1}{\hbar^2}\int_{0}^{\infty} K_F(\tau)d\tau \left[\hat{x},\left[\hat{x},\hat{w}\right]\right] \\ = -\frac{\pi}{\hbar^2} S_F(0)\left[\hat{x},\left[\hat{x},\hat{w}\right]\right] &= -\frac{k_B T}{\hbar^2}\eta \left[\hat{x},\left[\hat{x},\hat{w}\right]\right], \end{aligned} \tag{7.231}$$

and calculate the last double commutator in the Schrödinger picture. This may be done either using an explicit expression for the matrix elements of the coordinate operator or in a simpler way – using the same trick as at the derivation of the Ehrenfest theorem in Sec. 5.2. Namely, expanding an arbitrary function $f(p)$ into the Taylor series in p,

$$f(p) = \sum_{k=0}^{\infty} \frac{1}{k!}\frac{\partial^k f}{\partial p^k} p^k, \tag{7.232}$$

and using Eq. (240), we can write

$$\left[\hat{x}, f\right] = \sum_{k=0}^{\infty} \frac{1}{k!}\frac{\partial^k f}{\partial p^k}\left[\hat{x}, p^k\right] = \sum_{k=0}^{\infty} \frac{1}{k!}\frac{\partial^k f}{\partial p^k}\left(i\hbar k p^{k-1}\right) = i\hbar \sum_{k=1}^{\infty} \frac{1}{(k-1)!}\frac{\partial^{k-1}}{\partial p^{k-1}}\left(\frac{\partial f}{\partial p}\right) p^{k-1} = i\hbar \frac{\partial f}{\partial p}. \tag{7.233}$$

Now applying this result sequentially, first to w and then to the resulting commutator, we get

$$\left[\hat{x},\left[\hat{x}, w\right]\right] = \left[\hat{x}, i\hbar \frac{\partial w}{\partial p}\right] = i\hbar \frac{\partial}{\partial p}\left(i\hbar \frac{\partial w}{\partial p}\right) = -\hbar^2 \frac{\partial^2 w}{\partial p^2}. \tag{7.234}$$

It may look like the second integral in Eq. (181) might be simplified similarly. However, it vanishes at $p' \to p$, and $t' \to t$, so in order to calculate the first non-vanishing contribution from that integral for $p = p'$, we have to take into account the small difference $\tau \equiv t - t' \sim \tau_c$ between the arguments of the coordinate operators under that integral. This may be done using Eq. (169) with the free particle's Hamiltonian consisting of the kinetic-energy contribution alone:

$$\hat{x}(t') - \hat{x}(t) \approx -\tau \dot{\hat{x}} = -\tau \frac{1}{i\hbar}\left[\hat{x}, \hat{H}_s\right] = -\tau \frac{1}{i\hbar}\left[\hat{x}, \frac{\hat{p}^2}{2m}\right] = -\tau \frac{\hat{p}}{m}, \tag{7.235}$$

where the exact argument of the operator on the right-hand side is already unimportant and may be taken for t. As a result, we may use the last of Eqs. (136) to reduce the second term on the right-hand side of Eq. (181) to

$$-\frac{i}{2\hbar}\int_{-\infty}^{t} G(t-t')\left[\hat{x}(t),\{\hat{x}(t'),\hat{w}(t')\}\right]dt' \approx \frac{i}{2\hbar}\int_{0}^{\infty} G(\tau)\tau d\tau \left[\hat{x},\left\{\frac{\hat{p}}{m},\hat{w}\right\}\right] = \frac{\eta}{2i\hbar}\left[\hat{x},\left\{\frac{\hat{p}}{m},\hat{w}\right\}\right]. \quad (7.236)$$

In the momentum representation, the momentum operator and the density matrix w are just c-numbers and commute, so by applying Eq. (233) to the product pw, we get

$$\left[\hat{x},\left\{\frac{\hat{p}}{m},\hat{w}\right\}\right] = \left[\hat{x},2\frac{p}{m}w\right] = 2i\hbar\frac{\partial}{\partial p}\left(\frac{p}{m}w\right), \quad (7.237)$$

and may finally reduce the integro-differential equation Eq. (181) to a partial differential equation:

Fokker – Planck equation: free 1D particle

$$\frac{\partial w}{\partial t} = -\frac{\partial}{\partial p}(Fw) + \eta k_B T\frac{\partial^2 w}{\partial p^2}, \qquad \text{with } F \equiv -\eta\frac{p}{m}. \quad (7.238)$$

This is the 1D form of the famous *Fokker-Planck equation* describing the classical statistics of motion of a particle (in our particular case, of a free particle) in an environment providing a linear drag characterized by the coefficient η; it belongs to the same drift-diffusion type as Eq. (229). The first, *drift* term on its right-hand side describes the particle's deceleration due to the drag force (137), $F = -\eta p/m = -\eta v$, provided by the environment. The second, *diffusion* term on the right-hand side of Eq. (238) describes the effect of fluctuations: the particle momentum's random walk around its average (drift-affected, and hence time-dependent) value. The walk obeys the law similar to Eq. (85), but with the *momentum-space* diffusion coefficient

$$D_p = \eta k_B T. \quad (7.239)$$

This is the reciprocal-space version of the fundamental Einstein relation between the dissipation (friction) and fluctuations, in this classical limit represented by their thermal energy scale $k_B T$.[83]

The Fokker-Planck equation (238) may be readily generalized to the 3D motion of a particle under the effect of an additional external force,[84] and in this more general form is the basis for many important applications; however, due to its classical character, its discussion is also left for the SM part of this series.[85]

To summarize our discussion of the two alternative approaches to the analysis of quantum systems interacting with a thermally-equilibrium environment, described in the last five sections, let me emphasize again that they give different descriptions of the same phenomena, and are characterized by the same two functions $G(\tau)$ and $K_F(\tau)$. Namely, in the Heisenberg-Langevin approach, we describe the

[83] Note that Eq. (224), as well as the original Einstein's relation between the diffusion coefficient D in the direct space and temperature, may be derived much simpler by other means – for example, from the Nyquist formula (139). These issues are discussed in detail in SM Chapter 5.

[84] Moreover, Eq. (223) may be generalized to the motion of a *quantum* particle in an additional periodic potential $U(\mathbf{r})$. In this case, due to the band structure of the energy spectrum (which was discussed in Secs. 2.7 and 3.4), the coupling to the environment produces not only a continuous drift-diffusion of the probability density in the space of the quasimomentum $\hbar\mathbf{q}$ but also quantum transitions between different energy bands at the same $\hbar\mathbf{q}$ – see, e.g., K. Likharev and A. Zorin, *J. Low Temp. Phys.* **59**, 347 (1985).

[85] See SM Secs. 5.6-5.7. Some examples of *quantum* effects in dissipative systems with continuous spectra, mostly for particular models of the environment, may be found, e.g., in the monographs by U. Weiss, *Quantum Dissipative Systems*, 2nd ed., World Scientific, 1999, and H.-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems*, Oxford U. Press, 2007, and in references therein.

system by operators that change (fluctuate) in time, even in thermal equilibrium, while in the density-matrix approach, the system is described by deterministic probability functions, such as $W_n(t)$ or $w(p, t)$, which are stationary in equilibrium. In all cases when a problem may be solved analytically to the end by both methods (for example, for a harmonic oscillator), they give identical results.

7.10. Exercise problems

7.1. Calculate the density matrix of a two-level system whose Hamiltonian is described, in a certain basis, by the following matrix:

$$\mathrm{H} = \mathbf{c}\cdot\boldsymbol{\sigma} \equiv c_x\sigma_x + c_y\sigma_y + c_z\sigma_z ,$$

where σ_k are the Pauli matrices and c_j are c-numbers, in thermal equilibrium at temperature T.

7.2. In the usual z-basis, spell out the density matrix of a spin-½ with gyromagnetic ratio γ:

(i) in a pure state with the spin definitely directed along the z-axis,
(ii) in a pure state with the spin definitely directed along the x-axis,
(iii) in thermal equilibrium at temperature T, in a magnetic field directed along the z-axis, and
(iv) in thermal equilibrium at temperature T, in a magnetic field directed along the x-axis.

7.3. Calculate the Wigner function of a harmonic oscillator, with mass m and frequency ω_0, in thermodynamic equilibrium at temperature T. Discuss the relation between the result and the Gibbs distribution.

7.4. Calculate the Wigner function of a harmonic oscillator, with mass m and frequency ω_0:

(i) in the ground state,
(ii) in the first excited stationary state ($n = 1$),
(iii) in the Glauber state with an arbitrary dimensionless complex amplitude α, and
(iv) in the so-called *cat state*:[86] a linear superposition of two Glauber states with equal and opposite values of α.

In the last case, explore and interpret the behavior of the function near the origin at $|\alpha| >> 1$.

7.5.* A harmonic oscillator is weakly coupled to an Ohmic environment that is in thermal equilibrium at temperature T.

(i) Use the rotating-wave approximation to write the reduced equations of motion for the Heisenberg operators of the complex amplitude of oscillations.

(ii) Calculate the expectation values of the correlators of the fluctuation force operators participating in these equations, and express them via the average number n_e of thermally-induced excitations in equilibrium, given by the second of Eqs. (26b).

7.6. Calculate the average potential energy of the long-range electrostatic interaction between two similar isotropic 3D harmonic oscillators, each with the electric dipole moment $\mathbf{d} = q\mathbf{s}$, where $\mathbf{s}$ is the oscillator's displacement from its equilibrium position, at arbitrary temperature T.

[86] This state is frequently used to discuss the well-known *Schrödinger cat paradox* – see Sec. 10.1 below.

7.7. A semi-infinite string with mass μ per unit length is attached to a wall and stretched with a constant force (tension) $\mathcal{T}$. Calculate the spectral density of the transverse force exerted on the wall, in thermal equilibrium at temperature T.

7.8.* Calculate the low-frequency spectral density of small fluctuations of the voltage V across a Josephson junction shunted with an Ohmic conductor and biased with a dc external current $\bar{I} > I_c$.

Hint: You may use Eqs. (1.73)-(1.74) to describe the junction's dynamics, and assume that the shunting conductor remains in thermal equilibrium.

7.9. Prove that in the interaction picture of quantum dynamics, the expectation value of an arbitrary observable A may be indeed calculated using Eq. (167).

7.10. Show that the quantum-mechanical Golden Rule (6.149) and the master equation (196) give the same results for the rate of spontaneous quantum transitions $n' \to n$ in a system with a discrete energy spectrum, which is weakly coupled to a low-temperature heat bath (with $k_B T << \hbar\omega_{nn'}$).

Hint: You may establish the relation between the function $\chi''(\omega_{nn'})$ that participates in Eq. (196) and the density of states ρ_n that participates in the Golden Rule, by considering the particular case of sinusoidal classical oscillations in the system of interest.

7.11. A spin-½ with gyromagnetic ratio γ had been placed into a constant magnetic field with magnitude $\mathcal{B} >> k_B T/\hbar\gamma$, and let relax into its ground state. Then the direction of the field was suddenly changed by $\pi/2$ and kept constant after that. Taking into account the spin's weak coupling to a dissipative environment:

(i) calculate the time evolution of the spin's density matrix (in any basis you like), and

(ii) calculate the time evolution of the spin vector's expectation value $\langle \mathbf{S} \rangle$ and sketch its trajectory.

7.12. A spin-½ with gyromagnetic ratio γ is placed into the magnetic field $\mathcal{B}(t) = \mathcal{B}_0 + \tilde{\mathcal{B}}(t)$ with an arbitrary but relatively small time-dependent component, and is also weakly coupled to a dissipative environment in thermal equilibrium at temperature T. Derive the differential equations describing the time evolution of the expectation values of the spin's Cartesian components.

7.13. Use the Bloch equations derived in the previous problem to analyze the magnetic resonance[87] in a spin-½ which is weakly connected to a dissipative environment in thermal equilibrium. Use the result for a semi-quantitative discussion of the environmental broadening of arbitrary quantum transitions in systems with discrete energy spectra.

Hint: You may use the same rotating field model as in Problem 5.5.

7.14. Use the Bloch equations (see the solution of Problem 12) to analyze the dynamics of spin-½ with gyromagnetic ratio γ under the effect of an external ac magnetic field with a relatively low

[87] See the discussion in Sec. 5.2 and the solution of Problem 5.5.

frequency ω and/or large amplitude $\mathscr{B}_{max}$ (so that $|\gamma\mathscr{B}_{max}| >> \omega, 1/T_{1,2}$), assuming that the constants $T_{1,2}$ are field-independent.

7.15. Derive Eq. (220) from Eq. (222).

7.16. For a harmonic oscillator with weak Ohmic dissipation, use Eq. (220) to find the time evolution of the expectation value $\langle E \rangle$ of oscillator's energy for an arbitrary initial state, and compare the result with that following from the Heisenberg-Langevin approach.

7.17. Derive Eq. (234) in an alternative way – by using an expression dual to Eq. (4.244).

Chapter 8. Multiparticle Systems

This chapter provides a brief introduction to the quantum mechanics of systems of similar particles, with special attention to the case when they are indistinguishable. For such systems, theory predicts (and experiment confirms) very specific effects even in the case of negligible explicit ("direct") interactions between the particles, in particular, the Bose-Einstein condensation of bosons and the exchange interaction of fermions. In contrast, the last section of the chapter is devoted to quite a different topic – quantum entanglement of distinguishable systems, and attempts to use this effect for high-performance processing of information.

8.1. Distinguishable and indistinguishable particles

The importance of quantum systems of many similar particles is probably self-evident; just the very fact that most atoms include several/many electrons is sufficient to attract our attention. There are also important systems where the total number of electrons is much higher than in one atom; for example, a cubic centimeter of typical metal houses ~10^{23} conduction electrons that cannot be attributed to particular atoms, and have to be considered common parts of the system as the whole. Though quantum mechanics offers virtually no exact analytical results for systems of substantially interacting particles,[1] it reveals very important new quantum effects even in the simplest cases when particles do not interact, and least explicitly (*directly*).

If non-interacting particles are either different from each other by their nature, or physically similar but still distinguishable because of other reasons, everything is simple – at least, conceptually. Then, as was already discussed in Sec. 6.7, a system of two particles, 1 and 2, each in a pure quantum state, may be described by a state vector

$$\boxed{|\alpha\rangle = |\beta\rangle_1 \otimes |\beta'\rangle_2,} \tag{8.1a}$$

Distinguishable particles

which is a direct product of single-particle vectors, describing their states β and β' defined in different Hilbert spaces. (Below, I will frequently use, for this direct product, the following convenient shorthand:

$$\boxed{|\alpha\rangle = |\beta\beta'\rangle,} \tag{8.1b}$$

in which the particle's number is coded by the state symbol's position.) Hence the *permuted state*

$$\hat{\mathcal{P}}\,|\beta\beta'\rangle \equiv |\beta'\beta\rangle, \tag{8.2}$$

where $\hat{\mathcal{P}}$ is the *permutation operator* (which is defined by Eq. (2) itself), is different from the initial one.

The permutation operator may also be used for states of systems of *identical particles*. In physics, the last term may be used to describe:

[1] As was already noted in Sec. 7.3, for such systems of similar particles, the powerful methods discussed in the last chapter do not work well, because of the absence of a clear difference between some "system of interest" and elementary parts of its "environment".

© K. Likharev

(i) the "really elementary" particles like electrons, which (at least at this stage of development of physics) are considered as structure-less entities, and hence are all identical;

(ii) any objects (e.g., hadrons or mesons) that may be considered as a system of "more elementary" particles (e.g., quarks and gluons), but are placed in the same internal quantum state – most simply, though not necessarily, in the ground state.[2]

It is important to note that *identical* particles still may be *distinguishable* – say by their clear spatial separation. Such systems of similar but distinguishable particles (or subsystems) are broadly discussed nowadays in the context of quantum computing and encryption – see Sec. 5 below. This is why it is insufficient to use the term "identical particles" if we want to say that they are genuinely *indistinguishable*, so I below I will use the latter term, despite it being rather unpleasant grammatically.

It turns out that for a quantitative description of systems of indistinguishable particles, we need to use, instead of direct products of the type (1), linear combinations of such direct products; in the above example, of $|\beta\beta'\rangle$ and $|\beta'\beta\rangle$.[3] To see that, let us discuss the properties of the permutation operator defined by Eq. (2). Consider an observable A, and a system of eigenstates/eigenvalues of its operator:

$$\hat{A}\left|a_j\right\rangle = A_j\left|a_j\right\rangle. \tag{8.3}$$

If the particles are indistinguishable, the observable's expectation value should not be affected by their permutation. Hence the operators $\hat{A}$ and $\hat{\mathcal{P}}$ have to commute and share their eigenstates. This is why the eigenstates of the operator $\hat{\mathcal{P}}$ are so important: in particular, they include the eigenstates of the Hamiltonian, i.e. the stationary states of the system. Let us have a look at the action, on an elementary direct product, of the permutation operator squared:

$$\hat{\mathcal{P}}^2\left|\beta\beta'\right\rangle \equiv \hat{\mathcal{P}}\left(\hat{\mathcal{P}}\left|\beta\beta'\right\rangle\right) = \hat{\mathcal{P}}\left|\beta'\beta\right\rangle = \left|\beta\beta'\right\rangle, \tag{8.4}$$

i.e. $\hat{\mathcal{P}}^2$ brings the state back to its original form. Since any pure state of a two-particle system may be represented as a linear combination of such products, this result does not depend on the state, and may be represented as the following operator relation:

$$\hat{\mathcal{P}}^2 = \hat{I}. \tag{8.5}$$

Now let us find the possible eigenvalues $\mathcal{P}_j$ of the permutation operator. Acting by both sides of Eq. (5) on any of the eigenstates $|\alpha_j\rangle$ of the permutation operator, we get a very simple equation for its eigenvalues:

[2] Note that from this point of view, even complex atoms or molecules, in the same internal quantum state, may be considered on the same footing as the "really elementary" particles. For example, the already mentioned recent spectacular interference experiments by R. Lopes *et al*., which require particle identity, were carried out with couples of ^{4}He atoms in the same internal quantum state.

[3] A very legitimate question is why, in this situation, we need to introduce the particles' numbers to start with. A partial answer is that in this approach, it is much simpler to derive (or guess) the system's Hamiltonians from the correspondence principle – see, e.g., Eq. (27) below. Later in this chapter, we will discuss an alternative approach (the so-called "second quantization"), in which particle numbering is avoided. While that approach is more logical, writing adequate Hamiltonians (which, in particular, would avoid spurious self-interaction of the particles) within it is more challenging – see Sec. 3 below.

$$\mathcal{P}_j^2 = 1, \tag{8.6}$$

with two possible solutions:

$$\mathcal{P}_j = \pm 1. \tag{8.7}$$

Let us find the eigenstates of the permutation operator in the simplest case when each of the two particles can be only in one of two single-particle states – say, β and β'. Evidently, none of the simple products $|\beta\beta'\rangle$ and $|\beta'\beta\rangle$, taken alone, qualifies for such an eigenstate – unless the states β and β' are identical. This is why let us try their linear combination

$$|\alpha_j\rangle = a|\beta\beta'\rangle + b|\beta'\beta\rangle, \tag{8.8}$$

giving

$$\hat{\mathcal{P}}|\alpha_j\rangle = \mathcal{P}_j|\alpha_j\rangle = a|\beta'\beta\rangle + b|\beta\beta'\rangle. \tag{8.9}$$

For the case $\mathcal{P}_j = +1$ we have to require the states (8) and (9) to be the same, so $a = b$, giving the so-called *symmetric eigenstate*[4]

Symmetric entangled eigenstate

$$|\alpha_+\rangle = \frac{1}{\sqrt{2}}\left(|\beta\beta'\rangle + |\beta'\beta\rangle\right), \tag{8.10}$$

where the front coefficient guarantees the orthonormality of the two-particle state vectors, provided that the single-particle vectors are orthonormal. Similarly, for $\mathcal{P}_j = -1$ we get $a = -b$, i.e. an *antisymmetric eigenstate*

Anti-symmetric entangled eigenstate

$$|\alpha_-\rangle = \frac{1}{\sqrt{2}}\left(|\beta\beta'\rangle - |\beta'\beta\rangle\right). \tag{8.11}$$

These are the simplest (two-particle, two-state) examples of *entangled states*, defined as multiparticle system states whose vectors cannot be factored into direct products of single-particle vectors.

So far, our math does not preclude either sign of $\mathcal{P}_j$, in particular the possibility that the sign would depend on the state (i.e. on the index j). Here, however, comes a crucial fact: all indistinguishable particles fall into two groups:[5]

(i) *bosons*, particles with integer spin s, for whose states $\mathcal{P}_j = +1$, and

(ii) *fermions*, particles with half-integer spin, with $\mathcal{P}_j = -1$.

This fundamental connection between the particle's spin and parity ("statistics") can be proved using quantum field theory.[6] In non-relativistic quantum mechanics we are discussing now, it is usually considered experimental; however, our discussion of spin in Chapter 5 enables its following *interpretation*. In free space, the permutation of particles 1 and 2 may be viewed as a result of their pair's common rotation by angle $\phi = \pm\pi$ about an arbitrary z-axis. As we have seen in Sec. 5.7, at the

[4] As in many situations we have met before, the kets given by Eqs. (10) and (11) may be multiplied by the common factor $\exp\{i\varphi\}$ with an arbitrary real phase φ. However, until we discuss coherent superpositions of various states α, there is no good motivation for taking φ different from 0; that would only clutter the notation.

[5] This fact is often described as two different "statistics": the *Bose-Einstein statistics* of bosons and *Fermi-Dirac statistics* of fermions because their statistical distributions in thermal equilibrium are indeed different – see, e.g., SM Sec. 2.8. However, this difference is actually bigger: we are dealing with *two different quantum mechanics*.

[6] Such proofs were first offered by M. Fierz in 1939 and W. Pauli in 1940, and later refined by others.

rotation by this angle, the state vector $|\beta\rangle$ of a particle with a definite quantum number m_s acquires an extra factor $\exp\{\pm i m_s \pi\}$, where the quantum number m_s may be either an integer or a half-integer. As a result, for bosons, i.e. the particles with integer s, m_s can take only integer values, so $\exp\{\pm i m_s \pi\} = \pm 1$, and the product of two such factors in the state vector $|\beta\beta'\rangle$ is equal to +1. On the contrary, for the fermions with their half-integer s, all m_s are half-integer as well, so $\exp\{\pm i m_s \pi\} = \pm i$, and the product of two such factors in the state vector $|\beta\beta'\rangle$ is equal to $(\pm i)^2 = -1$.[7]

The most impressive corollaries of Eqs. (10) and (11) are for the case when the partial states of the two particles are the same: $\beta = \beta'$. The corresponding Bose state α_+ defined by Eq. (10) is possible; in particular, at sufficiently low temperatures, a set of many non-interacting Bose particles may be in the same ground state – the so-called *Bose-Einstein condensate* ("BEC").[8] The most fascinating feature of the condensates is that their dynamics is governed by quantum mechanical laws, which may show up in the behavior of their observables with virtually no quantum uncertainties[9] – see, e.g., Eqs. (1.73)-(1.74).

On the other hand, if we take $\beta = \beta'$ in Eq. (11), we see that the Fermi state α_- becomes the null state, i.e. *cannot exist* at all. This is the mathematical expression of *Pauli's exclusion principle*:[10] two indistinguishable fermions cannot be placed into the same quantum state. (As will be discussed below, this is true for systems with more than two fermions as well.) Perhaps, the key importance of this principle is obvious: if it were not valid for electrons (that are fermions), all electrons of each atom would condense in their ground (1*s*-like) state, and all the usual chemistry (and biochemistry, and biology, including dear us!) would not exist. Thus, the Pauli principle makes fermions *indirectly* interact even if they do not interact *directly*, in the usual sense of the word "interaction".

8.2. Singlets, triplets, and the exchange interaction

Now let us discuss possible approaches to quantitative analyses of identical particles, starting from a simple case of two spin-½ particles (say, electrons), whose explicit interaction with each other and the external world does not involve spin. The description of such a system may be based on factorable states with ket-vectors

$$|\alpha_-\rangle = |o_{12}\rangle \otimes |s_{12}\rangle, \tag{8.12}$$

with the orbital state vector $|o_{12}\rangle$ and the spin vector $|s_{12}\rangle$ belonging to different Hilbert spaces. It is frequently convenient to use the coordinate representation of such a state, sometimes called the *spinor*:

$$\langle \mathbf{r}_1, \mathbf{r}_2 | \alpha_- \rangle = \langle \mathbf{r}_1, \mathbf{r}_2 | o_{12} \rangle \otimes |s_{12}\rangle \equiv \psi(\mathbf{r}_1, \mathbf{r}_2) |s_{12}\rangle. \tag{8.13}$$

2-particle spinor

[7] Unfortunately, the simple generalization of these arguments to an arbitrary quantum state runs into problems, so they cannot serve as a strict proof of the universal relation between s and $\mathcal{P}_f$.

[8] For a quantitative discussion of the Bose-Einstein condensation, see, e.g., SM Sec. 3.4. Examples of such condensates include *superfluids* like helium, Cooper-pair condensates in superconductors, and BECs of weakly interacting atoms.

[9] For example, for a coherent condensate of $N >> 1$ particles of mass m, Heisenberg's uncertainty relation takes the form $\delta x \delta p = \delta x \delta(Nmv) \geq \hbar/2$, so its coordinate x and velocity v may be measured simultaneously with much higher precision than those of a single particle.

[10] It was first formulated for electrons by Wolfgang Pauli in 1925, on the background of less general rules suggested by Gilbert Lewis (1916), Irving Langmuir (1919), Niels Bohr (1922), and Edmund Stoner (1924) for the explanation of experimental spectroscopic data.

Since the spin-½ particles are fermions, the particle permutation has to change the spinor's sign:

$$\hat{\mathcal{P}}\psi(\mathbf{r}_1,\mathbf{r}_2)\left|s_{12}\right\rangle \equiv \psi(\mathbf{r}_2,\mathbf{r}_1)\left|s_{21}\right\rangle = -\psi(\mathbf{r}_1,\mathbf{r}_2)\left|s_{12}\right\rangle, \tag{8.14}$$

i.e. to change the sign of either its orbital factor or the spin factor.

In particular, in the case of symmetric orbital factor,

$$\psi(\mathbf{r}_2,\mathbf{r}_1) = \psi(\mathbf{r}_1,\mathbf{r}_2), \tag{8.15}$$

the spin factor has to obey the relation

$$\left|s_{21}\right\rangle = -\left|s_{12}\right\rangle. \tag{8.16}$$

Let us use the ordinary z-basis (where z, in the absence of an external magnetic field, is an arbitrary spatial axis) for both spins. In this basis, the ket-vector of any two spins-½ may be represented as a linear combination of the following four basis vectors:

$$\left|\uparrow\uparrow\right\rangle, \quad \left|\downarrow\downarrow\right\rangle, \quad \left|\uparrow\downarrow\right\rangle, \quad \text{and} \quad \left|\downarrow\uparrow\right\rangle. \tag{8.17}$$

The first two kets evidently do not satisfy Eq. (16), and cannot participate in the state. Applying to the remaining kets the same argumentation as has resulted in Eq. (11), we get

Singlet state

$$\boxed{\left|s_{12}\right\rangle = \left|s_-\right\rangle \equiv \frac{1}{\sqrt{2}}\left(\left|\uparrow\downarrow\right\rangle - \left|\downarrow\uparrow\right\rangle\right).} \tag{8.18}$$

Such an orbital-symmetric and spin-antisymmetric state is called the *singlet*.

The origin of this term becomes clear from the analysis of the opposite (orbital-antisymmetric and spin-symmetric) case:

$$\psi(\mathbf{r}_2,\mathbf{r}_1) = -\psi(\mathbf{r}_1,\mathbf{r}_2), \qquad \left|s_{12}\right\rangle = \left|s_{21}\right\rangle. \tag{8.19}$$

For the composition of such a symmetric spin state, the first two kets of Eq. (17) are completely acceptable (with arbitrary weights), and so is an entangled spin state that is the symmetric combination of the two last kets, similar to Eq. (10):

$$\left|s_+\right\rangle \equiv \frac{1}{\sqrt{2}}\left(\left|\uparrow\downarrow\right\rangle + \left|\downarrow\uparrow\right\rangle\right), \tag{8.20}$$

so the general spin state is a *triplet*:

Triplet state

$$\boxed{\left|s_{12}\right\rangle = c_+\left|\uparrow\uparrow\right\rangle + c_-\left|\downarrow\downarrow\right\rangle + c_0\frac{1}{\sqrt{2}}\left(\left|\uparrow\downarrow\right\rangle + \left|\downarrow\uparrow\right\rangle\right).} \tag{8.21}$$

Note that any such state (with any values of the coefficients c satisfying the normalization condition), corresponds to the same orbital wavefunction and hence the same energy. However, each of these three states has a specific value of the z-component of the net spin – evidently equal to, respectively, $+\hbar$, $-\hbar$, and 0. Because of this, even a small external magnetic field lifts their degeneracy, splitting the energy level in three; hence the term "triplet".

In the particular case when the particles do not interact directly, for example

$$\hat{H} = \hat{h}_1 + \hat{h}_2, \qquad \hat{h}_k = \frac{\hat{p}_k^2}{2m} + \hat{u}(\mathbf{r}_k), \qquad \text{with } k = 1,2, \tag{8.22}$$

the two-particle Schrödinger equation for the symmetrical orbital wavefunction (15) is obviously satisfied by the direct products,

$$\psi(\mathbf{r}_1,\mathbf{r}_2)=\psi_n(\mathbf{r}_1)\psi_{n'}(\mathbf{r}_2), \tag{8.23}$$

of single-particle eigenfunctions, with arbitrary sets n, n' of quantum numbers. For the particular but very important case $n = n'$, this means that the eigenenergy of the (only acceptable) singlet state,

$$\frac{1}{\sqrt{2}}\Big(\big|\uparrow\downarrow\big\rangle-\big|\downarrow\uparrow\big\rangle\Big)\psi_n(\mathbf{r}_1)\psi_n(\mathbf{r}_2), \tag{8.24}$$

is just $2\varepsilon_n$, where ε_n is the single-particle energy level.[11] In particular, for the ground state of the system, such singlet spin state gives the lowest energy $E_g = 2\varepsilon_g$, while any triplet spin state (19) would require one of the particles to be in a different orbital state, i.e. in a state of higher energy, so the total energy of the system would be also higher.

Now moving to the systems in which two indistinguishable spin-½ particles do interact, let us consider, as the simplest but important[12] example, the lower energy states of a neutral atom[13] of helium – more exactly, ^{4}He. Such an atom consists of a nucleus with two protons and two neutrons, with the total electric charge $q = +2e$, and two electrons "rotating" about the nucleus. Neglecting the small relativistic effects that were discussed in Sec. 6.3, the Hamiltonian describing the electron motion may be expressed as

$$\hat{H}=\hat{h}_1+\hat{h}_2+\hat{U}_{\rm int},\qquad \hat{h}_k=\frac{\hat{p}_k^2}{2m}-\frac{2e^2}{4\pi\varepsilon_0 r_k},\qquad \hat{U}_{\rm int}=\frac{e^2}{4\pi\varepsilon_0\left|\mathbf{r}_1-\mathbf{r}_2\right|}. \tag{8.25}$$

As with most problems of multiparticle quantum mechanics, the eigenvalue/eigenstate problem for this Hamiltonian does not have an exact analytical solution, so let us carry out its approximate analysis considering the electron-electron interaction $U_{\rm int}$ as a perturbation. As was discussed in Chapter 6, we have to start with the "0$^{\rm th}$-order" approximation in which the perturbation is ignored, so the Hamiltonian is reduced to the sum (22). In this approximation, the ground state of the atom is the singlet (24), with the orbital factor

$$\psi_g(\mathbf{r}_1,\mathbf{r}_2)=\psi_{100}(\mathbf{r}_1)\psi_{100}(\mathbf{r}_2), \tag{8.26}$$

and energy $E_g = 2\varepsilon_g$. Here each factor $\psi_{100}(\mathbf{r})$ is the single-particle wavefunction of the ground (1s) state of the hydrogen-like atom with $Z = 2$, with quantum numbers $n = 1$, $l = 0$, and $m = 0$ – hence the wavefunctions' indices. According to Eqs. (3.174) and (3.208),

$$\psi_{100}(\mathbf{r})=Y_0^0(\theta,\varphi)\mathcal{R}_{1,0}(r)=\frac{1}{\sqrt{4\pi}}\frac{2}{r_0^{3/2}}e^{-r/r_0},\quad \text{with } r_0=\frac{r_{\rm B}}{Z}=\frac{r_{\rm B}}{2}, \tag{8.27}$$

and according to Eqs. (3.191) and (3.201), in this approximation the total ground state energy is

[11] In this chapter, I try to use lowercase letters for all *single-particle* observables (in particular, ε for their energies), in order to distinguish them as clearly as possible from the *system's* observables (including the total energy E of the system), which are typeset in uppercase (capital) letters.

[12] Indeed, helium makes up more than 20% of all "ordinary" matter of our Universe.

[13] Note that the positive ion He^{+1} of this atom, with just one electron, is fully described by the hydrogen-like atom theory with $Z = 2$, whose ground-state energy, according to Eq. (3.191), is $-Z^2E_{\rm H}/2 = -2E_H \approx -55.4$ eV.

$$E_g^{(0)} = 2\varepsilon_g^{(0)} = 2\left(-\frac{\varepsilon_0}{2n^2}\right)_{n=1,Z=2} = 2\left(-\frac{Z^2 E_H}{2}\right)_{Z=2} = -4E_H \approx -109\,\text{eV}. \quad (8.28)$$

This is still somewhat far (though not terribly far!) from the experimental value $E_g \approx -78.8$ eV – see the bottom level in Fig. 1a.

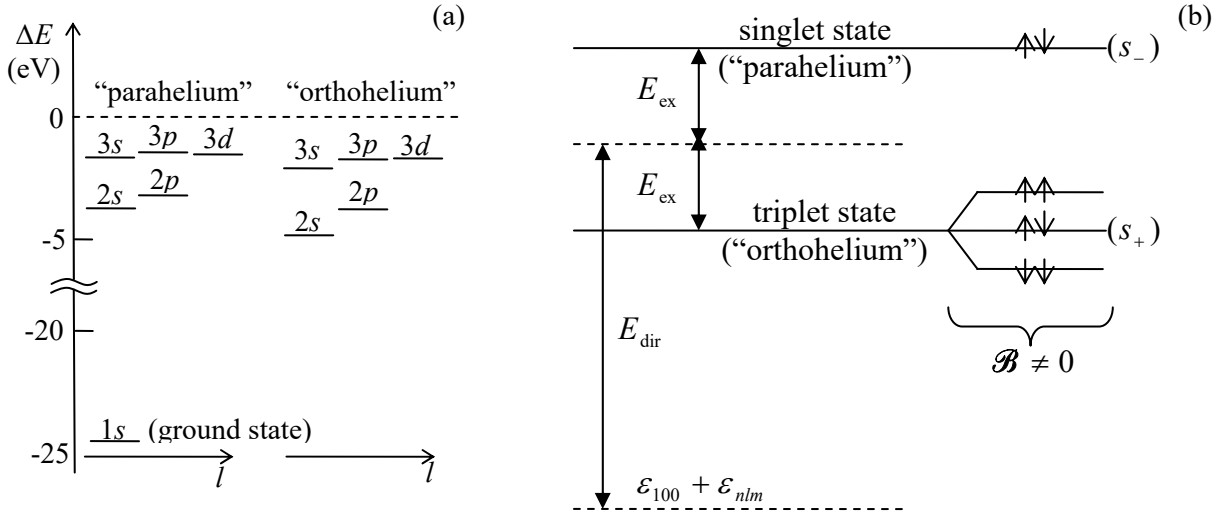


Fig. 8.1. The lower energy levels of a helium atom: (a) experimental data and (b) a schematic structure of an excited state in the first order of the perturbation theory. On panel (a), all energies are referred to that ($-2E_H \approx -55.4$ eV) of the ground state of the positive ion He^{+1}, so their magnitudes are the (readily measurable) energies of the atom's single ionization starting from the corresponding state of the neutral atom. Note that the "spin direction" nomenclature on panel (b) is rather crude: it does not reflect the difference between the entangled states s_+ and s_-.

Making a minor (but very useful) detour from our main topic, let us note that we can get a much better agreement with experiment by accounting for the electron interaction energy in the 1st order of the perturbation theory. Indeed, in application to our system, Eq. (6.14) reads

$$E_g^{(1)} = \langle g|\hat{U}_{int}|g\rangle = \int d^3r_1 \int d^3r_2\, \psi_g^*(\mathbf{r}_1,\mathbf{r}_2) U_{int}(\mathbf{r}_1,\mathbf{r}_2)\psi_g(\mathbf{r}_1,\mathbf{r}_2). \quad (8.29)$$

Plugging in Eqs. (25)-(27), we get

$$E_g^{(1)} = \left(\frac{1}{4\pi}\frac{4}{r_0^3}\right)^2 \int d^3r_1 \int d^3r_2 \frac{e^2}{4\pi\varepsilon_0|\mathbf{r}_1-\mathbf{r}_2|} \exp\left\{-\frac{2(r_1+r_2)}{r_0}\right\}. \quad (8.30)$$

As may be readily evaluated analytically (this exercise is left for the reader), this expression equals $(5/4)E_H$, so the corrected ground state energy,

$$E_g \approx E_g^{(0)} + E_g^{(1)} = (-4+5/4)E_H = -74.8\,\text{eV}, \quad (8.31)$$

is much closer to experiment.

There is still room here for a ready improvement by using the variational method discussed in Sec. 2.9. For our particular case of the ^{4}He atom, we may try to use, as the trial state, the orbital wavefunction given by Eqs. (26)-(27), but with the atomic number Z considered as an adjustable

parameter $Z_{ef} < Z = 2$ rather than a fixed number. The physics behind this approach is that the electric charge density $\rho(\mathbf{r}) = -e|\psi(\mathbf{r})|^2$ of each electron forms a negatively charged "cloud" that reduces the effective charge of the nucleus, as seen by the other electron, to $Z_{ef}e$, with some $Z_{ef} < 2$. As a result, the single-particle wavefunction spreads further in space (with the scale $r_0 = r_B/Z_{ef} > r_B/Z$), while keeping its functional form (27) nearly intact. Since the kinetic energy T in the system's Hamiltonian (25) is proportional to $r_0^{-2} \propto Z_{ef}^2$, while the potential energy is proportional to $r_0^{-1} \propto Z_{ef}^1$, we can write

$$E_g(Z_{ef}) = \left(\frac{Z_{ef}}{2}\right)^2 \langle T_g \rangle_{Z=2} + \frac{Z_{ef}}{2}\langle U_g \rangle_{Z=2}. \tag{8.32}$$

Now we can use the fact that according to Eq. (3.212), for any stationary state of a hydrogen-like atom (just as for the classical circular motion in the Coulomb potential), $\langle U \rangle = 2E$, and hence $\langle T \rangle = E - \langle U \rangle = -E$. Using Eq. (30), and adding the correction (31) to the potential energy, we get

$$E_g(Z_{ef}) = \left[4\left(\frac{Z_{ef}}{2}\right)^2 + \left(-8 + \frac{5}{4}\right)\frac{Z_{ef}}{2}\right]E_H. \tag{8.33}$$

This expression allows an elementary calculation of the optimal value of Z_{ef}, and the corresponding minimum of the function $E_g(Z_{ef})$:

$$(Z_{ef})_{opt} = 2\left(1 - \frac{5}{32}\right) = 1.6875, \qquad (E_g)_{min} \approx -2.85E_H \approx -77.5\ \text{eV}. \tag{8.34}$$

Given the trial state's crudeness, this number is in surprisingly good agreement with the experimental value cited above, with a difference of the order of 1%.

Now let us return to the main topic of this section – the effects of the particle (in this case, electron) indistinguishability. As we have just seen, the ground-level energy of the helium atom is not affected directly by this fact; the situation is different for its excited states – even the lowest ones. The reasonably good precision of the perturbation theory, which we have seen for the ground state, tells us that we can base our analysis of wavefunctions (ψ_e) of the lowest excited state orbitals, on products like $\psi_{100}(\mathbf{r}_k)\psi_{nlm}(\mathbf{r}_{k'})$, with $n > 1$. To satisfy the fermion permutation rule, $\mathcal{P}_j = -1$, we have to take the orbital factor of the state in either the symmetric or the antisymmetric form:

Orthohelium and parahelium: orbital wavefunctions

$$\psi_e(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\sqrt{2}}\left[\psi_{100}(\mathbf{r}_1)\psi_{nlm}(\mathbf{r}_2) \pm \psi_{nlm}(\mathbf{r}_1)\psi_{100}(\mathbf{r}_2)\right], \tag{8.35}$$

with the proper total permutation asymmetry provided by the corresponding spin factor (18) or (21), so the upper/lower sign in Eq. (35) corresponds to the singlet/triplet spin state. Let us calculate the expectation values of the total energy of the system in the first order of the perturbation theory. Plugging Eq. (35) into the 0^{th}-order expression

$$\langle E_e \rangle^{(0)} = \int d^3r_1 \int d^3r_2\, \psi_e^*(\mathbf{r}_1, \mathbf{r}_2)\left(\hat{h}_1 + \hat{h}_2\right)\psi_e(\mathbf{r}_1, \mathbf{r}_2), \tag{8.36}$$

we get two groups of similar terms that differ only by the particle index. We can merge the terms of each pair by changing the notation as ($\mathbf{r}_1 \to \mathbf{r}$, $\mathbf{r}_2 \to \mathbf{r}'$) in one of them, and ($\mathbf{r}_1 \to \mathbf{r}'$, $\mathbf{r}_2 \to \mathbf{r}$) in the counterpart term. Using Eq. (25), and the mutual orthogonality of the wavefunctions $\psi_{100}(\mathbf{r})$ and $\psi_{nlm}(\mathbf{r})$, we get the following result:

$$\langle E_e \rangle^{(0)} = \int \psi_{100}^*(\mathbf{r}) \left(-\frac{\hbar^2 \nabla_{\mathbf{r}}^2}{2m} - \frac{2e^2}{4\pi\varepsilon_0 r} \right) \psi_{100}(\mathbf{r}) d^3 r + \int \psi_{nlm}^*(\mathbf{r}') \left(-\frac{\hbar^2 \nabla_{\mathbf{r}'}^2}{2m} - \frac{2e^2}{4\pi\varepsilon_0 r'} \right) \psi_{nlm}(\mathbf{r}') d^3 r' \quad (8.37)$$
$$\equiv \varepsilon_{100} + \varepsilon_{nlm}, \qquad \text{with } n > 1.$$

It may be interpreted as the sum of eigenenergies of two separate single particles, one in the ground state 100, and another in the excited state *nlm* – although actually the electron states are entangled. Thus, in the 0th order of the perturbation theory, the electrons' entanglement does not affect their total energy.

However, the potential energy of the system also includes the interaction term U_{int}, which does not allow such separation. Indeed, in the 1st approximation of the perturbation theory, the total energy E_e of the system may be expressed as $\varepsilon_{100} + \varepsilon_{nlm} + E_{\text{int}}^{(1)}$, with

$$E_{\text{int}}^{(1)} = \langle U_{\text{int}} \rangle = \int d^3 r_1 \int d^3 r_2 \, \psi_e^*(\mathbf{r}_1, \mathbf{r}_2) U_{\text{int}}(\mathbf{r}_1, \mathbf{r}_2) \psi_e(\mathbf{r}_1, \mathbf{r}_2), \quad (8.38)$$

Plugging Eq. (35) into this result, using the symmetry of the function U_{int} with respect to the particle number permutation, and the same particle coordinate re-numbering as above, we get

$$E_{\text{int}}^{(1)} = E_{\text{dir}} \pm E_{\text{ex}}, \quad (8.39)$$

with the following, deceivingly similar expressions for the two components of this sum/difference:

Direct interaction energy

$$E_{\text{dir}} \equiv \int d^3 r \int d^3 r' \psi_{100}^*(\mathbf{r}) \psi_{nlm}^*(\mathbf{r}') U_{\text{int}}(\mathbf{r}, \mathbf{r}') \psi_{100}(\mathbf{r}) \psi_{nlm}(\mathbf{r}'), \quad (8.40)$$

Exchange interaction energy

$$E_{\text{ex}} \equiv \int d^3 r \int d^3 r' \psi_{100}^*(\mathbf{r}) \psi_{nlm}^*(\mathbf{r}') U_{\text{int}}(\mathbf{r}, \mathbf{r}') \psi_{nlm}(\mathbf{r}) \psi_{100}(\mathbf{r}'). \quad (8.41)$$

Since the single-particle orbitals can be always made real, both components are positive – or at least non-negative. However, their physics and magnitude are different. The integral (40), called the *direct interaction energy*, allows a simple semi-classical interpretation as the Coulomb energy of interacting electrons, each distributed in space with the electric charge density $\rho(\mathbf{r}) = -e\psi^*(\mathbf{r})\psi(\mathbf{r})$:[14]

$$E_{\text{dir}} = \int d^3 r \int d^3 r' \frac{\rho_{100}(\mathbf{r}) \rho_{nlm}(\mathbf{r}')}{4\pi\varepsilon_0 |\mathbf{r} - \mathbf{r}'|} \equiv \int \rho_{100}(\mathbf{r}) \phi_{nlm}(\mathbf{r}) d^3 r \equiv \int \rho_{nlm}(\mathbf{r}) \phi_{100}(\mathbf{r}) d^3 r, \quad (8.42)$$

where $\phi(\mathbf{r})$ are the electrostatic potentials created by the electron "charge clouds":[15]

$$\phi_{100}(\mathbf{r}) = \frac{1}{4\pi\varepsilon_0} \int d^3 r' \frac{\rho_{100}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}, \qquad \phi_{nlm}(\mathbf{r}) = \frac{1}{4\pi\varepsilon_0} \int d^3 r' \frac{\rho_{nlm}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}. \quad (8.43)$$

However, the integral (41), called the *exchange interaction energy*, evades a classical interpretation, and (as it is clear from its derivation) is the direct corollary of electrons' indistinguishability. The magnitude of E_{ex} is also very much different from E_{dir} because the function under the integral (41) disappears in the regions where the single-particle wavefunctions $\psi_{100}(\mathbf{r})$ and $\psi_{nlm}(\mathbf{r})$ do not overlap. This is in full agreement with the discussion in Sec. 1: if two particles are identical but well separated, i.e. their wavefunctions do not overlap, the exchange interaction disappears,

[14] See, e.g., EM Sec. 1.3, in particular Eq. (1.54).

[15] Note that the result for E_{dir} correctly reflects the basic fact that a charged particle does not interact with itself, even if its wavefunction is quantum-mechanically spread over a finite space volume. Unfortunately, this is not true for some popular approximate theories of multiparticle systems – see Sec. 4 below.

i.e. measurable effects of particle indistinguishability vanish. (In contrast, the integral (40) decreases with the growing separation of the electrons only slowly, due to their long-range Coulomb interaction.)

Figure 1b shows the structure of an excited energy level, with certain quantum numbers $n > 1$, l, and m, given by Eqs. (39)-(41). The upper, so-called *parahelium*[16] level, with the energy

$$E_{\text{para}} = (\varepsilon_{100} + \varepsilon_{nlm}) + E_{\text{dir}} + E_{\text{ex}} > \varepsilon_{100} + \varepsilon_{nlm}, \tag{8.44}$$

corresponds to the symmetric orbital state and hence to the singlet-spin state (18), while the lower, *orthohelium* level, with

$$E_{\text{orth}} = (\varepsilon_{100} + \varepsilon_{nlm}) + E_{\text{dir}} - E_{\text{ex}} < E_{\text{para}}, \tag{8.45}$$

corresponds to the degenerate triplet-spin state (21).

This degeneracy may be lifted by an external magnetic field, whose effect on the electron spins[17] is described by the following evident generalization of the Pauli Hamiltonian (4.163),

$$\hat{H}_{\text{field}} = -\gamma \hat{\mathbf{s}}_1 \cdot \mathscr{B} - \gamma \hat{\mathbf{s}}_2 \cdot \mathscr{B} \equiv -\gamma \hat{\mathbf{S}} \cdot \mathscr{B}, \qquad \text{with } \gamma = \gamma_{\text{e}} \equiv -\frac{e}{m_{\text{e}}} \equiv -2\frac{\mu_{\text{B}}}{\hbar}, \tag{8.46}$$

where

$$\hat{\mathbf{S}} \equiv \hat{\mathbf{s}}_1 + \hat{\mathbf{s}}_2, \tag{8.47}$$

is the operator of the (vector) sum of the system of two spins.[18] To analyze this effect, we need first to make one more detour, to address the general issue of *spin addition*. The main rule[19] here is that in a full analogy with the net spin of a single particle, defined by Eq. (5.170), the net spin operator (47) of *any* system of two spins, and its component $\hat{S}_z$ along the (arbitrarily selected) z-axis, obey the same commutation relations (5.168) as the component operators, and hence have the properties similar to those expressed by Eqs. (5.169) and (5.175):

$$\hat{S}^2 |S, M_S\rangle = \hbar^2 S(S+1) |S, M_S\rangle, \qquad \hat{S}_z |S, M_S\rangle = \hbar M_S |S, M_S\rangle, \qquad \text{with } -S \le M_S \le +S, \tag{8.48}$$

where the ket vectors correspond to the coupled basis of joint eigenstates of the operators of S^2 and S_z (but not necessarily all component operators – see again the Venn shown in Fig. 5.12 and its discussion, with the replacements $\mathbf{S}$, $\mathbf{L} \to \mathbf{s}_{1,2}$ and $\mathbf{J} \to \mathbf{S}$). Repeating the discussion of Sec. 5.7 with these replacements, we see that in both the coupled and the uncoupled bases, the net magnetic number M_S is simply expressed via those of the components

---

[16] This terminology reflects the historic fact that the observation of two different hydrogen-like spectra, corresponding to the opposite signs in Eq. (39), was first taken as evidence for two different species of ^{4}He, which were called, respectively, the "orthohelium" and the "parahelium".

[17] As we know from Sec. 6.4, the field also affects the orbital motion of the electrons, so the simple analysis based on Eq. (46) is strictly valid only for the s excited state ($l = 0$, and hence $m = 0$). However, the orbital effects of a weak magnetic field do not affect the triplet-level splitting we are analyzing now.

[18] Note that similarly to Eqs. (22) and (25), here the uppercase notation of the component spins is replaced with the lowercase notation, to avoid any possibility of confusion with the total spin of the system.

[19] Since we already know that the spin of a particle is physically nothing more than some (if specific) part of its angular momentum, the similarity of the properties (48) of the sum (47) of spins of different particles to those of the sum (5.170) of different spin components of the same particle it very natural, but still has to be considered as a new fact – confirmed by a vast body of experimental data.

$$M_S = (m_s)_1 + (m_s)_2 . \tag{8.49}$$

However, the net spin quantum number S (in contrast to the Nature-given spins $s_{1,2}$ of its elementary components) is not universally definite, and we may immediately say only that it has to obey the following analog of the relation $| l - s | \le j \le (l + s)$ discussed in Sec. 5.7:

$$|s_1 - s_2| \le S \le s_1 + s_2 . \tag{8.50}$$

What exactly S is (within these limits), depends on the spin state of the system.

For the simplest case of two spin-½ components, each with $s = ½$ and $m_s = \pm ½$, Eq. (49) gives three possible values of M_S, equal to 0 and ±1, while Eq. (50) limits the possible values of S to just either 0 or 1. Using the last of Eqs. (48), we see that the possible combinations of the quantum numbers are

$$\begin{cases} S = 0, \\ M_S = 0, \end{cases} \quad \text{and} \quad \begin{cases} S = 1, \\ M_S = 0, \pm 1. \end{cases} \tag{8.51}$$

It is virtually evident that the singlet spin state s_- belongs to the first class, while the simple (separable) triplet states ↑↑ and ↓↓ belong to the second class, with $M_S = +1$ and $M_S = -1$, respectively. However, for the entangled triplet state s_+, evidently with $M_S = 0$, the value of S is less obvious. Perhaps the easiest way to recover it[20] to use the "rectangular diagram", similar to that shown in Fig. 5.14, but redrawn for our case of two spins, i.e., with the replacements $m_l \to (m_s)_1 = \pm ½$, $m_s \to (m_s)_2 = \pm ½$ – see Fig. 2.

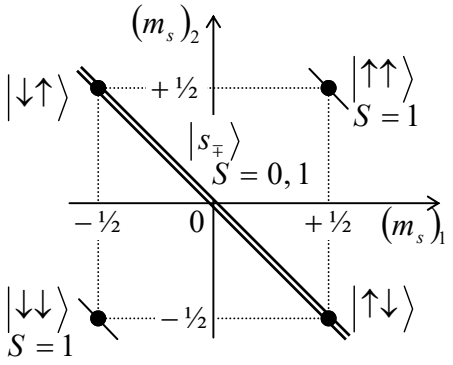


Fig. 8.2. The "rectangular diagram" showing the relation between the uncoupled-representation states (dots) and the coupled-representation states (straight lines) of a system of two spins-½ – cf. Fig. 5.14.

Just as at the addition of various angular momenta of a single particle, the top-right and bottom-left corners of this diagram correspond to the factorable triplet states ↑↑ and ↓↓, which participate in both the uncoupled-representation and coupled-representation bases, and have the largest value of S, i.e. 1. However, the entangled states $s_\pm$, which are linear combinations of the uncoupled-representation states ↑↓ and ↓↑, cannot have the same value of S, so for the triplet state s_+, S has to take the value different from that (0) of the singlet state, i.e. 1. With that, the first of Eqs. (48) gives the following expectation values for the square of the net spin operator:

$$\left\langle S^2 \right\rangle = \begin{cases} 2\hbar^2, & \text{for each triplet state,} \\ 0, & \text{for the singlet state.} \end{cases} \tag{8.52}$$

[20] Another, a bit longer but perhaps more prudent way is to directly calculate the expectation values of $\hat{S}^2$ for the states $s_\pm$, and then find S by comparing the results with the first of Eqs. (48); it is highly recommended to the reader as a useful exercise.

Note that for the entangled triplet state s_+, whose ket-vector (20) is a linear superposition of two kets of states with *opposite* spins, this result is highly counter-intuitive, and shows how careful we should be interpreting entangled quantum states. (As will be discussed in Chapter 10, quantum entanglement brings even more surprises for measurements.)

Now we may return to the particular issue of the magnetic field effect on the triplet state of the ^{4}He atom. Directing the z-axis along the field, we may reduce Eq. (46) to

$$\hat{H}_{\text{field}} = -\gamma_e \hat{S}_z \mathscr{B} \equiv 2\mu_B \mathscr{B} \frac{\hat{S}_z}{\hbar}. \tag{8.53}$$

Since all three triplet states (21) are eigenstates, in particular, of the operator $\hat{S}_z$, and hence of the Hamiltonian (53), we may use the second of Eqs. (48) to calculate their energy change simply as

$$\Delta E_{\text{field}} = 2\mu_B \mathscr{B} M_S = 2\mu_B \mathscr{B} \times \begin{cases} +1, & \text{for the factorable triplet state } \uparrow\uparrow, \\ 0, & \text{for the entangled triplet state } s_+, \\ -1, & \text{for the factorable triplet state } \downarrow\downarrow. \end{cases} \tag{8.54}$$

This splitting of the "orthohelium" level is schematically shown in Fig. 1b.[21]

8.3. Multiparticle systems

Leaving several other problems on two-particle systems for the reader's exercise, let me proceed to the discussion of systems with $N > 2$ indistinguishable particles, whose list notably includes atoms, molecules, and condensed-matter systems. In this case, Eq. (7) for fermions is generalized as

$$\hat{\mathscr{P}}_{kk'} \left| \alpha_- \right\rangle = -\left| \alpha_- \right\rangle, \quad \text{for all } k, k' = 1, 2, ..., N, \tag{8.55}$$

where the operator $\hat{\mathscr{P}}_{kk'}$ permutes particles with numbers k and k'. As a result, for systems with non-directly-interacting fermions, the Pauli principle forbids any state in which *any* two particles have similar single-particle wavefunctions. Nevertheless, it permits two fermions to have similar *orbital* wavefunctions, provided that their spins are in the singlet state (18), because this satisfies the permutation requirement (55). This fact is of paramount importance for the ground state of the systems whose Hamiltonians do not depend on spin because it allows the fermions to be in their orbital single-particle ground states, with two electrons of the spin singlet sharing the same orbital state. Hence, for the limited (but very important!) goal of finding ground-state energies of multi-fermion systems with negligible direct interaction, we may ignore the actual singlet spin structure, and reduce the Pauli

[21] It is interesting that another very important two-electron system, the hydrogen (H_2) molecule, which was briefly discussed in Sec. 2.6, also has two similarly named forms, *parahydrogen* and *orthohydrogen*. However, their difference is due to two possible (respectively, singlet and triplet) states of the system of two spins of the two hydrogen nuclei – protons, which are also spin-½ particles. The resulting ground-state energy of the parahydrogen is lower than that of the orthohydrogen by only ~15 meV per molecule – the difference lower than $k_B T$ at room temperature (~26 meV). As a result, at very low temperatures, hydrogen at equilibrium is dominated by parahydrogen, but at ambient conditions, the orthohydrogen is nearly three times more abundant, due to its triple nuclear spin degeneracy. Curiously, the theoretical prediction of this effect by W. Heisenberg (together with F. Hund) in 1927 was cited in his 1932 Nobel Prize award as the most noteworthy application of quantum theory.

exclusion principle to the rudimentary picture of single-particle orbital energy levels, each "occupied with two fermions".

As a very simple example, let us find the ground energy of five fermions confined in a hard-wall, cubic-shaped 3D volume of side a, ignoring their direct interaction. From Sec. 1.7, we know the single-particle energy spectrum of the system:

$$\varepsilon_{n_x,n_y,n_z} = \varepsilon_0\left(n_x^2 + n_y^2 + n_z^2\right), \qquad \text{with } \varepsilon_0 \equiv \frac{\pi^2\hbar^2}{2ma^2}, \quad \text{and } n_x, n_y, n_z = 1, 2, \ldots, \tag{8.56}$$

so the lowest-energy states are:

– one ground state with $\{n_x,n_y,n_z\} = \{1,1,1\}$, and energy $\varepsilon_{111} = (1^2+1^2+1^2)\varepsilon_0 = 3\varepsilon_0$, and

– three excited states, with $\{n_x,n_y,n_z\}$ equal to either $\{2,1,1\}$, or $\{1,2,1\}$, or $\{1,1,2\}$, with equal energies $\varepsilon_{211} = \varepsilon_{121} = \varepsilon_{112} = (2^2+1^2+1^2)\varepsilon_0 = 6\varepsilon_0$.

According to the above simple formulation of the Pauli principle, each of these orbital energy levels can accommodate up to two fermions. Hence the lowest-energy (ground) state of the five-fermion system is achieved by placing two of them on the ground level $\varepsilon_{111} = 3\varepsilon_0$, and the remaining three particles, in any of the degenerate "excited" states of energy $6\varepsilon_0$, so the ground-state energy of the system is

$$E_g = 2\times 3\varepsilon_0 + 3\times 6\varepsilon_0 \equiv 24\varepsilon_0 \equiv \frac{12\pi^2\hbar^2}{ma^2}. \tag{8.57}$$

Moreover, in many cases, relatively weak interaction between fermions does not blow up such a simple quantum state classification scheme qualitatively, and the Pauli principle allows tracing the order of single-particle state filling. This is exactly the simple approach that was used in our discussion of atoms in Sec. 3.7. Unfortunately, it does not allow for a more specific characterization of the ground states of most atoms, in particular the evaluation of the corresponding values of the quantum numbers S, L, and J that characterize the net angular momenta of the atom, and hence its response to an external magnetic field. These numbers are defined by relations similar to Eqs. (48), each for the corresponding vector operator of the net angular momenta:

$$\hat{\mathbf{S}} \equiv \sum_{k=1}^{N} \hat{\mathbf{s}}_k, \qquad \hat{\mathbf{L}} \equiv \sum_{k=1}^{N} \hat{\mathbf{l}}_k, \qquad \hat{\mathbf{J}} \equiv \sum_{k=1}^{N} \hat{\mathbf{j}}_k\ ; \tag{8.58}$$

note that these definitions are consistent with Eq. (5.170) applied both to the angular momenta $\mathbf{s}_k$, $\mathbf{l}_k$, and $\mathbf{j}_k$ of each particle, and to the full vectors $\mathbf{S}$, $\mathbf{L}$, and $\mathbf{J}$. When the numbers S, L, and J for a state are known, they are traditionally recorded in the form of the so-called *Russell-Saunders symbols*:[22]

$$^{2S+1}\mathscr{L}_J, \tag{8.59}$$

where S and J are the corresponding *values* of these quantum numbers, while $\mathscr{L}$ is a capital *letter*, encoding the quantum number L – via the same spectroscopic notation as for single particles (see Sec. 3.6): $\mathscr{L} = S$ for $L = 0$, $\mathscr{L} = P$ for $L = 1$, $\mathscr{L} = D$ for $L = 2$, etc. (The reason why the front superscript of the Russel-Saunders symbol lists $2S + 1$ rather than just S, is that according to the last of Eqs. (48), it

[22] Named after Henry Russell and Frederick Saunders, whose pioneering (circa 1925) processing of experimental spectral-line data has established the very idea of the vector addition of the electron spins, described by the first of Eqs. (58).

shows the number of possible values of the quantum number M_S, which characterizes the state's spin degeneracy, and is called its *multiplicity*.)

For example, for the simplest, hydrogen atom ($Z = 1$), with its single electron in the ground $1s$ state, $L = l = 0$, $S = s = ½$, and $J = S = ½$, so its Russell-Saunders symbol is ${}^2S_{1/2}$. Next, the discussion of the helium atom ($Z = 2$) in the previous section has shown that in its ground state $L = 0$ (because of the $1s$ orbital state of both electrons), and $S = 0$ (because of the singlet spin state), so the total angular momentum also vanishes: $J = 0$. As a result, the Russell-Saunders symbol for this state is ${}^1S_0$. The structure of the next atom, lithium ($Z = 3$) is also easy to predict, because, as was discussed in Sec. 3.7, its ground-state electron configuration is $1s^2 2s^1$, i.e. includes two electrons in the "helium shell", i.e. on the $1s$ orbitals (now we know that they are actually in an entangled singlet spin state), and one electron in the $2s$ state, of higher energy, also with zero orbital momentum, $l = 0$. As a result, the total L in this state is evidently equal to 0, and S is equal to ½, so $J = ½$, meaning that the Russell-Saunders symbol of the lithium's ground state is ${}^2P_{1/2}$. Even in the next atom, beryllium ($Z = 4$), with the ground-state configuration $1s^2 2s^2$, the symbol is readily predictable, because none of its electrons has non-zero orbital momentum, giving $L = 0$. Also, each electron pair is in the singlet spin state, i.e. we have $S = 0$, so $J = 0$ – the quantum number set described by the Russell-Saunders symbol ${}^1S_0$ – just as for helium.

However, for the next, boron atom ($Z = 5$), with its ground-state electron configuration $1s^2 2s^2 2p^1$ (see, e.g., Fig. 3.24), there is no obvious way to predict the result. Indeed, this atom has two pairs of electrons, with opposite spins, on its two lowest s-orbitals, giving zero contributions to the net S, L, and J. Hence these total quantum numbers may be only contributed by the last, fifth electron with $s = ½$ and $l = 1$, giving $S = ½$, $L = 1$. As was discussed in Sec. 5.7 for the single-particle case, the vector addition of the angular momenta **S** and **L** enables two values of the quantum number J: either $L + S = {}^3/_2$ or $L - S = ½$. Experiment shows that the difference between the energies of these two states of boron is very small (~2 meV), so at room temperature (with $k_B T \approx 26$ meV) they are both partly occupied, with the genuine ground state having $J = ½$, so its Russell-Saunders symbol is ${}^2P_{1/2}$.

Such energy differences, which become larger for heavier atoms, are determined both by the Coulomb and spin-orbit[23] interactions between the electrons. Their quantitative analysis is rather involved (see below), but the results tend to follow simple phenomenological *Hund rules*, with the following hierarchy:

Rule 1. For a given electron configuration, the ground state has the *largest* possible S, and hence the largest possible multiplicity $2S + 1$.

Rule 2. For a given S, the ground state has the *largest* possible L.

Rule 3. For given S and L, J has its *smallest* possible value, $| L - S |$, if the given sub-shell $\{n, l\}$ is filled not more than by half, while in the opposite case, J has its *largest* possible value, $L + S$.

Let us see how these rules work for the boron atom we have just discussed. For it, the Hund Rules 1 and 2 are satisfied automatically, while the sub-shell $\{n = 2, l = 1\}$, which can house up to $2\times(2l + 1) = 6$ electrons, is filled with just one $2p$ electron, i.e. by less than a half of the maximum value. As a result, Rule 3 predicts the ground state's value $J = ½$, in agreement with experiment. Generally, for

[23] In light atoms, the spin-orbit interaction is so weak that it may be reasonably well described as an interaction of the total momenta **L** and **S** of the system – the so-called *LS* (or "Russell-Saunders") *coupling*. On the other hand, in very heavy atoms, the interaction is effectively between the net momenta $\mathbf{j}_k = \mathbf{l}_k + \mathbf{s}_k$ of the individual electrons – the so-called *jj coupling*. This is the reason why in such atoms the Hund Rule 3 may be violated.

lighter atoms, the Hund rules are well obeyed. However, the lower down the Hund rule hierarchy, the less "powerful" the rules are, i.e. the more often they are violated in heavier atoms.

Now let us discuss possible approaches to a quantitative theory of multiparticle systems – not only atoms. As was discussed in Sec. 1, if fermions do not interact directly, the stationary states of the system have to be the antisymmetric eigenstates of the permutation operator, i.e. to satisfy Eq. (55). To understand how such states may be formed from the single-electron ones, let us return for a minute to the case of two electrons, and rewrite Eq. (11) in the following compact form:

$$\left|\alpha_{-}\right\rangle \equiv \frac{1}{\sqrt{2}}\left(\left|\beta\right\rangle\otimes\left|\beta'\right\rangle-\left|\beta'\right\rangle\otimes\left|\beta\right\rangle\right)=\frac{1}{\sqrt{2}}\begin{vmatrix}\left|\beta\right\rangle & \left|\beta'\right\rangle\\ \left|\beta\right\rangle & \left|\beta'\right\rangle\end{vmatrix}\begin{matrix}\text{state 1}\ \text{state 2}\\ \downarrow\quad\downarrow\\ \\ \leftarrow\text{particle number 1},\\ \leftarrow\text{particle number 2},\end{matrix} \tag{8.60a}$$

where, in the last form, the direct product signs are just implied. In this way, the Pauli principle is mapped on the well-known property of matrix determinants: if any two columns of a matrix coincide, its determinant vanishes. This *Slater determinant* approach[24] may be readily generalized to N fermions occupying any N (not necessarily the lowest-energy) single-particle states β, β', β'', etc:

Slater determinant

$$\left|\alpha_{-}\right\rangle=\frac{1}{(N!)^{1/2}}\underbrace{\begin{vmatrix}\left|\beta\right\rangle\left|\beta'\right\rangle\left|\beta''\right\rangle\ldots\\ \left|\beta\right\rangle\left|\beta'\right\rangle\left|\beta''\right\rangle\ldots\\ \left|\beta\right\rangle\left|\beta'\right\rangle\left|\beta''\right\rangle\ldots\\ \ldots\ \ldots\ \ldots\ \ldots\end{vmatrix}}_{N}\Bigg\}N\quad\begin{matrix}\text{state list}\rightarrow\\ \\ \text{particle}\\ \text{list}\\ \downarrow\end{matrix} \tag{8.60b}$$

The Slater determinant form is extremely nice and compact – in comparison with direct writing of a sum of $N!$ products, each of N ket factors. However, there are two major problems with using it for practical calculations:

(i) For the calculation of any bra-ket product (say, within the perturbation theory) we still need to spell out each bra- and ket-vector as a sum of component terms. Even for a limited number of electrons (say $N \sim 10^2$ in a typical atom), the number $N! \sim 10^{160}$ of terms in such a sum is impracticably large for any analytical or numerical calculation.

(ii) In the case of interacting fermions, the Slater determinant does not describe the eigenvectors of the system; rather the stationary state is a *superposition* of such basis functions, i.e. of the Slater determinants – each for a specific selection of N states from the full set of single-particle states – that is generally larger than N.

For atoms and simple molecules, whose filled-shell electrons may be excluded from an explicit analysis (by describing their effects, approximately, with effective *pseudo-potentials*), the effective number N may be reduced to a smaller number N_{ef} of the order of 10, so $N_{ef}! < 10^6$, and the Slater determinants may be used for numerical calculations – for example, in the Hartree-Fock theory – see the

[24] It was suggested in 1929 by John C. Slater.

next section. However, for condensed-matter systems, such as metals and semiconductors, with the number of free electrons is of the order of 10^{23} per cm^3, this approach is generally unacceptable, though with some smart tricks (such as using the crystal's periodicity) it may be still used for some approximate (also mostly numerical) calculations.

These challenges make the development of a more general theory that would not use particle numbers (which are superficial for indistinguishable particles to start with) a must for getting any final analytical results for multiparticle systems. The most effective formalism for this purpose, which avoids particle numbering at all, is called the *second quantization*.[25] Actually, we have already discussed a particular version of this formalism, for the case of the 1D harmonic oscillator, in Sec. 5.4. As a reminder, after the definition (5.65) of the "creation" and "annihilation" operators via those of the particle's coordinate and momentum, we have derived their key properties (5.89),

$$\hat{a}|n\rangle = n^{1/2}|n-1\rangle, \qquad \hat{a}^\dagger|n\rangle = (n+1)^{1/2}|n+1\rangle, \tag{8.61}$$

where n are the stationary (Fock) states of the oscillator. This property allows an interpretation of the operators' actions as the creation/annihilation of a single *excitation* with the energy $\hbar\omega_0$ – thus justifying the operator names. In the next chapter, we will show that such excitation of an electromagnetic field mode may be interpreted as a massless *boson* with $s = 1$, called the *photon*.

In order to generalize this approach to *arbitrary bosons*, not appealing to a specific system, we may use relations similar to Eq. (61) to *define* the creation and annihilation operators. The definitions look simple in the language of the so-called *Dirac states*, described by ket-vectors

$$\boxed{|N_1, N_2, \ldots N_j, \ldots\rangle}, \tag{8.62}$$

Dirac state

where N_j is the state *occupancy*, i.e. the number of bosons in the single-particle state j. Let me emphasize that here the indices 1, 2, …j,… number single-particle *states* (including their spin parts) rather than *particles*. Thus the very notion of an individual particle's number is completely (and for indistinguishable particles, very relevantly) absent from this formalism. Generally, the set of single-particle states participating in the Dirac state may be selected arbitrarily, provided that it is full and orthonormal in the sense

$$\langle N_1', N_2', \ldots, N_{j'}', \ldots | N_1, N_2 \ldots, N_j, \ldots\rangle = \delta_{N_1 N_1'} \delta_{N_2 N_2'} \ldots \delta_{N_j N_{j'}} \ldots, \tag{8.63}$$

though for systems of non- (or weakly) interacting bosons, using the stationary states of individual particles in the system under analysis is almost always the best choice.

Now we can define the *particle annihilation operator* as follows:

$$\boxed{\hat{a}_j |N_1, N_2, \ldots N_j, \ldots\rangle \equiv N_j^{1/2} |N_1, N_2, \ldots N_j - 1, \ldots\rangle}. \tag{8.64}$$

Boson annihilation operator

Note that the pre-ket coefficient, similar to that in the first of Eqs. (61), guarantees that any attempt to annihilate a particle in an initially unpopulated state gives the non-existing ("null") state:

[25] It was invented (first for photons and then for arbitrary bosons) by P. Dirac in 1927, and then (in 1928) adjusted for fermions by E. Wigner and P. Jordan. Note that the term "second quantization" is rather misleading for the non-relativistic case discussed here, but finds certain justification in the quantum field theory.

$$\hat{a}_j \left| N_1, N_2, 0_j, \ldots \right\rangle = 0, \tag{8.65}$$

where the symbol 0_j means zero occupancy of the j^{th} state. According to Eq. (63), an equivalent way to write Eq. (64) is

$$\left\langle N_1', N_2', \ldots, N_j', \ldots \right| \hat{a}_j \left| .N_1, N_2, \ldots, N_j, \ldots \right\rangle = N_j^{1/2} \delta_{N_1 N_1'} \delta_{N_2 N_2'} \ldots \delta_{N_j', N_j - 1} \ldots \tag{8.66}$$

According to the general Eq. (4.65), the matrix element of the Hermitian-conjugate operator $\hat{a}_j^\dagger$ is

$$\begin{aligned} \left\langle N_1', N_2', \ldots, N_j', \ldots \right| \hat{a}_j^\dagger \left| N_1, N_2, \ldots, N_j, \ldots \right\rangle &= \left\langle N_1, N_2, \ldots, N_j, \ldots \right| \hat{a}_j \left| N_1', N_2', \ldots, N_j', \ldots \right\rangle^* \\ = \left\langle N_1, N_2, \ldots, N_j, \ldots \right| \left(N_j' \right)^{1/2} \left| N_1', N_2', \ldots, N_j' - 1, \ldots \right\rangle &= \left(N_j' \right)^{1/2} \delta_{N_1 N_1'} \delta_{N_2 N_2'} \ldots \delta_{N_j, N_j' - 1} \ldots \\ = \left(N_j + 1 \right)^{1/2} \delta_{N_1 N_1'} \delta_{N_2 N_2'} \ldots \delta_{N_j + 1, N_j'} \ldots, & \end{aligned} \tag{8.67}$$

meaning that

Boson creation operator

$$\boxed{\hat{a}_j^\dagger \left| N_1, N_2, \ldots, N_j, \ldots \right\rangle = \left(N_j + 1 \right)^{1/2} \left| N_1, N_2, \ldots, N_j + 1, \ldots \right\rangle,} \tag{8.68}$$

in total compliance with the second of Eqs. (61). In particular, this *particle creation operator* allows a description of the generation of a single particle from the *vacuum* (not null!) *state* $|0, 0, \ldots\rangle$:

$$\hat{a}_j^\dagger \left| 0, 0, \ldots, 0_j, \ldots, 0 \right\rangle = \left| 0, 0, \ldots, 1_j, \ldots 0 \right\rangle, \tag{8.69}$$

and hence a product of such operators may create, from vacuum, a multiparticle state with an arbitrary set of occupancies: [26]

$$\underbrace{\hat{a}_1^\dagger \hat{a}_1^\dagger \ldots \hat{a}_1^\dagger}_{N_1 \text{ times}} \underbrace{\hat{a}_2^\dagger \hat{a}_2^\dagger \ldots \hat{a}_2^\dagger}_{N_2 \text{ times}} \ldots \left| 0, 0, \ldots \right\rangle = \left(N_1! N_2! \ldots \right)^{1/2} \left| N_1, N_2, \ldots \right\rangle. \tag{8.70}$$

Next, combining Eqs. (64) and (68), we get

$$\hat{a}_j^\dagger \hat{a}_j \left| N_1, N_2, \ldots N_j, \ldots \right\rangle = N_j \left| N_1, N_2, \ldots, N_j, \ldots \right\rangle, \tag{8.71}$$

so, just as for the particular case of the harmonic-oscillator excitations, the operator

Number-counting operator

$$\boxed{\hat{N}_j \equiv \hat{a}_j^\dagger \hat{a}_j} \tag{8.72}$$

"counts" the number of particles in the j^{th} single-particle state, while preserving the whole multiparticle state. Acting on a state by the creation-annihilation operators in the reverse order, we get

$$\hat{a}_j \hat{a}_j^\dagger \left| N_1, N_2, \ldots, N_j, \ldots \right\rangle = \left(N_j + 1 \right) \left| N_1, N_2, \ldots, N_j, \ldots \right\rangle. \tag{8.73}$$

Eqs. (71) and (73) show that for *any* state of a multiparticle system (which may be represented as a linear superposition of Dirac states with all possible sets of numbers N_j), we may write

[26] The resulting Dirac state is *not* an eigenstate of every multiparticle Hamiltonian. However, we will see below that for a set of non-interacting particles it *is* a stationary state, so the full set of such states may be used as a good basis in perturbation theories of systems of weakly interacting particles.

$$\hat{a}_j \hat{a}_j^\dagger - \hat{a}_j^\dagger \hat{a}_j \equiv \left[\hat{a}_j, \hat{a}_j^\dagger\right] = \hat{I}, \tag{8.74}$$

again in agreement with what we had for the 1D oscillator – cf. Eq. (5.68). According to Eqs. (63), (64), and (68), the creation and annihilation operators corresponding to different single-particle states do commute, so Eq. (74) may be generalized as

$$\left[\hat{a}_j, \hat{a}_{j'}^\dagger\right] = \hat{I}\delta_{jj'}, \tag{8.75}$$

Bosonic operators: commutation relations

while similar operators commute, regardless of which states they act upon:

$$\left[\hat{a}_j^\dagger, \hat{a}_{j'}^\dagger\right] = \left[\hat{a}_j, \hat{a}_{j'}\right] = \hat{0}. \tag{8.76}$$

As was mentioned earlier, a major challenge in the Dirac approach is to rewrite the Hamiltonian of a multiparticle system, that naturally carries particle numbers k (see, e.g., Eq. (22) for $k = 1, 2$), in the second quantization language, in which there are no these numbers. Let us start with *single-particle* components of such Hamiltonians, i.e. operators of the type

$$\hat{F} = \sum_{k=1}^{N} \hat{f}_k . \tag{8.77}$$

Single-particle operator

where all N operators $\hat{f}_k$ are similar, besides that each of them acts on one specific (k^{th}) particle, and N is the total number of particles in the system, which is evidently equal to the sum of single-particle state occupancies:

$$N = \sum_j N_j . \tag{8.78}$$

The most important examples of such operators are the kinetic energy of N similar single particles and their potential energy in an external field:

$$\hat{T} = \sum_{k=1}^{N} \frac{\hat{p}_k^2}{2m}, \qquad \hat{U} = \sum_{k=1}^{N} \hat{u}(\mathbf{r}_k). \tag{8.79}$$

For bosons, instead of the Slater determinant (60), we have to write a similar expression, but without the sign alternation at permutations:

$$\left|N_1, \ldots, N_j, \ldots\right\rangle = \left(\frac{N_1! \ldots N_j! \ldots}{N!}\right)^{1/2} \sum_P \left| \underbrace{\ldots \beta\beta'\beta'' \ldots}_{N \text{ operands}} \right\rangle, \tag{8.80}$$

sometimes called the *permanent*. Note again that the left-hand side of this relation is written in the Dirac notation (that does not use particle numbering), while on its right-hand side, just in formulas of Secs. 1 and 2, the particle numbers are coded with the positions of the single-particle states inside the state vectors, and the summation is over all different permutations of the states in the ket – cf. Eq. (10). (According to the basic combinatorics,[27] there are $N!/(N_1! \ldots N_j! \ldots)$ such permutations, so the front coefficient in Eq. (80) ensures the normalization of the Dirac state, provided that the single-particle

[27] See, e.g., MA Eq. (2.3).

states β, β',… are normalized.) Let us use Eq. (80) to spell out the following matrix element for a system with $(N-1)$ particles:

$$\begin{aligned}&\left\langle \dots N_j,\dots,N_{j'}-1,\dots\right|\hat{F}\left|\dots N_j-1,\dots,N_{j'},\dots\right\rangle \\ &=\frac{N_1!\dots(N_j-1)!\dots(N_{j'}-1)!\dots}{(N-1)!}\left(N_jN_{j'}\right)^{1/2}\sum_{P\left\langle N-1\right|}\sum_{P\left|N-1\right\rangle}\left\langle \dots\beta\beta'\beta''\dots\right|\sum_{k=1}^{N-1}\hat{f}_k\left|\dots\beta\beta'\beta''\dots\right\rangle,\end{aligned} \tag{8.81}$$

where all non-specified occupation numbers in the corresponding positions of the bra- and ket-vectors are equal to each other. Each single-particle operator $\hat{f}_k$ participating in the operator sum acts on the bra- and ket-vectors of states only in one (k^{th}) position, giving the following result, independent of the position number:

$$\left\langle \beta_j\right|_{\text{in }k^{\text{th}}\text{ position}}\hat{f}_k\left|\beta_{j'}\right\rangle_{\text{in }k^{\text{th}}\text{ position}}=\left\langle \beta_j\right|\hat{f}\left|\beta_{j'}\right\rangle\equiv f_{jj'}. \tag{8.82}$$

Since in both permutation sets participating in Eq. (81), with $(N-1)$ state vectors each, all positions are equivalent, we can fix the position (say, take the first one) and replace the sum over k with the multiplication by of the bracket by $(N-1)$. The fraction of permutations with the necessary bra-vector (with number j) in that position is $N_j/(N-1)$, while that with the necessary ket-vector (with number j') in the same position is $N_{j'}/(N-1)$. As a result, the permutation sum in Eq. (81) reduces to

$$(N-1)\frac{N_j}{N-1}\frac{N_{j'}}{N-1}f_{jj'}\sum_{P\left\langle N-2\right|}\sum_{P\left|N-2\right\rangle}\left\langle \dots\beta\beta'\dots\right|\left|\dots\beta\beta'\beta''\dots\right\rangle, \tag{8.83}$$

where our specific position k is now excluded from both the bra- and ket-vector permutations. Each of these permutations now includes only (N_j-1) states j and $(N_{j'}-1)$ states j', so using the state orthonormality, we finally arrive at a very simple result:

$$\begin{aligned}&\left\langle \dots N_j,\dots,N_{j'}-1,\dots\right|\hat{F}\left|\dots N_j-1,\dots N_{j'},\dots\right\rangle \\ &=\frac{N_1!\dots(N_j-1)!\dots(N_{j'}-1)!\dots}{(N-1)!}\left(N_jN_{j'}\right)^{1/2}(N-1)\frac{N_j}{N-1}\frac{N_{j'}}{N-1}f_{jj'}\frac{(N-2)!}{N_1!\dots(N_j-1)!\dots(N_{j'}-1)!\dots} \\ &\equiv\left(N_jN_{j'}\right)^{1/2}f_{jj'}.\end{aligned} \tag{8.84}$$

On the other hand, let us calculate the matrix elements of the following operator:

$$\sum_{j,j'}f_{jj'}\hat{a}_j^\dagger\hat{a}_{j'}. \tag{8.85}$$

A direct application of Eqs. (64) and (68) shows that the only non-vanishing elements are

$$\left\langle \dots N_j,\dots,N_{j'}-1,\dots\right|f_{jj'}\hat{a}_j^\dagger\hat{a}_{j'}\left|\dots N_j-1,\dots,N_{j'},\dots\right\rangle=\left(N_jN_{j'}\right)^{1/2}f_{jj'}. \tag{8.86}$$

But this is exactly the last form of Eq. (84), so in the basis of Dirac states, the operator (77) may be represented as

Single-particle operator in Dirac representation

$$\boxed{\hat{F}=\sum_{j,j'}f_{jj'}\hat{a}_j^\dagger\hat{a}_{j'}.} \tag{8.87}$$

This beautifully simple relation is the key formula of the second quantization theory and is essentially the Dirac-representation analog of Eq. (4.59) of the single-particle quantum mechanics. Each term of the sum (87) may be described by a very simple mnemonic rule: for each pair of single-particle states j and j', first, annihilate a particle in the state j', then create one in the state j, and finally weigh the result with the corresponding single-particle matrix element. One of the corollaries of Eq. (87) is that the expectation value of an operator whose eigenstates coincide with the Dirac states is

$$\langle F\rangle \equiv \langle \dots N_j,\dots|\hat{F}|\dots N_j,\dots\rangle = \sum_j f_{jj}N_j, \tag{8.88}$$

with an evident physical interpretation as the sum of single-particle expectation values over all states, weighed by the occupancy of each state.

Proceeding to *fermions*, which have to obey the Pauli principle, we immediately notice that any occupation number N_j may only take two values, 0 or 1. To account for that, and also make the key relation (87) valid for fermions as well, the creation-annihilation operators are defined by the following relations:

$$\hat{a}_j|N_1,N_2,\dots,0_j,\dots\rangle = 0, \qquad \hat{a}_j|N_1,N_2,\dots,1_j,\dots\rangle = (-1)^{\Sigma(1,j-1)}|N_1,N_2,\dots,0_j,\dots\rangle, \tag{8.89}$$

$$\hat{a}_j^\dagger|N_1,N_2,\dots,0_j,\dots\rangle = (-1)^{\Sigma(1,j-1)}|N_1,N_2,\dots,1_j,\dots\rangle, \qquad \hat{a}_j^\dagger|N_1,N_2,\dots,1_j,\dots\rangle = 0, \tag{8.90}$$

Fermion creation-annihilation operators

where the symbol $\Sigma(J, J')$ means the sum of all occupancy numbers in the states with numbers from J to J', including the border points:

$$\Sigma(J,J') \equiv \sum_{j=J}^{J'} N_j, \tag{8.91}$$

so the sum participating in Eqs. (89)-(90) is the total occupancy of all states with the numbers below j. (The states are supposed to be numbered in a fixed albeit arbitrary order.) As a result, these relations may be conveniently summarized in the following verbal form: if an operator replaces the j^{th} state's occupancy with the opposite one (1 with 0 and vice versa), it also changes the sign before the result if (and only if) the total number of particles in the states with $j' < j$ is odd.

Let us use this (perhaps somewhat counter-intuitive) sign alternation rule to spell out the ket-vector $|11\rangle$ of a completely filled two-state system, formed from the vacuum state $|00\rangle$ in two different ways. If we start by creating a fermion in state 1, we get

$$\hat{a}_1^\dagger|0,0\rangle = (-1)^0|1,0\rangle \equiv |1,0\rangle, \qquad \hat{a}_2^\dagger\hat{a}_1^\dagger|0,0\rangle = \hat{a}_2^\dagger|1,0\rangle = (-1)^1|1,1\rangle \equiv -|1,1\rangle, \tag{8.92a}$$

while if the operator order is different, the result is

$$\hat{a}_2^\dagger|0,0\rangle = (-1)^0|0,1\rangle \equiv |0,1\rangle, \qquad \hat{a}_1^\dagger\hat{a}_2^\dagger|0,0\rangle = \hat{a}_1^\dagger|0,1\rangle = (-1)^0|1,1\rangle \equiv |1,1\rangle, \tag{8.92b}$$

so

$$\left(\hat{a}_1^\dagger\hat{a}_2^\dagger + \hat{a}_2^\dagger\hat{a}_1^\dagger\right)|0,0\rangle = 0\,. \tag{8.93}$$

Since the action of any of these operator products on any initial state rather than the vacuum one also gives the null ket, we may write the following operator equality:

$$\hat{a}_1^\dagger \hat{a}_2^\dagger + \hat{a}_2^\dagger \hat{a}_1^\dagger \equiv \left\{ \hat{a}_1^\dagger, \hat{a}_2^\dagger \right\} = \hat{0}. \tag{8.94}$$

It is straightforward to check that this result is valid for Dirac vectors of an arbitrary length, and does not depend on the occupancy of other states, so we may generalize it as

$$\boxed{\left\{ \hat{a}_j^\dagger, \hat{a}_{j'}^\dagger \right\} = \left\{ \hat{a}_j, \hat{a}_{j'} \right\} = \hat{0};} \tag{8.95}$$

Fermionic operators: commutation relations

these equalities hold for $j = j'$ as well. On the other hand, an absolutely similar calculation shows that the mixed creation-annihilation commutators do depend on whether the states are different or not:[28]

$$\boxed{\left\{ \hat{a}_j, \hat{a}_{j'}^\dagger \right\} = \hat{I}\delta_{jj'}.} \tag{8.96}$$

These equations look very much like Eqs. (75)-(76) for bosons, "only" with the replacement of commutators with anticommutators. Since the core laws of quantum mechanics, including the operator compatibility (Sec. 4.5) and the Heisenberg equation (4.199) of operator evolution in time, involve commutators rather than anticommutators, one might think that all the behavior of bosonic and fermionic multiparticle systems should be dramatically different. However, the difference is not as big as one could expect; indeed, a straightforward check shows that the sign factors in Eqs. (89)-(90) just compensate those in the Slater determinant, and thus make the key relation (87) valid for the fermions as well. (Indeed, this is the very goal of the introduction of these factors.)

To illustrate this fact on the simplest example, let us examine what the second quantization formalism says about the dynamics of non-interacting particles in the system whose single-particle properties we have discussed repeatedly, namely two nearly similar potential wells, coupled by tunneling through the separating potential barrier – see, e.g., Figs. 2.21 or 7.4. If the coupling is so small that the states localized in the wells are only weakly perturbed, then in the basis of these states, the single-particle Hamiltonian of the system may be represented by the 2×2 matrix (5.3). With the energy reference selected in the middle between the energies of unperturbed states, the coefficient b vanishes, this matrix is reduced to

$$\mathrm{h} = \mathbf{c}\cdot\boldsymbol{\sigma} \equiv \begin{pmatrix} c_z & c_- \\ c_+ & -c_z \end{pmatrix}, \qquad \text{with } c_\pm \equiv c_x \pm i c_y, \tag{8.97}$$

and its eigenvalues to

$$\varepsilon_\pm = \pm c, \qquad \text{with } c \equiv |\mathbf{c}| \equiv \left(c_x^2 + c_y^2 + c_z^2 \right)^{1/2}. \tag{8.98}$$

Using the key relation (87) together with Eq. (97), we may represent the Hamiltonian of the whole system of particles in terms of the creation-annihilation operators:

$$\hat{H} = c_z \hat{a}_1^\dagger \hat{a}_1 + c_- \hat{a}_1^\dagger \hat{a}_2 + c_+ \hat{a}_2^\dagger \hat{a}_1 - c_z \hat{a}_2^\dagger \hat{a}_2, \tag{8.99}$$

where $\hat{a}_{1,2}^\dagger$ and $\hat{a}_{1,2}$ are the operators of creation and annihilation of a particle in the corresponding *potential well*. (Again, in the second quantization approach the *particles* are not numbered at all!) As

[28] A by-product of this calculation is proof that the operator defined by Eq. (72) counts the number of particles N_j (now equal to either 1 or 0), just as it does for bosons.

Eq. (72) shows, the first and the last terms of the right-hand side of Eq. (99) describe the particle energies $\varepsilon_{1,2} = \pm c_z$ in uncoupled wells,

$$c_z \hat{a}_1^\dagger \hat{a}_1 = c_z \hat{N}_1 \equiv \varepsilon_1 \hat{N}_1, \qquad -c_z \hat{a}_2^\dagger \hat{a}_2 = -c_z \hat{N}_2 \equiv \varepsilon_2 \hat{N}_2, \tag{8.100}$$

while the sum of the middle two terms is the second-quantization description of tunneling between the wells.

Now we can use the general Eq. (4.199) of the Heisenberg picture to spell out the equations of motion of the creation-annihilation operators. For example,

$$i\hbar \dot{\hat{a}}_1 = \left[\hat{a}_1, \hat{H}\right] = c_z \left[\hat{a}_1, \hat{a}_1^\dagger \hat{a}_1\right] + c_- \left[\hat{a}_1, \hat{a}_1^\dagger \hat{a}_2\right] + c_+ \left[\hat{a}_1, \hat{a}_2^\dagger \hat{a}_1\right] - c_z \left[\hat{a}_1, \hat{a}_2^\dagger \hat{a}_2\right]. \tag{8.101}$$

Since the Bose and Fermi operators satisfy different commutation relations, one could expect the right-hand side of this equation to be different for bosons and fermions. However, it is not so! Indeed, all commutators on the right-hand side of Eq. (101) have the following form:

$$\left[\hat{a}_j, \hat{a}_{j'}^\dagger \hat{a}_{j''}\right] \equiv \hat{a}_j \hat{a}_{j'}^\dagger \hat{a}_{j''} - \hat{a}_{j'}^\dagger \hat{a}_{j''} \hat{a}_j. \tag{8.102}$$

As Eqs. (74) and (94) show, the first pair product of operators on the right-hand side may be recast as

$$\hat{a}_j \hat{a}_{j'}^\dagger = \hat{I}\delta_{jj'} \pm \hat{a}_{j'}^\dagger \hat{a}_j, \tag{8.103}$$

where the upper sign pertains to bosons and the lower one to fermions, while according to Eqs. (76) and (95), the very last pair product in Eq. (102) is

$$\hat{a}_{j''} \hat{a}_j = \pm \hat{a}_j \hat{a}_{j''}, \tag{8.104}$$

with the same sign convention. Plugging these expressions into Eq. (102), we see that regardless of the particle type, there is a universal (and generally very useful) commutation relation

$$\left[\hat{a}_j, \hat{a}_{j'}^\dagger \hat{a}_{j''}\right] = \hat{a}_{j''} \delta_{jj'}, \tag{8.105}$$

valid for both bosons and fermions. As a result, the Heisenberg equation of motion for operator $\hat{a}_1$, and the equation for $\hat{a}_2$ (which may be obtained absolutely similarly), are also universal:[29]

$$\begin{aligned} i\hbar \dot{\hat{a}}_1 &= c_z \hat{a}_1 + c_- \hat{a}_2, \\ i\hbar \dot{\hat{a}}_2 &= c_+ \hat{a}_1 - c_z \hat{a}_2. \end{aligned} \tag{8.106}$$

This is a system of two coupled linear differential equations, which is similar to the equations for the *c*-number probability amplitudes of single-particle wavefunctions of a two-level system – see, e.g., Eq. (2.201) and the model solution of Problem 4.25. Their general solution is a linear superposition

$$\hat{a}_{1,2}(t) = \sum_{\pm} \hat{\alpha}_{1,2}^{(\pm)} \exp\{\lambda_\pm t\}. \tag{8.107}$$

[29] Equations of motion for the creation operators $\hat{a}_{1,2}^\dagger$ are just the Hermitian conjugates of Eqs. (106), and do not add any new information about the system's dynamics.

As usual, in order to find the exponents $\lambda_\pm$, it is sufficient to plug a particular solution $\hat{a}_{1,2}(t) = \hat{\alpha}_{1,2}\exp\{\lambda t\}$ into Eq. (106) and require that the determinant of the resulting linear system for the "coefficients" (actually, time-independent operators) $\hat{\alpha}_{1,2}$ equals zero. This gives us the following characteristic equation

$$\begin{vmatrix} c_z - i\hbar\lambda & c_- \\ c_+ & -c_z - i\hbar\lambda \end{vmatrix} = 0, \tag{8.108}$$

with two roots $\lambda_\pm = \pm i\Omega/2$, where $\Omega \equiv 2c/\hbar$ – cf. Eq. (5.20). Now plugging each of the roots, one by one, into the system of equations for $\hat{\alpha}_{1,2}$, we can find these operators, and hence the general solution of system (98) for arbitrary initial conditions.

Let us consider the simple case $c_y = c_z = 0$ (meaning in particular that the wells are exactly aligned, see Fig. 2.21), so $\hbar\Omega/2 \equiv c = c_x$; then the solution of Eq. (106) is

$$\hat{a}_1(t) = \hat{a}_1(0)\cos\frac{\Omega t}{2} - i\hat{a}_2(0)\sin\frac{\Omega t}{2}, \qquad \hat{a}_2(t) = -i\hat{a}_1(0)\sin\frac{\Omega t}{2} + \hat{a}_2(0)\cos\frac{\Omega t}{2}. \tag{8.109}$$

Multiplying the first of these relations by its Hermitian conjugate, and ensemble-averaging the result, we get

Quantum oscillations: second quantization form

$$\begin{aligned}\langle N_1\rangle \equiv \left\langle \hat{a}_1^\dagger(t)\hat{a}_1(t)\right\rangle = \left\langle \hat{a}_1^\dagger(0)\hat{a}_1(0)\right\rangle\cos^2\frac{\Omega t}{2} + \left\langle \hat{a}_2^\dagger(0)\hat{a}_2(0)\right\rangle\sin^2\frac{\Omega t}{2} \\ - i\left\langle \hat{a}_1^\dagger(0)\hat{a}_2(0) + \hat{a}_2^\dagger(0)\hat{a}_1(0)\right\rangle\sin\frac{\Omega t}{2}\cos\frac{\Omega t}{2}.\end{aligned} \tag{8.110}$$

Let the initial state of the system be a single Dirac state, i.e. have a definite number of particles in each well; in this case, only the two first terms on the right-hand side of Eq. (110) are different from zero, giving:[30]

$$\langle N_1\rangle = N_1(0)\cos^2\frac{\Omega t}{2} + N_2(0)\sin^2\frac{\Omega t}{2}. \tag{8.111}$$

For one particle, initially placed in either well, this gives us our old result (2.181) describing the usual quantum oscillations of the particle between two wells with the frequency Ω. However, Eq. (111) is valid for *any* set of initial occupancies; let us use this fact. For example, starting from two particles, with initially one particle in each well, we get $\langle N_1\rangle = 1$, regardless of time. So, the occupancies do not oscillate, and no experiment may detect the quantum oscillations, though their frequency Ω is still formally present in the time evolution equations. This fact may be interpreted as the simultaneous quantum oscillations of two particles between the wells, exactly in anti-phase. For bosons, we can go on to even larger occupancies by preparing the system, for example, in the state with $N_1(0) = N$, $N_2(0) = 0$. The result (111) says that in this case, we see that the quantum oscillation amplitude increases N-fold; this is a particular manifestation of the general fact that bosons can be (and in time, stay) in the same quantum state. On the other hand, for fermions we cannot increase the initial occupancies beyond 1, so the largest oscillation amplitude we can get is if we initially fill just one well.

[30] For the second well's occupancy, the result is complementary, $N_2(t) = N_1(0)\sin^2\Omega t + N_2(0)\cos^2\Omega t$, giving a good sanity check: $N_1(t) + N_2(t) = N_1(0) + N_2(0) = \text{const}$.

The Dirac approach may be readily generalized to more complex systems. For example, Eq. (99) implies that an arbitrary system of potential wells with weak tunneling coupling between the adjacent wells may be described by the Hamiltonian

$$\hat{H} = \sum_j \varepsilon_j \hat{a}_j^\dagger \hat{a}_j + \sum_{\{j,j'\}} \delta_{jj'} \hat{a}_j^\dagger \hat{a}_{j'} + \text{h.c.}, \tag{8.112}$$

where the symbol $\{j, j'\}$ means that the second sum is restricted to pairs of next-neighbor wells – see, e.g., Eq. (2.203) and its discussion. Note that this Hamiltonian is still a quadratic form of the creation-annihilation operators, so the Heisenberg-picture equations of motion of these operators are still linear, and its exact solutions, though possibly cumbersome, may be studied in detail. Due to this fact, the Hamiltonian (112) is widely used for the study of some phenomena, for example, the very interesting *Anderson localization* effects, in which a random distribution of the localized-site energies ε_j prevents tunneling particles, within a certain energy range, from spreading to unlimited distances.[31]

8.4. Perturbative approaches

The situation becomes much more difficult if we need to account for explicit interactions between the particles. Let us assume that the interaction may be reduced to that between their pairs (as in the case at the Coulomb forces and most other interactions[32]), so it may be described by the following "pair-interaction" Hamiltonian

$$\hat{U}_{\text{int}} = \frac{1}{2} \sum_{\substack{k,k'=1 \\ k \neq k'}}^{N} \hat{u}_{\text{int}}(\mathbf{r}_k, \mathbf{r}_{k'}), \tag{8.113}$$

Pair-interaction Hamiltonian: two forms

with the front factor of ½ compensating the double-counting of each particle pair by this double sum. The translation of this operator to the second-quantization form may be done absolutely similarly to the derivation of Eq. (87), and gives a similar (though naturally more involved) result

$$\hat{U}_{\text{int}} = \frac{1}{2} \sum_{j,j',l,l'} u_{jj'll'} \hat{a}_j^\dagger \hat{a}_{j'}^\dagger \hat{a}_{l'} \hat{a}_l, \tag{8.114}$$

where the two-particle matrix elements are defined similarly to Eq. (82):

$$u_{jj'll'} \equiv \left\langle \beta_j \beta_{j'} \middle| \hat{u}_{\text{int}} \middle| \beta_l \beta_{l'} \right\rangle. \tag{8.115}$$

The only new feature of Eq. (114) is a specific order of the indices of the creation operators. Note the mnemonic rule of writing this expression, similar to that for Eq. (87): each term corresponds to moving a pair of particles from states l and l' to states j' and j (in this order!) factored with the corresponding two-particle matrix element (115).

However, with the account of this term, the resulting Heisenberg equations of the time evolution of the creation/annihilation operators become nonlinear, so solving them and calculating observables from the results is usually impossible, at least analytically. The only case when some general results

[31] For a review of the 1D version of this problem, see, e.g., J. Pendry, *Adv. Phys.* **43**, 461 (1994).

[32] A simple but important example from the condensed matter theory is the so-called *Hubbard model*, in which particle repulsion limits their number on each of localized sites to either 0, or 1, or 2, with negligible interaction of the particles on different sites – though the next-neighbor sites are still connected by tunneling, as in Eq. (112).

may be obtained is the *weak interaction* limit. In this case, the unperturbed Hamiltonian contains only single-particle terms such as (79), and we can always (at least as a matter of principle :-) find such a basis of orthonormal single-particle states β_j in which that Hamiltonian is diagonal in the Dirac representation:

$$\hat{H}^{(0)} = \sum_j \varepsilon_j^{(0)} \hat{a}_j^\dagger \hat{a}_j \,. \tag{8.116}$$

Now we can use Eq. (6.14), in this basis, to calculate the interaction energy as a first-order perturbation:

$$\begin{aligned} E_{\rm int}^{(1)} &= \langle N_1, N_2, \ldots | \hat{U}_{\rm int} | N_1, N_2, \ldots \rangle = \frac{1}{2} \langle N_1, N_2, \ldots | \sum_{j,j',l,l'} u_{jj'll'} \hat{a}_j^\dagger \hat{a}_{j'}^\dagger \hat{a}_{l'} \hat{a}_l | N_1, N_2, \ldots \rangle \\ &\equiv \frac{1}{2} \sum_{j,j',l,l'} u_{jj'll'} \langle N_1, N_2, \ldots | \hat{a}_j^\dagger \hat{a}_{j'}^\dagger \hat{a}_{l'} \hat{a}_l | N_1, N_2, \ldots \rangle . \end{aligned} \tag{8.117}$$

Since, according to Eq. (63), the Dirac states with different occupancies are orthogonal, the last long bracket is different from zero only for three particular subsets of its indices:

(i) $j \neq j'$, $l = j$, and $l' = j'$. In this case, the four-operator product in Eq. (117) is equal to $\hat{a}_j^\dagger \hat{a}_{j'}^\dagger \hat{a}_{j'} \hat{a}_j$, and applying the proper commutation rules twice, we can bring it to the so-called *normal ordering*, with each creation operator standing to the right of the corresponding annihilation operator, thus forming the particle number operator (72):

$$\hat{a}_j^\dagger \hat{a}_{j'}^\dagger \hat{a}_{j'} \hat{a}_j = \pm \hat{a}_j^\dagger \hat{a}_{j'}^\dagger \hat{a}_j \hat{a}_{j'} = \pm \hat{a}_j^\dagger \left(\pm \hat{a}_j \hat{a}_{j'}^\dagger \right) \hat{a}_{j'} = \hat{a}_j^\dagger \hat{a}_j \hat{a}_{j'}^\dagger \hat{a}_{j'} = \hat{N}_j \hat{N}_{j'} , \tag{8.118}$$

with a similar sign of the final result for bosons and fermions.

(ii) $j \neq j'$, $l = j'$, and $l' = j$. In this case, the four-operator product is equal to $\hat{a}_j^\dagger \hat{a}_{j'}^\dagger \hat{a}_j \hat{a}_{j'}$, and bringing it to the form $\hat{N}_j \hat{N}_{j'}$ requires only one commutation:

$$\hat{a}_j^\dagger \hat{a}_{j'}^\dagger \hat{a}_j \hat{a}_{j'} = \hat{a}_j^\dagger \left(\pm \hat{a}_j \hat{a}_{j'}^\dagger \right) \hat{a}_{j'} = \pm \hat{a}_j^\dagger \hat{a}_j \hat{a}_{j'}^\dagger \hat{a}_{j'} = \pm \hat{N}_j \hat{N}_{j'} , \tag{8.119}$$

with the upper sign for bosons and the lower sign for fermions.

(iii) All indices are equal to each other, giving $\hat{a}_j^\dagger \hat{a}_{j'}^\dagger \hat{a}_{l'} \hat{a}_l = \hat{a}_j^\dagger \hat{a}_j^\dagger \hat{a}_j \hat{a}_j$. For fermions, such an operator (that "tries" to create or to kill two particles in a row, in the same state) immediately gives the null vector. In the case of bosons, we may use Eq. (74) to commute the internal pair of operators, getting

$$\hat{a}_j^\dagger \hat{a}_j^\dagger \hat{a}_j \hat{a}_j = \hat{a}_j^\dagger \left(\hat{a}_j \hat{a}_j^\dagger - \hat{I} \right) \hat{a}_j = \hat{N}_j (\hat{N}_j - \hat{I}) . \tag{8.120}$$

Note, however, that this expression formally covers the fermion case as well (always giving zero). As a result, Eq. (117) may be rewritten in the following universal form:

Particle interaction: energy correction

$$E_{\rm int}^{(1)} = \frac{1}{2} \sum_{\substack{j,j' \\ j \neq j'}} N_j N_{j'} \left(u_{jj'jj'} \pm u_{jj'j'j} \right) + \frac{1}{2} \sum_j N_j (N_j - 1) u_{jjjj} . \tag{8.121}$$

The corollaries of this important result are very different for bosons and fermions. In the former case, the last term usually dominates, because the matrix elements (115) are typically the largest when all basis functions coincide. Note that this term allows a very simple interpretation: the number of the diagonal matrix elements it sums up for each state (j) is just the number of interacting particle pairs residing in that state.

In contrast, for fermions, the last term is zero, and the interaction energy is proportional to the difference between the two terms inside the first parentheses. To spell them out, let us consider the case when there is no direct spin-orbit interaction. Then the vectors $|\beta\rangle_j$ of the single-particle state basis may be represented as direct products $|o_j\rangle\otimes|m_j\rangle$ of their orbital and spin-orientation parts. (Here, for the brevity of notation, I am using m instead of m_s.) For spin-½ particles, including electrons, m_j may equal only either +½ or –½; in this case, the spin part of the first matrix element proportional to $u_{jj'jj'}$ equals

$$\langle m|\otimes\langle m'\|m\rangle\otimes|m'\rangle, \tag{8.122}$$

where, as in the general Eq. (115), the position of a particular state vector in each direct product encodes the particle's number. Since the spins of different particles are defined in different Hilbert spaces, we may swap their state vectors to get

$$\langle m|\otimes\langle m'\|m\rangle\otimes|m'\rangle=\left(\langle m|m\rangle\right)_1\times\left(\langle m'|m'\rangle\right)_2=1, \tag{8.123}$$

for any pair of j and j'. On the other hand, the second matrix element, $u_{jj'j'j}$, is factored as

$$\langle m|\otimes\langle m'\|m'\rangle\otimes|m\rangle=\left(\langle m|m'\rangle\right)_1\times\left(\langle m'|m\rangle\right)_2=\delta_{mm'}. \tag{8.124}$$

In this case, it is convenient to rewrite Eq. (121) in the coordinate representation, by using single-particle wavefunctions called *spin-orbitals*

$$\psi_j(\mathbf{r})\equiv\langle\mathbf{r}|\beta_j\rangle=\left(\langle\mathbf{r}|o\rangle\otimes|m\rangle\right)_j. \tag{8.125}$$

Spin-orbital

They differ from the spatial parts of the usual orbital wavefunctions of the type (4.233) only in that their index j should be understood as the set of the orbital-state and the spin-orientation indices.[33] Also, due to the Pauli-principle restriction of the numbers N_j to either 0 or 1, Eq. (121) may be also rewritten without the explicit occupancy numbers, with the understanding that the summation is extended only over the pairs of occupied states. As a result, it becomes

$$E_{\rm int}^{(1)}=\frac{1}{2}\sum_{\substack{j,j'\\ j\neq j'}}\int d^3r\int d^3r'\left[\begin{array}{l}\psi_j^*(\mathbf{r})\psi_{j'}^*(\mathbf{r}')u_{\rm int}(\mathbf{r},\mathbf{r}')\psi_j(\mathbf{r})\psi_{j'}(\mathbf{r}')\\ -\psi_j^*(\mathbf{r})\psi_{j'}^*(\mathbf{r}')u_{\rm int}(\mathbf{r},\mathbf{r}')\psi_{j'}(\mathbf{r})\psi_j(\mathbf{r}')\end{array}\right]. \tag{8.126}$$

Energy correction due to fermion interaction

In particular, for a system of two electrons, we may limit the summation to just two states ($j, j' = 1, 2$). As a result, we return to Eqs. (39)-(41), with the bottom (minus) sign in Eq. (39), corresponding to the triplet spin states. Hence, Eq. (126) may be considered as the generalization of the direct and exchange interaction picture to an arbitrary number of orbitals and an arbitrary total number N of

[33] The spin-orbitals (125) are also close to spinors (13), besides that the former definition takes into account that the spin s of a single particle is fixed, so the spin-orbital may be indexed by the spin's orientation $m \equiv m_s$ only. Also, if an orbital index is used, it should be clearly distinguished from j, i.e. the set of the orbital and spin indices. This is why I believe that the frequently met notation of spin-orbitals as $\psi_{j,s}(\mathbf{r})$ may lead to confusion.

electrons. Note, however, that this formula cannot correctly describe the energy of the singlet spin states, corresponding to the plus sign in Eq. (39), and also of the entangled triplet states.[34] The reason is that the description of entangled spin states, given in particular by Eqs. (18) and (20), requires linear *superpositions* of different Dirac states. (Proof of this fact is left for the reader's exercise.)

Now comes a very important fact: the *approximate* result (126), added to the sum of unperturbed energies $\varepsilon_j^{(0)}$, equals the sum (over j) of *exact* eigenenergies of the so-called *Hartree-Fock equation*:[35]

Hartree-Fock equation

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + u(\mathbf{r})\right)\psi_j(\mathbf{r}) + \sum_{j'\neq j}\int \psi_{j'}^*(\mathbf{r}')u_{\text{int}}(\mathbf{r},\mathbf{r}')\left[\psi_j(\mathbf{r})\psi_{j'}(\mathbf{r}') - \psi_{j'}(\mathbf{r})\psi_j(\mathbf{r}')\right]d^3r' = \varepsilon_j\psi_j(\mathbf{r}), \tag{8.127}$$

where $u(\mathbf{r})$ is the external field's potential acting on each particle separately – see the second of Eqs. (79). An advantage of this equation in comparison with Eq. (126) is that it allows the (approximate) calculation of not only the energy spectrum of the system but, simultaneously, a more exact calculation of the corresponding spin-orbitals $\psi_j(\mathbf{r})$, which takes into account the electron-electron interaction. Of course, Eq. (127) describes a system of mutually coupled *integro-differential* equations. There are, however, efficient methods of numerical solution of such systems, typically based on iterative approaches. One more important practical trick is the exclusion of the filled internal electron shells (see Sec. 3.7) from the explicit calculations because the shell states are virtually unperturbed by the valence electrons involved in typical atomic phenomena and chemical reactions. In this approach, the Coulomb field of the shells, described by fixed pre-calculated *pseudo-potentials*, is added to that of the nuclei. This approach dramatically cuts the computing resources necessary for systems of relatively heavy atoms, enabling a pretty accurate simulation of electronic and chemical properties of rather complex molecules, with thousands of electrons.[36] As a result, the Hartree-Fock approximation has become the de facto baseline of all so-called *ab initio* ("first-principle") calculations in the very important field of quantum chemistry.[37]

In departures from this baseline, there are two opposite trends. For larger accuracy (and typically smaller systems), several "post-Hartree-Fock methods", notably including the *configuration interaction* method,[38] that are more complex but may provide higher accuracy, have been developed. There is also a strong opposite trend of extending such *ab initio* ("first-principle") methods to larger systems while sacrificing some of the results' accuracy and reliability. The ultimate limit of this trend is applicable when the single-particle wavefunction overlaps are small and hence the exchange interaction is

[34] Indeed, due to the condition $j' \neq j$, and Eq. (124), the calculated negative exchange interaction is limited to electron state pairs with the same spin direction – such as the factorable triplet states (↑↑ and ↓↓) of a two-electron system, in which the contribution of the E_{ex} given by Eq. (41), to the total energy is also negative.

[35] This equation was suggested in 1929 by Douglas Hartree for the direct interaction and extended to the exchange interaction by Vladimir Fock in 1930. It may be derived by variational methods, but to verify its compliance with Eq. (126), it is sufficient to multiply all terms of Eq. (127) by $\psi_j^*(\mathbf{r})$, integrate them over all $\mathbf{r}$-space (so the right-hand side would give ε_j), and then sum the single-particle energies over all occupied states j.

[36] For condensed-matter systems, this and other computational methods are applied to single elementary spatial cells, with a limited number of electrons in them, using cyclic boundary conditions.

[37] See, e.g., A. Szabo and N. Ostlund, *Modern Quantum Chemistry*, Revised ed., Dover, 1996.

[38] That method, in particular, allows the calculation of proper linear superpositions of the Dirac states (such as the entangled states for $N = 2$, discussed above) which are missing in the generic Hartree-Fock approach.

negligible. In this limit, the last term in the square brackets in Eq. (127) may be ignored and the multiplier $\psi_j(\mathbf{r})$ taken out of the integral, resulting in the Schrödinger equation for a single particle but in a self-consistent effective potential:

$$u_{\text{ef}}(\mathbf{r}) = u(\mathbf{r}) + u_{\text{dir}}(\mathbf{r}), \quad u_{\text{dir}}(\mathbf{r}) = \sum_{j' \neq j} \int \psi^*_{j'}(\mathbf{r}') u_{\text{int}}(\mathbf{r}, \mathbf{r}') \psi_{j'}(\mathbf{r}') d^3 r' . \tag{8.128}$$

Hartree approximation

This is the so-called *Hartree approximation* – which gives reasonable results for some systems,[39] especially those with low electron density.

However, in dense electron systems (such as typical atoms, molecules, and condensed matter), the exchange interaction described by the second term in the square brackets of Eqs. (126)-(127) may be as high as ~30% of the direct interaction, and frequently cannot be ignored. The tendency to take this interaction in the simplest possible form is currently dominated by the so-called *Density-Functional Theory*,[40] universally known by its acronym DFT. In this approach, the equation solved for each eigenfunction $\psi_j(\mathbf{r})$ is a Schrödinger-like *Kohn-Sham equation*

$$\left[-\frac{\hbar^2}{2m} \nabla^2 + u(\mathbf{r}) + u^{\text{KS}}_{\text{dir}}(\mathbf{r}) + u_{\text{xc}}(\mathbf{r}) \right] \psi_j(\mathbf{r}) = \varepsilon_j \psi_j(\mathbf{r}) , \tag{8.129}$$

Kohn-Sham equation

where

$$u^{\text{KS}}_{\text{dir}}(\mathbf{r}) = -e\phi(\mathbf{r}), \qquad \phi(\mathbf{r}) = \frac{1}{4\pi\varepsilon_0} \int d^3 r' \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}, \qquad \rho(\mathbf{r}) = -en(\mathbf{r}), \tag{8.130}$$

and $n(\mathbf{r})$ is the total electron density in a particular point, calculated self-consistently as

$$n(\mathbf{r}) \equiv \sum_j \psi^*_j(\mathbf{r}) \psi_j(\mathbf{r}). \tag{8.131}$$

The most important feature of the Kohn-Sham Hamiltonian is the simplified description of the exchange and correlation effects by the effective *exchange-correlation* potential $u_{\text{xc}}(\mathbf{r})$. This potential is calculated in various approximations, most of them valid only in the limit when the number of electrons in the system is very high. The simplest of them (proposed by Kohn *et al*. in the 1960s) is the *Local Density Approximation* (LDA) in which the effective exchange potential at each point $\mathbf{r}$ is a function only of the electron density n at the same point, taken from the theory of a *uniform* gas of free electrons.[41] However, for many tasks of quantum chemistry, the accuracy given by the LDA is insufficient because inside molecules, the density n typically changes very fast, so the DFT has become widely accepted in that field only after the introduction, in the 1980s, of more accurate though more cumbersome models for $u_{\text{xc}}(\mathbf{r})$, notably the so-called *Generalized Gradient Approximations* (GGAs). Due to its relative simplicity, the so-modified DFT enables calculation of some properties of much

[39] An example of the Hartree approximation is the *Thomas-Fermi model* of heavy atoms (with $Z >> 1$), in which the atom's electrons, at each distance r from the nucleus, are treated as an ideal, uniform Fermi gas, with a certain density $n(r)$ corresponding to the *local* value $u_{\text{ef}}(r)$, but a *global* value of their highest full single-particle energy, $\varepsilon = 0$, to ensure the equilibrium. (The analysis of this model is left for the reader's exercise.)

[40] It had been developed by Walter Kohn and his associates (notably Pierre Hohenberg) in 1965-66, and eventually (in 1998) was marked with a Nobel Prize in Chemistry for W. Kohn.

[41] Just for the reader's reference: for a uniform, degenerate Fermi-gas of electrons (with the Fermi energy $\varepsilon_F >> k_B T$), the most important, exchange part u_x of u_{xc} may be calculated analytically: $u_x = -(3/4\pi)\, e^2 k_F / 4\pi\varepsilon_0$, where the Fermi momentum $k_F = (2m_e\varepsilon_F)^{1/2}/\hbar$ is defined by the electron density: $n = 2(4\pi/3)k_F^3/(2\pi)^3 \equiv k_F^3/3\pi^2$.

larger systems than the methods based on the Hartree-Fock theory, with the same computing resources and reasonable precision. As a result, it has become a very popular tool for ab initio calculations. This popularity is enhanced by the availability of several advanced DFT software packages, some of them in the public domain.

Please note, however, that despite this undisputable success, this approach has its problems. From my personal point of view, the most offensive of them is the implicit assumption of unphysical Coulomb interaction of an electron with itself – by dropping, on the way from Eq. (128) to Eq. (130), the condition $j' \neq j$ at the calculation of $u_{\rm dir}^{\rm KS}$). As a result, all the available DFT packages I am aware of are either unable to account for some charge transfer effects or require substantial artificial tinkering.[42]

Unfortunately, because of a lack of time/space, for details I have to refer the interested reader to specialized literature.[43]

8.5. Quantum computation and cryptography

Now I have to review the emerging fields of *quantum computation and cryptography*.[44] These fields are currently the subject of intensive research and development efforts, which have already brought, besides much hype, some results of general importance. My coverage will focus on these results, referring the reader interested in details to special literature.[45] Because of the very active stage of the field, the style of this section is closer to a brief literature review than to a textbook's section.

Presently, most work on quantum computation and encryption is based on systems of spatially separated (and hence *distinguishable*) two-level systems – in this context, universally called *qubits*.[46] Due to this distinguishability, the issues that were the focus of the previous sections of this chapter, including the second quantization approach, are irrelevant here. On the other hand, systems of qubits have some interesting properties that have not been discussed in this course yet.

First of all, a system of $N >> 1$ qubits may contain much more information than the same number of N classical bits. Indeed, according to the discussions in Chapter 4 and Sec. 5.1, an arbitrary pure state of a single qubit may be represented by its ket vector (4.37) – see also Eq. (5.1):

$$\left|\alpha\right\rangle_{N=1} = \alpha_1 \left|u_1\right\rangle + \alpha_2 \left|u_2\right\rangle, \tag{8.132}$$

[42] For just a few examples, see N. Simonian *et al*., *J. Appl. Phys*. **113**, 044504 (2013); M. Medvedev *et al*., *Science* **335**, 49 (2017); A. Hutama *et al*., *J. Phys. Chem*. C **121**, 14888 (2017).

[43] See, e.g., either the monograph by R. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules*, Oxford U. Press, 1994, or the later textbook J. A. Steckel and D. Sholl, *Density Functional Theory*: *Practical Introduction*, Wiley, 2009. A popular review and references to more recent work in this still-developing field was given by A. Zangwill, *Phys. Today* **68**, 34 (July 2015).

[44] Since these fields are much related, they are often referred to under the common title of "quantum information science", though this term is rather misleading, de-emphasizing physical aspects of the topic.

[45] Despite the recent flood of new books on the field, one of its first surveys, by M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information*, Cambridge U. Press, 2000, is perhaps still the best one.

[46] In some texts, the term qubit (or "Qbit", or "Q-bit") is used instead for the *information contents* of a two-level system – very much like the classical bit of information (in this context, frequently called "Cbit" or "C-bit") describes the information contents of a classical bistable system – see, e.g., SM Sec. 2.2.

where $\{u_j\}$ is any orthonormal two-state basis. (It is natural and common to employ, as u_j, the eigenstates of the observable that is eventually measured in the particular physical implementation of the qubit.) It is also common to write the kets of these base states as $|0\rangle$ and $|1\rangle$, so Eq. (132) takes the form

$$|\alpha\rangle_{N=1} = a_0|0\rangle + a_1|1\rangle \equiv \sum_j a_j|j\rangle . \tag{8.133}$$

Qubit state's representation

(Here, and in the balance of this section, the letter j is used to denote an integer equal to either 0 or 1.) According to this relation, any state α of a qubit is completely defined by two complex c-numbers a_j, i.e. by 4 real numbers. Moreover, due to the normalization condition $|a_1|^2 + |a_2|^2 = 1$, we need just 3 independent real numbers – say, the Bloch sphere coordinates θ and φ (see Fig. 5.3), plus the common phase γ, which becomes important only when we consider states of a several-qubit system.

This is a good time to note that a qubit is very much different from any classical bistable system used to store single bits of information – such as two possible voltage states of the usual SRAM cell (essentially, a positive-feedback loop of two transistor-based inverters). Namely, the stationary states of a classical bistable system, due to its nonlinearity, are stable with respect to small perturbations, so they may be very robust to unintentional interactions with their environment. In contrast, the qubit's state may be disturbed (i.e. its representation point on the Bloch sphere shifted) by even minor perturbations, because it does not have such an internal state stabilization mechanism.[47] Due to this reason, qubit-based systems are rather vulnerable to environment-induced drifts, including the dephasing and relaxation discussed in the previous chapter, creating major experimental challenges – see below.

Now, if we have a system of two qubits, the vectors of its arbitrary pure state may be represented as a sum of $2^2 = 4$ terms,[48]

$$|\alpha\rangle_{N=2} = a_{00}|00\rangle + a_{01}|01\rangle + a_{10}|10\rangle + a_{11}|11\rangle \equiv \sum_{j_1,j_2} a_{j_1 j_2}|j_1 j_2\rangle , \tag{8.134}$$

with four complex coefficients, i.e. eight real numbers, subject to just one normalization condition that follows from the requirement $\langle\alpha|\alpha\rangle = 1$:

$$\sum_{j_{1,2}} \left| a_{j_1 j_2} \right|^2 = 1 . \tag{8.135}$$

The evident generalization of Eqs. (133)-(134) to an arbitrary pure state of an N-qubit system is a sum of 2^N terms:

$$|\alpha\rangle_N = \sum_{j_1, j_2, \cdots j_N}{}' a_{j_1 j_2 \ldots j_N}|j_1 j_2 \ldots j_N\rangle , \tag{8.136}$$

including all possible combinations of 0s and 1s for N indices j, so the state is fully described by 2^N complex numbers, i.e. $2\cdot 2^N \equiv 2^{N+1}$ real numbers, with only one constraint, similar to Eq. (135), imposed by the normalization condition. This exponential growth of the information contents would not be

[47] In this aspect as well, the information processing systems based on qubits are much closer to classical analog computers (which were popular once, but nowadays are used for a few special applications only) rather than classical digital ones.

[48] Here and in most instances below I use the same shorthand notation as was used at the beginning of this chapter – cf. Eq. (1b). In this short form, the qubit's number is coded by the order of its state index inside a full ket-vector, while in the long form, such as in Eq. (137), by the order of its single-qubit vector in a full direct product.

possible without the qubit state entanglement. Indeed, in the particular case when qubit states are not entangled, i.e. are factorable:

$$|\alpha\rangle_N = |\alpha_1\rangle|\alpha_2\rangle...|\alpha_N\rangle, \tag{8.137}$$

where each $|\alpha_n\rangle$ is described by an equality similar to Eq. (133) with its individual expansion coefficients, the system state description requires only $3N - 1$ real numbers – e.g., N sets $\{\theta, \varphi, \gamma\}$ less one common phase.

However, it would be wrong to project this exponential growth of information contents directly on the capabilities of quantum computation, because this process has to include the output information readout, i.e. qubit state measurements. Due to the fundamental intrinsic uncertainty of quantum systems, the measurement of a single qubit even in a pure state (133) generally may give either of two results, with probabilities $W_0 = |a_0|^2$ and $W_1 = |a_1|^2$. To comply with the general notion of computation, any quantum computer has to provide certain (or virtually certain) results, and hence the probabilities W_j have to be very close to either 0 or 1, so before the measurement, each measured qubit has to be in one of the basis states – either 0 or 1. This means that the computational system with N output qubits, just before their final readout, has to be in one of the factorable states

$$|\alpha\rangle_N = |j_1\rangle|j_2\rangle...|j_N\rangle \equiv |j_1 j_2 ... j_N\rangle, \tag{8.138}$$

which is a very small subset even of the set of all unentangled states (137), and whose maximum information contents is just N classical bits.

Now the reader may start thinking that this constraint strips quantum computations of any advantages over their classical counterparts, but such a view is also superficial. To show that, let us consider the scheme of the "baseline" type of quantum computation, shown in Fig. 3.

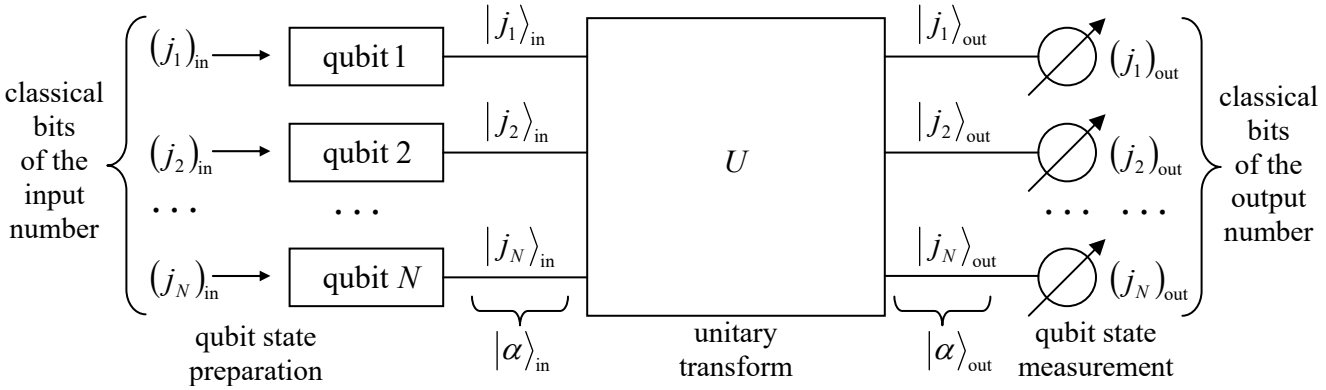


Fig. 8.3. The baseline scheme of quantum computation.

Here each horizontal line (sometimes called a "wire"[49]) corresponds to a single qubit, tracing its time evolution in the same direction as at the usual time function plots: from left to right. This means

[49] The notion of "wires" stems from the similarity between such quantum schemes and the schematics of classical computation circuits – see, e.g., Fig. 4a below. In the classical case, the lines may be indeed understood as physical wires connecting physical devices: logic gates and/or memory cells. In this context, note that classical computer components also have non-zero time delays, so even in that case, the left-to-right device ordering is useful to indicate the timing of (and frequently the causal relation between) the signals.

that the left column $|\alpha\rangle_{in}$ of ket-vectors describes the initial state of the qubits,[50] while the right column $|\alpha\rangle_{out}$ describes their final (but pre-measurement) state. The box labeled U represents the qubit evolution in time due to their specially arranged interactions between each other and/or external drive "forces". These forces are assumed to be noise-free, and the system, during this evolution, is supposed to be ideally isolated from any dephasing and energy-dissipating environment, so the process may be described by a unitary operator defined in the 2^N-dimensional Hilbert space of N qubits:

$$\left|\alpha\right\rangle_{out} = \hat{U}\left|\alpha\right\rangle_{in}. \tag{8.139}$$

With the condition that the input and output states have the simple form (138), this equality reads

$$\left|(j_1)_{out}(j_2)_{out}...(j_N)_{out}\right\rangle = \hat{U}\left|(j_1)_{in}(j_2)_{in}...(j_N)_{in}\right\rangle. \tag{8.140}$$

The art of quantum computer design consists of selecting such unitary operators $\hat{U}$ that would:

– satisfy Eq. (140),

– be physically implementable, and

– enable substantial performance advantages of the quantum computation over its classical counterparts with similar functionality, at least for some digital functions (algorithms).

I will have time/space to demonstrate the possibility of such advantages on just one, perhaps the simplest example – the so-called *Deutsch problem*,[51] discussing several common notions and issues of this field on the way. Let us consider the family of single-bit classical Boolean functions $j_{out} = f(j_{in})$. Since both j are Boolean variables, i.e. may take only values 0 and 1, there are evidently only $2\times2 = 4$ such functions – see the first four columns of the following table:

f	$f(0)$	$f(1)$	class	F	$f(1)–f(0)$
f_1	0	0	constant	0	0
f_2	0	1	balanced	1	+1
f_3	1	0	balanced	1	–1
f_4	1	1	constant	0	0

(8.141)

Of them, the functions f_1 and f_4, whose values are independent of their arguments, are called *constants*, while the functions f_2 (called "YES" or "IDENTITY") and f_3 ("NOT" or "INVERSION") are called *balanced*. The Deutsch problem is to determine the class of a single-bit function, implemented in a "black box", as being either constant or balanced, using just one experiment.

[50] As was discussed in Chapter 7, the preparation of a pure state (133) is (conceptually :-) straightforward. Placing a qubit into a weak contact with an environment of temperature $T << \Delta/k_B$, where Δ is the difference between energies of the eigenstates 0 and 1, we may achieve its relaxation into the lowest-energy state. Then, if the qubit must be set into a different pure state, it may be driven there by the application of a pulse of a proper external classical "force". In most physical implementations of qubits, the most practicable way for that step is to use the proper part of the Rabi oscillation period – see Sec. 6.5.

[51] It is named after David Elieser Deutsch, whose 1985 paper (motivated by an inspirational but not very specific publication by Richard Feynman in 1982) launched the whole field of quantum computation.

Classically, this is clearly impossible, and the simplest way to perform the function's classification involves two similar black boxes f – see Fig. 4a.[52] It also uses the so-called *exclusive-OR* (XOR for short) *gate* whose output is described by the following function F of its two Boolean arguments j_1 and j_2:[53]

$$F(j_1, j_2) = j_1 \oplus j_2 \equiv \begin{cases} 0, & \text{if } j_1 = j_2, \\ 1, & \text{if } j_1 \neq j_2. \end{cases} \tag{8.142}$$

In the particular circuit shown in Fig. 4a, the gate produces the following output:

$$F = f(0) \oplus f(1), \tag{8.143}$$

which is equal to 1 if $f(0) \neq f(1)$, i.e. if the function f is balanced, and to 0 in the opposite case – see column F in Eq. (141).

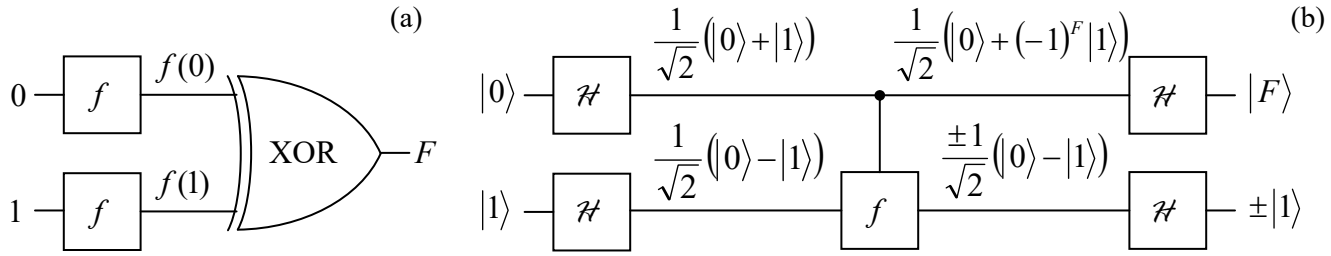


Fig. 8.4. The simplest (a) classical and (b) quantum ways to classify a single-bit Boolean function f.

On the other hand, as will be shown below, any of the four functions f may be implemented quantum-mechanically, for example (Fig. 5a) as a unitary transform of two input qubits, acting as follows on each basis component $|j_1 j_2\rangle \equiv |j_1\rangle|j_2\rangle$ of the general input state (134):

$$\hat{f}\,|j_1\rangle|j_2\rangle = |j_1\rangle|j_2 \oplus f(j_1)\rangle, \tag{8.144}$$

where f is the corresponding classical Boolean function – see the table in Eq. (141).

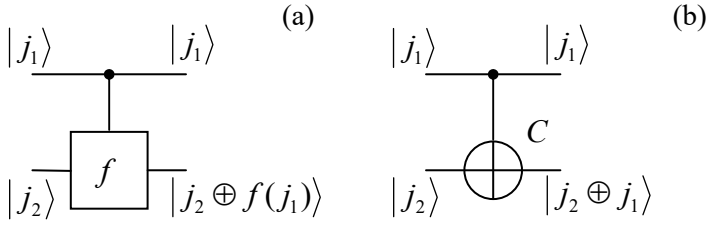


Fig. 8.5. Two-qubit quantum gates: (a) a two-qubit function f and (b) its particular case C (CNOT), and their actions on a basis state.

In the particular case when f in Eq. (144) is just the YES function: $f(j) = f_2(j) = j$, this "circuit" is reduced to the so-called *CNOT gate*, a key ingredient of many other quantum computation schemes, performing the following two-qubit transform:

[52] Alternatively, we may perform two sequential experiments on the same black box f, first recording, and then recalling the first experiment's result. However, the Deutsch problem calls for a single-shot experiment.

[53] The XOR sign $\oplus$ should be by no means confused with the sign $\otimes$ of the direct product of state vectors (which in this section is just implied).

$$\hat{C}\left|j_1 j_2\right\rangle = \left|j_1\right\rangle\left|j_2 \oplus j_1\right\rangle. \tag{8.145a}$$

CNOT function

Let us use Eq. (142) to spell out this function for all four possible input qubit combinations:

$$\hat{C}|00\rangle = |00\rangle, \quad \hat{C}|01\rangle = |01\rangle, \quad \hat{C}|10\rangle = |11\rangle, \quad \hat{C}|11\rangle = |10\rangle. \tag{8.145b}$$

In plain English, this means that acting on a basis state $j_1 j_2$, the CNOT gate leaves the state of the first, *source* qubit (shown by the upper horizontal line in Fig. 5) intact, but flips the state of the second, *target* qubit if the first one is in the basis state 1. In even simpler words, the state j_1 of the source qubit controls the NOT function acting on the target qubit; hence the gate's name CNOT – the semi-acronym of "Controlled NOT".

For the quantum function (144), with an arbitrary and unknown f, the Deutsch problem may be solved within the general scheme shown in Fig. 3, with the particular structure of the unitary-transform box U spelled out in Fig. 4b, which involves just one implementation of the function f. Here the single-qubit quantum gate $\hat{\mathcal{H}}$ performs the *Hadamard* (or "Walsh-Hadamard" or "Walsh") *transform*,[54] whose operator is defined by the following actions on the qubit's basis states:

$$\hat{\mathcal{H}}|0\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right), \qquad \hat{\mathcal{H}}|1\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle - |1\rangle\right), \tag{8.146}$$

Hadamard transform

– see also the two left state-label columns in Fig. 4b.[55] Since this operator has to be linear (to be quantum-mechanically realistic), it needs to perform the action (146) on the basis states even when they are parts of a linear superposition – as they are, for example, for the two right Hadamard gates in Fig. 4b. For example, as immediately follows from Eqs. (146) and the operator's linearity,

$$\hat{\mathcal{H}}\left(\hat{\mathcal{H}}|0\rangle\right) = \hat{\mathcal{H}}\left(\frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right)\right) = \frac{1}{\sqrt{2}}\hat{\mathcal{H}}\left(|0\rangle + \hat{\mathcal{H}}|1\rangle\right) = \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right) + \frac{1}{\sqrt{2}}\left(|0\rangle - |1\rangle\right)\right) = |0\rangle. \tag{8.147a}$$

Absolutely similarly, we may get[56]

$$\hat{\mathcal{H}}\left(\hat{\mathcal{H}}|1\rangle\right) = |1\rangle. \tag{8.147b}$$

Let us carry out a sequential analysis of the whole "circuit" shown in Fig. 4b. Since the input states of gate f in this particular circuit are described by Eqs. (146), its output state's ket is

$$\hat{f}\left(\hat{\mathcal{H}}|0\rangle\hat{\mathcal{H}}|1\rangle\right) = \hat{f}\left(\frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right)\frac{1}{\sqrt{2}}\left(|0\rangle - |1\rangle\right)\right) \equiv \frac{1}{2}\left(\hat{f}|00\rangle - \hat{f}|01\rangle + \hat{f}|10\rangle - \hat{f}|11\rangle\right). \tag{8.148}$$

Now we may apply Eq. (144) to each component in the parentheses:

---

[54] Named after mathematicians J. Hadamard (1865-1963) and J. Walsh (1895-1973). Note that to avoid any chance of confusion between the Hadamard transform's operator $\hat{\mathcal{H}}$ and the general Hamiltonian operator $\hat{H}$, in these notes, they are typeset using different fonts.

[55] Note that according to Eq. (146), the Hadamard operator does *not* belong to the class of transforms described by Eq. (140) – while the whole "circuit" shown in Fig. 4b, does – see below.

[56] Since the states 0 and 1 form a full basis of a single qubit, both Eqs. (147) may be summarized as operator equality $\hat{\mathcal{H}}^2 = \hat{I}$. It is also easy to verify that the Hadamard transform of an arbitrary state may be represented on the Bloch sphere (Fig. 5.3) as a π-rotation about the axis that bisects the angle between the x- and z-axes.

$$\begin{aligned}\hat{f}|00\rangle-\hat{f}|01\rangle+\hat{f}|10\rangle-\hat{f}|11\rangle &\equiv \hat{f}|0\rangle|0\rangle-\hat{f}|0\rangle|1\rangle+\hat{f}|1\rangle|0\rangle-\hat{f}|1\rangle|1\rangle \\ &= |0\rangle|0\oplus f(0)\rangle-|0\rangle|1\oplus f(0)\rangle+|1\rangle|0\oplus f(1)\rangle-|1\rangle|1\oplus f(1)\rangle \\ &\equiv |0\rangle\Big(|0\oplus f(0)\rangle-|1\oplus f(0)\rangle\Big)+|1\rangle\Big(|0\oplus f(1)\rangle-|1\oplus f(1)\rangle\Big).\end{aligned} \tag{8.149}$$

Note that the contents of the first parentheses of the last expression, characterizing the state of the target qubit, is equal to $(|0\rangle - |1\rangle) \equiv (-1)^0 (|0\rangle - |1\rangle)$ if $f(0) = 0$ (and hence $0\oplus f(0) = 0$ and $1\oplus f(0) = 1$), and to $(|1\rangle - |0\rangle) \equiv (-1)^1(|0\rangle - |1\rangle)$ in the opposite case $f(0) = 1$, so both cases may be described in one shot by rewriting the parentheses as $(-1)^{f(0)}(|0\rangle - |1\rangle)$. The second parentheses are absolutely similarly controlled by the value of $f(1)$, so the two outputs of gate f are unentangled:

$$\hat{f}\Big(\hat{\mathcal{H}}|0\rangle\hat{\mathcal{H}}|1\rangle\Big)=\frac{1}{2}\Big((-1)^{f(0)}|0\rangle+(-1)^{f(1)}|1\rangle\Big)\Big(|0\rangle-|1\rangle\Big)=\pm\frac{1}{\sqrt{2}}\Big(|0\rangle+(-1)^{F}|1\rangle\Big)\frac{1}{\sqrt{2}}\Big(|0\rangle-|1\rangle\Big), \tag{8.150}$$

where the last step has used the fact that the classical Boolean function F defined by Eq. (142) is equal to $\pm[f(1) - f(0)]$ – please compare the last two columns in Eq. (141). The front sign ± in the last form of Eq. (150) may be prescribed to any of the component ket-vectors – for example to that of the target qubit, as shown by the third column of state labels in Fig. 4b.

This intermediate result is already rather remarkable. Indeed, it shows that, despite the superficial impression one could get from Fig. 5, the gates f and C, being "controlled" by the source qubit, may change that qubit's state as well! This fact (partly reflected by the vertical direction of the control lines in Figs. 4 and 5, symbolizing the same stage of the system's time evolution) shows how careful one should be interpreting quantum-computational "circuits", thriving on qubits' entanglement: the "signals" on different sections of a horizontal "wire" may differ – see Fig. 4b again.

At the last stage of the circuit shown in Fig. 4b, the qubit components of the state (150) are fed into one more pair of Hadamard gates, whose outputs therefore are

$$\hat{\mathcal{H}}\frac{1}{\sqrt{2}}\Big(|0\rangle+(-1)^{F}|1\rangle\Big)=\frac{1}{\sqrt{2}}\Big(\hat{\mathcal{H}}|0\rangle+(-1)^{F}\hat{\mathcal{H}}|1\rangle\Big),\quad\text{and}\quad\hat{\mathcal{H}}\left(\pm\frac{1}{\sqrt{2}}\Big(|0\rangle-|1\rangle\Big)\right)=\pm\frac{1}{\sqrt{2}}\Big(\hat{\mathcal{H}}|1\rangle-\hat{\mathcal{H}}|0\rangle\Big). \tag{8.151}$$

Now using Eqs. (146) again, we see that the output state ket-vectors of the source and target qubits are, respectively,

$$\frac{1+(-1)^{F}}{2}|0\rangle+\frac{1-(-1)^{F}}{2}|1\rangle,\quad\text{and}\quad\pm|1\rangle. \tag{8.152}$$

Since, according to Eq. (142), the Boolean function F may take only values 0 or 1, the final state of the source qubit is always one of its basis states j, namely the one with $j = F$. Its measurement tells us whether the function f, participating in Eq. (144), is constant or balanced – see Eq. (141) again.[57]

Thus, the quantum circuit shown in Fig. 4b indeed solves the Deutsch problem in one shot. Reviewing our analysis, we may see that this is possible because the unitary transform performed by the quantum gate f is applied to the linear combinations (146) rather than to the basis states 0 and 1. Due to this trick, the quantum state components depending on $f(0)$ and $f(1)$ are processed simultaneously, in

[57] Note that the last Hadamard transform of the target qubit (i.e. the Hadamard gate shown in the lower right corner of Fig. 4b) is not necessary for the Deutsch problem's solution – though it should be included if we want the whole circuit to satisfy the condition (140).

parallel. This *quantum parallelism* may be extended to circuits with many ($N >> 1$) qubits and, for some tasks, provide a dramatic performance increase – for example, reducing the necessary circuit component number from $O(2^N)$ to $O(N^p)$, where p is a finite (and not very big) number.

However, this efficiency comes at a high price. Indeed, let us discuss the possible physical implementation of quantum gates, starting from the single-qubit case, on an example of the Hadamard gate (146). With the linearity requirement, its action on the arbitrary state (133) should be

$$\hat{\mathcal{H}}|\alpha\rangle = a_0\hat{\mathcal{H}}|0\rangle + a_1\hat{\mathcal{H}}|1\rangle = a_0\frac{1}{\sqrt{2}}\Big(|0\rangle + |1\rangle\Big) + a_1\frac{1}{\sqrt{2}}\Big(|0\rangle - |1\rangle\Big) = \frac{1}{\sqrt{2}}\big(a_0 + a_1\big)|0\rangle + \frac{1}{\sqrt{2}}\big(a_0 - a_1\big)|1\rangle, \quad (8.153)$$

meaning that the state probability amplitudes in the end ($t = \mathcal{T}$) and in the beginning ($t = 0$) of the qubit evolution in time have to be related as

$$a_0(\mathcal{T}) = \frac{a_0(0) + a_1(0)}{\sqrt{2}}, \quad a_1(\mathcal{T}) = \frac{a_0(0) - a_1(0)}{\sqrt{2}}. \quad (8.154)$$

This task may be again performed using the Rabi oscillations, which were discussed in Sec. 6.5, i.e. by applying to the qubit (a two-level system), for a limited time period $\mathcal{T}$, a weak sinusoidal external signal of frequency ω equal to the intrinsic quantum oscillation frequency $\omega_{nn'}$ defined by Eq. (6.85). The analysis of the Rabi oscillations was carried out in Sec. 6.5, even for non-vanishing (though small) detuning $\Delta = \omega - \omega_{nn'}$, but only for the particular initial conditions when at $t = 0$ the system was fully in one on the basis states (there labeled as n'), i.e. the counterpart state (there labeled n) was empty. For our current purposes, we need to find the amplitudes $a_{0,1}(t)$ for arbitrary initial conditions $a_{0,1}(0)$, subject only to the time-independent normalization condition $|a_0|^2 + |a_1|^2 = 1$. For the case of exact tuning, $\Delta = 0$, the solution of the system (6.94) is elementary,[58] and gives the following solution:[59]

$$\begin{aligned} a_0(t) &= a_0(0)\cos\Omega t - ia_1(0)e^{i\varphi}\sin\Omega t, \\ a_1(t) &= a_1(0)\cos\Omega t - ia_0(0)e^{-i\varphi}\sin\Omega t, \end{aligned} \quad (8.155)$$

where Ω is the Rabi oscillation frequency (6.99), in the exact-tuning case proportional to the amplitude $|A|$ of the external ac drive $A = |A|\exp\{i\varphi\}$ – see Eq. (6.86). Comparing these expressions with Eqs. (154), we see that for $t = \mathcal{T} = \pi/4\Omega$ and $\varphi = \pi/2$ they "almost" coincide, besides the opposite sign of $a_1(\mathcal{T})$. Conceptually the simplest way to correct this deficiency is to follow the *ac* "$\pi/4$-pulse", just discussed, by a short *dc* "π-pulse" of the duration $\mathcal{T}' = \pi/\delta$, which temporarily creates a small additional energy difference δ between the basis states 0 and 1. According to the basic Eq. (1.62), such difference creates an additional phase difference $\mathcal{T}'\delta/\hbar$ between the states, equal to π for the "π-pulse".

Another way (that may be also useful for two-qubit operations) is to use another, auxiliary state with energy E_2 whose distances from the basic levels E_1 and E_0 are significantly different from the difference $(E_1 - E_0)$ – see Fig. 6a. In this case, the weak external ac field tuned to any of the three potential quantum transition frequencies $\omega_{nn'} \equiv (E_n - E_{n'})/\hbar$ initiates such transitions between the corresponding states only, with a negligible perturbation of the third state. (Such transitions may be

[58] An alternative way to analyze the qubit evolution is to use the Bloch equation (5.21), with an appropriate function $\mathbf{\Omega}(t)$ describing the control field.

[59] To comply with our current notation, the coefficients $a_{n'}$ and a_n of Sec. 6.5 are replaced with a_0 and a_1.

again described by Eqs. (155), with the appropriate index changes.) For the Hadamard transform implementation, it is sufficient to apply (after the already discussed $\pi/4$-pulse of frequency ω_{10}, and with the initially empty level E_2), an additional π-pulse of frequency ω_{20}, with any phase φ. Indeed, according to the first of Eqs. (155), with the due replacement $a_1(0) \to a_2(0) = 0$, such pulse flips the sign of the amplitude $a_0(t)$, while the amplitude $a_1(t)$, not involved in this additional transition, remains unchanged.

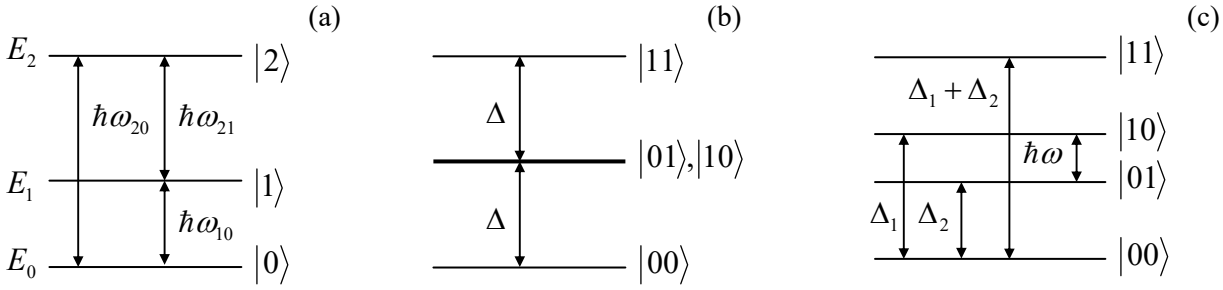


Fig. 8.6. Energy-level schemes used for unitary transformations of (a) single qubits and (b, c) two-qubit systems.

Now let me describe the conceptually simplest (though, for some qubit types, not the most practically convenient) scheme for the implementation of two-qubit gates, on an example of the CNOT gate whose operation is described by Eq. (145). For that, evidently, the involved qubits have to interact for some time $\mathcal{T}$. As was repeatedly discussed in the two last chapters, in most cases such interaction of two subsystems is factorable – see Eq. (6.145). For qubits, i.e. two-level systems, each of the component operators may be represented by a 2×2 matrix in the basis of states 0 and 1. According to Eq. (4.106), such a matrix may be always expressed as a linear combination ($b\mathrm{I} + \mathbf{c}\cdot\boldsymbol{\sigma}$), where b and three Cartesian components of the vector $\mathbf{c}$ are c-numbers. Let us consider the simplest form of such factorable interaction Hamiltonian:

$$\hat{H}_{\rm int}(t) = \begin{cases} \kappa\hat{\sigma}_z^{(1)}\hat{\sigma}_z^{(2)}, & \text{for } 0 < t < \mathcal{T}, \\ 0, & \text{otherwise}, \end{cases} \tag{8.156}$$

where the upper index is the qubit number and κ is a c-number constant.[60] According to Eq. (4.175), by the end of the interaction period, this Hamiltonian produces the following unitary transform:

$$\hat{U}_{\rm int} = \exp\left\{-\frac{i}{\hbar}\hat{H}_{\rm int}\mathcal{T}\right\} \equiv \exp\left\{-\frac{i}{\hbar}\kappa\hat{\sigma}_z^{(1)}\hat{\sigma}_z^{(2)}\mathcal{T}\right\}. \tag{8.157}$$

Since in the basis of unperturbed two-bit basis states $|j_1 j_2\rangle$, the product operator $\hat{\sigma}_z^{(1)}\hat{\sigma}_z^{(2)}$ is diagonal, so is the unitary operator (157), with the following action on these states:

[60] The assumption of simultaneous time independence of the basis state vectors and the interaction operator (within the time interval $0 < t < \mathcal{T}$) is possible only if the basis state energy difference Δ of both qubits is exactly the same. In this case, the simple Eq. (156) follows from Figs. 6b, which shows the spectrum of the total energy $E = E_1 + E_2$ of the two-bit system. In the absence of interaction (Fig. 6b), the energies of two basis states, $|01\rangle$ and $|10\rangle$, are equal, enabling even a weak qubit interaction to cause their substantial evolution in time – see Sec. 6.7. If the qubit energies are different (Fig. 6c), the interaction may still be reduced, in the rotating-wave approximation, to Eq. (156), by compensating the energy difference $(\Delta_1 - \Delta_2)$ with an external ac signal of frequency $\omega = (\Delta_1 - \Delta_2)/\hbar$ – see Sec. 6.5.

$$\hat{U}_{\rm int}\left|j_1 j_2\right\rangle = \exp\left\{i\theta\sigma_z^{(1)}\sigma_z^{(2)}\right\}\left|j_1 j_2\right\rangle, \tag{8.158}$$

where $\theta \equiv -\kappa\mathcal{T}/\hbar$, and σ_z are the eigenvalues of the Pauli matrix σ_z for the basis states of the corresponding qubit: $\sigma_z = +1$ for $|j\rangle = |0\rangle$, and $\sigma_z = -1$ for $|j\rangle = |1\rangle$. Let me, for clarity, spell out Eq. (158) for the particular case $\theta = -\pi/4$ (corresponding to the qubit coupling time $\mathcal{T} = \pi\hbar/4\kappa$):

$$\hat{U}_{\rm int}\left|00\right\rangle = e^{-i\pi/4}\left|00\right\rangle, \quad \hat{U}_{\rm int}\left|01\right\rangle = e^{i\pi/4}\left|01\right\rangle, \quad \hat{U}_{\rm int}\left|10\right\rangle = e^{i\pi/4}\left|10\right\rangle, \quad \hat{U}_{\rm int}\left|11\right\rangle = e^{-i\pi/4}\left|11\right\rangle. \tag{8.159}$$

In order to compensate for the undesirable parts of this joint phase shift of the basis states, let us now apply similar individual "rotations" of each qubit by angle $\theta' = +\pi/4$, using the following product of two independent operators, plus (just for the result's clarity) a common, and hence inconsequential, phase shift $\theta'' = -\pi/4$:[61]

$$\hat{U}_{\rm com} = \exp\left\{i\theta'\left(\hat{\sigma}_z^{(1)} + \hat{\sigma}_z^{(2)}\right) + i\theta''\right\} \equiv \exp\left\{i\frac{\pi}{4}\hat{\sigma}_z^{(1)}\right\}\exp\left\{i\frac{\pi}{4}\hat{\sigma}_z^{(2)}\right\}e^{-i\pi/4}. \tag{8.160}$$

Since this operator is also diagonal in the $|j_1 j_2\rangle$ basis, it is easy to calculate the change of the basis states by the total unitary operator $\hat{U}_{\rm tot} \equiv \hat{U}_{\rm com}\hat{U}_{\rm int}$:

$$\hat{U}_{\rm tot}\left|00\right\rangle = \left|00\right\rangle, \qquad \hat{U}_{\rm tot}\left|01\right\rangle = \left|01\right\rangle, \qquad \hat{U}_{\rm tot}\left|10\right\rangle = \left|10\right\rangle, \qquad \hat{U}_{\rm tot}\left|11\right\rangle = -\left|11\right\rangle. \tag{8.161}$$

This result already shows the main "miracle action" of two-qubit gates, such as the one shown in Fig. 4b: the source qubit is left intact (only if it is in one of the basis states!), while the state of the target qubit is altered. True, this change (of the sign) is still different from the CNOT operator's action (145), but may be readily used for its implementation by sandwiching the transform $U_{\rm tot}$ between two Hadamard transforms of the target qubit alone:

$$\hat{C} = \frac{1}{2}\hat{\mathcal{H}}^{(2)}\hat{U}_{\rm tot}\hat{\mathcal{H}}^{(2)}. \tag{8.162}$$

So, we have spent quite a bit of time on the discussion of the very simple CNOT gate,[62] and now I can reward the reader for their effort with a bit of good news: it has been proved that an *arbitrary* unitary transform that satisfies Eq. (140), i.e. may be used to implement the general scheme outlined in Fig. 3, may be decomposed into a set of CNOT gates, possibly augmented with simpler single-qubit gates.[63] Unfortunately, I have no time for a detailed discussion of more complex circuits.[64] The most

---

[61] As Eq. (4.175) shows, each of the component unitary transforms $\exp\{i\theta'\hat{\sigma}_z\}$ may be created by applying to each qubit, for time interval $\mathcal{T}' = \hbar\theta'/\kappa'$, a constant external field described by Hamiltonian $\hat{H} = -\kappa'\hat{\sigma}_z$. We already know that for a charged, spin-½ particle, such Hamiltonian may be created by applying a *z*-oriented external dc magnetic field – see Eq. (4.163). For most other physical implementations of qubits, the organization of such a Hamiltonian is also straightforward – see, e.g., Fig. 7.4 and its discussion.

[62] As was discussed above, this gate is identical to the two-qubit gate shown in Fig. 5a for $f = f_3$, i.e. $f(j) = j$. The implementation of the gate of f for 3 other possible functions f requires straightforward modifications, whose analysis is left for the reader's exercise.

[63] This fundamental importance of the CNOT gate was perhaps a major reason why David Wineland, the leader of the NIST group that had demonstrated its first experimental implementation in 1995 (following the theoretical suggestion by J. Cirac and P. Zoller), was awarded the 2012 Nobel Prize in Physics – shared with Serge Haroche, the leader of another group working towards quantum computation.

famous of them is the scheme for integer number factoring, suggested in 1994 by Peter Winston Shor.[65] Due to its potential practical importance for breaking the broadly used communication encryption schemes such as the RSA code,[66] this opportunity has incited much enthusiasm and triggered experimental efforts to implement quantum gates and circuits using a broad variety of two-level quantum systems. By now, the following experimental options have given the most significant results:[67]

(i) Trapped ions. The first experimental demonstrations of quantum state manipulation (including the already mentioned first CNOT gate) have been carried out using deeply cooled atoms in optical traps, similar to those used in frequency and time standards. Their total spins are natural qubits, whose states may be manipulated using the Rabi transfers excited by suitably tuned lasers. The spin interactions with the environment may be very weak, resulting in large dephasing times T_2 – up to a few seconds. Since the distances between ions in the traps are relatively large (of the order of a micron), their direct spin-spin interaction is even weaker, but the ions may be made effectively interacting either via their mechanical oscillations about the potential minima of the trapping field or via photons in external electromagnetic resonators ("cavities").[68] Perhaps the main challenge of using this approach to quantum computation is poor "scalability", i.e. the enormous experimental difficulty of creating and managing large ordered systems of individually addressable qubits. So far, only some few-qubit systems have been demonstrated.[69]

(ii) Nuclear spins are also typically very weakly connected to their environment, with dephasing times T_2 exceeding 10 seconds in some cases. Their eigenenergies E_0 and E_1 may be split by external dc magnetic fields (typically, of the order of 10 T), while the interstate Rabi transfers may be readily achieved by using the nuclear magnetic resonance, i.e. the application of external ac fields with frequencies $\omega = (E_1 - E_0)/\hbar$ – typically, of a few hundred MHz. The challenges of this option include the weakness of spin-spin interactions (typically mediated through molecular electrons), resulting in a very slow spin evolution, whose time scale $\hbar/\kappa$ may become comparable with T_2, and also very small level separations $E_1 - E_0$, corresponding to a few K, i.e. much smaller than the thermal energy at room temperature, creating a challenge of qubit state preparation.[70] Despite these challenges, the nuclear spin option was used for the first implementation of the Shor algorithm for factoring of a small number ($15 = 5\times3$) as early as 2001.[71] However, the extension of this success to larger systems, beyond the set of spins inside one molecule, is extremely challenging.

---

[64] For that, the reader may be referred to either the monographs by Nielsen-Chuang and Reiffel-Polak, cited above, or to a shorter (but much more formal) textbook by N. Mermin, *Quantum Computer Science*, Cambridge U. Press, 2007.

[65] A clear description of this algorithm may be found in several accessible sources, including *Wikipedia* – see the article *Shor's Algorithm*.

[66] Named after R. Rivest, A. Shamir, and L. Adleman, the authors of the first open publication of the code in 1977, but actually invented earlier (in 1973) by C. Cocks.

[67] For a discussion of other possible implementations (such as quantum dots and dopants in crystals) see, e.g., T. Ladd *et al*., *Nature* **464**, 45 (2010), and references therein.

[68] A brief discussion of such interactions (so-called *Cavity QED*) will be given in Sec. 9.4 below.

[69] See, e.g., S. Debnath *et al*., *Nature* **536**, 63 (2016). Note also related work on arrays of trapped neutral atoms – see, e.g., J. Perczel *et al*., *Phys. Rev. Lett*. **119**, 023603 (2017); D. Bluvstein *et al*., *Nature* **604**, 451 (2022).

[70] This challenge may be partly mitigated using ingenious spin manipulation techniques such as *refocusing* – see, e.g., either Sec. 7.7 in Nielsen and Chuang, or J. Keeler's monograph cited at the end of Sec. 6.5.

[71] B. Lanyon *et al*., *Phys. Rev. Lett*. **99**, 250505 (2001).

(iii) Josephson-junction devices. Much better scalability may be achieved with solid-state devices, especially using superconductor integrated circuits including weak contacts – Josephson junctions – see their brief discussion in Sec. 1.6. The qubits of this type are based on the fact that the energy U of such a junction is a highly nonlinear function of the Josephson phase difference φ – see Sec. 1.6. Indeed, combining Eqs. (1.73) and (1.74), we can readily calculate $U(\varphi)$ as the work $\mathscr{W}$ of an external circuit increasing the phase from, say, zero to some value φ:

$$U(\varphi)-U(0)=\int_{\varphi'=0}^{\varphi'=\varphi} d\mathscr{W}=\int_{\varphi'=0}^{\varphi'=\varphi} IVdt=\frac{2eI_c}{\hbar}\int_{\varphi'=0}^{\varphi'=\varphi}\sin\varphi'\frac{d\varphi'}{dt}dt=\frac{2eI_c}{\hbar}(1-\cos\varphi). \tag{8.163}$$

There are several options for using this nonlinearity for creating qubits;[72] currently, the leading option, called the *phase qubit*, is using the two lowest eigenstates localized in one of the potential wells of the periodic potential (163). A major problem with such qubits is that at the very bottom of this well the potential $U(\varphi)$ is almost quadratic, so the energy levels are nearly equidistant – cf. Eqs. (2.262), (6.16), and (6.23). This is even more true for the so-called "transmons" (and "Xmons", and "Gatemons", and several other very similar devices[73]) – the currently used phase qubits versions, where a Josephson junction is made a part of an external electromagnetic oscillator, making its relative total nonlineartity (anharmonism) even smaller. As a result, the external rf drive of frequency $\omega = (E_1 - E_0)/\hbar$, used to arrange the state transforms described by Eq. (155), may induce simultaneous undesirable transitions to (and between) higher energy levels. This effect may be mitigated by a reduction of the ac drive amplitude, but at a price of the proportional increase of the operation time and hence of dephasing effects – see below. (I am leaving a quantitative estimate of such an increase for the reader's exercise.)

Since the coupling of Josephson-junction qubits may be most readily controlled (and, very importantly, kept stable if so desired), they have been used to demonstrate the largest prototype quantum computing systems to date, despite quite modest dephasing times T_2 – for purely integrated circuits, in the tens of microseconds at best, even at operating temperatures in tens of mK. Several groups have announced chips with a few dozen of such qubits, but to the best of my knowledge, only their smaller subsets could be used for high-fidelity quantum operations so far.[74]

(iv) Optical systems, attractive because of their inherently enormous bandwidth, pose a special challenge for quantum computation: due to the virtual linearity of most electromagnetic media, the implementation of qubits requires relatively large components and high optical power.[75] In 2001, a very

[72] The "most quantum" option in this technology is to use Josephson junctions very weakly coupled to their dissipative environment (so the effective resistance shunting the junction is much higher than the quantum resistance unit $R_Q \equiv (\pi/2)\,\hbar/e^2 \sim 10^4\ \Omega$). In this case, the Josephson phase variable φ behaves as a coordinate of a 1D quantum particle moving in the 2π-periodic potential (163), forming the energy band structure $E_n(q)$ similar to those discussed in Sec. 2.7. Both theory and experiment show that in this case, the quantum states in adjacent Brillouin zones differ by the charge of one Cooper pair $2e$. (This is exactly the effect responsible for the Bloch oscillations of frequency (2.252).) These two states may be used as the basis states of *charge qubits*. Unfortunately, such qubits are rather sensitive to charged impurities, randomly located in the junction's vicinity, which may cause uncontrollable changes of its parameters, so currently, this option is not very actively pursued.

[73] For a recent review of these devices see, e.g., G. Wendin, *Repts. Progr. Phys*. **80**, 106001 (2017).

[74] See, e.g., C. Song *et al*., *Phys. Rev. Lett*. **119**, 180511 (2017); S. Krinner *et al*., *Nature* **605**, 669 (2022); R. Acharya *et al.*, *Nature* **614**, 676 (2023).

[75] For a state-of-the-art recent work in this direction see, e.g., X. Qiang *et al*., *Nature Photonics* **12**, 534 (2018).

smart way around this hurdle was invented.[76] In this *KLM scheme* (also called "linear optical quantum computing"), nonlinear elements are not needed at all, and quantum gates may be composed just of linear devices (such as optical waveguides, mirrors, and beam splitters), plus single-photon sources and detectors. However, estimates show that this approach requires many more physical components than those using nonlinear quantum systems such as usual qubits,[77] so right now it is not very popular.

So, despite three decades of large-scale (multi-billion-dollar) experimental and theoretical efforts, the progress of quantum computing development has been rather gradual. The main culprit here is the unintentional coupling of qubits to their environment, leading most importantly to their state dephasing, and eventually to errors. Let me discuss this major issue in detail.

Of course, some error probability exists in classical digital logic gates and memory cells as well.[78] However, in this case, there is no conceptual problem with the device state measurement, so the error may be detected and corrected in many ways. Conceptually,[79] the simplest of them is the so-called *majority voting logic* – using several similar logic circuits operating in parallel and fed with identical input data. Evidently, two such devices can detect a single error in one of them, while three devices in parallel may correct such an error, by taking two coinciding output signals for the genuine one.

For quantum computation, the general idea of using several devices (say, qubits) for coding the same information remains valid; however, there are two major complications. First, as we know from Chapter 7, the environment's dephasing effect may be described as a slow random drift of the probability amplitudes a_j, leading to the deviation of the output state α_{fin} from the required form (140), and hence to a non-vanishing probability of wrong qubit state readout – see Fig. 3. Hence the quantum error correction has to protect the result not against possible random state flips $0 \leftrightarrow 1$, as in classical digital computers, but against these "creeping" analog errors.

Second, the qubit state is impossible to copy exactly (*clone*) without disturbing it, as follows from the following simple calculation.[80] Cloning some state α of one qubit to another qubit that is initially in an independent state (say, the basis state 0), without any change of α, means the following transformation of the two-qubit ket: $|\alpha 0\rangle \to |\alpha\alpha\rangle$. If we want such transform to be performed by a real quantum system, whose evolution is described by a unitary operator $\hat{u}$, and to be correct for an arbitrary state α, it has to work not only for both basis states of the qubit:

$$\hat{u}|00\rangle = |00\rangle, \qquad \hat{u}|10\rangle = |11\rangle, \tag{8.164}$$

but also for their arbitrary linear combination (133). Since the operator $\hat{u}$ has to be linear, we may use that relation, and then Eq. (164) to write

[76] E. Knill *et al*., *Nature* **409**, 46 (2001).

[77] See, e.g., Y. Li *et al*., *Phys. Rev. X* **5**, 041007 (2015).

[78] In modern integrated circuits, such "soft" (runtime) errors are created mostly by the high-energy neutrons of cosmic rays, and also by the α-particles emitted by radioactive impurities in silicon chips and their packaging.

[79] Practically, the majority voting logic increases circuit complexity and power consumption, so it is used only in the most critical points. Since in modern digital integrated circuits, the bit error rate is very small ($< 10^{-5}$), in most of them, less radical but also less penalizing schemes are used – if used at all.

[80] Amazingly, this simple *no-cloning theorem* was discovered as late as 1982 (to the best of my knowledge, independently by W. Wooters and W. Zurek, and by D. Dieks), in the context of work toward quantum cryptography – see below.

$$\hat{u}|\alpha 0\rangle \equiv \hat{u}\left(a_0|0\rangle + a_1|1\rangle\right)|0\rangle \equiv a_0\hat{u}|00\rangle + a_1\hat{u}|10\rangle = a_0|00\rangle + a_1|11\rangle . \tag{8.165}$$

On the other hand, the desired result of state cloning is

$$|\alpha\alpha\rangle = \left(a_0|0\rangle + a_1|1\rangle\right)\left(a_0|0\rangle + a_1|1\rangle\right) \equiv a_0^2|00\rangle + a_0a_1\left(|10\rangle + |01\rangle\right) + a_1^2|11\rangle , \tag{8.166}$$

i.e. is evidently different, so, for an arbitrary state α, and an arbitrary unitary operator $\hat{u}$,

$$\boxed{\hat{u}|\alpha 0\rangle \neq |\alpha\alpha\rangle ,} \tag{8.167}$$

No-cloning theorem

meaning that the qubit state cloning is indeed impossible.[81] This problem may be, however, indirectly circumvented – for example, in the way shown in Fig. 7a.

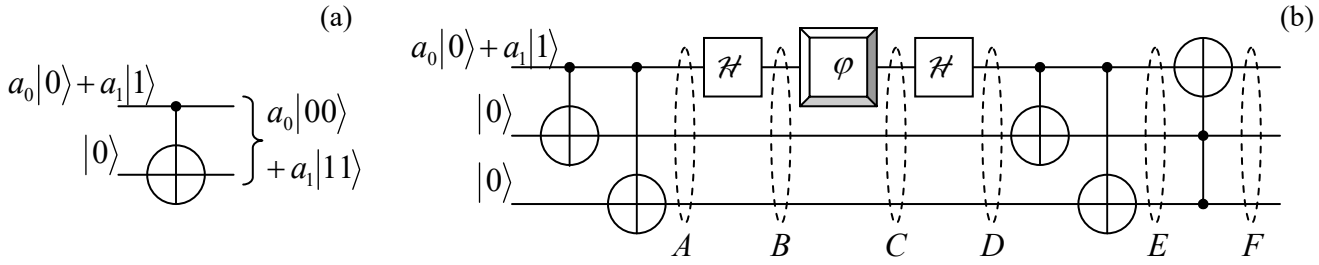


Fig. 8.7. (a) Quasi-cloning, and (b) detection and correction of dephasing errors in a single qubit.

Here the CNOT gate, whose action is described by Eq. (145), entangles an arbitrary input state (133) of the source qubit with a basis initial state of an ancillary target qubit – frequently called the *ancilla*. Using Eq. (145), we can readily calculate the output two-qubit state's vector:

$$\boxed{|\alpha\rangle_{N=2} = \hat{C}\left(a_0|0\rangle + a_1|1\rangle\right)|0\rangle \equiv a_0\hat{C}|00\rangle + a_1\hat{C}|10\rangle = a_0|00\rangle + a_1|11\rangle .} \tag{8.168}$$

Quasi-cloning

We see that this circuit does perform the desired operation (165), i.e. gives the initial source qubit's probability amplitudes a_0 and a_1 equally to two qubits, i.e. duplicates the input information. However, in contrast with "genuine" cloning, it changes the state of the source qubit as well, making it entangled with the target (ancilla) qubit. Such "quasi-cloning" is the key element of most suggested quantum error correction techniques.

Consider, for example, the three-qubit "circuit" shown in Fig. 7b, which uses two ancilla qubits – see the two lower "wires". At its first two stages, the double application of the quasi-cloning produces an intermediate state A with the following ket-vector:

$$|A\rangle = a_0|000\rangle + a_1|111\rangle , \tag{8.169}$$

which is an evident generalization of Eq. (168).[82] Next, subjecting the source qubit to the Hadamard transform (146), we get the three-qubit state B represented by the state vector

[81] Note that this does not mean that two (or several) qubits cannot be put into the same, arbitrary quantum state – theoretically, with arbitrary precision. Indeed, they may be first set into their lowest-energy stationary states, and then driven into the same arbitrary state (133) by exerting on them similar classical external fields. So, the no-cloning theorem pertains only to qubits in *unknown* states α – but this is exactly what we need for error correction – see below.

$$|B\rangle = a_0 \frac{1}{\sqrt{2}}\Big(|0\rangle + |1\rangle\Big)|00\rangle + a_1 \frac{1}{\sqrt{2}}\Big(|0\rangle - |1\rangle\Big)|11\rangle . \tag{8.170}$$

Now let us assume that at this stage, the source qubit comes into contact with a dephasing environment – in Fig. 7b, symbolized by the single-qubit "gate" φ. As we know from Chapter 7 (see Eq. (7.22) and its discussion, and also Sec. 7.3), its effect may be described by a random shift of the relative phase of two states:[83]

$$|0\rangle \to e^{i\varphi}|0\rangle, \quad |1\rangle \to e^{-i\varphi}|1\rangle . \tag{8.171}$$

As a result, for the intermediate state C (see Fig. 7b) we may write

$$|C\rangle = a_0 \frac{1}{\sqrt{2}}\Big(e^{i\varphi}|0\rangle + e^{-i\varphi}|1\rangle\Big)|00\rangle + a_1 \frac{1}{\sqrt{2}}\Big(e^{i\varphi}|0\rangle - e^{-i\varphi}|1\rangle\Big)|11\rangle . \tag{8.172}$$

At this stage of this simple theoretical model, the coupling with the environment is completely stopped (ahh, if this could be possible! we might have quantum computers by now :-), and the source qubit is fed into one more Hadamard gate. Using Eqs. (146) again, for state D after this gate we get

$$|D\rangle = a_0\Big(\cos\varphi|0\rangle + i\sin\varphi|1\rangle\Big)|00\rangle + a_1\Big(i\sin\varphi|0\rangle + \cos\varphi|1\rangle\Big)|11\rangle . \tag{8.173}$$

Next, the qubits are passed through the second, similar pair of CNOT gates – see Fig. 7b. Using Eq. (145), for the resulting state E we readily get the following expression:

$$|E\rangle = a_0\cos\varphi|000\rangle + a_0 i\sin\varphi|111\rangle + a_1 i\sin\varphi|011\rangle + a_1\cos\varphi|100\rangle , \tag{8.174a}$$

whose right-hand side may by evidently re-grouped as

$$|E\rangle = \Big(a_0|0\rangle + a_1|1\rangle\Big)\cos\varphi|00\rangle + \Big(a_1|0\rangle + a_0|1\rangle\Big)\, i\sin\varphi|11\rangle . \tag{8.174b}$$

This is already a rather remarkable result. It shows that if we measured the ancilla qubits at stage E, and both results corresponded to states 0, we might be 100% sure that the source qubit (which is not affected by these measurements!) is in its initial state even after the interaction with the environment. The only result of an increase of this unintentional interaction (as quantified by the r.m.s. magnitude of the random phase shift φ) is the growth of the probability,

$$W = \sin^2\varphi , \tag{8.175}$$

of getting the opposite result, which would signal a dephasing-induced error in the source qubit. Such implicit measurement, without disturbing the source qubit, is called *quantum error detection*.

An even more impressive result may be achieved by the last component of the circuit, the so-called *Toffoli* (or "CCNOT") *gate*, denoted by the rightmost symbol in Fig. 7b. This three-qubit gate is conceptually similar to the CNOT gate discussed above, besides that it flips the basis state of its target qubit only if *both* source qubits are in state 1. (In the circuit shown in Fig. 7b, the former role is played

[82] This state is also the three-qubit example of the so-called *Greenberger-Horne-Zeilinger* (GHZ) *states*, which are frequently called the "most entangled" states of a system of $N > 2$ qubits.

[83] Let me emphasize again that Eq. (171) is strictly valid only if the interaction with the environment is a pure dephasing, i.e. does not include the energy relaxation of the qubit or its thermal activation to the higher-energy eigenstate; however, it is a reasonable description of errors in the frequent case when $T_2 << T_1$.

by our source qubit, while the latter role, by the two ancilla qubits.) According to its definition, the Toffoli gate does not affect the first parentheses in Eq. (174b), but flips the source qubit's states in the second parentheses, so for the output three-qubit state F we get

$$|F\rangle = \left(a_0|0\rangle + a_1|1\rangle\right)\cos\varphi|00\rangle + \left(a_0|0\rangle + a_1|1\rangle\right) i\sin\varphi|11\rangle . \tag{8.176a}$$

Obviously, this result may be factored as

$$|F\rangle = \left(a_0|0\rangle + a_1|1\rangle\right)\left(\cos\varphi|00\rangle + i\sin\varphi|11\rangle\right), \tag{8.176b}$$

Quantum error correction

showing that now the source qubit is again fully unentangled from the ancilla qubits. Moreover, by calculating the norm squared of the second operand, we get

$$\left(\cos\varphi\langle 00| - i\sin\varphi\langle 11|\right)\left(\cos\varphi|00\rangle + i\sin\varphi|11\rangle\right) = \cos^2\varphi + \sin^2\varphi = 1, \tag{8.177}$$

so the final state of the source qubit *exactly* coincides with its initial state. This is the famous miracle of *quantum state correction*, taking place "automatically" – without any qubit measurements, and for any random phase shift φ.

The circuit shown in Fig. 7b may be further improved by adding Hadamard gate pairs, similar to that used for the source qubit, to the ancilla qubits as well. It is straightforward to show that if the dephasing is small in the sense that the W given by Eq. (175) is much less than 1, this modified circuit may provide a substantial error probability reduction (to $\sim W^2$) even if the ancilla qubits are also subjected to a similar dephasing and the source qubits, at the same stage – i.e. between the two Hadamard gates. Such perfect automatic correction of *any* error (not only of an inner dephasing of a qubit and its relaxation/excitation but also of the mutual dephasing between qubits) of *any* used qubit needs even more parallelism. The first circuit of that kind, based on nine qubits, which is a natural generalization of the three-qubit circuit discussed above, was invented in 1995 by the same P. Shor. Later, five-qubit circuits enabling similar error correction were suggested. (The further parallelism reduction has been proved impossible.)

However, all these results assume that the error correction circuits as such are perfect, i.e. completely isolated from the environment. In the real world, this cannot be done. Now the key question is what maximum level W_{max} of the error probability in each gate (including those in the used error correction scheme) can be automatically corrected, and how many qubits with $W < W_{max}$ would be required to implement quantum computers producing important practical results – first of all, factoring of large numbers.[84] To the best of my knowledge, estimates of these two related numbers have been made only for some very specific approaches, and they are rather pessimistic. For example, using the so-called *surface codes*, which employ many physical qubits for coding an informational one, and hence increase its fidelity, W_{max} may be increased to $\sim 10^{-3}$, but then we would need $\sim 10^8$ physical qubits for the Shor's algorithm implementation.[85] This is very far from what currently looks doable using the existing approaches.

Because of this hard situation, the current development of quantum computing is focused on finding at least *some* problems that could be within the reach of either the existing systems, or their immediate extensions, and simultaneously would present some practical interest. Currently, to the best

[84] In order to compete with the existing classical factoring algorithms, such numbers should have at least 10^3 bits.
[85] A. Fowler *et al.*, *Phys. Rev. A* **86**, 032324 (2012).

of my knowledge, all suggested problems of this kind address either specially crafted mathematical problems,[86] or properties of some simple physical systems – such as the molecular hydrogen[87] or the deuteron (the deuterium's nucleus, i.e. the proton-neutron system).[88] In the latter case, the interaction between the qubits of the computational system is organized so that the system's Hamiltonian is similar to that of the quantum system of interest. (For this work, *quantum simulation* is a more adequate name than "quantum computation".[89])

Such simulations are pursued by some teams using schemes different from that shown in Fig. 3. Of those, the most developed is the so-called *adiabatic quantum computation*,[90] which drops the hardest requirement of negligible interaction with the environment. In this approach, the qubit system is first prepared in a certain initial state and then is let to evolve on its own, with no effort to couple-uncouple qubits by external control signals during the evolution.[91] Due to the interaction with the environment, in particular the dephasing and the energy dissipation it imposes, the system eventually relaxes to a final incoherent state, which is then measured. (This reminds the scheme shown in Fig. 3, with the important difference that the transform U may be not fully unitary.) From numerous runs of such an experiment, the outcome statistics may be revealed. Thus, in this approach, the interaction with the environment is allowed to play a certain role in the system evolution, though every effort is made to reduce it, thus slowing down the relaxation process – hence the word "adiabatic" in the name of this approach. This slowness allows the system to exhibit *some* quantum properties, in particular, quantum tunneling[92] *through* the energy barriers separating close energy minima in the multi-dimensional space of states. This tunneling creates a substantial difference in the finite state statistics from that in purely classical systems, where such barriers may be overcome only by thermally-activated jumps *over* them.[93]

Due to technical difficulties of the organization and precise control of long-range interaction in multi-qubit systems, the adiabatic quantum computing demonstrations so far have been limited to a few simple arrays described by the so-called *extended quantum Ising* ("spin-glass") *model*

$$\hat{H} = -J \sum_{\{j,j'\}} \hat{\sigma}_z^{(j)} \hat{\sigma}_z^{(j')} - \sum_j h_j \hat{\sigma}_z^{(j)}, \tag{8.178}$$

where the curly brackets denote the summation over pairs of close (though not necessarily closest) neighbors. Though the Hamiltonian (178) is the traditional playground of phase transitions theory (see, e.g., SM Chapter 4), to the best of my knowledge there are not many practically important tasks that

[86] F. Arute *et al*., *Nature* **574**, 505 (2019). Note that the claim of the first achievement of "quantum supremacy", made in this paper, refers only to an artificial, specially crafted mathematical problem, and does not change my assessment of the current status of this technology.

[87] P. O'Malley *et al*., *Phys. Rev. X* **6**, 031007 (2016).

[88] E. Dumitrescu *et al*., *Phys. Lett. Lett*. **120**, 210501 (2018).

[89] To the best of my knowledge, this idea was first put forward by Yuri I. Malin in his book *Computable and Incomputable* published in 1980, i.e. before the famous 1982 paper by Richard Feynman. Unfortunately, since the book was in Russian, this suggestion was acknowledged by the international community only much later.

[90] Note that the qualifier "quantum" is important in this term, to distinguish this research direction from the *classical* adiabatic (or "reversible") computation – see, e.g., SM Sec. 2.3 and references therein.

[91] Some hybrids of this approach with the "usual" scheme of quantum computation have been demonstrated, in particular, using some control of inter-bit coupling during the relaxation process – see, e.g., R. Barends *et al*., *Nature* **534**, 222 (2016).

[92] As a reminder, this process was repeatedly discussed in this course, starting from Sec. 2.3.

[93] A quantitative discussion of such jumps may be found, for example, in SM Sec. 5.6.

could be achieved by studying the statistics of its solutions. Moreover, even for this limited task, the speed of the largest experimental adiabatic quantum "computers", with several hundreds of Josephson-junction qubits[94] is still comparable with that of classical, off-the-shelf semiconductor processors (with the dollar cost lower by many orders of magnitude), and no dramatic change of this comparison is predicted for realistic larger systems.

To summarize the current (circa mid-2024) situation with the quantum computation development, it faces a very hard challenge of mitigating the effects of unintentional coupling with the environment. This problem is exacerbated by the lack of algorithms, beyond Shor's factoring, that would give quantum computation a substantial advantage over the classical competition in solving real-world problems, and hence a much broader potential customer base that would provide the field with the necessary long-term motivation and resources. So far, even the leading experts in this field abstain from predictions on when quantum computation may become a self-supporting commercial technology.

There may be somewhat better prospects for another application of entangled qubit systems, namely to telecommunication cryptography.[95] The initial motivation here was much more modest: to replace the currently dominating classical encryption, based on the public-key RSA code mentioned above, which may be broken by factoring very large numbers, with a quantum encryption system that would be fundamentally unbreakable. The basis of this opportunity is the measurement postulate and the no-cloning theorem: if a message is carried over by a qubit, it is impossible for an eavesdropper (in cryptography, traditionally called *Eve*) to either measure or copy it faithfully, without also disturbing its state. However, as we have seen from the discussion of Fig. 7a, state *quasi*-cloning using entangled qubits is possible, so the issue is far from being simple, especially if we want to use a publicly distributed quantum key, in some sense similar to the classical public key used at the RSA encryption. Unfortunately, I would not have time/space to discuss various options for quantum cryptography and public distribution of quantum keys,[96] but cannot help demonstrating how inventive and counter-intuitive they may be, on the famous example of the so-called *quantum teleportation* (Fig. 8).[97]

Suppose that some party A (in cryptography, traditionally called *Alice*) wants to send to party B (*Bob*) the full information about the pure quantum state α of a qubit, unknown to either party. Instead of sending her qubit directly to Bob, Alice asks *him* to send *her* one qubit (β) of a pair of other qubits prepared in a certain entangled state, for example in the singlet state described by Eq. (11); in our current notation

$$|\beta\beta'\rangle = \frac{1}{\sqrt{2}}\Big(|01\rangle - |10\rangle\Big). \tag{8.179}$$

The initial state of the whole three-qubit system may be represented in the form

$$|\alpha\beta\beta'\rangle = \Big(a_0|0\rangle + a_1|1\rangle\Big)|\beta\beta'\rangle = \frac{a_0}{\sqrt{2}}|001\rangle - \frac{a_0}{\sqrt{2}}|010\rangle + \frac{a_1}{\sqrt{2}}|010\rangle - \frac{a_1}{\sqrt{2}}|111\rangle, \tag{8.180a}$$

[94] See, e.g., R. Harris *et al*., *Science* **361**, 162 (2018). Similar demonstrations with trapped-ion systems so far have been on a smaller scale, with a few tens of qubits – see, e.g., J. Zhang *et al*., *Nature* **551**, 601 (2017).

[95] This general field was pioneered in the 1970s by S. Wisener.

[96] Two of them are the BB84 suggested in 1984 by C. Bennett and G. Brassard, and the EPRBE suggested in 1991 by A. Ekert. For details, see, e.g., the review by N. Gizin *et al*., *Rev. Mod. Phys*. **74**, 145 (2002).

[97] This procedure had been first suggested in 1993 by the same Charles Bennett and then was repeatedly demonstrated experimentally – see, e.g., L. Steffen *et al*., *Nature* **500**, 319 (2013) and literature therein.

which may be equivalently rewritten as the following linear superposition,

$$\begin{aligned}|\alpha\beta\beta'\rangle = \frac{1}{2}|\alpha\beta\rangle_{\rm s}^{+}\bigl(-a_1|0\rangle + a_0|1\rangle\bigr) + \frac{1}{2}|\alpha\beta\rangle_{\rm s}^{-}\bigl(a_1|0\rangle + a_0|1\rangle\bigr) \\ + \frac{1}{2}|\alpha\beta\rangle_{\rm e}^{+}\bigl(-a_0|0\rangle + a_1|1\rangle\bigr) + \frac{1}{2}|\alpha\beta\rangle_{\rm e}^{-}\bigl(-a_0|0\rangle - a_1|1\rangle\bigr),\end{aligned} \tag{8.180b}$$

of the following four states of the qubit pair $\alpha\beta$:

$$|\alpha\beta\rangle_{\rm s}^{\pm} \equiv \frac{1}{\sqrt{2}}\bigl(|00\rangle \pm |11\rangle\bigr), \qquad |\alpha\beta\rangle_{\rm e}^{\pm} \equiv \frac{1}{\sqrt{2}}\bigl(|01\rangle \pm |10\rangle\bigr). \tag{8.181}$$

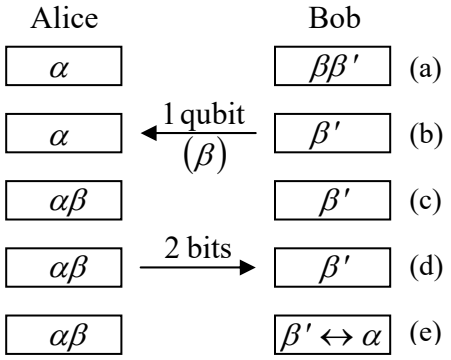


Fig. 8.8. Sequential stages of a "quantum teleportation" procedure: (a) the initial state with entangled qubits β and β', (b) the back transfer of the qubit β, (c) the measurement of the pair $\alpha\beta$, (d) the forward transfer of two classical bits with the measurement results, and (e) the final state, with the state of the qubit β' mirroring the initial state of the qubit α.

After having received qubit β from Bob, Alice measures which of these four states the pair $\alpha\beta$ has. This may be achieved, for example, by measurement of one observable represented by the operator $\hat{\sigma}_z^{(\alpha)}\hat{\sigma}_z^{(\beta)}$ and another one corresponding to $\hat{\sigma}_x^{(\alpha)}\hat{\sigma}_x^{(\beta)}$– cf. Eq. (156). (Since all four states (181) are eigenstates of both these operators, these two measurements do not affect each other and may be performed in any order.) The measured eigenvalue of the former operator enables distinguishing the couples of states (181) with different values of the lower index, while the latter measurement distinguishes the states with different upper indices.

Then Alice reports the measurement result (which may be coded with just two classical bits) to Bob over a classical communication channel. Since the measurement places the pair $\alpha\beta$ *definitely* into the corresponding state, the remaining Bob's bit β' is now *definitely* in the unentangled single-qubit state that is represented by the corresponding parentheses in Eq. (180b). Note that each of these parentheses contains both coefficients $a_{0,1}$, i.e. the whole information about the initial state that the qubit α had initially. If Bob likes, he may now use appropriate single-qubit operations, similar to those discussed earlier in this section, to move his qubit β' into the state *exactly* similar to the initial state of qubit α. This fact does not violate the no-cloning theorem (167) because the measurement has already changed the state of α.) This is, of course, a "teleportation" only in a very special sense of this term, but a good example of the importance of qubit entanglement's preservation at their spatial transfer.[98]

Returning for just a minute to quantum cryptography: since its most common quantum key distribution protocols require just a few simple quantum gates, whose experimental implementation is

[98] For this course, this was also a good primer for the forthcoming discussion of the *EPR paradox* and *Bell's inequalities* in Chapter 10.

already not a large technological challenge, the main focus of the current effort in the field is on decreasing the single-photon dephasing in long electromagnetic-wave transmission channels, with sufficiently high qubit transfer fidelity. The recent progress was rather impressive, with the demonstrated transfer of entangled qubits over landlines longer than 100 km,[99] and over at least one satellite-based line longer than 1,000 km;[100] and also a whole quantum key distribution over a comparable distance, though at a very low rate yet.[101, 102]

As an alternative, several classical encryption algorithms that are currently deemed quantum-resistant have already been developed[103] for a near-future replacement of the RSA. This is why the first practical applications of the long-range quantum information transfer are perhaps still to be found.[104]

8.6. Exercise problems

8.1. Prove that Eq. (30) indeed yields $E_g^{(1)} = (5/4)E_H$.

8.2. For a dilute gas of helium atoms in their ground state, with n atoms per unit volume, calculate its weak-field

(i) electric susceptibility χ_e, and
(ii) magnetic susceptibility χ_m,

and compare the results.

Hint: You may use the results of the variational description of the helium atom's ground state in Sec. 2, and the model solutions of Problems 6.8 and 6.15.

8.3. Calculate the expectation values of the observables $\mathbf{s}_1 \cdot \mathbf{s}_2$, $S^2 \equiv (\mathbf{s}_1 + \mathbf{s}_2)^2$, and $S_z \equiv s_{1z} + s_{2z}$, for the singlet and triplet states of the system of two spins-½, directly – without using the general Eq. (48). Compare the results with those for the system of two classical geometric vectors of length $\hbar/2$ each.

8.4. Discuss the factors $\pm 1/\sqrt{2}$ that participate in Eqs. (18) and (20) for the entangled states of the system of two spins-½, in terms of Clebsh-Gordan coefficients similar to those discussed in Sec. 5.7.

8.5.* Use the perturbation theory to calculate the so-called *hyperfine splitting* of the ground energy of the hydrogen atom,[105] due to the interaction between the spins of its nucleus (proton) and electron.

---

[99] See, e.g., T. Herbst *et al*., *Proc. Natl. Acad. Sci.* **112**, 14202 (2015), and references therein.
[100] J. Yin *et al*., *Science* **356**, 1140 (2017).
[101] H.-L. Yin *et al*., *Phys. Rev. Lett*. **117**, 190501 (2016).
[102] A comprehensive review of the quantum cryptography work was recently given by S. Pirandola *et al*., *Adv. Opt. Photon*. **12**, 1012 (2020).
[103] See, e.g., https://en.wikipedia.org/wiki/NIST_Post-Quantum_Cryptography_Standardization.
[104] For the progress of research in this direction, adequate measures of the informational capacity of quantum channels are clearly necessary. Due to lack of time/space, I have to refer the interested reader to literature – for example, to Chapter 12 in Nielsen and Chuang, or to S. Barnett, *Quantum Information*, Oxford, 2009.
[105] This effect was discovered by A. Michelson in 1881 and explained theoretically by W. Pauli in 1924, with the first quantitative calculation made in 1930 by E. Fermi.

Hint: The proton's magnetic moment operator is described by the same Eq. (4.115) as the electron, but with a positive gyromagnetic ratio $\gamma_p = g_p e/2m_p \approx 2.675\times10^8$ s^{-1}T^{-1}, whose magnitude is much smaller than that of the electron ($|\gamma_e| \approx 1.761\times10^{11}$ s^{-1}T^{-1}), due to the much higher mass, $m_p \approx 1.673\times10^{-27}$ kg $\approx 1{,}835\, m_e$. (The g-factor of the proton is also different, $g_p \approx 5.586$.[106])

8.6. In the simple case of just two similar spin-interacting particles, distinguishable by their spatial location, the famous *Heisenberg model* of ferromagnetism[107] is reduced to the following Hamiltonian:

$$H = -J\,\hat{\mathbf{s}}_1 \cdot \hat{\mathbf{s}}_2 - \gamma\mathscr{B} \cdot (\hat{\mathbf{s}}_1 + \hat{\mathbf{s}}_2),$$

where J is the spin interaction constant, γ is the gyromagnetic ratio of each particle, and $\mathscr{B}$ is the external magnetic field. Find the stationary states and energies of this system for spin-½ particles.

8.7. Two spins-½, different gyromagnetic ratios γ_1 and γ_2, are placed in an external magnetic field $\mathscr{B}$. In addition, the spins interact as in the Heisenberg model:

$$\hat{H}_{\text{int}} = -J\,\hat{\mathbf{s}}_1 \cdot \hat{\mathbf{s}}_2 .$$

Find the stationary states and energies of the system.

8.8. Two similar spins-½ with a gyromagnetic ratio γ, localized at two points separated by distance a, interact via the field of their magnetic dipole moments. Calculate the stationary states and energies of the system.

8.9. Consider the permutation of two identical particles, each of spin s. How many different symmetric and antisymmetric spin states can the system have?

8.10. For a system of two identical particles with $s = 1$:

(i) List all spin states forming the uncoupled-representation basis.

(ii) List all possible pairs $\{S, M_S\}$ of the quantum numbers describing the states of the coupled-representation basis – see Eq. (48).

(iii) Which of the $\{S, M_S\}$ pairs describe the states symmetric, and which the states antisymmetric, with respect to the particle permutation?

8.11. Represent the operators of the total kinetic energy and the total orbital angular momentum of a system of two particles, with masses m_1 and m_2, as combinations of the terms describing the center-of-mass motion and the relative motion. Use the results to calculate the energy spectrum of the so-called *positronium* – a metastable "atom"[108] consisting of one electron and its positively charged antiparticle, the positron.

---

[106] The relatively large value of the proton's g-factor results from the quark-gluon structure of this particle. (An exact calculation of g_p remains a challenge for quantum chromodynamics.)

[107] It was suggested in 1926 independently by W. Heisenberg and P. Dirac. A discussion of thermal effects in this and other similar systems (especially the Ising model of ferromagnetism) may be found in SM Chapter 4.

[108] Its lifetime (either 0.124 ns or 138 ns, depending on the parallel or antiparallel configuration of the component spins), is limited by the weak interaction of its components, which causes their annihilation with the emission of several gamma-ray photons.

8.12. Calculate the energy spectrum of the system of two identical spin-½ particles moving along the *x*-axis, which is described by the following Hamiltonian:

$$\hat{H} = \frac{\hat{p}_1^2}{2m_0} + \frac{\hat{p}_2^2}{2m_0} + \frac{m_0\omega_0^2}{2}\left(x_1^2 + x_2^2 + \varepsilon x_1 x_2\right),$$

and the degeneracy of each energy level.

8.13. Two particles with similar masses m and charges q are free to move along a planar circle of radius R. In the limit of very strong Coulomb interaction of the particles, find the lowest eigenenergies of the system, and sketch the system of its energy levels. Discuss possible effects of particle indistinguishability.

8.14. Low-energy spectra of many diatomic molecules may be well described by modeling the molecule as a system of two particles connected with a light and elastic but very stiff spring. Calculate the energy spectrum of a molecule within this model. Discuss possible effects of nuclear spins on spectra of the so-called *homonuclear* diatomic molecules formed by two similar atoms.

8.15. Two indistinguishable spin-½ particles are attracting each other at contact:

$$U(x_1, x_2) = -\mathcal{W}\delta(x_1 - x_2), \qquad \text{with } \mathcal{W} > 0,$$

but are otherwise free to move along the *x*-axis. Find the energy and the orbital wavefunction of the ground state of the system.

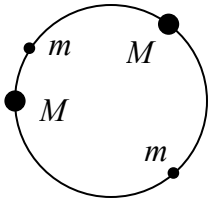
8.16. Two indistinguishable spin-½ particles are confined to move around a circle of radius R, and interact only at a very short arc distance $l = R(\varphi_1 - \varphi_2) \equiv R\varphi$ between them, so the interaction potential U may be well approximated with a delta function of φ. Find the ground state and its energy, for the cases of:

(i) the orbital (spin-independent) repulsion: $\hat{U} = \mathcal{W}\delta(\varphi)$,

(ii) the spin-spin interaction: $\hat{U} = -\mathcal{W}\,\hat{\mathbf{s}}_1 \cdot \hat{\mathbf{s}}_2\delta(\varphi)$,

both with $\mathcal{W} > 0$. Analyze the trends of your results in the limits $\mathcal{W} \to 0$ and $\mathcal{W} \to \infty$.

8.17. Two particles of mass M, separated by two much lighter particles of mass $m << M$, are placed on a circle of radius R – see the figure on the right. The particles fully repulse each other at contact, but otherwise, each of them is free to move along the circle. Calculate the lower part of the system's energy spectrum.



8.18. Two spin-½ particles are confined in a spherically symmetric potential well $U_0(\mathbf{r}) = m\omega_0^2 r^2/2$. In addition, they directly interact via a short-range potential that may be described as $U_{\text{int}} = \mathcal{W}\delta(\mathbf{r}_1 - \mathbf{r}_2)$. In the first approximation in small $\mathcal{W}$, calculate the energies of

(i) the ground state, and

(ii) the lowest excited states of the system.

8.19. N indistinguishable spin-½ particles are placed into the spherically-symmetric potential well $U(\mathbf{r}) = m\omega_0^2 r^2/2$. Neglecting the explicit interaction of the particles, find the ground-state energy of the system.

8.20. Use the Hund rules to evaluate the quantum numbers L, S, and J in the ground states of carbon and nitrogen atoms. Write down the Russell-Saunders symbols for these states.

8.21. $N >> 1$ indistinguishable, non-interacting quantum particles are placed into a hard-wall rectangular box with sides a_x, a_y, and a_z. Calculate the ground-state energy of the system and the average forces it exerts on each face of the box. Can we characterize the forces by certain pressure $\mathcal{P}$?

Hint: Consider separately the cases of bosons and fermions.

8.22. A system of three spins-½ is described by the Heisenberg Hamiltonian

$$\hat{H} = -J\left(\hat{\mathbf{s}}_1 \cdot \hat{\mathbf{s}}_2 + \hat{\mathbf{s}}_2 \cdot \hat{\mathbf{s}}_3 + \hat{\mathbf{s}}_3 \cdot \hat{\mathbf{s}}_1\right),$$

where J is a spin interaction constant (cf. Problems 6 and 7). Find the stationary states and energies of this system, and give an interpretation of your results.

8.23. For a system of three spins-½, find the common eigenstates and eigenvalues of the operators $\hat{S}_z$ and $\hat{S}^2$, where $\hat{\mathbf{S}} \equiv \hat{\mathbf{s}}_1 + \hat{\mathbf{s}}_2 + \hat{\mathbf{s}}_3$ is the vector operator of the total spin of the system. Do the corresponding quantum numbers S and M_S obey Eqs. (48)?

8.24. Explore basic properties of the Heisenberg model (whose few-spin versions were the subjects of Problems 6, 7, and 23), for a 1D chain of N spins-½:

$$\hat{H} = -J\sum_{\{j,j'\}} \hat{\mathbf{s}}_j \cdot \hat{\mathbf{s}}_{j'} - \gamma\mathcal{B} \cdot \sum_j \hat{\mathbf{s}}_j, \quad \text{with } J > 0,$$

where the summation is over all N spins, with the symbol $\{j, j'\}$ meaning that the first sum is only over the adjacent spin pairs. In particular, find the ground state of the system and its lowest excited states in the absence of external magnetic field $\mathcal{B}$, and also the dependence of their energies on the field.

Hint: For the sake of simplicity, you may assume that the first sum includes the term $\hat{\mathbf{s}}_N \cdot \hat{\mathbf{s}}_1$ as well. (Physically, this means that the chain is bent into a closed loop. [109])

8.25. Calculate commutators of the following operators:

$$\hat{\sigma}_+ \equiv \hat{a}_1^\dagger \hat{a}_2, \qquad \hat{\sigma}_- \equiv \hat{a}_2^\dagger \hat{a}_1, \qquad \hat{\sigma}_z \equiv \frac{1}{2}\left(\hat{a}_1^\dagger \hat{a}_1 - \hat{a}_2^\dagger \hat{a}_2\right),$$

where $\hat{a}_{1,2}^\dagger$ and $\hat{a}_{1,2}$ are the operators of the creation and annihilation of bosons in two different states.

[109] Note that for dissipative spin systems, differences between low-energy excitations of open-end and closed-end 1D chains may be substantial even in the limit $N \to \infty$ – see, e.g., SM Sec. 4.5. However, for our Hamiltonian (and hence dissipation-free) system, the differences are relatively small.

8.26. Compose the simplest model Hamiltonians, in terms of the second quantization formalism, for systems of indistinguishable particles placed in the following external potentials:

(i) two weakly coupled potential wells, with on-site particle interactions giving additional energy J per each pair of particles in the same potential well, and

(ii) a periodic 1D potential, with the same particle interactions, in the tight-binding limit.

8.27. For each of the Hamiltonians composed in the previous problem, derive the Heisenberg equations of motion for particle creation/annihilation operators.

8.28. Express the ket-vectors of all possible Dirac states of three indistinguishable

(i) bosons, and

(ii) fermions,

via those of the single-particle states β, β', and β'' they occupy.

8.29. Explain why the general perturbative result (8.126), when applied to the ^{4}He atom, gives the correct[110] expression (8.29) for the ground singlet state, and correct Eqs. (8.39)-(8.42) (with the minus sign in the first of these relations) for the excited triplet states, but cannot describe these results, with the plus sign in Eq. (8.39), for the excited singlet state.

8.30.* Explore the *Thomas-Fermi model*[111] of a heavy atom, with the nuclear charge $Q = Ze >> e$, in which the interaction between electrons is limited to their contribution to the common electrostatic potential $\phi(\mathbf{r})$. In particular, derive the ordinary differential equation obeyed by the radial distribution of the potential, and use it to estimate the effective radius of the atom.

8.31.* Use the Thomas-Fermi model explored in the previous problem to calculate the total binding energy of a heavy atom. Compare the result with that of the simpler model, in that the Coulomb electron-electron interaction is completely ignored.

8.32. Suggest and explore a simple model of dephasing in a system consisting of N similar, distinct, non-interacting components. How does the dephasing time scale with N?

8.33. The notion of the reduced density operator $\hat{w}$ defined by Eq. (7.160) is sometimes used for the characterization of entanglement in multi-qubit systems. Calculate $\hat{w}$ for one qubit of a two-qubit system that is in an arbitrary pure state, and analyze the result.

8.34. For a system of two distinct qubits (i.e. two-level systems), introduce a reasonable uncoupled-representation z-basis and write, in this basis, the 4×4 matrix of the operator that swaps their states.

8.35. Find a time-independent Hamiltonian that causes the qubit evolution described by Eqs. (155). Discuss the relation between your result and the time-dependent Hamiltonian (6.86).

[110] Correct in the sense of the first order of the perturbation theory.

[111] It was suggested in 1927, independently, by L. Thomas and E. Fermi.

Chapter 9. Introduction to Relativistic Quantum Mechanics

This chapter consists of two very different parts. Its first part is a discussion of the basic elements of the quantum theory of the electromagnetic field, usually called quantum electrodynamics (QED). We will see, in particular, that the QED may be viewed as the relativistic quantum theory of particles with zero rest mass – photons. The second part of the chapter is a brief review of the relativistic quantum theory of particles with non-zero rest mass, including the famous Dirac's theory of spin-½ particles such as electrons. These theories mark the point of entry into a more complete relativistic quantum theory – the quantum field theory – which is beyond the scope of this course.

9.1. Electromagnetic field quantization[1]

Classical physics gives us[2] the following general relativistic relation between the momentum **p** and energy E of a free particle with rest mass m, which may be simplified in two limits – non-relativistic and ultra-relativistic:

Free particle's relativistic energy

$$E = \left[(pc)^2 + (mc^2)^2\right]^{1/2} \to \begin{cases} mc^2 + p^2/2m, & \text{for } p << mc, \\ pc, & \text{for } p >> mc. \end{cases} \tag{9.1}$$

In both limits, the transfer from classical to quantum mechanics is easier than in the arbitrary case. Since all the previous contents of this course were committed to the first, non-relativistic limit, I will now jump to a brief discussion of the ultra-relativistic limit $p >> mc$, for a particular but very important system – the electromagnetic field. Since the excitations of this field, called *photons*, are currently believed to have zero rest mass m,[3] the ultra-relativistic relation $E = pc$ is exactly valid for any photon energy E, and the quantization scheme is rather straightforward.

As usual, the quantization has to be based on the classical theory of the system – in this case, the Maxwell equations. As the simplest case, let us consider the electromagnetic field inside a finite free-space volume limited by ideal walls, which reflect incident waves perfectly.[4] Inside the volume, the Maxwell equations give a simple wave equation[5] for the electric field

$$\nabla^2 \mathscr{E} - \frac{1}{c^2}\frac{\partial^2 \mathscr{E}}{\partial t^2} = 0, \tag{9.2}$$

and an absolutely similar equation for the magnetic field $\mathscr{B}$. We may look for the general solution of Eq. (2) in the variable-separating form

---

[1] The described approach was pioneered by the same P. A. M. Dirac as early as 1927.

[2] See, e.g., EM Chapter 9, in particular Eq. (9.78).

[3] By now, this fact has been verified experimentally with an accuracy of at least 10^{-18} eV/c^2, i.e. ~10^{-24} m_e – see, e.g. the review by C. Amsler *et al. (Particle Data Group), Phys. Lett. B* **667**, 1 (2008).

[4] In the case of finite energy absorption in the walls, or in the wave propagation media, the system is not energy-conserving (Hamiltonian), i.e. interacts with a dissipative environment as was discussed in Chapter 7. Specific cases of such interaction will be considered in Sections 2 and 3 below.

[5] See, e.g., EM Eq. (7.3), for the particular case $\varepsilon = \varepsilon_0$, $\mu = \mu_0$, so $v^2 \equiv 1/\varepsilon\mu = 1/\varepsilon_0\mu_0 \equiv c^2$.

© K. Likharev

$$\mathscr{E}(\mathbf{r},t)=\sum_j \mathscr{E}_j(\mathbf{r},t), \quad \text{with } \mathscr{E}_j(\mathbf{r},t)=p_j(t)\mathbf{e}_j(\mathbf{r}). \tag{9.3}$$

Physically, each term of this sum is a standing wave whose spatial distribution and polarization ("*mode*") are described by the vector function $\mathbf{e}_j(\mathbf{r})$, while its temporal dynamics follows the function $p_j(t)$. Plugging an arbitrary term of this sum into Eq. (2), and separating the variables exactly as we did, for example, in the Schrödinger equation in Sec. 1.5, we get

$$\frac{\nabla^2\mathbf{e}_j}{\mathbf{e}_j}=\frac{1}{c^2}\frac{\ddot{p}_j}{p_j}=\text{const}\equiv -k_j^2, \tag{9.4}$$

so the spatial distribution of the mode satisfies the following 3D Helmholtz equation:

$$\boxed{\nabla^2\mathbf{e}_j+k_j^2\mathbf{e}_j=0.} \tag{9.5}$$

Equation for spatial distribution

The set of solutions of this equation, with appropriate boundary conditions, determines the set of the functions $\mathbf{e}_j$, and simultaneously the spectrum of the wave number magnitudes k_j. The latter values determine the mode eigenfrequencies, following from Eq. (4):

$$\ddot{p}_j+\omega_j^2 p_j=0, \quad \text{with } \omega_j\equiv k_j c. \tag{9.6}$$

There is a big philosophical difference between the quantum-mechanical approach to Eqs. (5) and (6), despite their single origin (4). The first (Helmholtz) equation may be rather difficult to solve in realistic geometries,[6] but it remains intact in the basic quantum electrodynamics, with the scalar components of the vector functions $\mathbf{e}_j(\mathbf{r})$ still treated (at each point $\mathbf{r}$) as *c*-numbers. In contrast, the classical Eq. (6) is readily solvable (giving sinusoidal oscillations with frequency ω_j), but this is exactly where we can make the transfer to quantum mechanics because we already know how to quantize a mechanical 1D harmonic oscillator, which in classics obeys the same equation.

As usual, we need to start with the appropriate Hamiltonian – the operator corresponding to the classical Hamiltonian function H of the proper set of generalized coordinates and momenta. The electromagnetic field's Hamiltonian function (which in this case coincides with the field's energy) is[7]

$$H=\int d^3r\left(\frac{\varepsilon_0\mathscr{E}^2}{2}+\frac{\mathscr{B}^2}{2\mu_0}\right). \tag{9.7}$$

Let us represent the magnetic field in a form similar to Eq. (3),[8]

$$\mathscr{B}(\mathbf{r},t)=\sum_j \mathscr{B}_j(\mathbf{r},t), \quad \text{with } \mathscr{B}_j(\mathbf{r},t)=-\omega_j q_j(t)\mathbf{b}_j(\mathbf{r}). \tag{9.8}$$

[6] See, e.g., the cases discussed in EM Sec. 7.9.

[7] See, e.g., EM Sec. 9.8, in particular, Eq. (9.225). Here I am using SI units, with $\varepsilon_0\mu_0\equiv c^{-2}$; in the Gaussian units, the coefficients ε_0 and μ_0 disappear, but there is an additional common factor $1/4\pi$ in the formula H. However, if we modify the normalization conditions (see below) accordingly, all the subsequent results, starting from Eq. (10), look similar in any system of units.

[8] Here I am using the letter q_j, instead of x_j, for the generalized coordinate of the field oscillator, in order to emphasize the difference between the former variable and one of the Cartesian coordinates, i.e. one of the arguments of the *c*-number functions $\mathbf{e}_j(\mathbf{r})$ and $\mathbf{b}_j(\mathbf{r})$.

Since, according to the Maxwell equations, in our case, the magnetic field satisfies the equation similar to Eq. (2), the time-dependent amplitude q_j of each of its modes $\mathbf{b}_j(\mathbf{r})$ obeys an equation similar to Eq. (6), i.e. in the classical theory also changes in time sinusoidally, with the same frequency ω_j. Plugging Eqs. (3) and (8) into Eq. (7), we may recast it as

$$H = \sum_j \left[\frac{p_j^2}{2} \int \varepsilon_0 e_j^2(\mathbf{r}) d^3 r + \frac{\omega_j^2 q_j^2}{2} \int \frac{1}{\mu_0} b_j^2(\mathbf{r}) d^3 r \right]. \tag{9.9}$$

Now note that the distribution of constant factors between two operands in each product of Eq. (3) is so far arbitrary, so we may fix it by requiring the first integral in Eq. (9) to equal 1. Since according to the Maxwell equations, there is a specific relation between the field vector amplitudes, $\mathscr{B}_j/\mathscr{E}_j = (\varepsilon_0\mu_0)^{1/2} \equiv 1/c$,[9] this normalization makes the second integral in Eq. (9) equal 1 as well, and Eq. (9) becomes

$$H = \sum_j H_j, \qquad H_j = \frac{p_j^2}{2} + \frac{\omega_j^2 q_j^2}{2}. \tag{9.10}$$

Note that p_j is the legitimate generalized momentum corresponding to the generalized coordinate q_j, because it is equal to $\partial L / \partial \dot{q}_j$, where L is the Lagrangian function of the field – see EM Eq. (9.217):

$$L = \int d^3 r \left(\frac{\varepsilon_0 \mathscr{E}^2}{2} - \frac{\mathscr{B}^2}{2\mu_0} \right) = \sum_j L_j, \qquad L_j = \frac{p_j^2}{2} - \frac{\omega_j^2 q_j^2}{2}. \tag{9.11}$$

Hence we can carry out the standard quantization procedure, namely declare H_j, p_j, and q_j to be quantum-mechanical operators related as in Eq. (10),

Electro-magnetic mode's Hamiltonian

$$\hat{H}_j = \frac{\hat{p}_j^2}{2} + \frac{\omega_j^2 \hat{q}_j^2}{2}. \tag{9.12}$$

We see that this Hamiltonian coincides with that of a 1D harmonic oscillator with the mass m_j formally equal to 1,[10] and the frequency equal to ω_j. Moreover, since the Lagrangian function L has the same form (11) as in a set of independent harmonic oscillators with coordinates q_j and momenta p_j, the corresponding operators satisfy the following natural generalization of the commutation relations (2.14):

$$\left[\hat{q}_j, \hat{p}_{j'}\right] = i\hbar \delta_{jj'}, \tag{9.13}$$

As the reader already knows, Eqs. (12) and (13) open for us several alternative ways to proceed:

(i) Use the Schrödinger-picture wave mechanics based on wavefunctions $\Psi_j(q_j, t)$. As we know from Sec. 2.9, this way is inconvenient for most tasks, because the eigenfunctions of the harmonic oscillator are rather clumsy.

(ii) A substantially more efficient way is to write and use the Heisenberg-picture equations for the time evolution of the operators $\hat{q}_j(t)$ and $\hat{p}_j(t)$.

[9] See, e.g., EM Eqs. (7.6)-(7.7) with $\varepsilon = \varepsilon_0$, $\mu = \mu_0$, and $\mathscr{H} = \mathscr{B}/\mu_0$.

[10] Selecting a different normalization of the function $\mathbf{e}_j(\mathbf{r})$, we could readily arrange any value of m_j, so the choice corresponding to $m_j = 1$ is the best one just for the notation simplicity.

(iii) An even more convenient approach is to use equations similar to Eqs. (5.65) to decompose the Heisenberg operators $\hat{q}_j(t)$ and $\hat{p}_j(t)$ into the creation-annihilation operators $\hat{a}_j^\dagger(t)$ and $\hat{a}_j(t)$.

In this chapter, I will mostly use the last route. After the replacement of m with $m_j \equiv 1$, and ω_0 with ω_j, the last forms of Eqs. (5.65) become

$$\hat{a}_j = \left(\frac{\omega_j}{2\hbar}\right)^{1/2}\left(\hat{q}_j + i\frac{\hat{p}_j}{\omega_j}\right), \qquad \hat{a}_j^\dagger = \left(\frac{\omega_j}{2\hbar}\right)^{1/2}\left(\hat{q}_j - i\frac{\hat{p}_j}{\omega_j}\right). \tag{9.14}$$

Due to Eq. (13), these creation-annihilation operators obey the commutation similar to Eq. (5.68),

$$\left[\hat{a}_j, \hat{a}_{j'}^\dagger\right] = \hat{I}\delta_{jj'}. \tag{9.15}$$

As a result, according to Eqs. (3) and (8), the quantum-mechanical operators of the electric and magnetic fields may be represented as sums over all *field oscillators*:

$$\hat{\mathscr{E}}(\mathbf{r},t) = i\sum_j \left(\frac{\hbar\omega_j}{2}\right)^{1/2} \mathbf{e}_j(\mathbf{r})\left(\hat{a}_j^\dagger - \hat{a}_j\right), \tag{9.16a}$$

$$\hat{\mathscr{B}}(\mathbf{r},t) = \sum_j \left(\frac{\hbar\omega_j}{2}\right)^{1/2} \mathbf{b}_j(\mathbf{r})\left(\hat{a}_j^\dagger + \hat{a}_j\right), \tag{9.16b}$$

Electro-magnetic field operators

and Eq. (12) for the j^{th} mode's Hamiltonian becomes

$$\hat{H}_j = \hbar\omega_j\left(\hat{a}_j^\dagger\hat{a}_j + \frac{1}{2}\hat{I}\right) = \hbar\omega_j\left(\hat{n}_j + \frac{1}{2}\hat{I}\right), \quad \text{with } \hat{n}_j \equiv \hat{a}_j^\dagger\hat{a}_j, \tag{9.17}$$

absolutely similar to Eq. (5.72) for a mechanical oscillator.

Now comes a very important conceptual step. From Sec. 5.4, we know that the stationary (Fock) states n_j of the Hamiltonian (17) have energies

$$E_j = \hbar\omega_j\left(n_j + \frac{1}{2}\right), \quad n_j = 0,1,2,... \tag{9.18}$$

Electro-magnetic mode eigen-energies

and, according to Eq. (5.89), the operators $\hat{a}_j^\dagger$ and $\hat{a}_j$ act on the eigenkets of these partial states as

$$\hat{a}_j|n_j\rangle = (n_j)^{1/2}|n_j - 1\rangle, \qquad \hat{a}_j^\dagger|n_j\rangle = (n_j+1)^{1/2}|n_j+1\rangle, \tag{9.19}$$

regardless of the quantum states of other modes. These rules coincide with the definitions (8.64) and (8.68) of bosonic creation-annihilation operators, and hence their action may be considered as the creation/annihilation of certain bosons. Such a *quasiparticle* (actually, an *excitation*, with energy $\hbar\omega_j$, of the j^{th} field oscillator) is exactly what is, strictly speaking, called a *photon*. Note immediately that according to Eq. (16), such an excitation does not change the spatial distribution of the j^{th} mode of the field. So, such a "global" photon is an excitation created simultaneously at all points of the field confinement region.

If this picture is too contrary to the intuitive image of a particle, please recall that in Chapter 2, we discussed a similar situation with the fundamental solutions $\exp\{\pm ikx\}$ of the Schrödinger equation of a free non-relativistic particle: they represent sinusoidal de Broglie waves existing simultaneously at

all points of the particle confinement region. The (partial :-) reconciliation with the classical picture of a moving particle might be obtained by using the linear superposition principle to assemble a quasi-localized wave packet, as a group of sinusoidal waves with close wave numbers. Very similarly, we may form a similar wave packet using a linear superposition of the "global" photons with close values of $\mathbf{k}_j$ (and hence ω_j), to form a quasi-localized photon. An additional simplification here is that the dispersion relation for electromagnetic waves (at least in free space) is linear:

$$\frac{\partial \omega_j}{\partial k_j} = c = \text{const}, \qquad \text{i.e. } \frac{\partial^2 \omega_j}{\partial k_j{}^2} = 0, \tag{9.20}$$

so according to Eq. (2.39a), the electromagnetic wave packets (i.e. space-localized photons) do not spread out during their propagation. Note also that due to the fundamental classical relations $\mathbf{p} = \mathbf{n}E/c$ for the linear momentum of the traveling electromagnetic wave packet of energy E, propagating along the direction $\mathbf{n} \equiv \mathbf{k}/k$, and $\mathbf{S} = \pm\mathbf{n}E/\omega_j$ for its angular momentum,[11] such a photon may be prescribed the linear momentum $\mathbf{p} = \mathbf{n}\hbar\omega_j/c \equiv \hbar\mathbf{k}$ and the angular momentum (essentially, the spin) $\mathbf{S} = \pm\mathbf{n}\hbar$, with the sign depending on the direction of its circular polarization ("helicity").

This electromagnetic field quantization scheme should look very straightforward, but it raises an important conceptual issue of ground state energy. Indeed, Eq. (18) implies that the total ground-state (i.e., the lowest) energy of the field is

Ground-state energy of EM field

$$E_g = \sum_j (E_g)_j = \sum_j \frac{\hbar\omega_j}{2}. \tag{9.21}$$

Since for any realistic model of the field-confining volume, either infinite or not, the density of electromagnetic field modes only grows with frequency,[12] this sum diverges on its upper limit, leading to infinite ground-state energy per unit volume. This infinite-energy paradox cannot be dismissed by declaring the ground-state energy of field oscillators unobservable, because this would contradict numerous experimental observations – starting perhaps from the famous *Casimir effect*.[13]

The simplest implementation of this effect involves two parallel, perfectly conducting plates of area A, separated by a vacuum gap of thickness $t << A^{1/2}$ (Fig. 1).

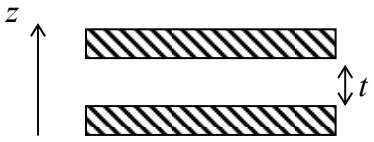


Fig. 9.1. The simplest geometry of the Casimir effect manifestation.

[11] See, e.g., EM Sections 7.7 and 9.8 (where the angular momentum of the field is denoted **L**).

[12] See, e.g., Eq. (1.1), which is similar to Eq. (1.90) for the de Broglie waves, derived in Sec. 1.7.

[13] This effect was predicted in 1948 by Hendrik Casimir and Dirk Polder, and confirmed semi-quantitatively in experiments by M. Sparnaay, *Nature* **180**, 334 (1957). After that and several other experiments, a decisive error bar reduction (to about ~5%), providing a quantitative confirmation of the Casimir formula (23), was achieved by S. Lamoreaux, *Phys. Rev. Lett.* **78**, 5 (1997) and by U. Mohideen and A. Roy, *Phys. Rev. Lett.* **81**, 004549 (1998). Note also that there are other experimental confirmations of the reality of the ground-state electromagnetic field, including, for example, the experiments by R. Koch *et al*. already discussed in Sec. 7.5, and the recent spectacular direct observations by C. Riek *et al*., *Science* **350**, 420 (2015).

Rather counterintuitively, the plates attract each other with a force F proportional to the area A and rapidly increasing with the decrease of t, even in the absence of any explicit electromagnetic field sources. The effect's explanation is that the energy of each electromagnetic field mode, including its ground-state energy, exerts average pressure,

$$\left\langle \mathcal{P}_j \right\rangle = -\frac{\partial E_j}{\partial V}, \tag{9.22}$$

on the walls constraining it to volume V. While the field's pressure on the external surfaces on the plates is due to the contributions (22) of all free-space modes, with arbitrary values of k_z (the z-component of the wave vector $\mathbf{k}_j$), in the gap between the plates, the spectrum of k_z is limited to the multiples of π/t, so the pressure on the internal surfaces is lower. This is why the net force exerted on the plates may be calculated as the sum of the contributions (22) from all "missing" low-frequency modes in the gap, with the minus sign. In the simplest model when the plates are made of an ideal conductor, which provides simple boundary conditions $\mathscr{E}_\tau = \mathscr{B}_n = 0$ on their surfaces,[14] such calculation is quite straightforward (and is hence left for the reader's exercise), and its result is

$$F = -\frac{\pi^2 A\hbar c}{240 t^4}. \tag{9.23}$$

Casimir effect

Note that for such calculation, the high-frequency divergence of Eq. (21) is not important, because it participates in the forces exerted on all surfaces of each plate, and cancels out from the net pressure. In this way, the Casimir effect not only confirms Eq. (21), but also teaches us an important lesson on how to deal with the divergences of such sums at $\omega_j \to \infty$. The lesson is: just get accustomed to the idea that the divergence exists, and ignore this fact while you can, i.e. if the final result you are interested in is finite. However, for some more complex problems of quantum electrodynamics (and the quantum theory of any other fields), this simplest approach becomes impossible, and then more complex, *renormalization* techniques become necessary. For their study, I have to refer the reader to a quantum field theory course – see the references at the end of this chapter.

9.2. Photon absorption and counting

As a matter of principle, the Casimir effect may be used to observe not only the electromagnetic field's ground state but also the field arriving from active sources – lasers, etc. However, usually, such studies may be done by simpler detectors, in which the absorption of a photon by a single atom leads to its ionization. This ionization, i.e. the emission of a free electron, triggers an avalanche-like chain reaction (e.g., an electric discharge in a Geiger-type counter), which may be readily registered using

[14] For realistic conductors, the reduction of t below ~1 μm causes significant deviations from this simple model, and hence from Eq. (23). The reason is that for gaps so narrow, the depth of field penetration into the conductors (see, e.g., EM Sec. 6.2), at the important frequencies $\omega \sim c/t$, becomes comparable with t, and an adequate theory of the Casimir effect has to involve a certain model of the penetration. (It is curious that in-depth analyses of this problem, pioneered in 1956 by E. Lifshitz, have revealed a deep relation between the Casimir effect and the London dispersion force which was the subject of Problems 3.16, 5.15, and 6.18 – for a review see, e.g., either I. Dzhyaloshinskii *et al.*, *Sov. Phys. Uspekhi* **4**, 153 (1961), or K. Milton, *The Casimir Effect*, World Scientific, 2001. Recent experiments in the 100 nm – 2 μm range of t, with an accuracy better than 1%, have enabled not only a clear observation of field penetration effects on the Casimir force but even a selection between some approximate models of the penetration – see D. Garcia-Sanchez *et al.*, *Phys. Rev. Lett.* **109**, 027202 (2012).

appropriate electronic circuitry. In good photon counters, the first step, the "trigger" atom ionization, is the bottleneck of the whole process (the *photon count*), so to analyze their statistics, it is sufficient to consider the field's interaction with just this atom.

Its ionization is a quantum transition from a discrete initial state of the atom to its final, ionized state with a continuous energy spectrum, induced by an external electromagnetic field. This is exactly the situation shown in Fig. 6.12, so we may apply to it the Golden Rule of quantum mechanics in the form (6.149), with system *a* associated with the electromagnetic field, and system *b* with the trigger atom. The atom's size is typically much smaller than the radiation wavelength, so the field-atom interaction may be adequately described in the electric-dipole approximation (6.146)

$$\hat{H}_{\rm int} = -\hat{\mathscr{E}} \cdot \hat{\mathbf{d}}, \tag{9.24}$$

where $\hat{\mathbf{d}}$ is the dipole moment's operator. Hence we may associate this operator with the operand $\hat{B}$ in Eqs. (6.145)-(6.149), while the electric field operator $\hat{\mathscr{E}}$ is associated with the operand $\hat{A}$ in those relations. First, let us assume that our field consists of only one mode of frequency ω. Then we can keep only one term in the sum (16a), and drop the index j, so Eq. (6.149) may be rewritten as

$$\begin{aligned} \Gamma &= \frac{2\pi}{\hbar}\left|\left\langle \text{ fin}\left|\hat{\mathscr{E}}(\mathbf{r},t)\right|\text{ ini}\right\rangle\right|^2 \left|\left\langle \text{ fin}\left|\hat{\mathbf{d}}(t)\cdot\mathbf{n}_e\right|\text{ ini}\right\rangle\right|^2 \rho_{\rm a} \\ &= \frac{2\pi}{\hbar}\frac{\hbar\omega}{2}\left|\left\langle \text{ fin}\left|\left[\hat{a}^\dagger(t)-\hat{a}(t)\right]e(\mathbf{r})\right|\text{ ini}\right\rangle\right|^2 \left|\left\langle \text{ fin}\left|\hat{\mathbf{d}}(t)\cdot\mathbf{n}_e\right|\text{ ini}\right\rangle\right|^2 \rho_{\rm a}, \end{aligned} \tag{9.25}$$

where $\mathbf{n}_e \equiv \mathbf{e}(\mathbf{r})/e(\mathbf{r})$ is the local direction of the vector $\mathbf{e}(\mathbf{r})$, symbols "ini" and "fin" denote the initial and final states of the corresponding subsystem (the electromagnetic field in the first long bracket, and the atom in the second bracket), and the density $\rho_{\rm a}$ of the atomic state continuum should be calculated at the final energy $E_{\rm fin} = E_{\rm ini} + \hbar\omega$.

As a reminder, in the Heisenberg picture of quantum dynamics, the initial and final states are time-independent, while the creation-annihilation operators are functions of time. In the Golden Rule formula (25), as in any perturbative result, this time dependence has to be calculated ignoring the perturbation – in this case, the field-atom interaction. For the field's creation-annihilation operators, this dependence coincides with that of the usual 1D oscillator – see Eq. (5.141), in which ω_0 should be, in our current notation, replaced with ω:

$$\hat{a}(t) = \hat{a}(0)e^{-i\omega t}, \quad \hat{a}^\dagger(t) = \hat{a}^\dagger(0)e^{+i\omega t}. \tag{9.26}$$

Hence Eq. (25) becomes

$$\Gamma = \pi\omega\left|\left\langle \text{ fin}\left|\left[\hat{a}^\dagger(0)e^{i\omega t} - \hat{a}(0)e^{-i\omega t}\right]e(\mathbf{r})\right|\text{ ini}\right\rangle\right|^2 \left|\left\langle \text{ fin}\left|\hat{\mathbf{d}}(t)\cdot\mathbf{n}_e\right|\text{ ini}\right\rangle\right|^2 \rho_{\rm a}. \tag{9.27a}$$

Now let us multiply the first long bracket by $\exp\{i\omega t\}$, and the second one by $\exp\{-i\omega t\}$:

$$\Gamma = \pi\omega\left|\left\langle \text{ fin}\left|\left[\hat{a}^\dagger(0)e^{2i\omega t} - \hat{a}(0)\right]e(\mathbf{r})\right|\text{ ini}\right\rangle\right|^2 \left|\left\langle \text{ fin}\left|\hat{\mathbf{d}}(t)\cdot\mathbf{n}_e e^{-i\omega t}\right|\text{ ini}\right\rangle\right|^2 \rho_{\rm a}. \tag{9.27b}$$

This, mathematically equivalent form of the previous relation shows more clearly that at resonant photon absorption, only the *annihilation* operator gives a significant time-averaged contribution to the

first bracket matrix element. (As a reminder, the quantum-mechanical Golden Rule for time-dependent perturbations is a result of averaging over a time interval much larger than $1/\omega$ – see Sec. 6.6.) Similarly, according to Eq. (4.199), the Heisenberg operator of the dipole moment corresponding to the *increase* of the atom's energy by $\hbar\omega$, has the Fourier components that differ in frequency from ω only by $\sim\Gamma << \omega$, so its time dependence virtually compensates the additional factor in the second bracket of Eq. (27b), and this bracket also may have a substantial time average. Hence, in the first bracket, we may neglect the fast-oscillating term, whose average over time interval $\sim 1/\Gamma$ is very close to zero.[15]

Now let us assume, first, that we use the same detector, characterized by the same matrix element of the quantum transition, i.e. the same second bracket in Eq. (27), and the same final state density ρ_a, for measurement of various electromagnetic fields – or just of the same field at different points **r**. Then we are only interested in the behavior of the first, field-related bracket, and may write

$$\Gamma \propto \left|\left\langle \text{fin}\left|\hat{a}e(\mathbf{r})\right|\text{ini}\right\rangle\right|^2 \equiv \left\langle \text{fin}\left|\hat{a}e(\mathbf{r})\right|\text{ini}\right\rangle\left\langle \text{fin}\left|\hat{a}e(\mathbf{r})\right|\text{ini}\right\rangle^* \equiv \left\langle \text{ini}\left|\hat{a}^\dagger e^*(\mathbf{r})\right|\text{fin}\right\rangle\left\langle \text{fin}\left|\hat{a}e(\mathbf{r})\right|\text{ini}\right\rangle, \quad (9.28)$$

where the creation-annihilation operators are implied to be taken at $t = 0$, i.e. in the Schrödinger picture, and the initial and final states are those of the field alone. Second, let us now calculate the *total* rate of transitions to *all* available final states of the given mode **e**(**r**). If such states formed a full and orthonormal set, we could use the closure relation (4.44) applied to the final states to write

$$\Gamma \propto \sum_{\text{fin}} \left\langle \text{ini}\left|\hat{a}^\dagger e^*(\mathbf{r})\right|\text{fin}\right\rangle\left\langle \text{fin}\left|\hat{a}e(\mathbf{r})\right|\text{ini}\right\rangle = \left\langle \text{ini}\left|\hat{a}^\dagger\hat{a}\right|\text{ini}\right\rangle e^*(\mathbf{r})e(\mathbf{r}) = \langle n\rangle_{\text{ini}}\left|e(\mathbf{r})\right|^2, \quad (9.29)$$

Photon counting rate

where, for a given field mode, $\langle n\rangle_{\text{ini}}$ is the expectation value of the operator $\hat{n} \equiv \hat{a}^\dagger\hat{a}$ for the initial state of the electromagnetic field. In the more realistic case of fields in relatively large volumes, $V >> \lambda^3$, with their virtually continuous spectrum of the final states, the middle equality in this relation is not strictly valid, but it is correct to a constant multiplier,[16] which we are currently not interested in. Note, however, that Eq. (29) may be substantially wrong for high-Q electromagnetic resonators ("cavities"), which may make just one (or a few) modes available for transitions. (Quantum electrodynamics of such cavities will be briefly discussed in Sec. 4 below.)

Let us apply Eq. (29) to several possible quantum states of the mode.

(i) First, as a sanity check, the ground initial state, $n = 0$, gives no photon absorption at all. The interpretation is easy: the ground state field, cannot emit a photon that would ionize an atom in the counter. Again, this does not mean that the ground-state "motion" is not observable (if you still think so, please review the Casimir effect discussion in Sec. 1), just that it cannot ionize the trigger atom – because it does not have any *spare* energy for doing that.

(ii) All other coherent states (Fock, Glauber, squeezed, etc.) of the field oscillator give the same counting rate, provided that their $\langle n\rangle_{\text{ini}}$ is the same. This result may be less evident if we apply Eq. (29) to the interference of two light beams from the same source – say, in the double-slit or the Bragg-scattering configurations. In this case, we may represent the spatial distribution of the field as a sum

[15] This is essentially the same rotating wave approximation (RWA) that was already used in Sec. 6.5 and beyond – see, e.g., the transition from Eq. (6.90) to the first of Eqs. (6.94).

[16] As the Golden Rule shows, this multiplier is proportional to the density ρ_f of the final states of the field.

$$e(\mathbf{r}) = e_1(\mathbf{r}) + e_2(\mathbf{r}) . \tag{9.30}$$

Here each term describes one possible wave path, so the operator product in Eq. (29) may be a rapidly changing function of the detector position. For this configuration, our result (29) means that the interference pattern (and its contrast) are independent of the particular state of the electromagnetic field's mode.

(iii) The last statement is also valid for a classical mixture of the different eigenstates of the same field mode, for example for its thermal-equilibrium state. Indeed, in this case, we need to average Eq. (29) over the corresponding classical ensemble, but it would only result in a different meaning of averaging n in that equation; the field part describing the interference pattern is not affected.

The last result may look a bit counter-intuitive because common sense tells us that the stochasticity associated with thermal equilibrium has to suppress the interference pattern contrast. These expectations are (partly :-) justified because a typical thermal source of radiation produces many field modes j, rather than one mode we have analyzed. These modes may have different wave numbers k_j and hence different field distribution functions $\mathbf{e}_j(\mathbf{r})$, resulting in shifted interference patterns. Their summation would indeed smear the interference, suppressing its contrast.

So the use of one photon detector is not the best way to distinguish different quantum states of an electromagnetic field mode. This task, however, may be achieved using the photon counting correlation technique shown in Fig. 2.[17]

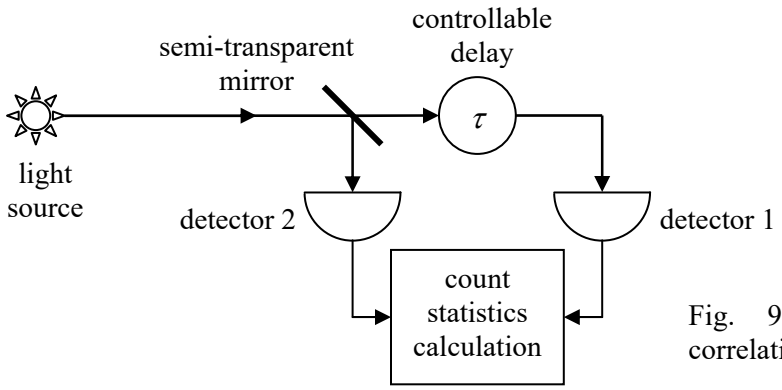


Fig. 9.2. Photon count correlation measurement.

In this experiment, the counting statistics may be characterized by the so-called *second-order correlation function* of the detector count rates,

Second-order correlation function

$$g^{(2)}(\tau) \equiv \frac{\langle \Gamma_1(t)\Gamma_2(t-\tau)\rangle}{\langle \Gamma_1(t)\rangle\langle \Gamma_2(t)\rangle}, \tag{9.31}$$

where the averaging may be carried out either over many similar experiments, or over a relatively long time interval $t >> \tau$, with usual field sources – due to their ergodicity.[18] Using the normalized correlation

[17] It was pioneered as early as the mid-1950s (i.e. before the advent of lasers), by Robert Hanbury Brown and Richard Twiss. Their second experiment was also remarkable for the rather unusual light source – the star Sirius! (Their work was an effort to improve astrophysics interferometry techniques.)

[18] This is why the argument t in Eq. (31) is just symbolic, and for the rates $\Gamma_{1,2}$, we can use Eq. (29), even though its derivation involved averaging over times much longer than $1/\omega$.

function (31) is very convenient because the characteristics of both detectors and the beam splitter (e.g., a semi-transparent mirror, see Fig. 2) drop out from this fraction.

Very unexpectedly for the mid-1950s, Hanbury Brown and Twiss discovered that the correlation function depends on the time delay τ in the way shown (schematically) with the solid line in Fig. 3. It is evident from Eq. (31) that if the counting events are completely independent, $g^{(2)}(\tau)$ should be equal to 1 – which is always the case in the limit $\tau \to \infty$. (As will be shown in the next section, the characteristic time of this approach is usually between 10^{-11}s and 10^{-8}s, so for its measurement, the delay time control may be provided just by moving one of the detectors by a human-scale distance between a few millimeters to a few meters.) Hence, the observed behavior at $\tau \to 0$ corresponds to a *positive* correlation of detector counts at small time delays, i.e. to a *higher* probability of the nearly simultaneous arrival of photons to both counters. This rather counterintuitive effect is called *photon bunching*.

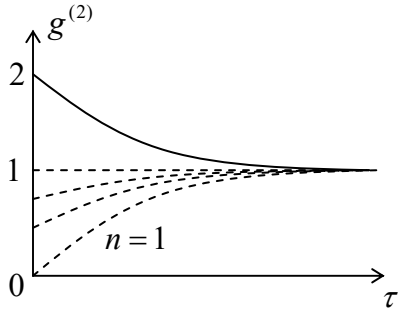


Fig. 9.3. Photon bunching (solid line) and antibunching for various n (dashed lines). The lines approach the level $g^{(2)} = 1$ at $\tau \to \infty$ (on the time scale depending on the light source).

Let us use our simple single-mode model to analyze this experiment. Now the elementary quantum process characterized by the numerator of Eq. (31) is the correlated, simultaneous ionization of two trigger atoms, at two spatial-temporal points $\{\mathbf{r}_1, t\}$ and $\{\mathbf{r}_2, t - \tau\}$, by the same field mode, so we need to make the following replacement in the first of Eqs. (25):

$$\hat{\mathscr{E}}(\mathbf{r},t) \to \text{const} \times \hat{\mathscr{E}}(\mathbf{r}_1,t)\hat{\mathscr{E}}(\mathbf{r}_2,t-\tau)\,. \tag{9.32}$$

Repeating all the manipulations done above for the single-counter case, we get

$$\langle \Gamma_1(t)\Gamma_2(t-\tau)\rangle \propto \left\langle \text{ini}\middle|\hat{a}(t)^\dagger \hat{a}(t-\tau)^\dagger \hat{a}(t-\tau)\hat{a}(t)\right|\text{ini}\right\rangle e^*(\mathbf{r}_1)e^*(\mathbf{r}_2)e(\mathbf{r}_1)e(\mathbf{r}_2). \tag{9.33}$$

Plugging this expression, as well as Eq. (29) for single-counter rates, into Eq. (31), we see that the field distribution factors (as well as the detector-specific brackets and the density of states ρ_a) cancel, giving a very simple final expression:

$$g^{(2)}(\tau) = \frac{\left\langle \hat{a}^\dagger(t)\hat{a}^\dagger(t-\tau)\hat{a}(t-\tau)\hat{a}(t)\right\rangle}{\left\langle \hat{a}^\dagger(t)\hat{a}(t)\right\rangle^2}, \tag{9.34}$$

where the averaging should be carried out, as before, over the initial state of the field.

Still, the calculation of this expression for arbitrary τ may be quite involved because in most practical cases, the relaxation of the correlation function to the asymptotic value $g^{(2)}(\infty)$ is due to the interaction of the light source with its environment, and hence requires the open-system techniques that were discussed in Chapter 7. However, the zero-delay value $g^{(2)}(0)$ may be calculated straightforwardly, because the time arguments of all operators are equal, so we may write

Zero-delay correlation

$$g^{(2)}(0) = \frac{\left\langle \hat{a}^\dagger \hat{a}^\dagger \hat{a}\hat{a} \right\rangle}{\left\langle \hat{a}^\dagger \hat{a} \right\rangle^2}. \tag{9.35}$$

Let us evaluate this ratio for the simplest states of the field.

(i) The n^{th} Fock state. In this case, it is convenient to act with the annihilation operators upon the ket-vectors, and by the creation operators, upon the bra-vectors, using Eqs. (19):

Photon anti-bunching

$$g^{(2)}(0) = \frac{\left\langle n \middle| \hat{a}^\dagger \hat{a}^\dagger \hat{a}\hat{a} \middle| n \right\rangle}{\left\langle n \middle| \hat{a}^\dagger \hat{a} \middle| n \right\rangle^2} = \frac{\left\langle n-2 \middle| \left[n(n-1)\right]^{1/2} \left[n(n-1)\right]^{1/2} \middle| n-2 \right\rangle}{\left\langle n-1 \middle| n^{1/2} n^{1/2} \middle| n-1 \right\rangle^2} = \frac{n(n-1)}{n^2} \equiv 1 - \frac{1}{n}. \tag{9.36}$$

We see that the correlation function at small delays is below 1, i.e. below the value for fully independent counts – see the dashed lines in Fig. 3. This *photon antibunching* effect has a very simple handwaving explanation: a single photon emitted by the wave source may be absorbed by just one of the detectors. For the initial state $n = 1$, this is the only option, and it is very natural that Eq. (36) predicts no simultaneous counts at $\tau = 0$. Despite this theoretical simplicity, reliable observations of the antibunching have not been carried out until 1977,[19] due to the experimental difficulty of driving electromagnetic field oscillators into their Fock states – see Sec. 4 below.

(ii) The Glauber state α. A similar procedure, but now using Eq. (5.124) and its Hermitian conjugate, $\left\langle \alpha \middle| \hat{a}^\dagger \right. = \left\langle \alpha \middle| \alpha^* \right.$, yields

Glauber field statistics

$$g^{(2)}(0) = \frac{\left\langle \alpha \middle| \hat{a}^\dagger \hat{a}^\dagger \hat{a}\hat{a} \middle| \alpha \right\rangle}{\left\langle \alpha \middle| \hat{a}^\dagger \hat{a} \middle| \alpha \right\rangle^2} = \frac{\alpha^* \alpha^* \alpha\alpha}{(\alpha^* \alpha)^2} \equiv 1, \tag{9.37}$$

for any parameter α. We see that the result is different from that for the Fock states, unless in the latter case $n \to \infty$. (We know that the Fock and Glauber properties should also coincide for the ground state, but at that state, the correlation function's value is uncertain because there are no photon counts at all.)

(iii) Classical mixture. From Chapter 7, we know that such statistical ensembles cannot be described by single-state vectors, and require the density matrix w for their description. Here, we may combine Eqs. (35) and (7.5) to write

$$g^{(2)}(0) = \frac{\mathrm{Tr}\left(\hat{w}\hat{a}^\dagger \hat{a}^\dagger \hat{a}\hat{a}\right)}{\left[\mathrm{Tr}\left(\hat{w}\hat{a}^\dagger \hat{a}\right)\right]^2}. \tag{9.38}$$

Spelling out this expression is easy for the field in thermal equilibrium at some temperature T, because its density matrix is diagonal in the basis of the Fock states n – see Eqs. (7.24):

$$w_{nn'} = W_n \delta_{nn'}, \quad W_n = \exp\left\{-\frac{E_n}{k_B T}\right\} \Big/ Z \equiv \lambda^n \Big/ \sum_{n=0}^{\infty} \lambda^n, \quad \text{where } \lambda \equiv \exp\left\{-\frac{\hbar\omega}{k_B T}\right\}. \tag{9.39}$$

[19] By H. J. Kimble *et al.*, *Phys. Rev. Lett.* **39**, 691 (1977). For a detailed review of photon antibunching, see, e.g., H. Paul, *Rev. Mod. Phys*. **54**, 1061 (1982).

So, for the operators in the numerator and denominator of Eq. (38) we also need just the diagonal terms of the operator products, which have already been calculated – see Eq. (36). As a result, we get

$$g^{(2)}(0) = \frac{\sum_{n=0}^{\infty} W_n n(n-1)}{\left(\sum_{n=0}^{\infty} W_n n\right)^2} = \frac{\sum_{n=0}^{\infty} \lambda^n n(n-1) \times \sum_{n=0}^{\infty} \lambda^n}{\left(\sum_{n=0}^{\infty} \lambda^n n\right)^2}. \tag{9.40}$$

One of the three series in the last expression is just the usual geometric progression,

$$\sum_{n=0}^{\infty} \lambda^n = \frac{1}{1-\lambda}, \tag{9.41}$$

and the remaining two series may be readily calculated by its differentiation over the parameter λ:

$$\begin{aligned} \sum_{n=0}^{\infty} \lambda^n n &\equiv \lambda \sum_{n=0}^{\infty} \lambda^{n-1} n = \lambda \frac{d}{d\lambda} \sum_{n=0}^{\infty} \lambda^n = \lambda \frac{d}{d\lambda} \frac{1}{1-\lambda} = \frac{\lambda}{(1-\lambda)^2}, \\ \sum_{n=0}^{\infty} \lambda^n n(n-1) &\equiv \lambda^2 \sum_{n=0}^{\infty} \lambda^{n-2} n(n-1) = \lambda^2 \frac{d^2}{d^2\lambda} \left(\sum_{n=0}^{\infty} \lambda^n \right) = \lambda^2 \frac{d^2}{d\lambda^2} \frac{1}{1-\lambda} = \frac{2\lambda^2}{(1-\lambda)^3}, \end{aligned} \tag{9.42}$$

and for the correlation function, we get an extremely simple result independent of the parameter λ and hence of temperature:

$$g^{(2)}(0) = \frac{\left[2\lambda^2/(1-\lambda)^3\right]\left[1/(1-\lambda)\right]}{\left[\lambda/(1-\lambda)^2\right]^2} \equiv 2. \tag{9.43}$$

Photon bunching

This is exactly the photon bunching effect first observed by Hanbury Brown and Twiss – see Fig. 3. We see that in contrast to antibunching, this is an essentially classical (statistical) effect. Indeed, Eq. (43) allows for a purely classical derivation. In the classical theory, the counting rate (of a single counter) is proportional to the wave intensity I, so Eq. (31) with $\tau = 0$ is reduced to

$$g^{(2)}(0) = \frac{\langle I^2 \rangle}{\langle I \rangle^2}, \quad \text{with } I \propto \overline{E^2(t)} \propto E_\omega E_\omega^*. \tag{9.44}$$

For a sinusoidal field, the intensity is constant, and $g^{(2)}(0) = 1$. (This is also evident from Eq. (37), because the classical state may be considered a Glauber state with $\alpha \to \infty$.) On the other hand, if the intensity fluctuates (either in time or from one experiment to another), the averages in Eq. (44) should be calculated as

$$\langle I^k \rangle = \int_0^\infty w(I) I^k dI, \quad \text{with } \int_0^\infty w(I) dI = 1, \quad \text{and } k = 1, 2, \tag{9.45}$$

where $w(I)$ is the probability density. For classical statistics, the probability is an exponential function of the electromagnetic field energy, and hence its intensity:

$$w(I) = Ce^{-\beta I}, \quad \text{where } \beta \propto 1/k_B T, \tag{9.46}$$

so Eqs. (45) yield:[20]

[20] See, e.g., MA Eq. (6.7c) with $n = 0$ and $n = 1$.

$$\int_0^\infty C\exp\{-\beta I\}dI \equiv C/\beta = 1, \quad \text{and hence } C = \beta,$$

$$\left\langle I^k \right\rangle = \int_0^\infty w(I) I^k dI = C\int_0^\infty \exp\{-\beta I\} I^k dI = \frac{1}{\beta^k}\int_0^\infty \exp\{-\xi\}\xi^k d\xi = \begin{cases} 1/\beta, & \text{for } k=1, \\ 2/\beta^2, & \text{for } k=2. \end{cases} \tag{9.47}$$

Plugging these results into Eq. (44), we get $g^{(2)}(0) = 2$, in complete agreement with Eq. (43).

For some field states, including the squeezed ground states ζ discussed at the end of Sec. 5.5, the values of $g^{(2)}(0)$ may be even higher than 2 – the so-called *super-bunching*. Analyses of two cases of such super-bunching are offered for the reader's exercise – see the problem list at the chapter's end.

9.3. Photon emission: spontaneous and stimulated

In the last section, we considered how the counter's trigger atom *absorbs* a photon. Now let us have a look at the opposite process of *spontaneous emission* of a photon by an atom at its transition from an excited quasi-stationary (metastable) state to a state with an energy lower by $\Delta E = \hbar\omega$.[21] By using the same electric-dipole approximation (24) for the atom-to-field interaction, we may still use the Golden Rule for the model depicted in Fig. 6.12, but now the roles of its two component systems change: we have to associate the operator $\hat{A}$ with the electric dipole moment of the atom, while the operator $\hat{B}$, with the electric field, so the continuous spectrum of the system *b* represents the plurality of the electromagnetic field states into which the spontaneous radiation may happen. Since now the transition *increases* the energy of the electromagnetic field, and *decreases* that of the atom, after the multiplication of the field bracket in Eq. (27a) by $\exp\{-i\omega t\}$, and the second, by $\exp\{+i\omega t\}$, we may keep only the photon *creation* operator whose time evolution (26) compensates this additional fast "rotation". As a result, the Golden Rule takes the following form:

Spontaneous photon emission rate

$$\Gamma_s = \pi\omega \left| \langle \text{fin} | \hat{a}^\dagger | 0 \rangle \right|^2 \left| \langle \text{fin} | \hat{\mathbf{d}} \cdot \mathbf{e}(\mathbf{r}) | \text{ini} \rangle \right|^2 \rho_f, \tag{9.48}$$

where all operators and states are time-independent, and ρ_f is the density of final states of the electromagnetic field – which in this problem plays the role of the atom's environment.[22] Here the electromagnetic field oscillator has been assumed to be initially in the ground state – the assumption that will be changed later in this section.

This relation, together with Eq. (19), shows that for the field's matrix element to be different from zero, the final field has to be its first excited Fock state, $n = 1$. (By the way, *this* is exactly the most practicable way of generating an excited Fock state of a field oscillator.) With that, Eq. (48) yields

[21] A straightforward Fourier transform of Eq. (6.114) (which is the inseparable counterpart of the Golden Rule we are going to use) shows that the emitted radiation has the Lorentzian line centered to the frequency ω, with the half-width equal to the transition rate Γ_s that we will calculate.

[22] Here the summation over all electromagnetic field modes *j* may be smuggled back. Since in the quasistatic approximation $k_j a << 1$, which is necessary for the validity of Eq. (24), the matrix elements in Eq. (48) are independent of the field vector direction, and their magnitudes are fixed by ω, this summation is reduced to the calculation of the total ρ_f for all modes, and the spatial averaging of $e^2(\mathbf{r})$ – see below.

$$\Gamma_s = \pi\omega \left| \langle \text{fin} | \hat{\mathbf{d}} \cdot \mathbf{e}(\mathbf{r}) | \text{ini} \rangle \right|^2 \rho_f \equiv \pi\omega \left| \langle \text{fin} | \hat{d} e_d(\mathbf{r}) | \text{ini} \rangle \right|^2 \rho_f , \tag{9.49}$$

where e_d is the Cartesian component of the vector $\mathbf{e}(\mathbf{r})$ along the electric dipole's direction, and ρ_f is now the density of *electromagnetic modes* at the frequency ω. The expression for it follows from our first formula in this course – see Eq. (1.1):

$$\rho_f \equiv \frac{dN}{dE} = V \frac{\omega^2}{\pi^2 \hbar c^3}, \tag{9.50}$$

where the volume V should be large enough to ensure the spectrum's virtual continuity:[23] $V >> \lambda^3 = (2\pi c/\omega)^3$. Because of that, in the normalization condition used to simplify Eq. (9), we may consider $e^2(\mathbf{r})$ constant. Let us represent this square as a sum of squares of the three Cartesian components of the vector $\mathbf{e}(\mathbf{r})$, with one of those (e_d) aligned with the dipole's direction; due to the space isotropy we may write

$$e^2 \equiv e_d^2 + e_{\perp 1}^2 + e_{\perp 2}^2 = 3e_d^2. \tag{9.51}$$

As a result, the normalization condition yields

$$e_d^2 = \frac{1}{3\varepsilon_0 V}. \tag{9.52}$$

and Eq. (49) gives the famous (and very important) formula[24]

$$\Gamma_s = \frac{1}{4\pi\varepsilon_0} \frac{4\omega^3}{3\hbar c^3} \left| \langle \text{fin} | \hat{\mathbf{d}} | \text{ini} \rangle \right|^2 \equiv \frac{1}{4\pi\varepsilon_0} \frac{4\omega^3}{3\hbar c^3} \langle \text{fin} | \hat{\mathbf{d}} | \text{ini} \rangle \cdot \langle \text{ini} | \hat{\mathbf{d}} | \text{fin} \rangle^* . \tag{9.53}$$

Free-space spontaneous emission rate

Leaving a comparison of this formula with the classical theory of radiation,[25] and the exact evaluation of Γ_s for a particular transition in the hydrogen atom, for the reader's exercises, let me just estimate its order of magnitude. Assuming that $d \sim e r_B \equiv e\hbar^2/m_e(e^2/4\pi\varepsilon_0)$ and $\hbar\omega \sim E_H \equiv m_e(e^2/4\pi\varepsilon_0)^2/\hbar^2$, and taking into account the definition (6.62) of the fine-structure constant $\alpha \approx 1/137$, we get

$$\frac{\Gamma_s}{\omega} \sim \left(\frac{e^2}{4\pi\varepsilon_0 \hbar c} \right)^3 \equiv \alpha^3 \sim 3\times 10^{-7}. \tag{9.54}$$

This estimate shows that the emission lines at atomic transitions are typically very sharp. With the present-day availability of high-speed electronics, it also makes sense to evaluate the time scale $\tau = 1/\Gamma$ of a typical quantum transition: for a typical optical frequency $\omega \sim 3\times 10^{15}$ s^{-1}, it is close to 1 ns. This is exactly the time constant that determines the time-delay dependence of the photon counting statistics of the spontaneously emitted radiation – see Fig. 3. Very colloquially, this is the temporal scale of the

[23] In the opposite case when the same atom is placed into a high-Q resonant cavity with a discrete spectrum (see, e.g., EM 7.9), the rate of its photon emission is strongly suppressed at frequencies between the cavity resonances (where $\rho_f \to 0$) – see, e.g., the review by S. Haroche and D. Klepner, *Phys. Today* **42**, 24 (Jan. 1989). On the other hand, the emission is strongly (by a factor $\sim (\lambda^3/V)Q$, where V is the cavity's volume) enhanced at resonance frequencies – the so-called *Purcell effect*, discovered by E. Purcell in the 1940s. For a brief discussion of this and other quantum electrodynamic effects in cavities, see the next section.

[24] This was the breakthrough result obtained by P. Dirac in 1927, which jumpstarted the whole field of quantum electrodynamics. An equivalent expression was obtained from more formal arguments in 1930 by V. Weisskopf and E. Wigner, so sometimes Eq. (53) is (very unfairly) called the "Weisskopf-Wigner formula".

[25] See, e.g., EM Sec. 8.2, in particular Eq. (8.29).

photon emitted by an atom. Note that the scale $c\tau$ of the spatial extension of the corresponding wave packet is surprisingly macroscopic – of the order of 30 cm. Such a "human" size of spontaneously emitted photons makes the usual optical table, with its 1-cm-scale components, the key equipment for many optical experiments – see, e.g., Fig. 2.

Note, however, that the above estimate of τ is only valid for a transition with a non-zero electric-dipole matrix element. If it equals zero, i.e. the transition does not satisfy the *selection rules*,[26] – say, due to the initial and final state symmetry – it is "forbidden". The "forbidden" transition may still take place due to a different, smaller interaction (say, via a magnetic dipole field of the atom, or its quadrupole electric field[27]), but takes much longer. In some cases the increase of τ is rather dramatic – sometimes to hours! Such long-lasting radiation is called *fluorescence*. (If the initial excitation is followed first by a series of non-radiative transitions down the energy level ladder, which delay the final radiative transition, the resulting radiation is called either *phosphorescence* or *luminescence*.)

Now let us consider a more general case when the electromagnetic field mode of frequency ω is initially in an arbitrary Fock state n, and from it may either get energy $\hbar\omega$ from the atomic system (*photon emission*) or, vice versa, give such energy back to the atom (*photon absorption*). For the photon emission rate, an evident generalization of Eq. (48) gives

$$\frac{\Gamma_e}{\Gamma_s} \equiv \frac{\Gamma_{n\to \text{fin}}}{\Gamma_{0\to 1}} = \frac{\left|\left\langle \text{fin}\left|\hat{a}^\dagger\right|n\right\rangle\right|^2}{\left|\left\langle 1\left|\hat{a}^\dagger\right|0\right\rangle\right|^2}, \tag{9.55}$$

where both brackets should be calculated in the Schrödinger picture, and Γ_s is the spontaneous emission rate (48) of the same atomic system. According to the second of Eqs. (19), at the photon emission, the final field state has to be the Fock state with $n' = n + 1$, and Eq. (55) yields

Stimulated photon emission rate

$$\Gamma_e = (n+1)\Gamma_s . \tag{9.56}$$

Thus the initial field increases the photon emission rate; this effect is called the *stimulated emission of radiation*. Note that the spontaneous emission may be considered a particular case of the stimulated emission for $n = 0$, and hence interpreted as the emission stimulated by the ground state of the electromagnetic field – one more manifestation of the non-trivial nature of this "vacuum" state.

On the other hand, following the arguments of Sec. 2,[28] for the description of radiation *absorption*, the photon creation operator has to be replaced with the annihilation operator, giving the rate ratio

---

[26] As was already discussed in Sec. 5.6, for a single spin-less particle moving in a spherically symmetric potential (e.g., a hydrogen-like atom), the orbital selection rules are simple: the only allowed electric-dipole transitions are those with $\Delta l \equiv l_{\text{fin}} - l_{\text{ini}} = \pm 1$ and $\Delta m \equiv m_{\text{fin}} - m_{\text{ini}} = 0$ or ± 1. The simplest example of the transition that does *not* satisfy this rule, i.e. is "forbidden", is that between the s-states ($l = 0$) with $n = 2$ and $n = 1$; because of that, the lifetime of the lowest excited s-state of a hydrogen atom is as long as ~0.15 s.

[27] See, e.g., EM Sec. 8.9.

[28] Mind, however, a major difference between the rate Γ discussed in Sec. 2, and Γ_a in Eq. (57). In our current case, the atomic transition is still between two *discrete* energy levels (see Fig. 4 below), so the rate Γ_a is proportional to ρ_f, the density of final states of the electromagnetic *field*, i.e. the same density as in Eq. (48) and

$$\frac{\Gamma_a}{\Gamma_s} = \frac{\left|\langle \text{fin}|\, \hat{a}\, |\, n\rangle\right|^2}{\left|\langle 1\,|\hat{a}^\dagger|\,0\rangle\right|^2}. \tag{9.57}$$

According to this relation and the first of Eqs. (19), the final state of the field at the photon absorption has to be the Fock state with $n' = n - 1$, and Eq. (57) yields

$$\boxed{\Gamma_a = n\Gamma_s}. \tag{9.58}$$

Photon absorption rate

The results (56) and (58) are sometimes formulated in terms of relations between the *Einstein coefficients* A and B defined in the way shown in Fig. 4, where the two energy levels are those of the atom, Γ_a is the rate of photon absorption from the electromagnetic field, Γ_e is that of photon emission into the field, and u is the electromagnetic field density in thermal equilibrium – see Eq. (1.3). Since per Eq. (7.26b), the ratio $u/\langle n\rangle$ is a constant independent of $\langle n\rangle$ (and temperature), the statistical averaging of Eqs. (56) and (58) yields[29]

$$A_{21} = B_{21}\frac{u}{\langle n\rangle} = B_{12}\frac{u}{\langle n\rangle}, \tag{9.59}$$

Einstein coefficients' relation

because each of these expressions equals the spontaneous emission rate Γ_s.

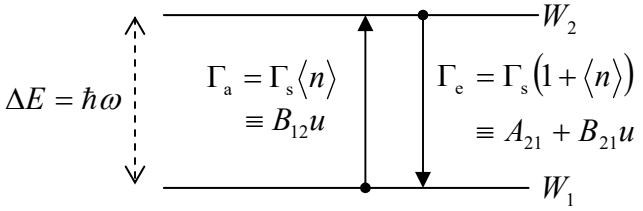


Fig. 9.4. The Einstein coefficients on the atomic quantum transition diagram – cf. Fig. 7.6.

I cannot resist the temptation to use this point for a small detour – an alternative derivation of the Bose-Einstein statistics for photons. Indeed, in the thermodynamic equilibrium we have just discussed, the average probability flows between levels 1 and 2 (see Fig. 4 again) should be equal:[30]

$$W_2\Gamma_e = W_1\Gamma_a, \tag{9.60}$$

where W_1 and W_2 are the probabilities for the atomic system to occupy the corresponding levels, so Eqs. (56) and (58) yield

$$W_2\Gamma_s\big(\langle n\rangle + 1\big) = W_1\Gamma_s\langle n\rangle, \qquad \text{i.e. } \frac{W_2}{W_1} = \frac{\langle n\rangle}{\langle n\rangle + 1}. \tag{9.61}$$

But, on the other hand, for an atomic subsystem only weakly coupled to its electromagnetic environment, we ought to have the Gibbs distribution of these probabilities:

$$\frac{W_2}{W_1} = \frac{\exp\{-E_2/k_BT\}}{\exp\{-E_1/k_BT\}} = \exp\left\{-\frac{\Delta E}{k_BT}\right\} = \exp\left\{-\frac{\hbar\omega}{k_BT}\right\}. \tag{9.62}$$

---

beyond, while the rate (27) is proportional to ρ_a, the density of final (ionized) states of the photon counter's "trigger" *atom* – more exactly, of it's the electron released at its ionization.

[29] These relations were conjectured, from general statistical arguments, by Albert Einstein as early as 1916.

[30] This is just a particular embodiment of the detailed balance equation (7.198).

By requiring Eqs. (61) and (62) to give the same result for the probability ratio, we readily get the Bose-Einstein distribution for the electromagnetic field in thermal equilibrium:

$$\langle n \rangle = \frac{1}{\exp\{\hbar\omega / k_B T\} - 1}, \tag{9.63}$$

i.e. the same result (7.26b) as that obtained in Sec. 7.1 by other means.

Now returning to the discussion of Eqs. (56) and (58), their very important implication is the possibility to achieve the stimulated emission of coherent radiation using the level *occupancy inversion*. Indeed, if the ratio W_2/ W_1 is larger than that given by Eq. (62), the net power flow from the atomic system into the electromagnetic field,

$$\text{power } = \hbar\omega \times \Gamma_s \left[W_2 \left(\langle n \rangle + 1 \right) - W_1 \langle n \rangle \right], \tag{9.64}$$

may be positive. The necessary inversion may be produced using several ways, notably by intensive quantum transitions to level 2 from an even higher energy level (which, in turn, is populated, e.g., by absorption of external radiation, usually called *pumping*, at a higher frequency.)

A less obvious but crucial feature of the stimulated emission is spelled out by Eq. (55): as was mentioned above, it shows that the final state of the field after the absorption of energy $\hbar\omega$ from the atom is a pure (coherent) Fock state $(n + 1)$. Colloquially, one may say that the new, $(n + 1)^{st}$ photon emitted from the atom is automatically in phase with the n photons that had been in the field mode initially, i.e. joins them coherently.[31] The idea of stimulated emission of coherent radiation using population inversion[32] was first implemented in the early 1950s in the microwave range (*masers*) and in 1960 in the optical range (*lasers*). Nowadays, lasers are ubiquitous components of almost all high-tech systems and constitute one of the cornerstones of our technological civilization.

A quantitative discussion of laser operation is well beyond the framework of this course, and I have to refer the reader to special literature,[33] but still would like to briefly mention two key points:

(i) In a typical laser, each generated electromagnetic field mode is in its Glauber (rather than the Fock) state, so Eqs. (56) and (58) are applicable only for the n averaged over the Fock-state decomposition of the Glauber state – see Eq. (5.134).

(ii) Since in a typical laser $\langle n \rangle >> 1$, its operation may be well described using quasiclassical theories that use Eq. (64) to describe the electromagnetic energy balance (with the addition of a term describing the energy loss due to field absorption in external components of the laser, including the useful load), plus the equation describing the balance of occupancies $W_{1,2}$ due to all interlevel transitions – similar to Eq. (60), but including also the contribution(s) from the particular population inversion mechanism used in the laser. In this approach, the role of quantum mechanics in laser science is essentially reduced to the calculation of the parameter Γ_s for a particular system.

This role becomes more prominent when one needs to describe *fluctuations* of the laser field. Here two approaches are possible, following the two options discussed in Chapter 7. If the fluctuations

[31] It is straightforward to show that this fact is also true if the field is initially in the Glauber state – which is more typical for modes in practical lasers.

[32] This idea has been traced back to an obscure 1939 publication by V. Fabrikant.

[33] I can recommend, for example, P. Milloni and J. Eberly, *Laser Physics*, 2nd ed., Wiley, 2010, and a less technical text by A. Yariv, *Quantum Electronics*, 3rd ed., Wiley, 1989.

are relatively small, one can linearize the Heisenberg equations of motion of the field oscillator operators near their stationary-lasing "values", with the Langevin "forces" (also time-dependent operators) describing the fluctuation sources, and use these Heisenberg-Langevin equations to calculate the radiation fluctuations, just as was described in Sec. 7.5. On the other hand, near the lasing threshold, the field fluctuations are relatively large, smearing the phase transition between the no-lasing and lasing states. Here the linearization is not an option, but one can use the density-matrix approach described in Sec. 7.6, for the fluctuation analysis.[34] Note that while the laser fluctuations may look like a peripheral issue, the pioneering research in that field has led to the development of the general theory of open quantum systems, which was discussed in Chapter 7.

9.4. Cavity QED

Now I have to discuss, at least in passing, the field of *cavity quantum electrodynamics* (usually called *cavity QED* for short) – the art and science of creating and using the entanglement between quantum states of an atomic system (either an atom, or an ion, or a molecule, etc.) and the electromagnetic field in a macroscopic volume called the *resonant cavity* (or just "resonator", or just "cavity"). This field is very popular nowadays, especially in the context of the quantum computation and communication research reviewed in Sec. 8.5.[35]

Our discussion in the previous section was based on the implicit assumption that the energy spectrum of the electromagnetic field interacting with an atomic subsystem is essentially continuous, so its final state is spread among many field modes, effectively losing its coherence with the quantum state of the atomic subsystem. This assumption has justified our use of the quantum-mechanical Golden Rule for the calculation of the spontaneous and stimulated transition rates. However, the assumption becomes invalid if the electromagnetic field is contained inside a relatively small volume, with its linear size comparable with the radiation wavelength. If the walls of such a cavity mostly reflect, rather than absorb, radiation, then in the 0^{th} approximation, the energy dissipation may be disregarded, and the particular solutions $\mathbf{e}_j(\mathbf{r})$ of the Helmholtz equation (5) correspond to discrete, well-separated mode wave numbers k_j and hence well-separated field oscillation frequencies ω_j.[36] Due to the energy conservation, an atomic transition corresponding to the energy $\Delta E = | E_{\rm ini} - E_{\rm fin} |$ may be effective only if the corresponding quantum transition frequency $\Omega \equiv \Delta E/\hbar$ is close to one of these resonance frequencies.[37] As a result of such resonant interaction, the quantum states of the atomic system and the resonant electromagnetic mode may become entangled.

A very popular approximation for the quantitative description of this effect is the so-called *Rabi model*,[38] in which the atom is treated as a two-level system interacting with a single electromagnetic field mode of the resonant cavity. (As was shown in Sec. 6.5, this model is justified, e.g., if transitions

[34] This path has been developed (also in the mid-1960s), by several researchers, notably including W. Lamb and M. Sully – see, e.g., M. Sargent III, M. Scully, and W. Lamb, Jr., *Laser Physics,* Westview, 1977.

[35] This popularity was demonstrated, for example, by the award of the 2012 Nobel Prize in Physics to cavity QED experimentalists S. Haroche and D. Wineland.

[36] The calculation of such modes and corresponding frequencies for several simple cavity geometries is the subject of EM Sec. 7.8 of this series.

[37] On the contrary, if Ω is far from any ω_j, the interaction is suppressed; in particular, the spontaneous emission rate may be much lower than that given by Eq. (53) – so this result is not as fundamental as it may look.

[38] After the pioneering work by I. Rabi in 1936-37.

between all other energy level pairs have considerably different frequencies.) As the reader knows well from Chapters 4-6 (see in particular Sec. 5.1), any two-level system may be described, just as a spin-½, by the Hamiltonian $b\hat{I} + \mathbf{c}\cdot\hat{\boldsymbol{\sigma}}$. Since we may always select such energy origin that $b = 0$, and such coordinate system that $\mathbf{c} = c\mathbf{n}_z$, and ignore the atomic subsystem's degrees of freedom that do not participate in the interaction with the field, its Hamiltonian may be taken in the form

$$\hat{H}_{\rm a} = c\hat{\sigma}_z \equiv \frac{\hbar\Omega}{2}\hat{\sigma}_z, \tag{9.65}$$

where $\hbar\Omega \equiv 2c = \Delta E$ is the difference between the energy levels in the absence of interaction with the cavity field. Next, according to Eq. (17), ignoring the constant ground-state energy $\hbar\omega/2$ (which may be always added to the energy at the end – if necessary), the contribution of a single field mode of frequency ω to the total Hamiltonian of the system is

$$\hat{H}_{\rm f} = \hbar\omega\hat{a}^\dagger\hat{a}. \tag{9.66}$$

Finally, according to Eq. (16a), the electric field of the mode may be represented as

$$\hat{\mathscr{E}}(\mathbf{r},t) = \frac{1}{i}\left(\frac{\hbar\omega}{2}\right)^{1/2}\mathbf{e}(\mathbf{r})\left(\hat{a} - \hat{a}^\dagger\right), \tag{9.67}$$

so in the electric-dipole approximation (24), the cavity-atom interaction may be represented as a product of the field by some (say, y-) Cartesian component[39] of the Pauli spin-½ operator:

$$\hat{H}_{\rm int} = {\rm const}\times\hat{\sigma}_y\times\hat{\mathscr{E}} = {\rm const}\times\hat{\sigma}_y\times\left(\frac{\hbar\omega}{2}\right)^{1/2}\frac{1}{i}\left(\hat{a} - \hat{a}^\dagger\right) \equiv i\hbar\kappa\hat{\sigma}_y\left(\hat{a} - \hat{a}^\dagger\right), \tag{9.68}$$

where κ is a coupling constant (with the dimension of frequency). The sum of these three terms,

Rabi Hamiltonian

$$\hat{H} \equiv \hat{H}_{\rm a} + \hat{H}_{\rm f} + \hat{H}_{\rm int} = \frac{\hbar\Omega}{2}\hat{\sigma}_z + \hbar\omega\hat{a}^\dagger\hat{a} + i\hbar\kappa\hat{\sigma}_y\left(\hat{a} - \hat{a}^\dagger\right), \tag{9.69}$$

giving a very reasonable description of the system, is called the *Rabi Hamiltonian*.

Despite the apparent simplicity of Eq. (69), using this Hamiltonian for calculations is not that straightforward.[40] Only in the case when the electromagnetic field is large and hence may be treated classically, the results following from Eq. (69) are reduced to Eqs. (6.94) describing, in particular, the Rabi oscillations discussed in Sec. 6.3. The situation becomes simpler in the most important case when the frequencies Ω and ω are very close, enabling an effective interaction between the cavity field and the atom even if the coupling constant κ is relatively small. Indeed, if both the κ and the so-called *detuning* (defined similarly to the parameter Δ used in Sec. 6.5),

$$\xi \equiv \Omega - \omega, \tag{9.70}$$

[39] The exact component is not important for final results, while intermediate formulas are simpler if the interaction is proportional to either pure $\hat{\sigma}_x$ or pure $\hat{\sigma}_y$.

[40] For example, an exact quasi-analytical expression for its eigenenergies (as zeros of a Taylor series in the parameter κ, with coefficients determined by a recurrence relation) was found only recently – see D. Braak, *Phys. Rev. Lett.* **107**, 100401 (2011).

are much smaller than $\Omega \approx \omega$, the Rabi Hamiltonian may be simplified using the rotating-wave approximation, already used several times in this course.

For this, it is convenient to use the *spin ladder operators* defined similarly for those of the orbital angular momentum – see Eqs. (5.153):

$$\hat{\sigma}_{\pm} \equiv \hat{\sigma}_x \pm i\hat{\sigma}_y, \qquad \text{so} \quad \hat{\sigma}_y = \frac{\hat{\sigma}_+ - \hat{\sigma}_-}{2i}. \tag{9.71}$$

From Eq. (4.105), it is very easy to find the matrices of these operators in the standard z-basis,

$$\sigma_+ = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}, \qquad \sigma_- = \begin{pmatrix} 0 & 0 \\ 2 & 0 \end{pmatrix}, \tag{9.72}$$

and their commutation rules – which turn out to be naturally similar to Eqs. (5.154):

$$\left[\hat{\sigma}_+, \hat{\sigma}_-\right] = 4\hat{\sigma}_z, \qquad \left[\hat{\sigma}_z, \hat{\sigma}_\pm\right] = \pm 2\hat{\sigma}_\pm. \tag{9.73}$$

In this notation, the Rabi Hamiltonian becomes

$$\hat{H} = \frac{\hbar\Omega}{2}\hat{\sigma}_z + \hbar\omega\hat{a}^\dagger\hat{a} + \frac{\hbar\kappa}{2}\left(\hat{\sigma}_+ - \hat{\sigma}_-\right)\left(\hat{a} - \hat{a}^\dagger\right), \tag{9.74}$$

and it is straightforward to use Eq. (4.199) and (73) to derive the Heisenberg-picture equations of motion for the involved operators. (Doing this, we have to remember that operators of the "spin" subsystem, on one hand, and of the field mode, on the other hand, are defined in different Hilbert spaces and hence commute – at least at coinciding time moments.) The result (so far, exact!) is

$$\begin{aligned} &\dot{\hat{a}} = -i\omega\hat{a} + \frac{i\kappa}{2}\left(\hat{\sigma}_+ - \hat{\sigma}_-\right), \qquad \dot{\hat{a}}^\dagger = i\omega\hat{a}^\dagger + \frac{i\kappa}{2}\left(\hat{\sigma}_+ - \hat{\sigma}_-\right), \\ &\dot{\hat{\sigma}}_\pm = \pm i\Omega\hat{\sigma}_\pm + 2i\kappa\left(\hat{a} - \hat{a}^\dagger\right)\hat{\sigma}_z, \qquad \dot{\hat{\sigma}}_z = i\kappa\left(\hat{a}^\dagger - \hat{a}\right)\left(\hat{\sigma}_+ + \hat{\sigma}_-\right). \end{aligned} \tag{9.75}$$

At negligible coupling, $\kappa \to 0$, these equations have simple solutions,

$$\hat{a}(t) \propto e^{-i\omega t}, \quad \hat{a}^\dagger(t) \propto e^{i\omega t}, \quad \hat{\sigma}_\pm(t) \propto e^{\pm i\Omega t}, \quad \hat{\sigma}_z(t) \approx \text{const}, \tag{9.76}$$

and the small terms proportional to κ on the right-hand sides of Eqs. (75) cannot affect these time evolution laws dramatically even if κ is not exactly zero. Of those terms, ones with frequencies close to the "basic" frequency of each variable would act in resonance and hence may have a substantial impact on the system's dynamics, while non-resonant terms may be ignored. In this rotating-wave approximation, Eqs. (75) are reduced to a much simpler system of equations:

$$\begin{aligned} &\dot{\hat{a}} = -i\omega\hat{a} - \frac{i\kappa}{2}\hat{\sigma}_-, \qquad \dot{\hat{a}}^\dagger = i\omega\hat{a}^\dagger + \frac{i\kappa}{2}\hat{\sigma}_+, \\ &\dot{\hat{\sigma}}_+ = i\Omega\hat{\sigma}_+ + 2i\kappa\hat{a}^\dagger\hat{\sigma}_z, \qquad \dot{\hat{\sigma}}_- = -i\Omega\hat{\sigma}_- - 2i\kappa\hat{a}\hat{\sigma}_z, \qquad \dot{\hat{\sigma}}_z = i\kappa\left(\hat{a}^\dagger\hat{\sigma}_- - \hat{a}\hat{\sigma}_+\right). \end{aligned} \tag{9.77}$$

Alternatively, these equations of motion may be obtained *exactly* from the Hamiltonian (74) if it is preliminary cleared of the terms proportional to $\hat{\sigma}_+\hat{a}^\dagger$ and $\hat{\sigma}_-\hat{a}$, that oscillate fast and hence self-average to produce virtually zero effect:

Jaynes-Cummings Hamiltonian

$$\hat{H} = \frac{\hbar\Omega}{2}\hat{\sigma}_z + \hbar\omega\hat{a}^\dagger\hat{a} + \frac{\hbar\kappa}{2}\left(\hat{\sigma}_+\hat{a} + \hat{\sigma}_-\hat{a}^\dagger\right), \qquad \text{at } \kappa, |\xi| << \omega, \Omega\,. \tag{9.78}$$

This is the famous *Jaynes-Cummings Hamiltonian*,[41] which is the basic model used in the cavity QED and its applications.[42] To find its eigenstates and eigenenergies, let us note that at negligible interaction ($\kappa \to 0$), the spectrum of the total energy E of the system, which is the sum of two independent contributions from the atomic and cavity-field subsystems, $\pm\hbar\Omega/2$ and $\hbar\Omega n$, forms close pairs (Fig. 5)

$$E\big|_{\kappa=0} = E_n \pm \frac{\hbar\xi}{2}, \tag{9.79}$$

centered to values[43]

$$E_n \equiv \hbar\omega\left(n - \frac{1}{2}\right), \quad \text{with } n = 1, 2, 3\ldots,. \tag{9.80}$$

(At the exact resonance $\omega = \Omega$, i.e. at $\xi = 0$, each pair merges into one double-degenerate level E_n.)

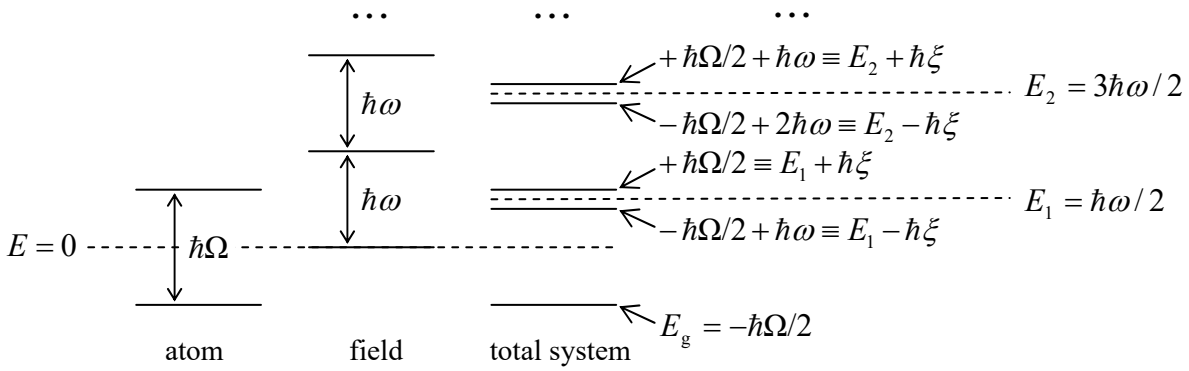


Fig. 9.5. The energy spectrum (79) of the Jaynes-Cummings Hamiltonian in the limit $\kappa << |\xi|$. Note again that the energy is referred to the ground-state energy $\hbar\omega/2$ of the cavity field.

Since at $\kappa \to 0$ the two subsystems do not interact, the eigenstates corresponding to the two levels of the n^{th} pair may be represented by direct products of their independent state vectors:

$$|+\rangle \equiv |\uparrow\rangle \otimes |n-1\rangle \quad \text{and} \quad |-\rangle \equiv |\downarrow\rangle \otimes |n\rangle, \tag{9.81}$$

where the first ket of each product represents the state of the two-level (spin-½-like) atomic subsystem, and the second ket, that of the field oscillator.

As we know from Chapter 6, weak interaction may lead to strong coherent mixing[44] of quantum states with close energies (in this case, the two states (81) within each pair with the same n), even if

[41] It was first proposed and analyzed in 1963 by two engineers, E. Jaynes and F. Cummings, in a paper published in *Proc. IEEE*, and it took the physics community a while to recognize and acknowledge the fundamental importance of that work.

[42] For most applications, the baseline Hamiltonian (78) has to be augmented by additional term(s) describing, for example, the incoming radiation and/or the system's coupling to the environment, for example, due to the electromagnetic energy loss in a finite-Q-factor cavity – see Eq. (7.68).

[43] Only the ground state level $E_g = -\hbar\Omega/2$ is non-degenerate – see Fig. 5.

[44] In some fields, especially chemistry, such mixing is frequently called *hybridization*.

their mixing with the states with farther energies is still negligible. Hence, at $0 < \kappa$, $|\xi| << \omega \approx \Omega$, a good approximation of the eigenstate with $E \approx E_n$ is given by a linear superposition of the states (81):

$$\left|\alpha_n\right\rangle = c_+\left|+\right\rangle + c_-\left|-\right\rangle \equiv c_+\left|\uparrow\right\rangle \otimes \left|n-1\right\rangle + c_-\left|\downarrow\right\rangle \otimes \left|n\right\rangle, \tag{9.82}$$

Jaynes-Cummings eigenstates

with certain c-number coefficients $c_\pm$. This relation describes the entanglement of the atomic eigenstates $\uparrow$ and $\downarrow$ with the Fock states number n and $n - 1$ of the field mode. Let me leave the (straightforward) calculation of the coefficients ($c_\pm$) for the reader's exercise. (The result for the corresponding two eigenenergies $(E_n)_\pm$ may be again represented by the same anticrossing diagram as shown in Figs. 2.29 and 5.1, now with the detuning ξ as the argument.) This calculation shows, in particular, that at $\xi = 0$ (i.e. at $\omega = \Omega$), $|c_+| = |c_-| = 1/\sqrt{2}$ for both states of the pair. This fact may be interpreted as a (coherent!) equal sharing of an energy quantum $\hbar\omega = \hbar\Omega$ by the atom and the cavity field at the exact resonance.

As a (hopefully, self-evident) by-product of the calculation of $c_\pm$ is the fact that the dynamics of the state α_n described by Eq. (82) is similar to that of the generic two-level system that was repeatedly discussed in this course – the first time in Sec. 2.6 and then in Chapters 4-6. In particular, if the composite system had been initially prepared to be in one component state, for example $|\uparrow\rangle\otimes|0\rangle$ (i.e. with the atom excited, while the cavity in its ground state), and then allowed to evolve on its own, after some time interval $\Delta t \sim 1/\kappa$ it may be found definitely in the counterpart state $|\downarrow\rangle\otimes|1\rangle$, including the first excited Fock state $n = 1$ of the field mode. If the process is allowed to continue, after the equal time interval Δt, the system returns to the initial state $|\uparrow\rangle\otimes|0\rangle$, etc. This most striking prediction of the Jaynes-Cummings model was directly observed, by G. Rempe *et al.*, only in 1987, although less directly this model was repeatedly confirmed by numerous experiments carried out in the 1960s and 1970s.

This quantized-field version of the Rabi oscillations can only persist in time if the inevitable electromagnetic energy losses (not described by the basic Jaynes-Cummings Hamiltonian) are somehow compensated – for example, by passing a beam of particles, externally excited into the higher-energy state $\uparrow$, through the cavity. If the losses become higher, the dissipation suppresses quantum coherence, in our case the coherence between two components of each pair (82), as was discussed in Chapter 7. As a result, the transition from the higher-energy atomic state $\uparrow$ to the lower-energy state $\downarrow$, giving the energy quantum $\hbar\omega$ to the cavity ($n - 1 \to n$), which is then rapidly drained into the environment, becomes incoherent, so the system's dynamics is reduced to the Purcell effect mentioned in Sec. 3. A quantitative analysis of this effect is left for the reader's exercise.

The number of interesting games one can play with such systems – say, by adding external sources of radiation at a frequency close to ω and Ω, in particular with manipulated time-dependent amplitude and/or phase, is almost unlimited.[45] Unfortunately, my time/space allowance for the cavity QED is over, and for further discussion, I have to refer the interested reader to special literature.[46]

9.5. The Klein-Gordon and relativistic Schrödinger equations

Now let me switch gears and discuss the basics of relativistic quantum mechanics of particles with a *non-zero* rest mass *m.* In the ultra-relativistic limit $pc >> mc^2$, the quantization scheme of such

[45] Most of them may be described by adding the corresponding terms to the basic Jaynes-Cummings Hamiltonian.
[46] I can recommend, for example, either C. Gerry and P. Knight, *Introductory Quantum Optics*, Cambridge U. Press, 2005, or G. Agarwal, *Quantum Optics*, Cambridge U. Press, 2012.

particles may be essentially the same as for electromagnetic waves, but for the intermediate energy range, $pc \sim mc^2$, a more general approach is necessary. Historically, the first attempts[47] to extend the non-relativistic wave mechanics into the relativistic energy range were based on performing the same transition from the classical observables to their quantum-mechanical operators as in the non-relativistic limit:

$$\mathbf{p} \to \hat{\mathbf{p}} = -i\hbar\nabla, \qquad E \to \hat{H} = i\hbar\frac{\partial}{\partial t}. \tag{9.83}$$

The substitution of these operators, acting on the Schrödinger-picture wavefunction $\Psi(\mathbf{r},t)$, into the classical relation (1) between the energy E and momentum $\mathbf{p}$ (for of a free particle) leads to the following formulas:

Table 9.1. Deriving the Klein-Gordon equation for a free relativistic particle.[48]

	Non-relativistic limit	Relativistic case
Classical mechanics	$E = \frac{1}{2m}p^2$	$E^2 = c^2p^2 + \left(mc^2\right)^2$
Wave mechanics	$i\hbar\frac{\partial}{\partial t}\Psi = \frac{1}{2m}\left(-i\hbar\nabla\right)^2\Psi$	$\left(i\hbar\frac{\partial}{\partial t}\right)^2\Psi = c^2\left(-i\hbar\nabla\right)^2\Psi + (mc^2)^2\Psi$

The resulting equation for the non-relativistic limit, in the left-bottom cell of the table, is just the usual Schrödinger equation (1.28) for a free particle. Its relativistic generalization, in the right-bottom cell, usually rewritten as

Klein-Gordon equation

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right)\Psi + \mu^2\Psi = 0, \qquad \text{with } \mu \equiv \frac{mc}{\hbar}, \tag{9.84}$$

is called the *Klein-Gordon* (or sometimes "Klein-Gordon-Fock") *equation*. The fundamental solutions of this equation are the same plane monochromatic waves

$$\Psi(\mathbf{r},t) \propto \exp\left\{i\left(\mathbf{k}\cdot\mathbf{r} - \omega t\right)\right\}. \tag{9.85}$$

as in the non-relativistic case. Indeed, such waves are the eigenstates of the operators (83), with eigenvalues, respectively,

$$\mathbf{p} = \hbar\mathbf{k}, \qquad \text{and } E = \hbar\omega, \tag{9.86}$$

so their substitution into Eq. (84) immediately returns us to Eq. (1) with the replacements (86):

$$E_{\pm} = \hbar\omega_{\pm} = \pm\left[\left(\hbar ck\right)^2 + \left(mc^2\right)^2\right]^{1/2}. \tag{9.87}$$

[47] This approach was suggested in 1926-1927, i.e. virtually simultaneously, by (at least) V. Fock, E. Schrödinger, O. Klein and W. Gordon, J. Kudar, T. de Donder, and F.-H. van der Dungen, and L. de Broglie.

[48] Note that in the left, non-relativistic column of this table, the energy is measured from the rest energy mc^2, while in its right, relativistic column, it is measured from zero – see Eq. (1).

Though one may say that this dispersion relation is just a simple combination of the classical relation (1) and the same basic quantum-mechanical relations (86) as in non-relativistic limit, it attracts our attention to the fact that the energy $\hbar\omega$ as a function of the momentum $\hbar\mathbf{k}$ has two branches, with $E_-(\mathbf{p}) = -E_+(\mathbf{p})$ – see Fig. 6a. Historically, this fact has played a very important role in spurring the fundamental idea of *particle-antiparticle pairs.* In this idea (very similar to the concept of electrons and holes in semiconductors, which was discussed in Sec. 2.8), what we call the "vacuum" actually corresponds to all quantum states of the lower branch, with energies $E_-(\mathbf{p}) < 0$, being completely filled, while the states on the upper branch, with energies $E_+(\mathbf{p}) > 0$, being empty. Then an externally supplied energy,

$$\Delta E = E_+ - E_- \equiv E_+ + \left(-E_-\right) \geq 2mc^2 > 0, \tag{9.88}$$

may bring the system from the lower branch to the upper one (Fig. 6b). The resulting excited state is interpreted as a combination of a particle (formally, of the infinite spatial extension) with an energy $E_+ > 0$ and the corresponding momentum $\mathbf{p}$, and a "hole" (antiparticle), also of a *positive* energy ($-E_-$) and the opposite momentum $-\mathbf{p}$. This idea[49] has led to a search for, and discovery of the positron: the electron's antiparticle with charge $q = +e$, in 1932, and later of the antiproton and other antiparticles.

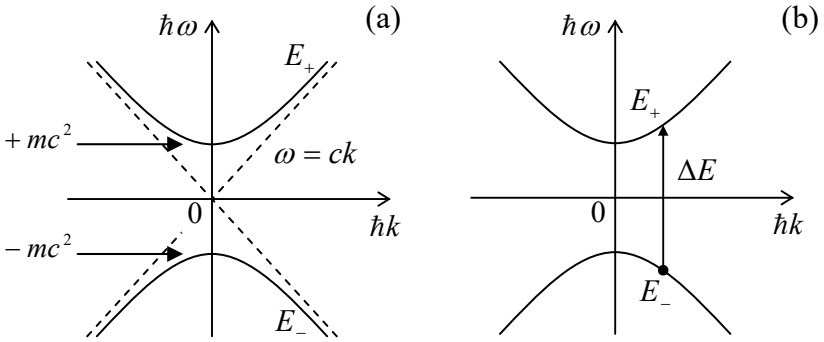


Fig. 9.6. (a) The free-particle dispersion relation resulting from the Klein-Gordon and Dirac equations, and (b) the scheme of creation of a particle-antiparticle pair from the vacuum.

Free particles of a finite spatial extension may be described, in this approach, just as in the non-relativistic Schrödinger equation, by wave packets, i.e. linear superpositions of the de Broglie waves (85) with close wave vectors $\mathbf{k}$, and the corresponding values of ω given by Eq. (87), with the positive sign for the "usual" particles, and negative sign for antiparticles – see Fig. 6a above. Note that to form, from a particle's wave packet, a similar wave packet for the antiparticle, with the same phase and group velocities (2.33a) in each direction, we need to change the sign not only before ω, but also before $\mathbf{k}$, i.e. to replace all component wavefunctions (85), and hence the full wavefunction, with their complex conjugates.

Of more formal properties of Eq. (84), it is straightforward (and hence left for the reader's exercise) to prove that its solutions satisfy the same continuity equation (1.52), with the probability current density $\mathbf{j}$ still given by Eq. (1.47), but a different expression for the probability density w – which becomes very similar to that for $\mathbf{j}$:

$$w = \frac{i\hbar}{2mc^2}\left(\Psi^* \frac{\partial \Psi}{\partial t} - \text{c.c.}\right), \qquad \mathbf{j} = \frac{i\hbar}{2m}\left(\Psi \nabla \Psi^* - \text{c.c.}\right) \tag{9.89}$$

[49] Due to the same P. A. M. Dirac!

– very much in the spirit of the relativity theory, which treats space and time on equal footing. (In the non-relativistic limit $p/mc \to 0$, Eq. (84) allows the reduction of this expression for w to the non-relativistic Eq. (1.22): $w \to \Psi\Psi^*$.)

The Klein-Gordon equation may be readily generalized to describe a particle moving in external fields; for example, the electromagnetic field effects on a particle with charge q may be described by the same replacement as in the non-relativistic limit (see Sec. 3.1):

$$\hat{\mathbf{p}} \to \hat{\mathbf{P}} - q\mathbf{A}(\mathbf{r},t), \qquad \hat{H} \to \hat{H} - q\phi(\mathbf{r},t), \tag{9.90}$$

where $\hat{\mathbf{P}} = -i\hbar\nabla$ is the canonical momentum operator (3.25), and the vector- and scalar potentials, $\mathbf{A}$ and ϕ, should be treated appropriately – either as c-number functions if the electromagnetic field quantization is not important for a particular problem, or as operators (see Secs. 1-4 above) if it is.

However, the practical value of the resulting *relativistic Schrödinger equation* is rather limited, for two main reasons. First of all, it does not give the correct description of particles with spin. For example, for the hydrogen-like atom/ion problem, i.e. the motion of an electron with the electric charge $-e$, in the Coulomb central field of an immobile nucleus with charge $+Ze$, the equation may be readily solved exactly[50] and yields the following spectrum of (doubly degenerate) energy levels:

$$E = mc^2\left(1+\frac{Z^2\alpha^2}{\lambda^2}\right)^{-1/2}, \quad \text{with } \lambda \equiv n + \left[\left(l+\tfrac{1}{2}\right)^2 - Z^2\alpha^2\right]^{1/2} - \left(l+\tfrac{1}{2}\right), \tag{9.91}$$

where $n = 1, 2, \ldots$ and $l = 0, 1, \ldots, n-1$ are the same quantum numbers as in the non-relativistic theory (see Sec. 3.6), and $\alpha \approx 1/137$ is the fine structure constant (6.62). The three leading terms of the Taylor expansion of this result in the small parameter $Z\alpha$ are as follows:

$$E \approx mc^2\left[1 - \frac{Z^2\alpha^2}{2n^2} - \frac{Z^4\alpha^4}{2n^4}\left(\frac{n}{l+\tfrac{1}{2}} - \frac{3}{4}\right)\right]. \tag{9.92}$$

The first of these terms is just the rest energy of the particle. The second term,

$$E_n = -mc^2\frac{Z^2\alpha^2}{2n^2} \equiv -\frac{mZ^2e^4}{(4\pi\varepsilon_0)^2\hbar^2}\frac{1}{2n^2} \equiv -\frac{E_0}{2n^2}, \quad \text{with } E_0 = Z^2E_{\rm H}, \tag{9.93}$$

reproduces the non-relativistic Bohr's formula (3.201). Finally, the third term,

$$-mc^2\frac{Z^4\alpha^4}{2n^4}\left(\frac{n}{l+\tfrac{1}{2}} - \frac{3}{4}\right) \equiv -\frac{2E_n^2}{mc^2}\left(\frac{n}{l+\tfrac{1}{2}} - \frac{3}{4}\right), \tag{9.94}$$

is just the perturbative kinetic-relativistic contribution (6.51) to the fine structure of the Bohr levels (93). However, as we already know from Sec. 6.3, for a spin-½ particle such as the electron, the spin-orbit interaction (6.55) gives an additional contribution to the fine structure, of the same order, so the net result, confirmed by experiment, is given by Eq. (6.60), i.e. is different from Eq. (94). This is very natural because the relativistic Schrödinger equation does not have the very notion of spin.

Second, even for massive spinless particles (such as the Z^0 bosons), for which this equation is believed to be valid, the most important problems are related to particle interactions at high energies of

[50] This task is left for the reader's exercise.

the order of $\Delta E \sim 2mc^2$ and beyond – see Eq. (88). Due to the possibility of creation and annihilation of particle-antiparticle pairs at such energies, the number of particles participating in such interactions is typically considerable (and variable), and the adequate description of the system is given not by the relativistic Schrödinger equation (which is formulated in single-particle terms), but by the quantum field theory – to which I will devote only a few sentences at the very end of this chapter.

9.6. Dirac's theory

The real breakthrough toward the quantum relativistic theory of electrons (and other spin-½ fermions) was achieved in 1928 by P. A. M. Dirac. For that time, the structure of his theory was highly nontrivial. Namely, while formally preserving, in the coordinate representation, the same Schrödinger-picture equation of quantum dynamics as in the non-relativistic quantum mechanics,[51]

$$i\hbar \frac{\partial \Psi}{\partial t} = \hat{H}\Psi , \tag{9.95}$$

it postulates that the wavefunction Ψ it describes is not a scalar complex function of time and coordinates, but a *four-component* column-vector (sometimes called the *bispinor*) of such functions, its Hermitian-conjugate bispinor $\Psi^\dagger$ being a four-component row-vector of their complex conjugates:

$$\Psi = \begin{pmatrix} \Psi_1(\mathbf{r},t) \\ \Psi_2(\mathbf{r},t) \\ \Psi_3(\mathbf{r},t) \\ \Psi_4(\mathbf{r},t) \end{pmatrix}, \qquad \Psi^\dagger = \left(\Psi_1^*(\mathbf{r},t), \quad \Psi_2^*(\mathbf{r},t), \quad \Psi_3^*(\mathbf{r},t), \quad \Psi_4^*(\mathbf{r},t)\right), \tag{9.96}$$

and that the Hamiltonian participating in Eq. (95) is a 4×4 matrix defined in the Hilbert space of bispinors Ψ. For a free particle, the postulated Hamiltonian looks amazingly simple:[52]

$$\boxed{\hat{H} = c\hat{\boldsymbol{\alpha}} \cdot \hat{\mathbf{p}} + \hat{\beta} mc^2 .} \tag{9.97}$$

Free-particle's Hamiltonian

[51] After the "naturally-relativistic" form of the Klein-Gordon equation (84), this apparent return to the non-relativistic Schrödinger equation may look very counter-intuitive. However, it becomes a bit less surprising taking into account the fact (whose proof is left for the reader's exercise) that Eq. (84) may be also recast into the form (95) for a *two-component* column-vector Ψ (sometimes called *spinor*), with a Hamiltonian which may be represented by a 2×2 matrix – and hence expressed via the Pauli matrices (4.105) and the identity matrix I.

[52] Moreover, if the time derivative participating in Eq. (95), and the three coordinate derivatives participating (via the momentum operator) in Eq. (97), are merged into one 4-vector operator $\partial/\partial x_k \equiv \{\nabla, \partial/\partial(ct)\}$, the Dirac equation (95) may be rewritten in an even simpler, manifestly Lorentz-invariant 4-vector form (with the implied summation over the repeated index $k = 1, ..., 4$ – see, e.g., EM Sec. 9.4):

$$\left(\hat{\gamma}_k \frac{\partial}{\partial x_k} + \mu\right)\Psi = 0, \quad \text{where } \hat{\boldsymbol{\gamma}} \equiv \{\hat{\gamma}_1, \hat{\gamma}_2, \hat{\gamma}_3\} = \begin{pmatrix} 0 & -i\hat{\boldsymbol{\sigma}} \\ i\hat{\boldsymbol{\sigma}} & 0 \end{pmatrix}, \quad \hat{\gamma}_4 = \hat{\beta},$$

where $\mu \equiv mc/\hbar$ – just as in Eq. (84). Note also that, very counter-intuitively, the Dirac Hamiltonian (97) is *linear* with respect to the momentum, while the non-relativistic Hamiltonian of a particle, as well as the relativistic Schrödinger equation, are *quadratic* in $\mathbf{p}$. In my humble opinion, Dirac's theory (including the concept of antiparticles it has inspired) may compete for the title of the most revolutionary theoretical idea in physics of all times, despite such strong contenders as Newton's laws, Maxwell's equations, Gibbs' statistical distribution, Bohr's theory of the hydrogen atom, and Einstein's general relativity.

where $\hat{\mathbf{p}} = -i\hbar\nabla$ is the same 3D vector operator of momentum as in the non-relativistic case, while the operators $\hat{\boldsymbol{\alpha}}$ and $\hat{\beta}$ may be represented in the following shorthand 2×2 form:

Dirac operators

$$\hat{\boldsymbol{\alpha}} \equiv \begin{pmatrix} \hat{0} & \hat{\boldsymbol{\sigma}} \\ \hat{\boldsymbol{\sigma}} & \hat{0} \end{pmatrix}, \qquad \hat{\beta} \equiv \begin{pmatrix} \hat{I} & \hat{0} \\ \hat{0} & -\hat{I} \end{pmatrix}. \tag{9.98a}$$

The operator $\hat{\boldsymbol{\alpha}}$, composed of the Pauli vector operators $\hat{\boldsymbol{\sigma}}$, is also a vector in the usual 3D space, with each of its 3 Cartesian components being a 4×4 matrix. The particular form of the 2×2 matrices corresponding to the operators $\hat{\boldsymbol{\sigma}}$ and $\hat{I}$ in Eq. (98a) depends on the basis selected for the spin state representation; for example, in the standard z-basis, in which the Cartesian components of $\hat{\boldsymbol{\sigma}}$ are represented by the Pauli matrices (4.105), the 4×4 matrix form of Eq. (98a) is

$$\alpha_x = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad \alpha_y = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix}, \quad \alpha_z = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \quad \beta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{9.98b}$$

It is straightforward to use Eqs. (98) to verify that the matrices α_x, α_y, α_z and β satisfy the following relations:

$$\alpha_x^2 = \alpha_y^2 = \alpha_z^2 = \beta^2 = \mathrm{I}, \tag{9.99}$$

$$\alpha_x\alpha_y + \alpha_y\alpha_x = \alpha_y\alpha_z + \alpha_z\alpha_y = \alpha_z\alpha_x + \alpha_x\alpha_z = \alpha_x\beta + \beta\alpha_x = \alpha_y\beta + \beta\alpha_y = \alpha_z\beta + \beta\alpha_z = 0, \tag{9.100}$$

i.e. anticommute.

Using these commutation relations, and acting essentially as in Sec. 1.4, it is straightforward to show that any solution to the Dirac equation obeys the probability conservation law, i.e. the continuity equation (1.52), with the probability density:

$$w = \Psi^\dagger\Psi, \tag{9.101}$$

and the probability current,

$$\mathbf{j} = \Psi^\dagger c\hat{\boldsymbol{\alpha}}\Psi, \tag{9.102}$$

looking *almost* as in the non-relativistic wave mechanics – cf. Eqs. (1.22) and (1.47). Note, however, the Hermitian conjugation used in these formulas instead of the complex conjugation, in order to form the scalars w, j_x, j_y, and j_z from the 4-component state vectors (96).

This close similarity is extended to the fundamental plane-wave solutions of the Dirac equations in free space. Indeed, plugging such a solution, in the form

$$\Psi = \mathrm{u}e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} \equiv \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}, \tag{9.103}$$

into Eqs. (95) and (97), we see that they are indeed satisfied, provided that a system of four coupled linear algebraic equations for four complex c-number amplitudes $u_{1,2,3,4}$ is satisfied. The condition of its

consistency yields the same dispersion relation (87), i.e. the same two-branch diagram shown in Fig. 6, as follows from the Klein-Gordon equation. The difference is that plugging each value of ω, given by Eq. (87), back into the system of the linear equations for the four amplitudes u, we get two solutions for their vector $\mathbf{u} \equiv (u_1, u_2, u_3, u_4)$ for each of the two energy branches – see Fig. 6 again. In the standard z-basis of spin operators, they may be represented as follows:

$$\text{for } E = E_+ > 0: \quad \mathrm{u}_{+\uparrow} = c_{+\uparrow}\begin{pmatrix} 1 \\ 0 \\ \dfrac{cp_z}{E_+ + mc^2} \\ \dfrac{cp_+}{E_+ + mc^2} \end{pmatrix}, \quad \mathrm{u}_{+\downarrow} = c_{+\downarrow}\begin{pmatrix} 0 \\ 1 \\ \dfrac{cp_-}{E_+ + mc^2} \\ \dfrac{-cp_z}{E_+ + mc^2} \end{pmatrix}, \tag{9.104a}$$

$$\text{for } E = E_- < 0: \quad \mathrm{u}_{-\uparrow} = c_{-\uparrow}\begin{pmatrix} \dfrac{cp_z}{E_- - mc^2} \\ \dfrac{cp_+}{E_- - mc^2} \\ 1 \\ 0 \end{pmatrix}, \quad \mathrm{u}_{-\downarrow} = c_{-\downarrow}\begin{pmatrix} \dfrac{cp_-}{E_- - mc^2} \\ \dfrac{-cp_z}{E_- - mc^2} \\ 0 \\ 1 \end{pmatrix}, \tag{9.104b}$$

where $p_\pm \equiv p_x \pm ip_y$, and $c_\pm$ are the normalization coefficients determined by initial conditions.

The simplest interpretation of these solutions is that Eq. (103), with the vectors $\mathbf{u}_+$ given by Eq. (104a), represents a spin-½ particle (say, an electron), while with the vectors $\mathbf{u}_-$ given by Eq. (104b), it represents an antiparticle (a positron), and the two solutions for each particle, indexed with opposite arrows, correspond to two possible directions of the spin: $\sigma_z = \pm 1$, i.e. $S_z = \pm\hbar/2$. This interpretation is indeed solid in the non-relativistic limit, when the two last components of the vector (104a), and two first components of the vector (104b) are negligibly small:

$$\mathrm{u}_{+\uparrow} \to \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \mathrm{u}_{+\downarrow} \to \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathrm{u}_{-\uparrow} \to \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathrm{u}_{-\downarrow} \to \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \qquad \text{for } \frac{p_{x,y,z}}{mc} \to 0. \tag{9.105}$$

However, at arbitrary energies, the physical picture is more complex. To show this, let us use the Dirac equation to calculate the Heisenberg-picture law of time evolution of the operator of some Cartesian component of the orbital angular momentum $\mathbf{L} \equiv \mathbf{r}\times\mathbf{p}$, for example of $L_x = yp_z - zp_y$, taking into account the fact that the Dirac operators (98a) commute with those of $\mathbf{r}$ and $\mathbf{p}$, and also the Heisenberg commutation relations (2.14):

$$i\hbar\frac{\partial \hat{L}_x}{\partial t} = \left[\hat{L}_x, \hat{H}\right] = c\hat{\boldsymbol{\alpha}}\cdot\left[\left(\hat{y}\hat{p}_z - \hat{z}\hat{p}_y\right), \hat{\mathbf{p}}\right] = -i\hbar c\left(\hat{\alpha}_z\hat{p}_y - \hat{\alpha}_y\hat{p}_z\right), \tag{9.106}$$

with similar relations for two other Cartesian components. Since the right-hand side of these equations is different from zero, the orbital momentum is generally *not* conserved – even for a free particle! Let us, however, consider the following vector operator,

Spin operator in Dirac's theory

$$\hat{\mathbf{S}} \equiv \frac{\hbar}{2}\begin{pmatrix} \hat{\boldsymbol{\sigma}} & \hat{0} \\ \hat{0} & \hat{\boldsymbol{\sigma}} \end{pmatrix}. \tag{9.107a}$$

According to Eqs. (4.105), its Cartesian components, in the z-basis, are represented by the 4×4 matrices

$$\mathrm{S}_x = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad \mathrm{S}_y = \frac{\hbar}{2}\begin{pmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix}, \quad \mathrm{S}_z = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{9.107b}$$

Let us calculate the Heisenberg-picture law of time evolution of these components, for example

$$i\hbar \frac{\partial \hat{S}_x}{\partial t} = \left[\hat{S}_x, \hat{H}\right] = c\left[\hat{S}_x, \left(\hat{\alpha}_x \hat{p}_x + \hat{\alpha}_y \hat{p}_y + \hat{\alpha}_z \hat{p}_z\right)\right]. \tag{9.108}$$

A direct calculation of the commutators of the matrices (98) and (107) yields

$$\left[\hat{S}_x, \hat{\alpha}_x\right] = 0, \quad \left[\hat{S}_x, \hat{\alpha}_y\right] = i\hbar\hat{\alpha}_z, \quad \left[\hat{S}_x, \hat{\alpha}_z\right] = -i\hbar\hat{\alpha}_y, \tag{9.109}$$

so we finally get

$$i\hbar \frac{\partial \hat{S}_x}{\partial t} = i\hbar c\left(\hat{\alpha}_z \hat{p}_y - \hat{\alpha}_y \hat{p}_z\right), \tag{9.110}$$

with similar expressions for the other two components of the operator. Comparing this result with Eq. (106), we see that any Cartesian component of the operator defined similarly to Eq. (5.170),

$$\hat{\mathbf{J}} \equiv \hat{\mathbf{L}} + \hat{\mathbf{S}}, \tag{9.111}$$

is an integral of motion,[53] so this operator may be interpreted as the one representing the total angular momentum of the particle. Hence, the operator (107) *may* be interpreted as the spin operator of a spin-½ particle (e.g., electron). As it follows from the last of Eq. (107b), in the non-relativistic limit, the columns (105) represent the eigenkets of the z-component of that operator, with eigenvalues $S_z = \pm\hbar/2$, the sign corresponding to the arrow index. So, the Dirac theory provides a justification for spin-½ – or, somewhat more humbly, replaces the Pauli Hamiltonian postulate (4.163) with that of a simpler (and hence more plausible) Lorentz-invariant Hamiltonian (97).

Note, however, that this simple interpretation, fully separating a particle from its antiparticle, is not valid for the exact solutions (103)-(104), so generally the eigenstates of the Dirac Hamiltonian are certain linear (coherent) superpositions of the components describing the particle and its antiparticle – each with both directions of spin. This fact leads to several interesting effects, including the so-called *Klien paradox* at the reflection of a relativistic electron from a potential barrier.[54]

[53] It is straightforward to show that this result remains valid for a particle in any central field $U(r)$.
[54] See, e.g., A. Calogeracos and N. Dombey, *Contemp. Phys.* **40**, 313 (1999).

9.7. Low-energy limit

The generalization of Dirac's theory to the case of a (spin-½) particle with an electric charge q, moving in a classically-described electromagnetic field, may be obtained using the same replacement (90). As a result, Eq. (95) turns into

$$\left[c\hat{\boldsymbol{\alpha}}\cdot\left(-i\hbar\nabla-q\mathbf{A}\right)+mc^2\hat{\beta}+\left(q\phi-\hat{H}\right)\right]\Psi=0\,, \tag{9.112}$$

Dirac equation in EM field

where the Hamiltonian operator $\hat{H}$ is understood in the sense of Eq. (95), i.e. as the partial time derivative with the multiplier $i\hbar$. Let us prepare this equation for a low-energy approximation by acting on its left-hand side by a similar square bracket but with the opposite sign before the last parentheses – also an operator! Using Eqs. (99) and (100), and the fact that the space- and time-independent operators $\hat{\boldsymbol{\alpha}}$ and $\hat{\beta}$ commute with the spin-independent c-number functions $\mathbf{A}(\mathbf{r},t)$ and $\phi(\mathbf{r},t)$, as well as with the Hamiltonian operator $i\hbar\partial/\partial t$, the result is

$$\left\{c^2\left[\hat{\boldsymbol{\alpha}}\cdot\left(-i\hbar\nabla-q\mathbf{A}\right)\right]^2+\left(mc^2\right)^2-c\left[\hat{\boldsymbol{\alpha}}\cdot\left(-i\hbar\nabla-q\mathbf{A}\right),\left(q\phi-\hat{H}\right)\right]-\left(q\phi-\hat{H}\right)^2\right\}\Psi=0\,. \tag{9.113}$$

A direct calculation of the first square bracket, using Eqs. (98) and (107), yields

$$\left[\hat{\boldsymbol{\alpha}}\cdot\left(-i\hbar\nabla-q\mathbf{A}\right)\right]^2\equiv\left(-i\hbar\nabla-q\mathbf{A}\right)^2-2q\hat{\mathbf{S}}\cdot\nabla\times\mathbf{A}\,. \tag{9.114}$$

But the last vector product on the right-hand side is just the magnetic field – see, e.g., Eqs. (3.21):

$$\mathscr{B}=\nabla\times\mathbf{A}\,. \tag{9.115}$$

Similarly, we may use the first of Eqs. (3.21), for the electric field,

$$\mathscr{E}=-\nabla\phi-\frac{\partial\mathbf{A}}{\partial t}\,, \tag{9.116}$$

to simplify the commutator participating in Eq. (9.113):

$$\left[\hat{\boldsymbol{\alpha}}\cdot\left(-i\hbar\nabla-q\mathbf{A}\right),\left(q\phi-\hat{H}\right)\right]\equiv-q\hat{\boldsymbol{\alpha}}\cdot\left[\hat{H},\mathbf{A}\right]-i\hbar q\hat{\boldsymbol{\alpha}}\cdot\left[\nabla,\phi\right]\equiv-i\hbar q\frac{\partial\mathbf{A}}{\partial t}-i\hbar\hat{\boldsymbol{\alpha}}\cdot\nabla\phi\equiv i\hbar q\hat{\boldsymbol{\alpha}}\cdot\mathscr{E}\,. \tag{9.117}$$

As a result, Eq. (113) becomes

$$\left\{c^2\left(-i\hbar\nabla-q\mathbf{A}\right)^2+\left(q\phi-\hat{H}\right)^2-\left(mc^2\right)^2-2qc^2\hat{\mathbf{S}}\cdot\mathscr{B}+i\hbar cq\hat{\boldsymbol{\alpha}}\cdot\mathscr{E}\right\}\Psi=0\,. \tag{9.118}$$

So far, this is an exact result equivalent to Eq. (112), but it is more convenient for an analysis of the low-energy limit, in which not only the energy offset $E-mc^2$ (which is just the energy used in the non-relativistic mechanics), but also the electrostatic energy of the particle, $|q\langle\phi\rangle|$, are much smaller than the rest energy mc^2. In this limit, the second and third terms of Eq. (118) almost cancel, and introducing the rest-energy-offset Hamiltonian

$$\hat{\tilde{H}}\equiv\hat{H}-mc^2\hat{I}\,. \tag{9.119}$$

we may approximate their difference, up to the first non-zero term, as

$$\left(q\phi\hat{I}-\hat{H}\right)^2-\left(mc^2\right)^2\hat{I}\equiv\left(q\phi\hat{I}-mc^2\hat{I}-\hat{\tilde{H}}\right)^2-\left(mc^2\right)^2\hat{I}\approx 2mc^2\left(\hat{\tilde{H}}-q\phi\hat{I}\right). \tag{9.120}$$

As a result, after the division of all terms by $2mc^2$, Eq. (118) may be approximated as

Low-energy Hamiltonian

$$\hat{\tilde{H}}\Psi = \left[\frac{1}{2m}\left(-i\hbar\nabla - q\mathbf{A}\right)^2 + q\phi - \frac{q}{m}\hat{\mathbf{S}}\cdot\mathscr{B} + \frac{i\hbar q}{2mc}\hat{\boldsymbol{\alpha}}\cdot\mathscr{E}\right]\Psi . \tag{9.121}$$

Let us discuss this important result. The first two terms in the square brackets give the non-relativistic Hamiltonian (3.26), which was extensively used in Chapter 3 for the discussion of charged particle motion. Note again that the contribution of the vector potential **A** into that Hamiltonian is essentially relativistic, in the following sense: the magnetic interaction of two charged particles, due to their *orbital* motion with speed $v << c$, is a factor of $(v/c)^2$ smaller than the electrostatic interaction of the particles.[55] The reason why we did discuss the effects of **A** in Chapter 3 was that is was used there to describe *external* magnetic fields, keeping our analysis valid even for the cases when that field is strong because of being produced by relativistic effects – such as aligned spins of a permanent magnet.

The next, third term in the square brackets of Eq. (121) should be also familiar to the reader: this is the Pauli Hamiltonian – see Eqs. (4.3), (4.5), and (4.163). When justifying this form of interaction in Chapter 4, I referred mostly to the results of Stern-Gerlach-type experiments, but it is extremely pleasing that this result[56] follows from such a fundamental relativistic treatment as Dirac's theory. As we already know from the discussion of the Zeeman effect in Sec. 6.4, the magnetic field effects on the orbital motion of an electron (described by the orbital angular momentum **L**) and its spin **S** are of the same order, though quantitatively different.

Finally, the last term in the square brackets of Eq. (121) is also not quite new for us: in particular, it describes the spin-orbit interaction. Indeed, in the case of a classical, spherical-symmetric electric field $\mathscr{E}$ corresponding to the potential $\phi(r) = U(r)/q$, this term may be reduced to Eq. (6.56):

Spin-orbit coupling

$$\hat{H}_{so} = \frac{1}{2m^2c^2}\hat{\mathbf{S}}\cdot\hat{\mathbf{L}}\frac{1}{r}\frac{dU}{dr} \equiv -\frac{q}{2m^2c^2}\hat{\mathbf{S}}\cdot\hat{\mathbf{L}}\frac{1}{r}\mathscr{E} . \tag{9.122}$$

The proof of this correspondence requires a bit of additional work.[57] Indeed, in Eq. (121), the term responsible for the spin-orbit interaction acts on 4-component wavefunctions, while the Hamiltonian (122) is supposed to act on non-relativistic state vectors with an account of spin, whose coordinate representation may be given by 2-component spinors:[58]

---

[55] This difference may be traced by classical means – see, e.g., EM Sec. 5.1.

[56] Note that in this result, the *g*-factor of the particle is still equal to exactly 2 – see Eq. (4.115) and its discussion in Sec. 4.4. In order to describe the small deviation of g_e from 2, the electromagnetic field should be quantized (just as this was discussed in Secs. 1-4 of this chapter), and its potentials **A** and ϕ, participating in Eq. (121), should be treated as operators – rather than as *c*-number functions as was assumed above.

[57] The only facts immediately evident from Eq. (121) are that the term we are discussing is proportional to the electric field, as required by Eq. (122), and that it is of the proper order of magnitude. Indeed, Eqs. (101)-(102) imply that in the Dirac theory, $c\hat{\boldsymbol{\alpha}}$ plays the role of the velocity operator, so the expectation values of the term are of the order of $\hbar qv\mathscr{E}/2mc^2$. Since the expectation values of the operators participating in the Hamiltonian (122) scale as $S \sim \hbar/2$ and $L \sim mvr$, the spin-orbit interaction energy has the same order of magnitude.

[58] In this course, the notion of spinor (popular in some textbooks) was not used much; it was introduced earlier only for two-particle states – see Eq. (8.13). For a single particle, such definition is reduced to $\psi(\mathbf{r})|s\rangle$, whose representation in a particular spin-½ basis is the column (123). Note that such spinors may be used as a basis for an expansion of the spin-orbitals $\psi_j(\mathbf{r})$ defined by Eq. (8.125), where the index *j* is used for numbering both the spin's orientation (i.e. the particular component of the spinor's column) and the orbital eigenfunction.

$$\psi = \begin{pmatrix} \psi_\uparrow \\ \psi_\downarrow \end{pmatrix}. \tag{9.123}$$

The simplest way to prove the equivalence of these two expressions is not to use Eq. (121) directly, but to return to the Dirac equation (112), for the particular case of motion in a static electric field but no magnetic field, when Dirac's Hamiltonian is reduced to

$$\hat{H} = c\hat{\boldsymbol{\alpha}} \cdot \hat{\mathbf{p}} + \hat{\beta} mc^2 + U(\mathbf{r}), \qquad \text{with } U = q\phi. \tag{9.124}$$

Since this Hamiltonian is time-independent, we may look for its 4-component eigenfunctions in the form

$$\Psi(\mathbf{r},t) = \begin{pmatrix} \psi_+(\mathbf{r}) \\ \psi_-(\mathbf{r}) \end{pmatrix} \exp\left(-i\frac{E}{\hbar}t\right), \tag{9.125}$$

where each of $\psi_\pm$ is a 2-component column of the type (123), representing two spin states of the particle (index +) and its antiparticle (index –). Plugging Eq. (125) into Eq. (95) with the Hamiltonian (124), and using Eq. (98a), we get the following system of two linear equations:

$$\left[E - mc^2 - U(\mathbf{r})\right]\psi_+ - c\hat{\boldsymbol{\sigma}} \cdot \hat{\mathbf{p}}\,\psi_- = 0, \qquad \left[E + mc^2 - U(\mathbf{r})\right]\psi_- - c\hat{\boldsymbol{\sigma}} \cdot \hat{\mathbf{p}}\,\psi_+ = 0. \tag{9.126}$$

Expressing ψ_- from the latter equation, and plugging the result into the former one, we get the following single equation for the particle's spinor:

$$\left[E - mc^2 - U(\mathbf{r}) - c^2\hat{\boldsymbol{\sigma}} \cdot \hat{\mathbf{p}} \frac{1}{E + mc^2 - U(\mathbf{r})} \hat{\boldsymbol{\sigma}} \cdot \hat{\mathbf{p}}\right]\psi_+ = 0. \tag{9.127}$$

So far, this is an exact equation for eigenstates and eigenvalues of the Hamiltonian (124), but it may be substantially simplified in the low-energy limit when both the potential energy[59] and the non-relativistic eigenenergy

$$\widetilde{E} \equiv E - mc^2 \tag{9.128}$$

are much lower than mc^2. Indeed, in this case, the expression in the denominator of the last term in the brackets of Eq. (127) is very close to $2mc^2$. Since $\boldsymbol{\sigma}^2 = 1$, with that replacement, Eq. (127) is reduced to the non-relativistic Schrödinger equation, similar for both spin components of ψ_+, and hence giving spin-degenerate energy levels. To recover small relativistic and spin-orbit effects, we need a slightly more accurate approximation:

$$\frac{1}{E + mc^2 - U(\mathbf{r})} \equiv \frac{1}{2mc^2 + \widetilde{E} - U(\mathbf{r})} \equiv \frac{1}{2mc^2}\left[1 + \frac{\widetilde{E} - U(\mathbf{r})}{2mc^2}\right]^{-1} \approx \frac{1}{2mc^2}\left[1 - \frac{\widetilde{E} - U(\mathbf{r})}{2mc^2}\right], \tag{9.129}$$

in which Eq. (127) is reduced to

$$\left[\widetilde{E} - U(\mathbf{r}) - \frac{\hat{p}^2}{2m} + \hat{\boldsymbol{\sigma}} \cdot \hat{\mathbf{p}} \frac{\widetilde{E} - U(\mathbf{r})}{(2mc^2)^2} \hat{\boldsymbol{\sigma}} \cdot \hat{\mathbf{p}}\right]\psi_+ = 0. \tag{9.130}$$

As Eqs. (5.34) shows, the operators of the momentum and of a function of coordinates commute as

$$\left[\hat{\mathbf{p}}, U(\mathbf{r})\right] = -i\hbar\nabla U, \tag{9.131}$$

[59] Strictly speaking, this requirement is imposed on the expectation values of $U(\mathbf{r})$ in the eigenstates to be found.

so the last term in the square brackets of Eq. (130) may be rewritten as

$$\hat{\boldsymbol{\sigma}}\cdot\hat{\mathbf{p}}\,\frac{\widetilde{E}-U(\mathbf{r})}{(2mc)^2}\,\hat{\boldsymbol{\sigma}}\cdot\hat{\mathbf{p}}\equiv\frac{\widetilde{E}-U(\mathbf{r})}{(2mc)^2}\,\hat{p}^2-\frac{i\hbar}{(2mc)^2}(\hat{\boldsymbol{\sigma}}\cdot\nabla U)(\hat{\boldsymbol{\sigma}}\cdot\hat{\mathbf{p}}). \tag{9.132}$$

Since in the low-energy limit, both terms on the right-hand side of this relation are much smaller than the three leading terms of Eq. (130), we may replace the first term's numerator with its non-relativistic approximation $\hat{p}^2/2m$. With this replacement, the term coincides with the first relativistic correction to the kinetic energy operator – see Eq. (6.47). The second term, proportional to the electric field $\mathscr{E}=-\nabla\phi=-\nabla U/q$, may be transformed further on, using a readily verifiable identity

$$(\hat{\boldsymbol{\sigma}}\cdot\nabla U)(\hat{\boldsymbol{\sigma}}\cdot\hat{\mathbf{p}})\equiv(\nabla U)\cdot\hat{\mathbf{p}}+i\hat{\boldsymbol{\sigma}}\cdot[(\nabla U)\times\hat{\mathbf{p}}]. \tag{9.133}$$

Of the two terms on the right-hand side of this relation, only the second one depends on the spin,[60] giving the following spin-orbital interaction contribution to the Hamiltonian,

$$\hat{H}_{\rm so}=\frac{\hbar}{(2mc)^2}\hat{\boldsymbol{\sigma}}\cdot[(\nabla U)\times\hat{\mathbf{p}}]\equiv\frac{q}{2m^2c^2}\hat{\mathbf{S}}\cdot[(\nabla\phi)\times\hat{\mathbf{p}}]. \tag{9.134}$$

For a central potential $\phi(r)$, its gradient has only the radial component: $\nabla\phi=(d\phi/dr)\mathbf{r}/r=-\mathscr{E}\mathbf{r}/r$, and with the angular momentum definition (5.147), Eq. (134) is (finally!) reduced to Eq. (122).

As was shown in Sec. 6.3, the perturbative treatment of Eq. (122), together with the kinetic-relativistic correction (6.47), in the hydrogen-like atom/ion problem, leads to the fine structure of each Bohr level E_n, given by Eq. (6.60):

$$\Delta E_{\rm fine}=\frac{E_n}{2mc^2}\left(3-\frac{4n}{j+1/2}\right)\equiv-mc^2\frac{Z^4\alpha^4}{2n^4}\left(\frac{n}{j+1/2}-\frac{3}{4}\right). \tag{9.135}$$

This result receives a confirmation from the surprising fact that for the hydrogen-like atom/ion problem, the Dirac equation may be solved *exactly* – without any assumptions. I would not have time/space to reproduce the solution,[61] and will only list the final result for the energy spectrum:

H-like atom/ion: energies

$$\frac{E}{mc^2}=\left(1+\frac{Z^2\alpha^2}{\left\{n+\left[(j+1/2)^2-Z^2\alpha^2\right]^{1/2}-(j+1/2)\right\}^2}\right)^{-1/2}. \tag{9.136}$$

Here n = 1, 2, … is the same principal quantum number as in Bohr's theory, while j is the quantum number specifying the eigenvalues (5.175) of J^2, in our case of a spin-½ particle taking half-integer values: $j = l \pm 1/2$ = 1/2, 3/2, 5/2, … – see Eq. (5.189). This is natural because due to the spin-orbit interaction, the orbital momentum and spin are not conserved, while their vector sum, $\mathbf{J}=\mathbf{L}+\mathbf{S}$, is. Each energy level (136), besides that of the ground state with n = 1 and j + ½ = 1,[62] is doubly degenerate, with two eigenstates representing two directions of the spin. (In the low-energy limit, we may say: corresponding to two values of $l = j \mp 1/2$, at fixed j.)

[60] The first term gives a small spin-independent energy shift, which is very difficult to verify experimentally.

[61] Good descriptions of the solution are available in several textbooks – see, e.g., Sec. 24.9 in E. Merzbacher, *Quantum Mechanics*, 3rd ed., Wiley (1998).

[62] Note that for that level, the theory yields $E_g/mc^2=(1-Z^2\alpha^2)^{1/2}$, so it is consistent only for $Z<1/\alpha\approx137$.

Speaking of that limit (when $E - mc^2 \sim E_H << mc^2$): since according to Eq. (1.13) for E_H, the square of the fine-structure constant $\alpha \equiv e^2/4\pi\varepsilon_0\hbar c$ may be represented as the ratio E_H/mc^2, we may follow this limit by expanding Eq. (136) into the Taylor series in $(Z\alpha)^2 << 1$. The result,

$$E \approx mc^2\left[1-\frac{Z^2\alpha^2}{2n^2}-\frac{Z^4\alpha^4}{2n^4}\left(\frac{n}{j+½}-\frac{3}{4}\right)\right], \tag{9.137}$$

has the same structure, and allows the same interpretation as Eq. (92), but with the last term coinciding with Eq. (135) – and with experimental results. Historically, this correct description of the fine structure of the atomic levels provided decisive proof of Dirac's theory.

However, even such an impressive theory does not have too many direct applications. The main reason for that was already mentioned at the end of Sec. 5: due to the possibility of creation and annihilation of particle-antiparticle pairs by an energy influx higher than $2mc^2$, the number of particles participating in high-energy interactions is not fixed. An adequate general description of such situations is given by the quantum field theory, in which the particle's wavefunction is treated as a field to be quantized, using so-called *field operators* $\hat{\Psi}(\mathbf{r},t)$– very much similar to the electromagnetic field operators (16). The Dirac equation follows from such a theory in the single-particle approximation.

As was mentioned earlier on several occasions, the quantum field theory is beyond the time/space limits of this course, and I have to stop here, referring the interested reader to one of several excellent textbooks on this discipline.[63] However, I would strongly encourage the students going in this direction to start by playing with the field operators on their own, taking clues from Eqs. (16) but replacing the creation/annihilation operators $\hat{a}_j^\dagger$ and $\hat{a}_j$ of the electromagnetic field oscillators with those of the general second quantization formalism outlined in Sec. 8.3.

9.8. Exercise problems

9.1. Prove the Casimir formula (23) by calculating the net force $F = \mathcal{P}A$ exerted by the electromagnetic field, in its ground state, on two perfectly conducting parallel plates of area A, separated by a free-space gap of width $t << A^{1/2}$.

Hint: Calculate the field's energy in the gap's volume with and without the account of the plate effect, and then apply the Euler-Maclaurin formula[64] to the difference between these two results.

9.2. Electromagnetic radiation by some single-mode quantum sources may have such a high degree of coherence that it is possible to observe the interference of waves from two independent sources with virtually the same frequency, incident on one detector.

(i) Generalize Eq. (29) to this case.

(ii) Use this generalized expression to show that incident waves in different Fock states do not create an interference pattern.

[63] For a gradual introduction see, e.g., either L. Brown, *Quantum Field Theory*, Cambridge U. Press (1994) or R. Klauber, *Student Friendly Quantum Field Theory*, Sandtrove (2013). On the other hand, M. Srednicki, *Quantum Field Theory*, Cambridge U. Press (2007) and A. Zee, *Quantum Field Theory in a Nutshell*, 2nd ed., Princeton (2010), among others, offer steeper learning curves.

[64] See, e.g., MA Eq. (2.12a).

9.3. Calculate the zero-delay value $g^{(2)}(0)$ of the second-order correlation function of a single-mode electromagnetic field in the so-called *cat state* (see Problem 7.4): a coherent superposition of two Glauber states with equal but sign-opposite parameters α and a certain phase shift between them.

9.4. Calculate the zero-delay value $g^{(2)}(0)$ of the second-order correlation function of a single-mode electromagnetic field in the squeezed ground state ζ defined by Eq. (5.142).

9.5. Calculate the rate of spontaneous photon emission (into unrestricted free space) by a hydrogen atom, initially in the 2*p* state ($n = 2$, $l = 1$) with $m = 0$. Would the result be different for $m = \pm 1$? for the 2*s* state ($n = 2$, $l = 0$, $m = 0$)? Discuss the relation between these quantum-mechanical results and those given by the classical theory of radiation for the simplest classical model of the atom.

9.6. An electron has been placed on the lowest excited level of a spherically symmetric, quadratic potential well $U(\mathbf{r}) = m_e\omega^2 r^2/2$. Calculate the rate of its relaxation to the ground state, with the emission of a photon (into unrestricted free space). Compare the rate with that for a similar transition of the hydrogen atom, for the case when the radiation frequencies of these two systems are equal.

9.7. Derive an analog of Eq. (53) for the spontaneous photon emission (into free space) due to a change of the *magnetic* dipole moment **m** of a small-size system.

9.8. A spin-½ particle with a gyromagnetic ratio γ is in its orbital ground state in a static magnetic field $\mathscr{B}_0$. Calculate the rate of its spontaneous transition from the higher to the lower energy level, with the emission of a photon into unrestricted free space. Evaluate this rate for an electron in a field of 10 T, and discuss the implications of this result for laboratory experiments with electron spins.

9.9. Calculate the rate of spontaneous transitions between the two sublevels of the ground state of a hydrogen atom, formed as a result of its hyperfine splitting. Discuss the implications of the result for the width of the 21-cm spectral line of hydrogen.

9.10. Find the eigenstates and eigenvalues of the Jaynes-Cummings Hamiltonian (78), and discuss their behavior near the resonance point $\omega = \Omega$.

9.11. Analyze the Purcell effect, mentioned in Secs. 3 and 4, quantitatively. In particular, calculate the so-called *Purcell factor* F_P defined as the ratio of the rate Γ_s of an atom's spontaneous emission into a resonant cavity tuned exactly to the quantum transition frequency, to that into free space.

9.12. Use Eq. (84) to prove the continuity relation (1.52) with the probability density w and current density **j** given by Eqs. (89).

9.13. Prove that the Klein-Gordon equation (84) may be rewritten in a form similar to the non-relativistic Schrödinger equation (1.25) but for a two-component wavefunction, with the Hamiltonian represented (in the usual *z*-basis) by the following 2×2-matrix:

$$\mathrm{H} = -\left(\sigma_z + i\sigma_y\right)\frac{\hbar^2}{2m}\nabla^2 + mc^2\sigma_z .$$

Use your solution to discuss the physical meaning of the wavefunction's components.

9.14. Calculate and discuss the energy spectrum of a relativistic, spinless, charged particle placed into an external uniform, time-independent magnetic field $\mathscr{B}$. Use the result to formulate the condition of validity of the non-relativistic theory for this case.

9.15. Prove Eq. (91) for the energy spectrum of a hydrogen-like atom/ion, starting from the relativistic Schrödinger equation.

Hint: A mathematical analysis of Eq. (3.193) shows that its eigenvalues are given by Eq. (3.201), $\varepsilon_n = -1/2n^2$, with $n = l + 1 + n_r$, where $n_r = 0, 1, 2, \ldots$, even if the parameter l is not an integer.

9.16. Derive a general expression for the differential cross-section of elastic scattering of a spinless relativistic particle by a static potential $U(\mathbf{r})$, in the Born approximation, and formulate the conditions of its validity. Use these results to calculate the differential cross-section of scattering of a particle with the electric charge $-e$ by the Coulomb electrostatic potential $\phi(\mathbf{r}) = Ze/4\pi\varepsilon_0 r$.

9.17. Starting from Eqs. (95)-(98), prove that the probability density w given by Eq. (101) and the probability current density $\mathbf{j}$ defined by Eq. (102) do indeed satisfy the continuity equation (1.52): $\partial w/\partial t + \nabla\cdot\mathbf{j} = 0$.

9.18. Calculate the commutator of the operator $\hat{L}^2$ and the Dirac Hamiltonian of a free particle. Compare the result with that for the non-relativistic Hamiltonian, and interpret the difference.

9.19. Calculate commutators of the operators $\hat{S}^2$ and $\hat{J}^2$ with the Dirac Hamiltonian (97) and give an interpretation of the results.

9.20. In the Heisenberg picture of quantum dynamics, derive an equation describing the time evolution of a free electron's velocity in the Dirac theory. Solve the equation for the simplest state with definite energy and momentum, and interpret the oscillations (so-called *Zitterbewegung* or "trembling motion") appearing in the solution.

9.21. Calculate the energy spectrum of a relativistic spin-½ particle with an electric charge q, placed into a time-independent uniform external magnetic field $\mathscr{B}$. Compare the calculated spectrum with that following from the non-relativistic theory and the relativistic Schrödinger equation.

9.22.* Following the discussion at the end of Section 7, introduce quantum field operators $\hat{\psi}$ that would be related to the usual wavefunctions ψ just as the electromagnetic field operators (16) are related to the classical electromagnetic fields, and explore the basic properties of these operators. (For this preliminary study, consider the fixed-time situation.)

Chapter 10. Making Sense of Quantum Mechanics

This (rather brief) chapter addresses the conceptually important issues of quantum measurements and quantum state interpretation. Please note that some of these issues are still subjects of debate[1] *– fortunately not affecting the quantum mechanics' practical results discussed in the previous chapters.*

10.1. Quantum measurements

The knowledge base outlined in the previous chapters gives us a sufficient background for a (by necessity, very brief) discussion of *quantum measurements*.[2] Let me start by reminding the reader of the only quantum theory's postulate that relates it to experiment. In the simplest case when the system is in a coherent (pure) quantum state, its ket-vector may be represented as a linear superposition

$$|\alpha\rangle = \sum_j \alpha_j |a_j\rangle, \tag{10.1}$$

where a_j are the eigenstates of the operator of an observable A, related to its eigenvalues A_j by Eq. (4.68):

$$\hat{A}|a_j\rangle = A_j|a_j\rangle. \tag{10.2}$$

In such a state, the outcome of a single measurement (at this stage, meaning a *perfect* measurement) of the observable A may be uncertain but is restricted to the set of eigenvalues A_j, with the j^{th} outcome probability equal to

$$W_j = |\alpha_j|^2. \tag{10.3}$$

As was discussed in Chapter 7, the state of the system (or rather of the statistical ensemble of macroscopically similar systems we are using for this particular series of similar experiments) may be mixed rather than pure, and hence even more uncertain than the state described by Eq. (1). Hence, the measurement postulate means that even if the system is in its least uncertain state, the measurement outcomes are *still* probabilistic.[3]

If we believe that each particular measurement may be done perfectly, and do not worry too much about how exactly, we are subscribing to the *mathematical* notion of measurement, that was, rather reluctantly, used in these notes – up to this point. However, the actual (*physical*) measurements are *always* imperfect, first of all, because of the huge gap between the energy-time scale $\hbar \sim 10^{-34}$ J·s of the quantum phenomena in "microscopic" systems such as atoms, and the "macroscopic" scale of the direct human perception, so the role of the instruments bridging this gap (Fig. 1), is highly nontrivial.

[1] For an excellent review of these controversies, as presented in a few leading textbooks, I highly recommend J. Bell's paper published in the collection by A. Miller (ed.), *Sixty-Two Years of Uncertainty*, Plenum, 1989.

[2] "Quantum measurements" is a rather unfortunate and misleading term; it would be more sensible to speak about "measurements of observables in quantum mechanical systems". However, the former term is so common and compact that I will use it – albeit rather reluctantly.

[3] The measurement outcomes become definite only in the trivial case when the system is definitely in one of the eigenstates a_j, say a_0; then $\alpha_j = \delta_{j,0}\exp\{i\varphi\}$, and $W_j = \delta_{j,0}$.

© K. Likharev

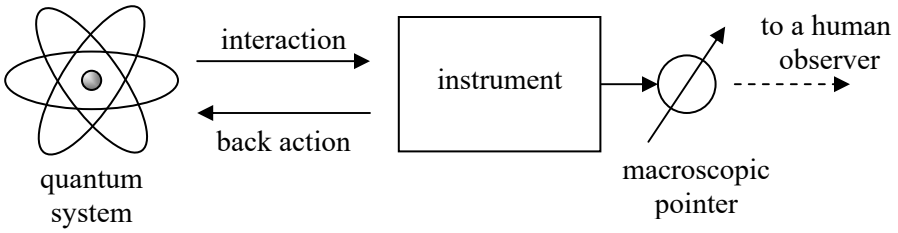


Fig.10.1. General scheme of a quantum measurement.

Besides the famous Bohr-Einstein discussion in the mid-1930s, which will be briefly reviewed in Sec. 3, the founding fathers of quantum mechanics have not paid much attention to these issues, apparently because of the following reason. At that time it looked like the experimental instruments (at least the best of them :-) were doing exactly what the measurement postulate was telling. For example, the *z*-oriented Stern-Gerlach experiment (Fig. 4.1) turns two complex coefficients $\alpha_\uparrow$ and $\alpha_\downarrow$ describing the spin state of the incoming particles, into a set of particle-counter clicks, with the rates proportional to, respectively, $|\alpha_\uparrow|^2$ and $|\alpha_\downarrow|^2$. The complex internal nature of these instruments makes more detailed questions unnatural. For example, each click of a Geiger counter involves an effective disappearance of the observed particle in a zillion-particle electric discharge avalanche it has triggered. A century ago, it looked much more important to extend the newborn quantum mechanics to more complex systems (such as atomic nuclei, etc.) than to think about the physics of such instruments.

However, since that time the experimental techniques, notably including high-vacuum and low-temperature systems, micro- and nano-fabrication, and low-noise electronics, have improved quite dramatically. In particular, we now may observe the quantum-mechanical behavior of more and more macroscopic objects – such as the mechanical oscillators mentioned in Sec. 2.9. Moreover, some "macroscopic quantum systems" (in particular, special systems of Josephson junctions, see below) have properties enabling their use as essential parts of measurement setups. Such developments are making the line separating the "micro" and "macro" worlds finer and finer, so more inquisitive questions about the physical nature of quantum measurements are not so hopeless now. In my personal scheme of these developments, the main questions may be grouped as follows:

(i) Does a quantum measurement involve any laws besides those of quantum mechanics? In particular, should it necessarily involve a human/intelligent observer? (The last question is not as laughable as it may look – see below.)

(ii) What is the state of the measured system *just after* a *single-shot measurement* – meaning a measurement process limited to a time interval much shorter than the time scale of the measured system's evolution? (This question is a necessary part of any discussion of *repeated measurements* and of their ultimate form – *continuous monitoring* of a certain observable.)

(iii) If a measurement of an observable A has produced a certain outcome A_j, what statements may be made about the state of the system *just before* the measurement? (This question is most closely related to various interpretations of quantum mechanics.)

Let me discuss these issues in the listed order. First of all, I am happy to report that there is a virtual consensus of physicists on some aspects of these issues. According to this consensus, any reasonable quantum measurement needs to result in a certain, distinguishable state of a *macroscopic* output component of the measurement instrument – see Fig. 1. (Traditionally, its component is called a *pointer*, though its role may be played by a printer or a plotter, an electronic circuit sending out the

result as a number, etc.). This requirement implies that the measurement process should have the following features:

– provide a large "signal gain", i.e. some means of mapping the quantum state with its $\hbar$-scale of action (i.e. of the energy-by-time product) onto a macroscopic position of the pointer with a much larger action scale, and

– if we want to approach the fundamental limit of uncertainty, the instrument should introduce as little additional fluctuation ("noise") as permitted by the laws of physics.

Both these requirements *are* fulfilled in a well-designed Stern-Gerlach experiment – see Fig. 4.1 again. Indeed, the magnetic field gradient, splitting the particle beam, turns the minuscule (microscopic) energy difference (4.167) between two spin-polarized states into a macroscopic difference between the final positions of two output beams, where two particle detectors may be located. However, as was noted above, the internal physics of the particle detectors (say, Geiger counters) at this measurement is rather complicated, and would not allow us to discuss some aspects of the measurement, in particular. to answer the second question we are working on.

This is why let me describe the scheme of an almost similar single-shot measurement of a two-level quantum system, which shares the simplicity, high gain, and low internal noise of the Stern-Gerlach apparatus, but has an advantage that at its certain hardware implementations,[4] the measurement process allows a thorough, quantitative theoretical description. Let us consider a 1D particle confined in a double-well potential (Fig. 2), where x is some continuous generalized coordinate – not necessarily a mechanical displacement. Let the particle be initially in a pure quantum state, with the energy close to the wells' bottom. Then, as we know from the discussion of such systems in Secs. 2.6 and 5.1, the state may be represented by a ket-vector similar to that of spin-½:

$$\left|\alpha\right\rangle = \alpha_{\rightarrow}\left|\rightarrow\right\rangle + \alpha_{\leftarrow}\left|\leftarrow\right\rangle, \tag{10.4}$$

where the component states $\rightarrow$ and $\leftarrow$ are described by wavefunctions localized near the potential well bottoms at $x \approx \pm x_0$ – see the blue lines in Fig. 2. Our goal is to measure in which well the particle resides at a certain time instant, say at $t = 0$. For that, let us rapidly change, at that moment, the potential profile of the system, so at $t > 0$, near the origin, it may be well described by an inverted parabola:

$$U(x) \approx -\frac{m\lambda^2}{2}x^2, \qquad \text{for } t > 0, \quad \left|x\right| << x_{\text{f}}. \tag{10.5}$$

[4] The scheme may be implemented, for example, using a simple Josephson-junction circuit called the *balanced comparator* – see, e.g., T. Walls *et al*., *IEEE Trans. on Appl. Supercond*. **17**, 136 (2007), and references therein. (RF devices with similar properties are currently called "bifurcation detectors" or bifurcation amplifiers".) Experiments have demonstrated that this system may have a measurement variance dominated by the theoretically expected quantum-mechanical uncertainty, at quite practicable experimental conditions (at temperatures of the order of 1K). A conceptual advantage of this system is that it is based on externally-shunted Josephson junctions, i.e. the devices whose quantum-mechanical model, including the coupling to the environment, is in a quantitative agreement with experiment – see, e.g., D. Schwartz *et al*., *Phys. Rev. Lett*. **55**, 1547 (1985). Colloquially, the balanced comparator is a high-gain instrument with a "well-documented Hamiltonian", eliminating the need for speculations about the environmental effects. In particular, the dephasing process in it, and its time T_2, are well described by Eqs. (7.89) and (7.142), with the coefficient η equal to the Ohmic conductance G of the shunt.

It is straightforward to verify that the Heisenberg equations of motion in such an inverted potential describe an exponential growth of the operator $\hat{x}$ in time (proportional to $\exp\{\lambda t\}$) and hence a similar, proportional growth of the expectation value $\langle x \rangle$ and its r.m.s. uncertainty δx.[5] At this "inflation" stage, the coherence between the two component states $\rightarrow$ and $\leftarrow$ is still preserved, i.e. the time evolution of the system is, in principle, reversible.

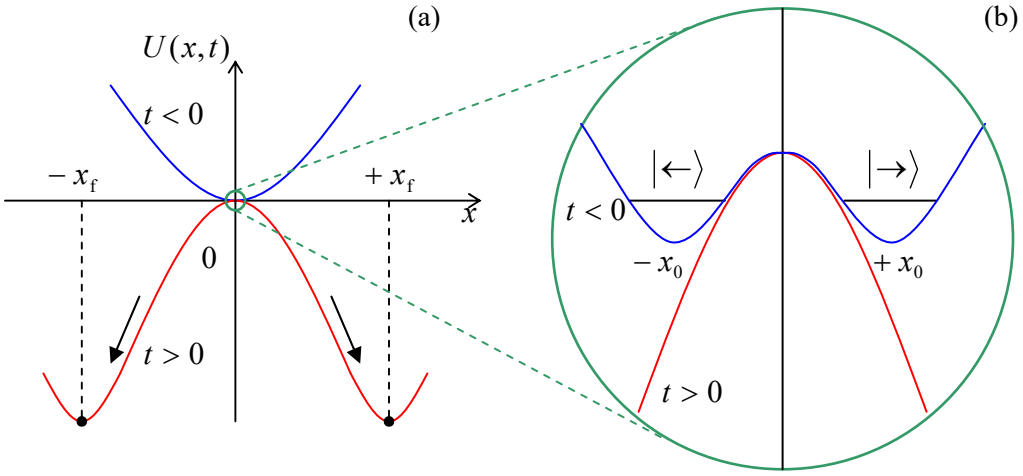


Fig. 10.2. The confining potential's inversion, as viewed on the (a) "macroscopic" and (b) "microscopic" scales of the generalized coordinate x.

Now let the system be weakly coupled, also at $t > 0$, to a dissipative (e.g., Ohmic) environment. As we know from Chapter 7, such coupling ensures the state's dephasing on some time scale T_2. If

$$x_0 << x_0 \exp\{\lambda T_2\}, x_f, \tag{10.6}$$

then the process, after the potential inversion, consists of two stages, well separated in time:

– the already discussed "inflation" stage, preserving the state components' coherence, and

– the dephasing stage, at which the coherence of the component states $\rightarrow$ and $\leftarrow$ is gradually suppressed as described by Eq. (7.89), i.e. the density matrix of the system is gradually reduced to the diagonal form describing a classical mixture of two probability packets with the probabilities (3) equal to, respectively, $W_{\rightarrow} = |\alpha_{\rightarrow}|^2$ and $W_{\leftarrow} = |\alpha_{\leftarrow}|^2 \equiv 1 - |\alpha_{\rightarrow}|^2$.

Besides dephasing, the environment gives the motion certain kinematic friction, with the drag coefficient η (7.141), so the system eventually settles to rest at one of the macroscopically separated minima $x = \pm x_f$ of the inverted potential (Fig. 2a), thus ensuring a high "signal gain" $x_f/x_0 >> 1$. As a result, the final probability density distribution $w(x)$ along the x-axis has two narrow, well-separated peaks. But this is just the situation that was discussed in Sec. 2.5 – see, in particular, Fig. 2.17. Since that discussion is very important, let me repeat – or rather rephrase it. The final state of the system is a classical mixture of two well-separated states, with the respective probabilities $W_{\leftarrow}$ and $W_{\rightarrow}$, whose sum equals 1. Now let us use some detector to test whether the system is in one of these states – say the right

[5] Somewhat counter-intuitively, the latter growth *improves* the measurement's fidelity. Indeed, it does not affect the intrinsic "signal-to-noise ratio" $\delta x/\langle x \rangle$, while making the intrinsic (say, quantum-mechanical) uncertainty much larger than the possible noise contribution by the later measurement stage(s).

one. (If x_f is sufficiently large, the noise contribution of this detector to the measurement uncertainty is negligible,[6] and its physics is unimportant.) If the system has been found at this location (again, the probability of this outcome is $W_{\rightarrow} = |\alpha_{\rightarrow}|^2$), the probability of finding it at the counterpart (left) location at a consequent detection turns to zero.

This probability "reduction" is a purely classical (or if you like, mathematical) effect of the statistical ensemble's re-definition: $W_{\leftarrow}$ equals zero not in the initial ensemble of all similar experiments (where is equals $|\alpha_{\leftarrow}|^2$), but only in the re-defined ensemble of experiments in that the system had been found at the right location. Of course, which ensemble to use, i.e. what probabilities to register/publish is a purely accounting decision, which should be made by a human (or otherwise intelligent :-) observer. If we are only interested in an objective recording of the results of a pre-fixed sequence of experiments (i.e. the members of a pre-defined, fixed statistical ensemble), there is no need to include such an observer in any discussion. In any case, this detection/registration process, very common in classical statistics, leaves no space for any mysterious "wave packet reduction" – understood as a hypothetical process that would not obey the regular laws of quantum mechanical evolution.

The state dephasing and ensemble re-definition at measurements are also at the core of several paradoxes, of which the so-called *quantum Zeno paradox* is perhaps the most spectacular.[7] Let us return to a two-level system with the unperturbed Hamiltonian given by Eq. (4.166), the quantum oscillation period $2\pi/\Omega$ much longer than the single-shot measurement time, and the system initially (at $t = 0$) definitely in one of the partial quantum states – for example, a certain potential well of the double-well potential. Then, as we know from Secs. 2.6 and 4.6, the probability to find the system in this initial state at time $t > 0$ is

$$W(t) = \cos^2 \frac{\Omega t}{2} \equiv 1 - \sin^2 \frac{\Omega t}{2}. \tag{10.7}$$

If the time is small enough ($t = dt << 1/\Omega$), we may use the Taylor expansion to write

$$W(dt) \approx 1 - \frac{\Omega^2 dt^2}{4}. \tag{10.8}$$

Now, let us use some good measurement scheme (say, the potential inversion discussed above) to measure whether the system is still in this initial state. If it is (as Eq. (8) shows, the probability of such an outcome is nearly 100%), then the system, after the measurement, is in the same state. Let us allow it to evolve again, following the same Hamiltonian. Then the evolution of W will follow the same law as in Eq. (7). Thus, when the system is measured again at time $2dt$, the probability to find it in the same state both times is

$$W(2dt) \approx W(dt)\left(1 - \frac{\Omega^2 dt^2}{4}\right) = \left(1 - \frac{\Omega^2 dt^2}{4}\right)^2. \tag{10.9}$$

[6] At the balanced-comparator implementation mentioned above, the final state detection may be readily performed using a "SQUID" magnetometer based on the same Josephson junction technology – see, e.g., EM Sec. 6.5. In this case, the distance between the potential minima $\pm x_f$ is close to one superconducting flux quantum (3.38), while the additional uncertainty induced by the SQUID may be as low as a few millionths of that amount.

[7] This name, coined by E. Sudarshan and B. Mishra in 1997 (though the paradox had been discussed in detail by A. Turing in 1954) is due to its superficial similarity to the classical paradoxes by the ancient Greek philosopher Zeno of Elea.

After repeating this cycle N times (with the total time $t = Ndt$ still much less than $N^{1/2}/\Omega$), the probability that the system is still in its initial state is

$$W(Ndt) \equiv W(t) \approx \left(1 - \frac{\Omega^2 dt^2}{4}\right)^N = \left(1 - \frac{\Omega^2 t^2}{4N^2}\right)^N \approx 1 - \frac{\Omega^2 t^2}{4N}. \tag{10.10}$$

Comparing this result with Eq. (7), we see that the process of the system's transfer to the opposite partial state has been slowed down rather dramatically, and in the limit $N \to \infty$ (at fixed t), its evolution is virtually stopped by the measurement process. There is of course nothing mysterious here; the evolution slowdown is due to the state dephasing and the statistical ensemble re-definition at each measurement.

This may be the only acceptable occasion for me to mention, very briefly, one more famous – or rather infamous *Schrödinger cat paradox*, so much overplayed in popular publications.[8] For this thought experiment, there is no need to discuss the (rather complicated :-) physics of the cat. As soon as the charged particle produced at the radioactive decay reaches the Geiger counter, the initial coherent superposition of the two possible quantum states ("the decay has happened"/"the decay has not happened") of the system is rapidly dephased, i.e. reduced to their classical mixture, leading, correspondingly, to the classical mixture of the final macroscopic states "cat dead"/"cat alive". So, despite attempts by numerous authors lacking a proper physics background to represent this situation as a mystery whose discussion needs the involvement of professional philosophers, hopefully, the reader of these notes knows enough about dephasing from Chapter 7 to ignore all this babble.

10.2. QND measurements

I hope that the above discussion has sufficiently illuminated the issues of group (i), so let me proceed to question group (ii), in particular to the general issue of the *back action* of the instrument upon the system under measurement – symbolized with the back arrow in Fig. 1. In the instruments like the Geiger counter, such back action is large: the instrument essentially destroys ("demolishes") the state of the system under measurement. Even the "cleaner" potential-inversion measurement illustrated by Fig. 2 fully destroys the initial coherence of the system's states, i.e. perturbs it rather substantially.

However, in the 1970s it was understood that this is not really necessary. For example, in Sec. 7.3, we have already discussed an example of a two-level system coupled with its environment and described by the Hamiltonian (7.68)-(7.70):

$$\hat{H} = \hat{H}_s + \hat{H}_{\rm int} + \hat{H}_e\{\lambda\}, \quad \text{with } \hat{H}_s = c_z\hat{\sigma}_z, \text{ and } \hat{H}_{\rm int} = -f\{\lambda\}\hat{\sigma}_z, \tag{10.11}$$

so

$$\left[\hat{H}_s, \hat{H}_{\rm int}\right] = 0. \tag{10.12}$$

Comparing this equality with Eq. (4.199), applied to the explicitly-time-independent Hamiltonian $\hat{H}_s$,

$$i\hbar\dot{\hat{H}}_s = \left[\hat{H}_s, \hat{H}\right] \equiv \left[\hat{H}_s, \left(\hat{H}_s + \hat{H}_{\rm int} + \hat{H}_e\{\lambda\}\right)\right] = \left[\hat{H}_s, \hat{H}_{\rm int}\right] = 0, \tag{10.13}$$

[8] I fully agree with S. Hawking who has been quoted to say, "When I hear about the Schrödinger cat, I reach for my gun." The only good aspect of this popularity is that the formulation of this paradox should be so well known to the reader that I do not need to waste time/space repeating it.

we see that in the Heisenberg picture, the Hamiltonian operator (and hence the energy) of the system of our interest does not change in time. On the other hand, if the "environment" in this discussion is the instrument used for the measurement (see Fig. 1 again), then the interaction can change its state, so it may be used to measure the system's energy – or another observable whose operator commutes with the interaction Hamiltonian. Such a trick is called the *quantum non-demolition* (QND), or sometimes "back-action-evading" measurements.[9] Due to the lack of back action of the instrument on the corresponding variable, such measurements allow its continuous monitoring. Let me present a fine example of an actual measurement of this kind – see Fig. 3.[10]

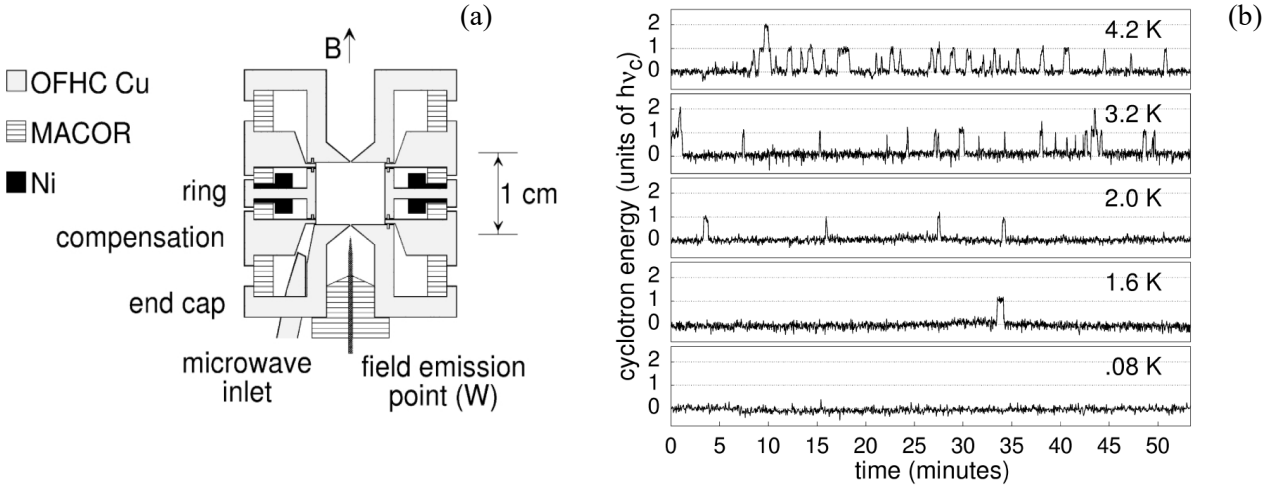


Fig. 10.3. QND measurements of single electron's energy by Peil and Gabrielse: (a) the experimental setup's core, and (b) a record of the thermal excitation and spontaneous relaxation of the Fock states. © 1999 APS; reproduced with permission.

In this experiment, a single electron is captured in a *Penning trap* – a combination of a (virtually) uniform magnetic field $\mathscr{B}$ and a quadrupole electric field.[11] This electric field stabilizes the cyclotron orbits but does not have any noticeable effect on electron motion in the plane normal to the magnetic field, and hence on its Landau level energies – see Eq. (3.50):

$$E_n = \hbar\omega_c\left(n+\frac{1}{2}\right), \quad \text{with } \omega_c = \frac{e\mathscr{B}}{m_e}. \tag{10.14}$$

(In the cited work, with $\mathscr{B} \approx 5.3$ T, the cyclic frequency $\omega_c/2\pi$ was about 147 GHz, so the Landau level splitting $\hbar\omega_c$ was close to 10^{-22} J, i.e. corresponded to $k_B T$ at $T \sim 10$ K, while the physical temperature of the system might be reduced well below that, down to 80 mK). Now note that the analogy between a Landau-level particle and a harmonic oscillator goes beyond the energy spectrum (14). Indeed, since the Hamiltonian of a 2D particle in a perpendicular magnetic field may be reduced to Eq. (3.47) similar to that of a 1D oscillator, we may repeat all procedures of Sec. 5.4 and rewrite this effective Hamiltonian in terms of the creation-annihilation operators – see Eq. (5.72):

[9] For a more detailed discussion, see, e.g., V. Braginsky and F. Y. Khalili, *Rev. Mod. Phys*. **68**, 1 (1996).
[10] S. Peil and G. Gabrielse, *Phys. Rev. Lett.* **83,** 1287 (1999).
[11] It is similar to the 2D system discussed in EM Sec. 2.7 but with additional rotation about one of the axes.

$$\hat{H}_s = \hbar\omega_c\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right). \tag{10.15}$$

In the Peil and Gabrielse experiment, the trapped electron had one more degree of freedom – along the magnetic field. The electric field of the Penning trap created a very soft confining potential along this direction (vertical in Fig. 3a; I will take it for the *z*-axis), so small electron oscillations along that axis could be well described as those of a 1D harmonic oscillator of a much lower frequency, in that particular experiment with $\omega_z/2\pi \approx 64$ MHz. This frequency could be measured very accurately (with an error of ~1 Hz) by sensitive electronics whose electric field does affect the *z*-motion of the electron, but not its motion in the perpendicular plane. In an exactly uniform magnetic field, the two modes of electron motion would be completely uncoupled. However, the experimental setup included two special superconducting rings made of niobium (see Fig. 3a), which slightly distorted the magnetic field and created an interaction between the modes, which might be well approximated by the Hamiltonian[12]

$$\hat{H}_{\rm int} = \text{const}\times\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right)\hat{z}^2, \tag{10.16}$$

so the main condition (12) of a QND measurement was very closely satisfied. At the same time, the coupling (16) ensured that a change of the Landau level number *n* by 1 changed the *z*-oscillation eigenfrequency by ~12.4 Hz. Since this shift was substantially larger than the used electronics' noise, rare spontaneous changes of *n* (due to a weak coupling of the electron to the environment) could be readily measured – moreover, continuously monitored – see Fig. 3b. The record shows spontaneous excitations of the electron to higher Landau levels, with its sequential relaxation, just as described by Eqs. (7.208)-(7.210). The detailed data statistics analysis showed that there was virtually no effect of the measuring instrument on these processes – at least on the scale of minutes, i.e. as many as $\sim 10^{13}$ cyclotron orbit periods.[13]

It is important, however, to note that any measurement – QND or not – cannot avoid the uncertainty relations between incompatible variables; in the particular case described above, continuous monitoring of the Landau state number *n* does not allow the simultaneous monitoring of its quantum phase (which may be defined exactly as in the harmonic oscillator). In this context, it is natural to wonder whether the QND measurement concept may be extended from quadratic-form variables like energy to "usual" observables such as coordinates and momenta whose uncertainties are bound by the ordinary Heisenberg's relation (1.35). The answer is yes, but the required methods are a bit more tricky.

For example, let us place an electrically charged particle into a uniform electric field $\boldsymbol{\mathscr{E}} = \mathbf{n}_x\mathscr{E}(t)$ of an instrument, so their interaction Hamiltonian is

$$\hat{H}_{\rm int} = -q\hat{\mathscr{E}}(t)\hat{x}. \tag{10.17}$$

Such interaction may certainly pass the information on the time evolution of the coordinate *x* to the instrument. However, in this case, Eq. (12) is *not* satisfied – at least for the kinetic-energy part of the

[12] Here I am simplifying the real situation a bit. Actually, in that experiment, there was an electron spin's contribution to the interaction Hamiltonian as well, but since the used high magnetic field polarized the spins quite reliably, their only effect was a constant shift of the frequency ω_z, which is not important for our discussion.

[13] See also the conceptually similar experiments, performed by different means: G. Nogues *et al*., *Nature* **400**, 239 (1999).

particle's Hamiltonian; as a result, the interaction distorts its time evolution. Indeed, by writing the Heisenberg equation (4.199) for the x-component of the momentum, we get

$$\dot{\hat{p}} - \dot{\hat{p}}\Big|_{\mathscr{E}=0} = q\hat{\mathscr{E}}(t)\,. \tag{10.18}$$

On the other hand, integrating Eq. (5.139) for the coordinate operator evolution,[14] we get the expression

$$\hat{x}(t) = \hat{x}(t_0) + \frac{1}{m}\int_{t_0}^{t} \hat{p}(t')dt'\,, \tag{10.19}$$

which shows that the perturbations (18) of the momentum eventually find their way to the coordinate evolution, not allowing its unperturbed sequential measurements.

However, for such an important particular system as a harmonic oscillator, the following trick is possible. For this system, Eqs. (5.139) with the addition (18) may be readily combined to give a second-order differential equation for the coordinate operator, that is absolutely similar to the classical equation of motion of the system, and has a similar solution:[15]

$$\hat{x}(t) = \hat{x}(t)\Big|_{\mathscr{E}=0} + \frac{q}{m\omega_0}\int_{-\infty}^{t} \hat{\mathscr{E}}(t')\sin\omega_0(t-t')dt'\,. \tag{10.20}$$

This formula confirms that generally, the external field $\mathscr{E}(t)$ (in our case, the sensing field of the measurement instrument) affects the time evolution law – of course. However, Eq. (20) shows that if the field is applied only at moments t'_n separated by intervals $\mathcal{T}/2$, where $\mathcal{T} \equiv 2\pi/\omega_0$ is the oscillation period, its effect on coordinate vanishes at similarly spaced observation instants $t_n = t_{n'} + (m + 1/2)\mathcal{T}$. This is the idea of *stroboscopic* QND measurements. Of course, according to Eq. (18), even such measurement perturbs the oscillator momentum, so even if the values x_n are measured with high accuracy, Heisenberg's uncertainty relation is not violated.

A direct implementation of stroboscopic measurements is technically complicated, but this initial idea has opened ways to more practicable solutions. For example, it is straightforward to use the Heisenberg equations of motion to show that if the coupling of two harmonic oscillators, with coordinates x and X, and unperturbed frequencies ω and Ω, is modulated in time as

$$\hat{H}_{\text{int}} \propto \hat{x}\hat{X}\cos\omega t\,\cos\Omega t\,, \tag{10.21}$$

then the process in one of the oscillators (say, that with frequency Ω) does not affect the dynamics of one of the *quadrature components* of the counterpart oscillator, defined by the following relations:[16]

[14] This simple relation is limited to 1D systems with Hamiltonians of the type (1.41), but by now the reader certainly knows enough to understand that this discussion may be readily generalized to many other systems.

[15] Note in particular that the function $\sin\omega_0\tau$ (with $\tau \equiv t - t'$) under the integral, divided by ω_0, is nothing more than the temporal Green's function $G(\tau)$ of a loss-free harmonic oscillator – see, e.g., CM Sec. 5.1.

[16] The physical sense of these relations should be clear from Fig. 5.8: they define a system of coordinates rotating clockwise with the angular velocity equal to ω, so the point representing unperturbed classical oscillations with that frequency is at rest in this rotating frame. The reader familiar with the classical theory of oscillations may notice that the observables x_1 and x_2 so defined are just the *Poincaré plane* coordinates ("RWA variables") – see, e.g., CM Sec. 5.3-5.6, and especially Fig. 5.9, where these coordinates are denoted as u and v.

$$\hat{x}_1 \equiv \hat{x}\cos\omega t - \frac{\hat{p}}{m\omega}\sin\omega t, \qquad \hat{x}_2 \equiv \hat{x}\sin\omega t + \frac{\hat{p}}{m\omega}\cos\omega t, \tag{10.22}$$

while this component's motion does affect the dynamics of one of the quadrature components of the counterpart oscillator. (For the counterpart quadrature components' couple, the information transfer goes in the opposite direction.) This scheme has been successfully used for QND measurements.[17]

Note that the last two QND measurement examples are based on the idea of a periodic change of a certain parameter in time – either in the short-pulse form or the sinusoidal form. If the only goal of a QND measurement is a sensitive measurement of a weak *classical force* acting on a quantum *probe system*, i.e. a 1D oscillator of eigenfrequency ω_0, it may be implemented much simpler – just by modulating an oscillator's parameter with a frequency $\omega \approx 2\omega_0$. From the classical dynamics, we know that if the depth of such modulation exceeds a certain threshold value, it results in the excitation of the degenerate parametric oscillations of frequency $\omega/2 \approx \omega_0$, with one of two opposite phases.[18] In the language of Eq. (22), the parametric excitation means an exponential growth of one of the quadrature components (with its sign depending on initial conditions), while the counterpart component is suppressed. Close to, but below the excitation threshold, the parameter modulation boosts all fluctuations of the almost-excited component, including its quantum-mechanical uncertainty, and suppresses (*squeezes*) those of the counterpart component. The result is a squeezed state, already discussed in Sec. 5.5 of this course (see in particular Eqs. (5.143) and Fig. 5.8), which allows one to notice the effect of an external force on the oscillator on the backdrop of a quantum uncertainty much smaller than the standard quantum limit (5.99).

In electrical engineering, this fact may be conveniently formulated in terms of *noise parameter* Θ_N of a *linear amplifier* – essentially the tool for continuous monitoring of an input "signal" – e.g., a microwave or optical waveform.[19] Namely, the Θ_N of "usual" (say, transistor or maser) amplifiers that are equally sensitive to both quadrature components of the signal, Θ_N has the minimum value $\hbar\omega/2$, due to the quantum uncertainty pertinent to the quantum state of the amplifier itself (which therefore plays the role of its "quantum noise") – the fact that was recognized in the early 1960s.[20] On the other hand, a degenerate parametric amplifier, sensitive to just one quadrature component, may have Θ_N well below $\hbar\omega/2$, due to its ground state's squeezing.[21]

Let me note that the parameter-modulation schemes of the QND measurements are not limited to harmonic oscillators, and may be applied to other important quantum systems, notably including two-

[17] The first, initially imperfect QND experiments were carried out by R. Slusher *et al*., *Phys. Rev. Lett.* **55**, 2409 (1985), and other groups soon after this, using nonlinear interactions of optical waves. Later, the results were much improved – see, e.g., P. Grangier *et al*., *Nature* **396**, 537 (1998), and references therein. Recently, such experiments were extended to mechanical systems – see, e.g., F. Lecocq *et al*., *Phys. Rev. X* **5**, 041037 (2015).

[18] See, e.g., CM Sec. 5.5, and also Fig. 5.8 and its discussion in Sec. 5.6.

[19] For a quantitative definition of the latter parameter, suitable for the quantum sensitivity range ($\Theta_N \sim \hbar\omega$) as well, see, e.g., I. Devyatov *et al*., *J. Appl. Phys*. **60**, 1808 (1986). In the classical noise limit ($\Theta_N >> \hbar\omega$), it coincides with $k_B T_N$, where T_N is a more popular measure of electronics' noise, called the *noise temperature*.

[20] See, e.g., H. Haus and J. Mullen, *Phys. Rev*. **128**, 2407 (1962).

[21] See, e.g., the spectacular experiments by B. Yurke *et al*., *Phys. Rev. Lett*. **60**, 764 (1988). Note also that the squeezed ground states of light are now used to improve the sensitivity of interferometers in gravitational wave detectors – see, e.g., the review by R. Schnabel, *Phys. Repts*. **684**, 1 (2017), and the later papers by F. Acernese *et al*., *Phys. Rev. Lett*. **123**, 231108 (2019) and D. Ganapathy *et al*., *Phys. Rev. X* **13**, 041021 (2023).

level (i.e. spin-½-like) systems.[22] Such measurements may be an important tool for the further progress of quantum computation and cryptography.[23]

Finally, let me mention that the composite systems consisting of a quantum subsystem and a classical subsystem performing its continuous weakly-perturbing measurement and using its results for providing specially crafted feedback to the quantum subsystem, may have some curious properties, in particular mock a quantum system detached from the environment.[24]

10.3. Hidden variables and local reality

Now we are ready to proceed to the discussion of the last, hardest group (iii) of the questions posed in Sec. 1, namely on the state of a quantum system *just before* its measurement. After a very important but inconclusive discussion of this issue by Albert Einstein and his collaborators on one side, and Niels Bohr on the other side, in the mid-1930s, such discussions resumed in the 1950s.[25] They have led to a key contribution by John Stewart Bell in the early 1960s, summarized as so-called *Bell's inequalities*, and then to experimental work on better and better verification of these inequalities. (Besides that continuing work, the recent progress, in my humble view, has been rather marginal.)

The central question may be formulated as follows: what *had been* the "real" state of a quantum-mechanical system *just before* a virtually perfect single-shot measurement was performed on it, and gave a certain documented outcome? To be specific, let us focus again on the example of Stern-Gerlach measurements of spin-½ particles – because of their conceptual simplicity.[26] For a single-component system (in this case a single spin-½), the answer to the posed question may look evident. Indeed, as we know, if the spin is in a pure (least-uncertain) state α, i.e. its ket-vector may be expressed in the form similar to Eq. (4),

$$\left|\alpha\right\rangle = \alpha_{\uparrow}\left|\uparrow\right\rangle + \alpha_{\downarrow}\left|\downarrow\right\rangle, \tag{10.23}$$

where, as usual, $\uparrow$ and $\downarrow$ denote the states with definite spin orientations along the z-axis, the probabilities of the corresponding outcomes of the z-oriented Stern-Gerlach experiment are $W_{\uparrow} = |\alpha_{\uparrow}|^2$ and $W_{\downarrow} = |\alpha_{\downarrow}|^2$. Then it looks natural to suggest that if a particular experiment gave the outcome corresponding to the state $\uparrow$, the spin had been in that state just before the experiment. For a classical system such an answer would be certainly correct, and the fact that the probability $W_{\uparrow} = |\alpha_{\uparrow}|^2$ defined for the statistical ensemble of *all* experiments (regardless of their outcome), may be less than 1, would merely reflect our ignorance about the *real state* of this particular system before the measurement – which just *reveals* the real situation.

However, as was first argued in the famous *EPR paper* published in 1935 by A. Einstein, B. Podolsky, and N. Rosen, such an answer becomes impossible in the case of an entangled quantum system, if only one of its components is measured with an instrument. The original EPR paper discussed

[22] See, e.g., D. Averin, *Phys. Rev. Lett*. **88**, 207901 (2002).
[23] See, e.g., G. Jaeger, *Quantum Information*: *An Overview*, Springer, 2006.
[24] See, e.g., the monograph by H. Wiseman and G. Milburn, *Quantum Measurement and Control*, Cambridge U. Press (2009), more recent experiments by R. Vijay *et al*., *Nature* **490**, 77 (2012), and references therein.
[25] See, e.g., the collection by J. Wheeler and W. Zurek (eds.), *Quantum Theory and Measurement*, Princeton U. Press, 1983.
[26] As was discussed in Sec. 1, the Stern-Gerlach-type experiments may be readily made virtually perfect, provided that we do not care about the evolution of the system *after* the single-shot measurement.

thought experiments with a pair of 1D particles prepared in a quantum state in that both the *sum* of their momenta and the *difference* of their coordinates simultaneously have definite values: $p_1 + p_2 = 0$, $x_1 - x_2 = a$.[27] However, usually this discussion is recast into an equivalent Stern-Gerlach experiment shown in Fig. 4a.[28] A source emits rare pairs of spin-½ particles propagating in opposite directions. The particle spin states are random, but with the net spin of the pair is definitely equal to zero. After the spatial separation of the particles has become sufficiently large (see below), the spin state of each of them is measured with a Stern-Gerlach detector, with one of them (in Fig. 1, SG_1) somewhat closer to the particle source, so it makes the measurement first, at a time $t_1 < t_2$.

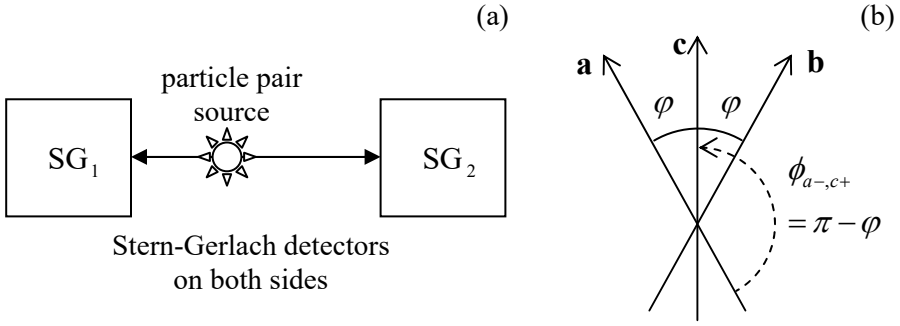


Fig. 10. 4. (a) General scheme of two-particle Stern-Gerlach experiments, and (b) the orientation of the detectors, assumed at Wigner's deviation of Bell's inequality (36).

First, let the detectors be oriented say along the same direction, say the z-axis. Evidently, the probability of each detector giving any of the values $s_z = \pm\hbar/2$ is 50%. However, if the first detector had given the result $S_z = -\hbar/2$, then even before the second detector's measurement, we know that it will give the result $S_z = +\hbar/2$ with 100% probability. So far, this situation still allows for a classical interpretation, just as for the single-particle measurements: we may fancy that the second particle has a definite spin before the measurement, and the first measurement just removes our ignorance about that reality. In other words, the change of the probability of the outcome $S_z = +\hbar/2$ at the second detection from 50% to 100% is due to the statistical ensemble re-definition: the 50% probability of this detection belongs to the ensemble of all experiments, while the 100% probability, to the sub-ensemble of experiments with the $S_z = -\hbar/2$ outcome of the first experiment.

However, let the source generate the spin pairs in the singlet state (8.18):

$$\left|s_{12}\right\rangle = \frac{1}{\sqrt{2}}\left(\left|\uparrow\downarrow\right\rangle - \left|\downarrow\uparrow\right\rangle\right). \tag{10.24}$$

As was discussed in Sec. 8.2, this state satisfies the above assumptions: the probability of each value of S_z of any particle is 50%, and the sum of both S_z is definitely zero, so if the first detector's result is $S_z = -\hbar/2$, then the state of the remaining particle is ↑, with zero uncertainty.[29] Now let us use Eqs. (4.123) to represent the same state (24) in a different form:

[27] This is possible because the corresponding operators commute: $\left[\hat{p}_1 + \hat{p}_2, \hat{x}_1 - \hat{x}_2\right] = \left[\hat{p}_1, \hat{x}_1\right] - \left[\hat{p}_2, \hat{x}_2\right] = 0$.

[28] This version was first proposed by D. Bohm in 1951. Another equivalent and experimentally more convenient (and as a result, frequently used) technique is the degenerate parametric excitation of entangled optical photon pairs – see, e.g., the publications cited at the end of this section.

[29] Here we assume that both detectors are perfect in the sense of their readout fidelity. As was discussed in Sec. 1, this condition may be closely approached in practical SG experiments.

$$|s_{12}\rangle = \frac{1}{\sqrt{2}}\left[\frac{1}{\sqrt{2}}\left(|\rightarrow\rangle + |\leftarrow\rangle\right)\frac{1}{\sqrt{2}}\left(|\rightarrow\rangle - |\leftarrow\rangle\right) - \frac{1}{\sqrt{2}}\left(|\rightarrow\rangle - |\leftarrow\rangle\right)\frac{1}{\sqrt{2}}\left(|\rightarrow\rangle + |\leftarrow\rangle\right)\right]. \quad (10.25)$$

Opening the parentheses (carefully, without swapping the ket-vector order, which encodes the particle numbers!), we get an expression similar to Eq. (24), but now for the *x*-basis:

$$|s_{12}\rangle = \frac{1}{\sqrt{2}}\left(|\rightarrow\leftarrow\rangle - |\leftarrow\rightarrow\rangle\right). \quad (10.26)$$

Hence if we use the first detector (closest to the particle source) to measure S_x rather than S_z, then after it had given a certain result (say, $S_x = -\hbar/2$), we know for sure, before the second particle spin's measurement, that its S_x component definitely equals $+\hbar/2$.

So, depending on the experiment performed on the first particle, the second particle, before its measurement, may be in one of two states – either with a definite component S_z or with a definite component S_x, in each case with zero uncertainty. Evidently, this situation cannot be interpreted in classical terms – if the particles do not interact during the measurements. A. Einstein was deeply unhappy with this situation because it did not satisfy what, in his view, was the general requirement to any theory, which nowadays is called the *local reality*. His definition of this requirement was as follows: "The real factual situation of system 2 is independent of what is done with system 1 that is spatially separated from the former". (Here the term "spatially separated" is not defined, but from the context, it is clear that Einstein meant the detector separation by a *superluminal interval*, i.e. by distance

$$|\mathbf{r}_1 - \mathbf{r}_2| > c|t_1 - t_2|, \quad (10.27)$$

where the difference between the measurement times on the right-hand side includes the measurement duration.) In Einstein's view, since quantum mechanics did not satisfy the local reality condition, it could not be considered a complete theory of Nature.

This situation naturally raises the question of whether *something* (usually called *hidden variables*) may be added to the quantum-mechanical description to enable it to satisfy the local reality requirement. The first definite statement in this regard was John von Neumann's "proof"[30] (first famous, then infamous :-) that such variables cannot be introduced; for a while, his work satisfied the quantum mechanics' practitioners, who apparently did not pay much attention.[31] A major new contribution to the problem was made only in the 1960s by John Bell.[32] First of all, he has found an elementary (in his words, "foolish") error in von Neumann's logic, which voids his "proof". Second, he has demonstrated that Einstein's local reality condition is *incompatible* with conclusions of quantum mechanics – that had been, by that time, confirmed by too many experiments to be seriously questioned.

Let me describe a particular version of Bell's result (suggested by E. Wigner), using the same EPR pair experiment (Fig. 4a) where each SG detector may be oriented in any of three directions: **a**, **b**, or **c** – see Fig. 4b. As we already know from Chapter 4, if a fully-polarized beam of spin-½ particles is

[30] In his very early book J. von Neumann, *Mathematische Grundlagen der Quantenmechanik* [Mathematical Foundations of Quantum Mechanics], Springer, 1932. (The first English translation was published only in 1955.)
[31] Perhaps it would not satisfy A. Einstein, but reportedly he did not know about von Neumann's publication before signing the EPR paper.
[32] See, e.g., either J. Bell, *Rev. Mod. Phys*. **38**, 447 (1966) or J. Bell, *Foundations of Physics* **12**, 158 (1982).

passed through a Stern-Gerlach apparatus forming angle ϕ with the polarization axis, the probabilities of two alternative outcomes of the experiment are

$$W(\phi_+) = \cos^2\frac{\phi}{2}, \qquad W(\phi_-) = \sin^2\frac{\phi}{2}. \tag{10.28}$$

Let us use this formula to calculate all joint probabilities of measurement outcomes, starting from the detectors 1 and 2 oriented, respectively, in the directions **a** and **c**. Since the angle between the *negative* direction of the **a**-axis and the *positive* direction of the **c**-axis is $\phi_{\mathbf{a}-,\mathbf{c}+} = \pi - \varphi$ (see the dashed arrow in Fig. 4b), we get

$$W(a_+ \wedge c_+) \equiv W(a_+)W(c_+|a_+) = W(a_+)W\left(\phi_{a-,c+}\right) = \frac{1}{2}\cos^2\frac{\pi-\varphi}{2} \equiv \frac{1}{2}\sin^2\frac{\varphi}{2}, \tag{10.29}$$

where $W(x \wedge y)$ is the joint probability of both outcomes x and y, while $W(x|y)$ is the conditional probability of the outcome x, provided that the outcome y has happened.[33] Absolutely similarly,

$$W(c_+ \wedge b_+) \equiv W(c_+)W(b_+|c_+) = \frac{1}{2}\sin^2\frac{\varphi}{2}, \tag{10.30}$$

$$W(a_+ \wedge b_+) \equiv W(a_+)W(b_+|a_+) = \frac{1}{2}\cos^2\frac{\pi-2\varphi}{2} \equiv \frac{1}{2}\sin^2\varphi. \tag{10.31}$$

Now note that for any angle φ smaller than $\pi/2$ (as in the case shown in Fig. 4b), trigonometry gives

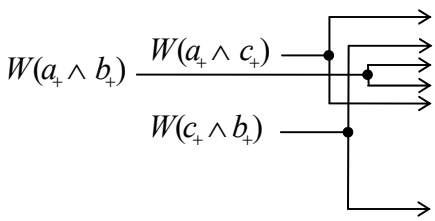
$$\frac{1}{2}\sin^2\varphi \geq \frac{1}{2}\sin^2\frac{\varphi}{2} + \frac{1}{2}\sin^2\frac{\varphi}{2} \equiv \sin^2\frac{\varphi}{2}. \tag{10.32}$$

(For example, for $\varphi \to 0$, the left-hand side of this inequality tends to $\varphi^2/2$, while the right-hand side, to $\varphi^2/4$.) Hence the quantum-mechanical result gives, in particular,

$$\boxed{W(a_+ \wedge b_+) \geq W(a_+ \wedge c_+) + W(c_+ \wedge b_+), \qquad \text{for } |\varphi| \leq \pi/2.} \tag{10.33}$$

Quantum-mechanical result

On the other hand, we can get a different inequality for these probabilities without calculating them from any particular theory, but using the local reality assumption. For that, let us prescribe some probability to each of $2^3 = 8$ possible outcomes of a set of three spin measurements. (Due to zero net spin of particle pairs, the probabilities of the sets shown in both columns of the table have to be equal.)



Detector 1	Detector 2	Probability
$a_+ \wedge b_+ \wedge c_+$	$a_- \wedge b_- \wedge c_-$	W_1
$a_+ \wedge b_+ \wedge c_-$	$a_- \wedge b_- \wedge c_+$	W_2
$a_+ \wedge b_- \wedge c_+$	$a_- \wedge b_+ \wedge c_-$	W_3
$a_+ \wedge b_- \wedge c_-$	$a_- \wedge b_+ \wedge c_+$	W_4
$a_- \wedge b_+ \wedge c_+$	$a_+ \wedge b_- \wedge c_-$	W_5
$a_- \wedge b_+ \wedge c_-$	$a_+ \wedge b_- \wedge c_+$	W_6
$a_- \wedge b_- \wedge c_+$	$a_+ \wedge b_+ \wedge c_-$	W_7
$a_- \wedge b_- \wedge c_-$	$a_+ \wedge b_+ \wedge c_+$	W_8

[33] The first equality in Eq. (29) is the well-known identity of the basic probability theory.

From the local reality point of view, these measurement options are independent, so we may write (see the arrows on the left of the table):

$$W(a_+ \wedge c_+) = W_2 + W_4, \qquad W(c_+ \wedge b_+) = W_3 + W_7, \qquad W(a_+ \wedge b_+) = W_3 + W_4. \tag{10.34}$$

On the other hand, since no probability may be negative (by its very definition), we may always write

$$W_3 + W_4 \le (W_2 + W_4) + (W_3 + W_7). \tag{10.35}$$

Plugging into this inequality the values of these two parentheses, given by Eq. (34), we get

Bell's inequality (local-reality theory)

$$W(a_+ \wedge b_+) \le W(a_+ \wedge c_+) + W(c_+ \wedge b_+). \tag{10.36}$$

This is *Bell's inequality*, which has to be satisfied by *any* local-reality theory; it directly contradicts the quantum-mechanical result (33) – opening the issue to direct experimental testing. Such tests were started in the late 1960s, but the first results were vulnerable to two criticisms:

(i) The detectors were not fast enough and not far enough to have the relation (27) satisfied. This is why, as a matter of principle, there was a chance that information on the first measurement outcome had been transferred (by some, mostly implausible) means to particles before the second measurement – the so-called *locality loophole*.

(ii) The particle/photon detection efficiencies were too low to have sufficiently small error bars for both parts of the inequality – the *detection loophole*.

Gradually, these loopholes have been closed.[34] As expected, substantial violations of the Bell inequalities (36) (or their equivalent forms) have been proved, essentially rejecting any possibility to reconcile quantum mechanics with Einstein's local reality requirement.

10.4. Interpretations of quantum mechanics

The fact that quantum mechanics is incompatible with local reality, makes its reconciliation with our (classically bred) "common sense" rather challenging. Here is a brief list of the major interpretations of quantum mechanics, that try to provide at least a partial reconciliation of this kind.

(i) The so-called *Copenhagen interpretation* – to which most physicists adhere. This "interpretation" does not really interpret anything; it just accepts the intrinsic stochasticity of measurement results in quantum mechanics and the absence of local reality, essentially saying: "Do not worry; this is just how it is; live with it". I generally subscribe to this school of thought, with the following qualification. While the Copenhagen interpretation implies statistical ensembles (otherwise, how would you define the probability? – see Sec. 1.3), its most frequently stated formulations[35] do not put sufficient emphasis on their role, in particular on the ensemble re-definition as the only point of human observer's involvement in a nearly-perfect measurement process – see Sec. 1 above. The most

[34] Important milestones in that way were the experiments by A. Aspect *et al., Phys. Rev. Lett.* **49**, 91 (1982) and M. Rowe *et al.*, *Nature* **409**, 791 (2001). Detailed reviews of the experimental situation were given, for example, by M. Genovese, *Phys. Repts.* **413**, 319 (2005) and A. Aspect, *Physics* **8**, 123 (2015); see also the later paper by J. Handsteiner *et al.*, *Phys. Rev. Lett.* **118**, 060401 (2017). Presently, a high-fidelity demonstration of the Bell inequality violation has become a standard test in virtually every experiment with entangled qubits used for quantum encryption research – see Sec. 8.5, and in particular, the paper by J. Lin cited there.

[35] With certain pleasant exceptions – see, e.g. L. Ballentine, *Rev. Mod. Phys*. **42**, 358 (1970).

famous objection to the Copenhagen interpretation belongs to A. Einstein: "God does not play dice." OK, when Einstein speaks, we all should listen, but perhaps when God speaks (through random results of the same experiment), we have to pay even more attention.

(ii) *Non-local reality*. After the dismissal of J. von Neumann's "proof" by J. Bell, to the best of my knowledge, there has been *no* proof that hidden parameters *could not* be introduced, provided that they do not imply the local reality. Of constructive approaches, perhaps the most notable contribution was made by David Bohm,[36] who developed the initial Louis de Broglie's interpretation of the wavefunction as a "pilot wave", making it quantitative. In the wave-mechanics version of this concept, the wavefunction governed by the Schrödinger equation just guides a "real", point-like classical particle whose coordinates serve as hidden variables. However, this concept does not satisfy the notion of local reality. For example, the measurement of the particle's coordinate at a certain point $\mathbf{r}_1$ has to *instantly* change the wavefunction everywhere in space, including the points $\mathbf{r}_2$ in the superluminal range (27). After A. Einstein's private criticism, D. Bohm essentially abandoned his theory.

(iii) The *many-world interpretation* that was introduced in 1957 by Hugh Everett and popularized in the 1960s and 1970s by Bruce de Witt. In this interpretation, *all* possible measurement outcomes *do* happen, splitting the Universe into the corresponding number of "parallel multiverses", so from one of them, other multiverses and hence other outcomes cannot be observed. Let me leave to the reader an estimate of the rate at which the parallel multiverses have to be constantly generated (say, per second), taking into account that such generation should take place not only at explicit lab experiments but at every irreversible process – such as fission of every atomic nucleus or an absorption/emission of every photon, everywhere in each multiverse – whether its result is formally recorded or not. Nicolaas van Kampen has called this a "mind-boggling fantasy".[37] Even the main proponent of this interpretation, B. de Witt has confessed: "The idea is not easy to reconcile with common sense." I agree.

To summarize, as far as I know, neither of these interpretations has yet provided a suggestion on how it might be tested experimentally to exclude the other ones. On the other hand, quantum mechanics makes correct (if sometimes probabilistic) predictions that do not contradict any reliable experimental results we are aware of. Maybe, this is not that bad for a scientific theory.[38]

10.5. Exercise problem

10.1.* The original (circa 1964) J. Bell's inequality was derived for the results of SG measurements performed on two non-interacting particles with zero net spin, by using the following local-reality-based assumption: the result of each single-particle measurement is uniquely determined (besides the experimental setup) by some *c*-number hidden parameter λ that may be random, i.e. change from experiment to experiment. Derive such inequality for the experiment shown in Fig. 4 and compare it with the corresponding quantum-mechanical result for the singlet state (24).

[36] D. Bohm, *Phys. Rev.* **85**, 165; 180 (1952).

[37] N. van Kampen, *Physica A* **153**, 97 (1988). By the way, I highly recommend the very reasonable summary of the quantum measurement issues, given in this paper, though believe that the quantitative theory of dephasing, discussed in Chapter 7 of this course, might give additional clarity to some of van Kampen's statements.

[38] For the reader who is not satisfied with this "positivistic" approach and wants to improve the situation, my earnest advice is to start not from square one but from reading what other (including some very clever!) people thought about it. The review collection by J. Wheeler and W. Zurek, cited above, may be a good starting point.