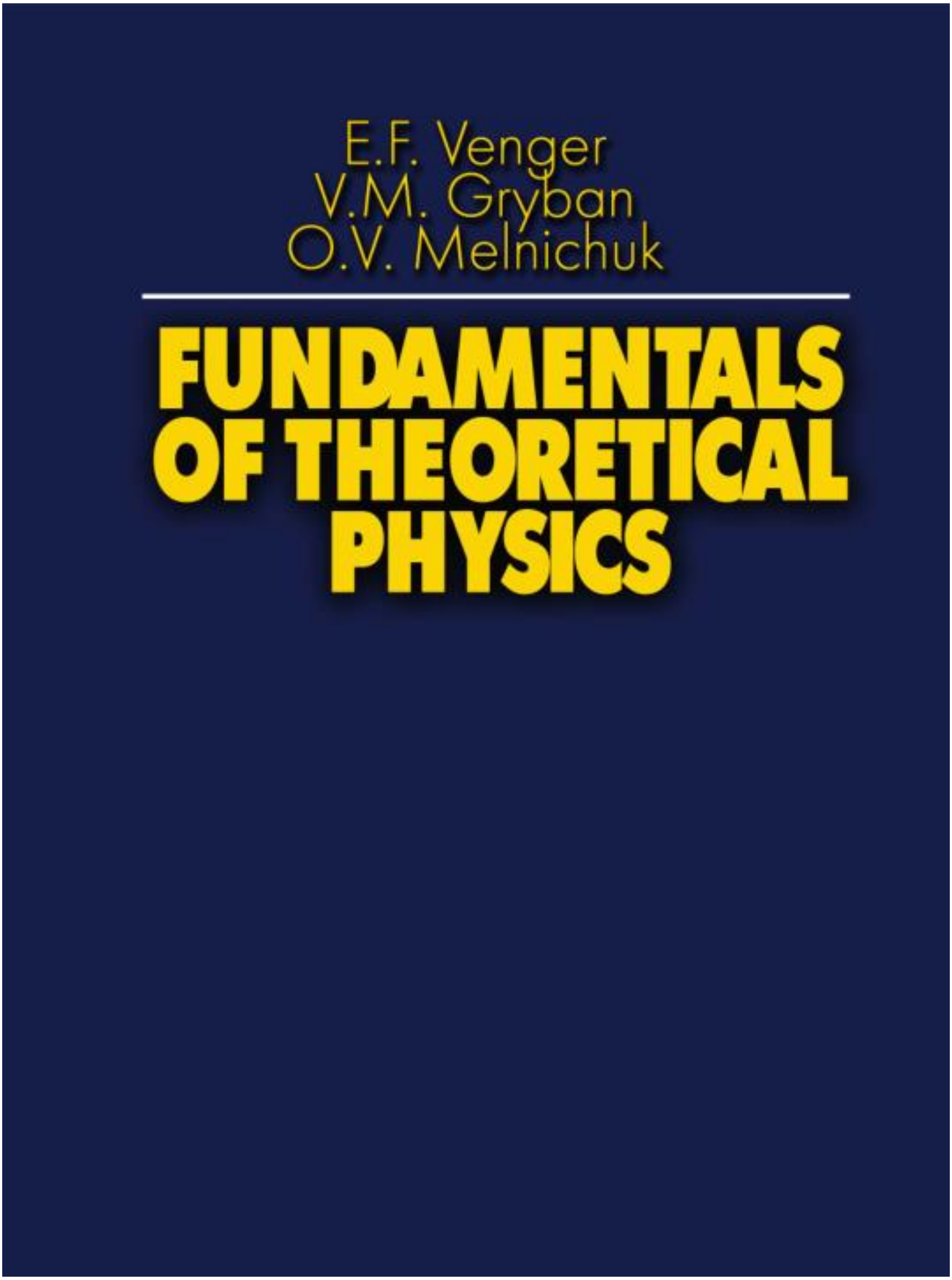

E.F. Venger
V.M. Gryban
O.V. Melnichuk

FUNDAMENTALS OF THEORETICAL PHYSICS

NATIONAL ACADEMY OF SCIENCES OF UKRAINE

E.F. Venger
V.M. Gryban
O.V. Melnichuk

FUNDAMENTALS OF THEORETICAL PHYSICS

A tutorial for high school students

Recommended by the Ministry
of Education and Science of Ukraine

Translated from Ukrainian
by *O.I. Voitenko*

KYIV · AKADEMPERIODYKA · 2016

UDC 530.1/8:514.82(075.8)
LBC 22.31
V38

Referees:

Dr.Sci. in physics and mathematics, Prof. *O.I. Anisimov*
(Taras Shevchenko National University of Kyiv);

Dr.Sci. in physics and mathematics, Prof. *P.I. Baranskyi*
(V.E. Lashkaryov Institute of Semiconductor Physics);

Dr.Sci. in physics and mathematics, Prof. *V.Z. Lozovskyi*
(Taras Shevchenko National University of Kyiv);

Dr.Sci. in physics and mathematics, Prof. *I.I. Kachuryk*
(Oleksandr Dovzhenko National University of Glukhiv)

Classified by the Ministry of Education and Science of Ukraine
(letter of 27 December 2010 No. 1/11 – 8966)

Venger E.F.
V38 Fundamentals of theoretical physics: A tutorial / E.F. Venger, V.M. Gryban, O.V. Melnichuk. – Kyiv: Akademperiodyka, 2016. – 491 p.: figs.

ISBN 978-966-360-310-0

In the brief course of theoretical physics, the concepts, formulations, and laws of classical electrodynamics, relativity theory, thermodynamics, and statistical physics are expounded.

The course is intended for high school students.

UDC 530.1/8:514.82(075.8)
LBC 22.31

ISBN 978-966-360-310-0

CONTENTS

I Electrodynamics and theory of relativity 1

A. Electromagnetic field in vacuum

B. Electromagnetic field in the medium

C. Principles of relativity theory

II Quantum mechanics 213

Appendices

PART I

ELECTRODYNAMICS AND THEORY OF RELATIVITY

Electrodynamics studies the electromagnetic field that provides interaction between electric charges, i.e. quantitative parameters that describe the ability of elementary particles and macroscopic objects to interact with one another electromagnetically. In other words, electrodynamics studies electromagnetic interaction, the source of which is the electric charge.

Modern physics distinguishes four types of fundamental interactions, and all interactions known in the micro- and macroworlds can be correspondingly classified. These are gravitational, weak, electromagnetic, and strong interactions. Their intensities are related as $10^{-38} : 10^{-10} : 10^{-2} : 1$, in that order, and the interaction radius, i.e. the maximum distance R at which the interaction differs from zero, equals infinity for the gravitational and electromagnetic interactions, 10^{-18} m for weak interaction, and 10^{-15} m for strong interaction. Each of those interactions matters very much in nature.

Weak interaction plays an important role in the microworld, e.g., in the course of elementary particle transformations. In particular, no thermonuclear reactions in the interior of the sun and other stars could be possible without it. However, owing to the extremely small value of its interaction radius, weak interaction is not observed in the macroworld.

Strong interaction, together with the weak one, is also of great importance in reactions of thermonuclear synthesis. Besides, it makes the existence of atomic nuclei and, as a consequence, atoms and molecules possible. However, due to its short-range character, strong interaction is not observed at distances $r > 10^{-15}$ m.

Gravitational interaction is long range, but it is the least intensive. Therefore, this kind of interaction manifests itself only if the body's mass is large, thus governing the structure of planetary systems, galaxies, and the

Universe in general. However, in the world of elementary particles, their mutual transformations associated with gravitation are still unknown.

Electromagnetic interaction is rather powerful and long range. Therefore, it plays an important role both in the microworld, being responsible for a series of elementary particle transformations and the very existence of atoms and molecules, and in the macroworld. Most forces observed in macroscopic phenomena belong to this class of interaction: elastic forces, friction force, surface tension forces, and so on. Electromagnetic interaction governs the properties of substances in different aggregate states and a number of optical phenomena. It is widely used in electro- and radio-engineering, quantum electronics, and so forth.

In general, electrodynamics is divided into classic and quantum-mechanical ones. Classic electrodynamics examines the electromagnetic field using such concepts as electromagnetic waves and the vectors of electric field strength and magnetic field induction. In quantum electrodynamics, the electromagnetic field is identified with field quanta, photons, emitted and absorbed by electric charges. Classic electrodynamics, in turn, is divided into macro- and microscopic ones. In the former (electrodynamics of continuous media), matter is regarded as a continuous medium, and the electric charge is considered to be distributed continuously. The latter (the electron theory) takes into account the discreteness of the matter and electric charge structure.

First information about electric and magnetic phenomena appeared as early as in ancient times. However, the emergence of electrodynamics as a science is dated to 1820, when the interconnection between electric and magnetic phenomena was discovered. That year, the Danish physicist H. Ørsted revealed the action of electric current on a magnetic needle, the French physicist A.-M. Ampère established the law of current interaction and put forward the hypothesis that magnetization is associated with currents in molecules, the French physicists J.B. Biot and F. Savart discovered the law later formulated in the general form by P. Laplace, which determines the magnetic field created by an element of current.

Of great importance for the development of electrodynamics were the works by the English physicist M. Faraday. He discovered the principle of electromagnetic induction (1831), the laws of electrolysis (1833–1834), the polarization phenomenon in dielectrics (1837), para- and diamagnetism (1845), and the magnetically induced rotation of the light polarization plane (1845). Faraday's ideas concerning the existence of electric and magnetic fields as real physical objects that act on electric charges and currents at the points where they are located deserve special attention. The concept of short-range interaction became dominating in the development of electrodynamics.

The foundation of classic macroscopic electrodynamics was completed in 1861–1873 in the works of the English physicist J.C. Maxwell. The latter generalized the phenomenological laws of Coulomb, Ampère, and Biot–Savart, and using the hypothesis that the magnetic field is generated by the oscillating electric field, derived a system of equations describing the electromagnetic field. The equations led to a conclusion about the electromagnetic nature of light and predicted the existence of electromagnetic waves.

The discovery of electron by J.J. Thomson in 1897 stimulated the creation of the fundamentals of classic electron theory by the Dutch physicist H. Lorentz. In his theory, matter is considered as an aggregate of charged particles that interact with the electromagnetic field, being simultaneously the source of this field; and the resulting field is determined by averaging.

However, while studying electromagnetic phenomena in moving media, there emerged some difficulties. Their overcoming resulted in the creation of the special theory of relativity (A. Einstein, 1905) and, on its basis, the formulation of relativistic electrodynamics. The creation of quantum mechanics and special theory of relativity stimulated the development of quantum electrodynamics and quantum field theory, which will not be considered here. Nevertheless, classic electrodynamics and electron theory did not lose their actuality till now and remain a basis for radio- and electro-engineering, optics, plasma physics, electronics (but for quantum electronics), and other domains.

A. Electromagnetic field in vacuum

1

ELECTROSTATIC FIELD

1.1. Electric charge. Coulomb's law

The source of electromagnetic field and, hence, electromagnetic interaction is the **electric charge**. All electric and magnetic phenomena are manifestations of the existence, motion, and interaction of electric charges. Elementary particles—electrons, protons, and others—are material carries of electric charge.

There is a minimum electric charge e, with the charge of every particle or body being multiple of it. All known elementary particles have electric charges equal to 0, $\pm e$, and $\pm 2e$, where $|e| = 1.602 \times 10^{-19}$ C. In particular, the charge of electron equals $-e$, and that of proton $+e$.

The division of charges into positive and negative is conditional: the charge of a glass stick electrified by friction is called positive, and that of an amber stick, negative. The whole body of experimental data testifies that the algebraic sum of electric charges in a closed system remains constant at all interactions and transformations in it. This is the essence of the charge conservation law, which has a universal character.

In classic macroscopic electrodynamics, the electric charge of macroscopic objects is assumed to be continuously distributed with the density

$$\rho(\vec{r}\,) = \frac{dq(\vec{r}\,)}{dV}, \tag{1.1}$$

where dV is a physically infinitesimal volume, which, nevertheless, contains plenty of charged particles ($dV \gg \delta V_0$, V_0 is the atom's volume), and dq is an physically infinitesimal charge in the volume dV ($dq \gg e$). Then, the complete charge in the volume V equals

$$q = \int_V \rho(\vec{r}\,)\, dV. \tag{1.2}$$

If the electric charge is distributed in a thin layer on the body's surface with the surface density

$$\sigma(\vec{r}\,) = \frac{dq(\vec{r}\,)}{dS}, \tag{1.3}$$

then,

$$q = \int_S \sigma(\vec{r}\,)\, dS; \tag{1.4}$$

and, if it is distributed along a wire or a thread, we obtain the linear charge density

$$\eta(\vec{r}\,) = \frac{dq(\vec{r}\,)}{dl}, \tag{1.5}$$

so that

$$q = \int_L \eta(\vec{r}\,)\, dl. \tag{1.6}$$

If the charge q occupies a volume that can be considered as a material point, it is called the **point charge**. In this case, to provide the validity of relation (1.2), the charge density is written down in the form

$$\rho(\vec{r}\,) = q\delta(\vec{r} - \vec{r}_1), \tag{1.7}$$

where $\vec{r}_1$ is the radius vector of charge q, and $\delta(\vec{r} - \vec{r}_1)$ is the Dirac delta-function (see Appendix A).

For a system of point charges,

$$\rho(\vec{r}\,) = \sum_i q_i \delta(\vec{r} - \vec{r}_i), \tag{1.8}$$

The charge

$$dq = \rho(\vec{r}\,)dV, \tag{1.9}$$

which follows from equation (1.1), is also considered to be point charge.

The unit of electric charge in SI is **coulomb** (C). One coulomb is the electric charge that flows through the transverse cross-section of a conductor in one second, provided that the stationary current strength in the conductor equals one ampere (1 A). Similar (of the same sign) electric charges repulse and opposite (of the different signs) ones attract one another. The **law of point charge interaction** was established by the French physicist Ch. Coulomb:

- the absolute value of interaction force between two fixed point charges q_1 and q_2 in vacuum is proportional to the product of their charge magnitudes, $|q_1|$ and $|q_2|$, and reciprocal to the squared distance between them, r;

namely,

$$|F_{12}| = k\frac{|q_1|\,|q_2|}{r^2}, \tag{1.10}$$

where $k = 9 \times 10^9$ N $\times$ m^2/C^2 $= 9 \times 10^9$ m/F. In order to simplify differential equations in electrodynamics, the coefficient k is written in the form

$$k = \frac{1}{4\pi}\varepsilon_0, \tag{1.11}$$

where ε_0 is the electric constant equal to

$$\varepsilon_0 = \frac{1}{36\pi \times 10^9}\frac{\mathrm{F}}{\mathrm{m}} \approx 8.85 \times 10^{-12}\frac{\mathrm{F}}{\mathrm{m}}. \tag{1.12}$$

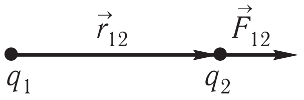

Fig. 1.1.

Taking expression (1.11) into account, **Coulomb's law** (1.10) in the vector representation is written as follows:

$$\vec{F}_{12} = \frac{q_1 q_2}{4\pi\varepsilon_0 r_{12}^2}\frac{\vec{r}_{12}}{r_{12}}. \tag{1.13}$$

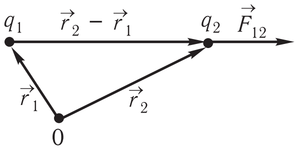

Fig. 1.2.

In the case of similar charges, the vectors $\vec{F}_{12}$ and $\vec{r}_{12}$ are shown in Fig. 1.1, where $\vec{F}_{12}$ is the force, with which the charge q_1 acts on the charge q_2. The force, with which the charge q_2 acts on the charge q_1, is written in the form

$$\vec{F}_{21} = \frac{q_1 q_2}{4\pi\varepsilon_0 r_{21}^2}\frac{\vec{r}_{21}}{r_{21}}, \vec{r}_{21} = -\vec{r}_{12}. \tag{1.14}$$

Introducing the charge coordinates, $q_1 = q_1(\vec{r}_1)$ and $q_2 = q_2(\vec{r}_2)$ (see Fig. 1.2), equation (1.13) can be rewritten in the form

$$\vec{F}_{12} = \frac{q_1(\vec{r}_1)q_2(\vec{r}_2)}{4\pi\varepsilon_0\,|\vec{r}_2 - \vec{r}_1|^2}\frac{\vec{r}_2 - \vec{r}_1}{|\vec{r}_2 - \vec{r}_1|}. \tag{1.15}$$

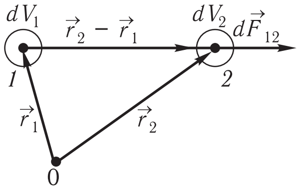

Fig. 1.3.

Now, let us consider the method allowing the interaction force between two non-point electric charge distributions to be calculated. Let the charges be distributed in the volumes V_1 and V_2 with the densities $\rho_1(\vec{r})$ and $\rho_2(\vec{r})$, respectively. Around points 1 and 2 located inside the volumes V_1 and V_2, respectively, we select physically infinitesimal volumes dV_1 and dV_2 (Fig. 1.3). The latter contain the charges $dq_1 = \rho_1(\vec{r}_1)dV_1$ and $dq_2 = \rho_2(\vec{r}_2)dV_2$. Then, according to (1.15), the force

$$\begin{aligned} d\vec{F}_{12} &= \frac{dq_1 dq_2}{4\pi\varepsilon_0 \left|\vec{r}_2 - \vec{r}_1\right|^2} \frac{\vec{r}_2 - \vec{r}_1}{\left|\vec{r}_2 - \vec{r}_1\right|} \\ &= \frac{\rho_1(\vec{r}_1)dV_1\rho_2(\vec{r}_2)dV_2}{4\pi\varepsilon_0 \left|\vec{r}_2 - \vec{r}_1\right|^2} \frac{\vec{r}_2 - \vec{r}_1}{\left|\vec{r}_2 - \vec{r}_1\right|} \end{aligned} \tag{1.16}$$

acts on the charge dq_2 from the charge dq_1. The total force, with which the charge q_1 acts on the charge q_2, equals

$$\vec{F}_{12} = \frac{1}{4\pi\varepsilon_0} \int_{V1} \int_{V2} \frac{\rho_1(\vec{r}_1)\rho_2(\vec{r}_2)}{\left|\vec{r}_2 - \vec{r}_1\right|^2} \frac{\vec{r}_2 - \vec{r}_1}{\left|\vec{r}_2 - \vec{r}_1\right|} dV_1 dV_2. \tag{1.17}$$

The calculation of integrals in formula (1.17) is usually connected with considerable difficulties. The overcoming of those difficulties can be facilitated by introducing the concept of electric field strength.

1.2. Electrostatic field. Electrostatic field strength

The electric charge changes the properties of surrounding space. This modification is defined as the creation of an electric field by the charge. The main property of induced filed is its action on the other charge. In other words, electric charges interact by means of the electric field, which, in the case of fixed charges, is called the **electrostatic field**. The charge interaction in this case is called the **electrostatic interaction**.

It should be noted that the electrostatic field acts on the electric charge with a certain force $\vec{F}$. Let us consider the action of the field on the charge q_0, which will be considered point-like, positive, and so small, that it cannot distort the examined field by its own one (the **test charge**). Experimental data demonstrate that, in the general case, the force $\vec{F}(\vec{r}\,)$ is different at different points of the field, but the ratio $\vec{F}(\vec{r}\,)/q_0$ at the definite point does not depend on the charge value q_0. Therefore, this ratio is a force characteristic of the field at the given point, which is called the **electrostatic field strength**,

$$\vec{E}(\vec{r}\,) = \frac{\vec{F}(\vec{r}\,)}{q_0}. \tag{1.18}$$

According to (1.18), the unit of electric (electrostatic) field strength is 1 N/C (or 1 V/m, see section 1.3).

If the electrostatic field is created by a point charge q located at the coordinate origin, then, in accordance with (1.13),

$$\vec{F}(\vec{r}\,) = \frac{qq_0}{4\pi\varepsilon_0 r^2}\frac{\vec{r}}{r}, \tag{1.19}$$

where $\vec{r}$ is the radius vector of observation point. In this case,

$$\vec{E}(\vec{r}\,) = \frac{q}{4\pi\varepsilon_0 r^2}\frac{\vec{r}}{r}. \tag{1.20}$$

For a point charge q located at an arbitrary point with the coordinate $\vec{r}\,'$ (see (1.15)),

$$\vec{E}(\vec{r}\,) = \frac{q}{4\pi\varepsilon_0 \left|\vec{r}-\vec{r}\,'\right|^2}\frac{\vec{r}-\vec{r}\,'}{\left|\vec{r}-\vec{r}\,'\right|}. \tag{1.21}$$

Consider a system of point charges $q_1, q_2, \ldots, q_N$ located at the points with the coordinates $\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N$, respectively. Experimental data testify that, in this case, the electrostatic field strength at an arbitrary point with the coordinate $\vec{r}$ is a vector sum of strengths generated by each of the charges taken separately,

$$\vec{E}(\vec{r}\,) = \sum_{i=1}^{N}\vec{E}_i(\vec{r}\,) = \frac{1}{4\pi\varepsilon_0}\sum_{i=1}^{N}\frac{q_i\left(\vec{r}-\vec{r}_i\right)}{\left|\vec{r}-\vec{r}_i\right|^3}. \tag{1.22}$$

This statement is called the **superposition principle** for electrostatic fields. It claims that electrostatic fields do not interact with one another.

For the charge continuously distributed with the density $\rho(\vec{r}')$ in the volume V', formula (1.22) should be rewritten in the form

$$\vec{E}(\vec{r}) = \frac{1}{4\pi\varepsilon_0} \int_{V'} \frac{\rho(\vec{r}')\,(\vec{r}-\vec{r}')}{|\vec{r}-\vec{r}'|^3} dV', \tag{1.23}$$

and for the charge distributed with the density $\sigma(\vec{r}')$ on the surface S', in the form

$$\vec{E}(\vec{r}) = \frac{1}{4\pi\varepsilon_0} \int_{S'} \frac{\sigma(\vec{r}')\,(\vec{r}-\vec{r}')}{|\vec{r}-\vec{r}'|^3} dS'. \tag{1.24}$$

For visualization, the electrostatic field is represented by force lines. They are such curves that the vector $\vec{E}$ is directed tangentially at every curve point. Force lines start on positive charges and end on negative ones. In the idealized case of a single charge, they start or end at infinity.

The flux of vector $\vec{E}$, i.e. the number of force lines that pass through the surface S, is calculated by the formula

$$\Delta N = \int_{S'} E_n \, dS' = \int_S \vec{E}\, d\vec{S}, \tag{1.25}$$

where

$$d\vec{S} = \vec{n}\, dS', \tag{1.26}$$

and $\vec{n}$ is the unit vector directed normally to the surface element dS'.

According to (1.23) and (1.25), the flux of vector $\vec{E}$ created by the charge q located in the volume V' through the closed surface S that surrounds the charge equals

$$\oint_{S'} \vec{E}\, d\vec{S} = \frac{1}{4\pi\varepsilon_0} \oint_{S'} \int_{V'} \frac{\rho(\vec{r}')\,(\vec{r}-\vec{r}')\, dV'}{|\vec{r}-\vec{r}'|^3} dS'. \tag{1.27}$$

Since

$$\frac{(\vec{r}-\vec{r}')\, d\vec{S}}{|\vec{r}-\vec{r}'|^3} = \frac{(\vec{r}-\vec{r}')\, \vec{n} dS'}{|\vec{r}-\vec{r}'|^3} = \frac{dS_0}{|\vec{r}-\vec{r}'|^2} = d\Omega,$$

where dS_0 is the projection of the area dS on a spherical surface centered in the volume element dV', and $d\Omega$ is the element of the solid angle corresponding to the area dS_0, we obtain

$$\oint_{S'} \vec{E}\, d\vec{S} = \frac{1}{4\pi\varepsilon_0} \oint_S d\Omega \int_{V'} \rho(\vec{r}')\, dV' = \frac{1}{4\pi\varepsilon_0} 4\pi q = \frac{q}{\varepsilon_0}.$$

The relation

$$\oint_{S'} \vec{E}\, d\vec{S} = \frac{q}{\varepsilon_0}, \tag{1.28}$$

in which q is the electric charge located inside the closed surface S is called **Gauss's theorem**.

According to (1.2),

$$q = \int_V \rho(\vec{r}\,)\, dV,$$

where V is the volume confined by the closed surface S. Besides (see Appendix B),

$$\oint_{S'} \vec{E} d\vec{S} = \int_V \operatorname{div} \vec{E}\ dV. \tag{1.29}$$

Therefore, instead of (1.29), the following relation can be written:

$$\int_V \operatorname{div} \vec{E}\ dV = \frac{1}{\varepsilon_0} \int_V \rho dV,$$

whence it follows that

$$\operatorname{div} \vec{E} = \frac{1}{\varepsilon_0} \rho.$$

Equations (1.28) in the integral form and (1.29) in the differential one state that the sources of electrostatic field are electric charges.

Equation (1.29) can also be rewritten as follows:

$$\operatorname{div} \vec{D} = \rho. \tag{1.30}$$

The vector

$$\vec{D} = \varepsilon_0 \vec{E} \tag{1.31}$$

is called the **vector of electric induction** in vacuum. Then, instead of (1.28), we may write

$$\oint \vec{D}\, d\vec{S} = q. \tag{1.32}$$

Gauss's theorem makes it possible to substantially simplify the calculation of electric field strength in a number of cases. For example, let us determine the electrostatic field strength for a uniformly charged plane with the surface charge density σ (Fig. 1.4). If the distance from the observation point to the plane is considerably smaller than the point's linear size, the plane can be regarded as infinite, the field as homogeneous, and the force lines as directed perpendicularly to the plane. The closed surface S is selected as the surface of a cube with the edge a, as is shown in Fig. 1.4. In this case, it is easy to verify that, on the basis of (1.29), the flux of vector $\vec{E}$ through the cube surface equals

$$2Ea^2 = \frac{\sigma a^2}{\varepsilon_0},$$

so that

$$E = \frac{\sigma}{2\varepsilon_0}.$$

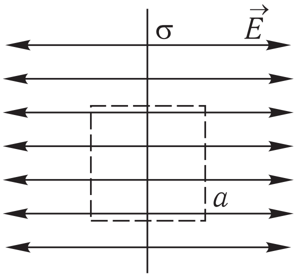

Fig. 1.4.

1.3. Electrostatic potential

Taking into account the relation

$$\nabla\frac{1}{|\vec{r}-\vec{r}\,'|} = \frac{\partial}{\partial\vec{r}}\frac{1}{|\vec{r}-\vec{r}\,'|} = -\frac{1}{|\vec{r}-\vec{r}\,'|^2}\,\frac{\vec{r}-\vec{r}\,'}{|\vec{r}-\vec{r}\,'|},$$

electrostatic field strength (1.23) can be written in the form

$$\vec{E}(\vec{r}\,) = \frac{1}{4\pi\varepsilon_0}\int\frac{\rho(\vec{r}\,')\,(\vec{r}-\vec{r}\,')}{|\vec{r}-\vec{r}\,'|^3}dV' = -\nabla\varphi(\vec{r}\,), \tag{1.33}$$

where

$$\varphi(\vec{r}\,) = \frac{1}{4\pi\varepsilon_0}\int\frac{\rho(\vec{r}\,')dV'}{|\vec{r}-\vec{r}\,'|}. \tag{1.34}$$

Then, in accordance with (1.18) and (1.33), the force that acts on the charge q in the electrostatic field is written as follows:

$$\vec{F}(\vec{r}\,) = q\vec{E}(\vec{r}\,) = -\nabla(q\varphi) = -\nabla U(\vec{r}\,) = \frac{\partial U}{\partial\vec{r}},$$

and the work executing by the field when the charge q moves from point 1 to point 2 as follows:

$$A = \int_1^2\vec{F}d\vec{r} = -\int_1^2 dU(\vec{r}\,) = U(\vec{r}_1) - U(\vec{r}_{2-} = q\left[\varphi(\vec{r}_1) - \varphi(\vec{r}_2)\vec{\,}\right]. \tag{1.35}$$

Hence, the work executed by the electrostatic field is determined only by the initial and final coordinates of moving charge. If the charge moves along the closed path, the work done by the field equals zero. Such fields are called **potential**. Hence, the electrostatic field is potential, and function (1.34) is called the **electrostatic potential**. The quantity

$$U(\vec{r}) = q\varphi(\vec{r}) \tag{1.36}$$

is the **potential energy** of the charge q in the electrostatic field.

The work done by the electrostatic field to move the charge q along the closed path (contour) L can be written as follows:

$$A = \oint_L \vec{F} d\vec{l} = q \oint_L \vec{E} d\vec{l},$$

where $d\vec{l}$ is the path element. Since $A = 0$ in this case, we obtain the equation

$$\oint_L \vec{E} d\vec{l} = 0. \tag{1.37}$$

On the other hand, on the basis of Stokes theorem (see Appendix B),

$$\oint_L \vec{E} d\vec{l} = \int_S \text{rot } \vec{E} d\vec{S},$$

where S is the surface confined by the contour L. Then, taking (1.37) into account, we obtain

$$\text{rot } \vec{E} = 0. \tag{1.38}$$

Equation (1.37) in the integral form and equation (1.38) in the differential one are an indication that the electrostatic field is potential.

Formula (1.33) establishes a relationship between the strength and the potential of electrostatic field (the **relation equation**),

$$\vec{E}(\vec{r}) = \nabla\varphi(\vec{r}) \tag{1.39}$$

On its basis, it is possible to write that

$$\varphi(\vec{r}) = -\int \vec{E}(\vec{r}) d\vec{r}. \tag{1.40}$$

The integral in (1.40) is indefinite; therefore, the potential is an ambiguous function. However, this circumstance is not substantial, because the quantity that is measured and has a physical meaning is the potential difference (see (1.35)). In this connection, the integration constant in (1.40) is chosen for convenience reasons.

For instance, if the charge q is a point charge, then, on the basis of (1.21) and (1.40), it is easy to be convinced that

$$\varphi(\vec{r}\,) = \frac{q}{4\pi\varepsilon_0\,|\vec{r} - \vec{r}\,'|} + C, \tag{1.41}$$

where C is the integration constant. Imposing the condition $\varphi(\infty) = 0$, we obtain

$$\varphi(\vec{r}\,) = \frac{q}{4\pi\varepsilon_0\,|\vec{r} - \vec{r}\,'|}. \tag{1.42}$$

If the charge q is located at the coordinate origin ($\vec{r}\,' = 0$),

$$\varphi(\vec{r}\,) = \frac{q}{4\pi\varepsilon_0 r}. \tag{1.43}$$

In the case of surface charge distribution, instead of (1.34), we should write

$$\varphi(\vec{r}\,) = \frac{1}{4\pi\varepsilon_0}\int\frac{\sigma(\vec{r}\,')dS'}{|\vec{r} - \vec{r}\,'|}. \tag{1.44}$$

In accordance with (1.35),

$$\varphi_1 - \varphi_2 = \frac{A}{q}.$$

Therefore, the unit of potential difference—one **volt** (1 V)—is the potential difference between two points of electrostatic field, if the work executed while moving an electric charge of 1 C between them equals 1 J.

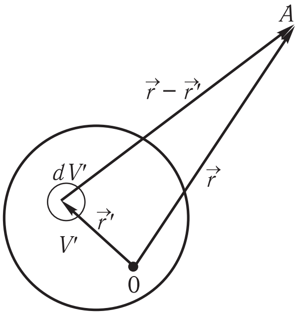

Fig. 1.5.

Let us consider a case when the electrostatic field is created by charges removed from the observation point A (Fig. 1.5). Let us designate the linear dimensions of volume V' containing the charges as a and consider potential (1.34) under the condition $a \ll r$ supposing that the coordinate origin (point O) is located in the volume

V'. Then, the inequality $r' \ll r$ is satisfied, and the function $|\vec{r}-\vec{r}\,'|^{-1}$ can be expanded in the Taylor series of the vector $\vec{r}\,'$ projections, which we designate as x_1', x_2', and x_3', up to the first three terms,

$$\begin{aligned}\frac{1}{|\vec{r}-\vec{r}\,'|} &= \frac{1}{r}+\sum_{i=1}^{3}\left(\frac{\partial\,|\vec{r}-\vec{r}\,'|^{-1}}{\partial x_i'}\right)_0 x_i' \\ &+\frac{1}{2}\sum_{i,j=1}^{3}\left(\frac{\partial\,|\vec{r}-\vec{r}\,'|^{-1}}{\partial x_i'\partial x_j'}\right)_0 x_i'x_j'. \qquad (1.45)\end{aligned}$$

The zero subscripts in (1.45) mean that, after calculating the derivative, the coordinates x_i' and x_j' must be put equal to zero. Carrying out the corresponding calculations (see section 1.8 in [10]), we obtain

$$\frac{1}{|\vec{r}-\vec{r}\,'|}=\frac{1}{r}+\frac{\vec{r}\,\vec{r}\,'}{r^3}+\frac{1}{2r^5}\sum_{i,j=1}^{3}x_ix_j\left(-r'^2\delta_{ij}+3x_i'x_j'\right), \qquad (1.46)$$

and potential (1.34) reads

$$\begin{aligned}\varphi(\vec{r}) &= \frac{1}{4\pi\varepsilon_0 r}\int\rho(\vec{r}\,')dV'+\frac{\vec{r}}{4\pi\varepsilon_0 r^3}\int\rho(\vec{r}\,')\vec{r}\,'dV' \\ &+\frac{1}{8\pi\varepsilon_0 r^5}\sum_{i,j=1}^{3}x_ix_j\int\rho(\vec{r}\,')\left(-r'^2\delta_{ij}+3x_i'x_j'\right)dV'. \qquad (1.47)\end{aligned}$$

Let three summands in (1.47) be designated as φ_1, φ_2, and φ_3. In the first summand, the integral is equal to the total charge q in the volume V', therefore

$$\varphi_1(r)=\frac{q}{4\pi\varepsilon_0 r}, \qquad (1.48)$$

i.e. this summand is the potential of the point charge q located at the coordinate origin. In the second summand, the quantity

$$\vec{d}=\int\rho(\vec{r}\,')\vec{r}\,'dV' \qquad (1.49)$$

is called the **electric dipole moment** of the system of charges. Then, this term can be written as follows:

$$\varphi_2(r)=\frac{\vec{r}\vec{d}}{4\pi\varepsilon_0 r^3}. \qquad (1.50)$$

The third summand in (1.47) contains the quantity

$$D_{ij} = \frac{1}{2} \int \rho(\vec{r}\,') \left(-r'^2 \delta_{ij} + 3x'_i x'_j \right) dV', \tag{1.51}$$

which is called the **tensor of electric quadrupole moment** of the system of charges. Therefore,

$$\varphi_3(r) = \frac{1}{4\pi\varepsilon_0 r^5} \sum_{i,j=1}^{3} x_i x_j D_{ij} \tag{1.52}$$

Hence, any system of charges, at a large distance from it, can be considered as composed of a point charge, an electric dipole, and an electric quadrupole, which are called the **multipole moments** of the system of charges and located at the coordinate origin. In this case, the potential is a sum of potentials created by multipoles,

$$\varphi = \varphi_1 + \varphi_2 + \varphi_3, \tag{1.53}$$

with

$$\varphi_1 \sim \frac{1}{r}, \varphi_2 \sim \frac{1}{r^2}, \varphi_3 \sim \frac{1}{r^3}.$$

Therefore, if the total charge q differs from zero, the field of this point charge dominates at large distances. If the system is electrically neutral but has an electric dipole moment, the field of electric dipole plays the main role. And so on.

The simplest electrically neutral system of charges with electric dipole moment consists of two point charges $q_1 = q$ and $q_2 = -q$ with the coordinates $\vec{r}_1$ and $\vec{r}_2$, respectively, with the distance $\left|\vec{l}\right| = |\vec{r}_1 - \vec{r}_2|$ between them being small ($l \ll r$). Then, according to (1.8),

$$\rho(\vec{r}\,') = q\delta(\vec{r}\,' - \vec{r}_1) - q\delta(\vec{r}\,' - \vec{r}_2), \tag{1.54}$$

and the electric dipole moment (1.49) for this systems equals

$$\begin{aligned} \vec{d} &= q \int \vec{r}\,'\delta(\vec{r}\,' - \vec{r}_1)dV' - q \int \vec{r}\,'\delta(\vec{r}\,' - \vec{r}_2)dV' \\ &= q\left(\vec{r}_1 - \vec{r}_2\right) = q\vec{l}. \end{aligned} \tag{1.55}$$

Hence, the electric dipole moment of electrically neutral system of charges does not depend on the choice of coordinate origin.

According to (1.39) and (1.50), the strength of electrostatic field created by the electric dipole can be determined as follows:

$$\begin{aligned}\vec{E}_{dip}(\vec{r}\,) &= -\nabla\varphi_2(\vec{r}\,) = -\frac{\partial}{\partial\vec{r}}\frac{\vec{r}\vec{d}}{4\pi\varepsilon_0 r^3} \\ &= \frac{1}{4\pi\varepsilon_0}\left[-\frac{\vec{d}}{r^3} + 3\frac{\left(\vec{r}\vec{d}\right)\vec{r}}{r^5}\right]. \end{aligned} \tag{1.56}$$

An example of electrically neutral system of charges with no electric dipole moment but with an electric quadrupole one is a system of two dipoles identical by magnitude but with oppositely directed dipole moments that are located at a small distance from each other; e.g., four alternating positive and negative charges arranged at the corners of a square.

1.4. Electric dipole in electrostatic field

Let us consider the behavior of electric dipole in the electrostatic field. If the coordinate of charge $-q$ is designated as $\vec{r}$, then the coordinate of charge q can be written down as $\vec{r}+\vec{l}$, where the vector $\vec{l}$ is directed from the charge $-q$ to the charge $+q$. According to (1.36), the potential energy of dipole in the electrostatic field equals

$$U(\vec{r}) = q\varphi(\vec{r}+\vec{l}) - q\varphi(\vec{r}). \tag{1.57}$$

The vector $\vec{l}$ is supposed to be small. Therefore, we may write

$$\varphi(\vec{r}+\vec{l}) \approx \varphi(\vec{r}) + \vec{l}\,\nabla\varphi(\vec{r}). \tag{1.58}$$

Substituting (1.58) into (1.57), we obtain

$$U(\vec{r}) = q\vec{l}\,\nabla\varphi(\vec{r}) = -\vec{d}\vec{E}(\vec{r}). \tag{1.59}$$

The force, with which the fields acts on the dipole, is determined as

$$\vec{F} = -\nabla U = \nabla(\vec{d}\vec{E}). \tag{1.60}$$

The gradient of scalar product $(\vec{d}\vec{E})$ looks like (see Appendix B)

$$\nabla(\vec{d}\vec{E}) = \left[\vec{d} \operatorname{rot} \vec{E}\right] + \left[\vec{E} \operatorname{rot} \vec{d}\right] + (\vec{d}\nabla)\vec{E} + (\vec{E}\nabla)\vec{d}. \tag{1.61}$$

Since the electric dipole moment $\vec{d}$ does not depend on the coordinates and, in accordance with (1.38), $\operatorname{rot} \vec{E} = 0$, we obtain

$$\vec{F} = (\vec{d}\nabla)\vec{E}. \tag{1.62}$$

According to this formula, the force $\vec{F}$ that acts on the electric dipole in the electrostatic field differs from zero if the field is non-uniform.

The electric dipole in the electrostatic field is also subjected to the action of the moment of forces (the **torque**). If (1.59) is written in the form

$$U = -dE\cos\theta, \tag{1.63}$$

the absolute value of the torque vector is determined by the formula

$$\left|\vec{M}\right| = -\frac{\partial U}{\partial \theta} = dE\sin\theta, \tag{1.64}$$

or, in the vector notation,

$$\vec{M} = \left[\vec{d}\vec{E}\right]. \tag{1.65}$$

The torque $\vec{M}$ attempts to orient the dipole moment $\vec{d}$ along the vector $\vec{E}$. It differs from zero (if $\theta \neq 0$) in both uniform and non-uniform electrostatic fields.

1.5. Equation for electrostatic field in vacuum. Boundary conditions

The electrostatic field strength $\vec{E}$ must satisfy equations (1.28) and (1.30), according to which the sources of field are electric charges, and equations (1.37) and (1.38) associated with the potential character of electrostatic field. Let us express those equations in the following form:

$$\text{(a)} \oint_S \vec{E}\, d\vec{S} = \frac{q}{\varepsilon_0}, \qquad \text{(b)} \operatorname{div} \vec{E} = \frac{\rho}{\varepsilon_0}, \tag{1.66}$$

$$\text{(a)} \oint_L \vec{E}\, d\vec{l} = 0, \qquad \text{(b)} \operatorname{rot} \vec{E} = 0. \tag{1.67}$$

Equations for the electrostatic potential can be found by substituting relation equation (1.39) into (1.66b). Then, taking into consideration that div grad = Δ (see Appendix B), we obtain the equation

$$\Delta\varphi(\vec{r}) = -\frac{1}{\varepsilon_0}\rho(\vec{r}), \tag{1.68}$$

which is called the **Poisson equation**. Its general solution is expressed by formula (1.34),

$$\varphi(\vec{r}) = -\frac{1}{4\pi\varepsilon_0}\int_{V'}\frac{\rho(\vec{r}\,')dV'}{|\vec{r}-\vec{r}\,'|}. \tag{1.69}$$

At those points in the space where the electric charge is absent ($\rho(\vec{r}) = 0$), equation (1.68) takes the form

$$\Delta\varphi(\vec{r}) = 0. \tag{1.70}$$

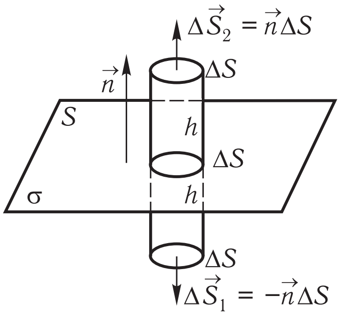

Fig. 1.6.

This equation is called the **Laplace equation**.

Equations (1.66)–(1.68) and (1.70) have to be supplemented with boundary conditions, which determine the behavior of the vector $\vec{E}$ and the potential φ across the charged surface.

Let the surface S (Fig. 1.6) have the surface charge density $\sigma(\vec{r})$. Consider an element ΔS, which is so small that the charge density σ within its limits can be considered constant, and use it as a cross-section of a straight cylinder of the height $2h$ and the bases ΔS. Now, let us apply Gauss's theorem (1.28) to the closed surface of the cylinder,

$$\oint \vec{E}d\vec{S} = \frac{\sigma\Delta S}{\varepsilon_0}. \tag{1.71}$$

At $h \to 0$, the flux of vector $\vec{E}$ through the lateral surface of the cylinder tends to zero. Then,

$$\begin{aligned}\oint_S \vec{E}d\vec{S} &= \vec{E}_2\Delta\vec{S}_2 + \vec{E}_1\Delta\vec{S}_1\\ &= \vec{n}\vec{E}_2\Delta S - \vec{n}\vec{E}_1\Delta S\\ &= (E_{2n} - E_{1n})\,\Delta S. \end{aligned}\tag{1.72}$$

From (1.71) and (1.72), it follows that

$$E_{2n} - E_{1n} = \frac{\sigma}{\varepsilon_0}, \tag{1.73}$$

i.e. while crossing the charged surface, the normal component of field strength varies in a jump-like manner by σ/ε_0.

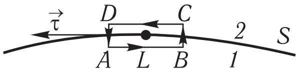

Fig. 1.7.

In order to elucidate the behavior of the tangential component of vector $\vec{E}$, we should construct the path L (Fig. 1.7) in the form of a small rectangle $ABCD$, the plane of which is perpendicular to the surface S and the sides AB and CD are on the opposite side of the surface. Figure 1.7 demonstrates the unit vector $\vec{\tau}$, which is tangent to the surface S and is located in the plane of rectangle $ABCD$, as well as the path-tracing direction. Since the contour L is small, we may write (see (1.67a))

$$\oint_L \vec{E}d\vec{l} = \oint_L E_l dl = E_{AB}AB + E_{BC}BC + E_{CD}CD + E_{DA}DA = 0.$$

From Fig. 1.7, it follows that $E_{AB} = -E_{1\tau}$, $E_{CD} = -E_{2\tau}$, and $CD = AB$. Therefore,

$$\oint_L \vec{E}d\vec{l} = (E_{2\tau} - E_{1\tau})\, AB + E_{BC}BC + E_{DA}DA = 0. \tag{1.74}$$

Moving sections AB and CD closer until they coincide on the surface S, i.e. in the limiting case $BC \to 0$ and $DA \to 0$, we obtain from (1.74) the following relation:

$$E_{2\tau} = E_{1\tau}. \tag{1.75}$$

It means that the tangential component of vector $\vec{E}$ does not change while crossing the charged surface.

Formulas (1.73) and (1.75) are boundary conditions for vector $\vec{E}$. On their basis, taking into account that

$$E_n = \vec{n}\vec{E} = \vec{n}\,(-\nabla\varphi) = -\frac{\partial\varphi}{\partial n}, \tag{1.76}$$

$$E_\tau = \vec{\tau}\vec{E} = \vec{\tau}\,(-\nabla\varphi) = -\frac{\partial\varphi}{\partial\tau}, \tag{1.77}$$

the behavior of potential derivatives across the charged surface can be determined,

$$\left(\frac{\partial\varphi}{\partial n}\right)_1 - \left(\frac{\partial\varphi}{\partial n}\right)_2 = \frac{\sigma}{\varepsilon_0}, \tag{1.78}$$

$$\left(\frac{\partial\varphi}{\partial\tau}\right)_1 = \left(\frac{\partial\varphi}{\partial\tau}\right)_2. \tag{1.79}$$

1.6. Energy of electrostatic field in vacuum

Let us consider a system of two point charges q_1 and q_2. The charge q_1 creates an electrostatic field, the potential of which at the point where the charge q_2 is located equals φ_2. Then, the potential energy of charge q_2 in the field of charge q_1, in accordance with (1.36), can be written in the form

$$U_{12} = q_2\varphi_2. \tag{1.80}$$

Analogously, the potential energy of charge q_1 in the field of charge q_2 should be written as

$$U_{21} = q_1\varphi_1. \tag{1.81}$$

where φ_1 is the potential created by the charge q_2 at the point where the charge q_1 is located. It is easy to understand that $U_{12} = U_{21}$. Hence, the potential energy of interaction between charges can be written in the symmetric form

$$U = \frac{1}{2}q_1\varphi_1 + \frac{1}{2}q_2\varphi_2. \tag{1.82}$$

Then, the energy of interaction of n point charges looks like

$$U = \frac{1}{2}\sum_{i=1}^{n} q_i\varphi_i, \tag{1.83}$$

where φ_i is the field potential created at the point where the charge q_i is located by other charges.

However, the electric point charge which is located at a geometric point is no more than a limiting, fictitious case. The real charge occupies some volume, i.e. there is interaction between its components, but this interaction is not taken into account in formula (1.83). Therefore, the charge

q_i in formula (1.83) should be substituted by $dq = \rho dV$, the potential φ_i by the potential φ created in the volume dV by the components of all charges, and the sum sign by the integral. Then, we obtain the formula

$$U = \frac{1}{2} \int \rho\varphi dV, \tag{1.84}$$

which makes allowance for the energy of interaction of both the components of different charges and the components of every charge with one another, i.e., in the latter case, the own energies of charges.

From (1.67b), it follows that $\rho = \varepsilon_0 \mathrm{div}\vec{E}$. Therefore,

$$U = \frac{1}{2}\varepsilon_0 \int \varphi \mathrm{div}\vec{E} dV. \tag{1.85}$$

On the basis of the formula (see Appendix B)

$$\mathrm{div}(\psi\vec{a}) = \psi \mathrm{div}\vec{a} + \vec{a}\nabla\psi,$$

we may write

$$\varphi \mathrm{div}\vec{E} = \mathrm{div}\left(\varphi\vec{E}\right) - \vec{E}\nabla\varphi.$$

Taking into account that $\nabla\varphi = -\vec{E}$, we obtain

$$U = \frac{1}{2}\varepsilon_0 \int \mathrm{div}\left(\varphi\vec{E}\right) dV + \frac{1}{2}\varepsilon_0 \int E^2 dV. \tag{1.86}$$

According to Gauss's theorem,

$$\int \mathrm{div}\vec{E}\left(\varphi\vec{E}\right) dV = \oint \varphi\vec{E}d\vec{S}, \tag{1.87}$$

When integrating in the left hand side in (1.87) over the whole electrostatic field volume, the integration surface in the right hand side is extended to infinity. Then, bearing in mind that $\varphi \sim r^{-1}$, $E \sim r^{-2}$, and $dS \sim r^2$, we arrive at the conclusion that the integral equals zero. As a result, instead of (1.86), we obtain the formula

$$U = \frac{1}{2}\varepsilon_0 \int E^2 dV, \tag{1.88}$$

which is equivalent to formula (1.84), but allows the energy of a system of charges to be interpreted as the energy of electrostatic field distributed in space with the bulk density

$$w = \frac{1}{2}\varepsilon_0 E^2. \tag{1.89}$$

2

MAGNETIC FIELD OF STATIONARY CURRENT

2.1. Electric current

The electrostatic field provides interaction between motionless electric charges. This interaction is also called **Coulomb interaction**. Let us proceed to the consideration of interaction between moving electric charges, i.e. currents.

The **electric current** is a directed motion of electric charges—electrons, ions, and so forth—that can move over macroscopic distances. Those charges are called **charge carriers**, and the electric current itself, the **conduction current**, unlike the **bias current** (see section 3.2).

In macroscopic electrodynamics, the electric charge is assumed to be distributed continuously. Therefore, the current is considered as the motion of electrically charged fluid. In what follows, under the term electric current, the current in metal conductors is meant.

The main characteristic of electric current is the **current strength** I, which is defined as follows:

$$I = \frac{dq}{dt}, \tag{2.1}$$

where dq is the electric charge transferred within the time interval dt through the transverse cross-section ΔS of conductor. If $I = \text{const}$, the current is called the **stationary current**, and if $I = I(t)$, the **alternating current**. In this section, we will consider only stationary currents.

The unit of current strength is the **amper** (1 A), which is defined on the basis of magnetic interaction between currents (see section 2.4).

The direction of positive charge motion is taken as the electric current direction. At the same time, the current strength does not take into account the current direction. Moreover, it is an integrated quantity that concerns the whole conductor cross-section ΔS. In this connection, the **vector of electric current density** $\vec{j}$, the absolute value of which is defined as follows:

$$\left|\vec{j}\right| = \frac{dI}{dS}, \tag{2.2}$$

and the direction coincides with that of the current, i.e. with the direction of motion of positively charged particles. Hence, the vector $\vec{j}$ can be determined at every point.

If the cross-section ΔS is not transverse, i.e. it is not perpendicular to the current direction (Fig. 2.1), we have

$$\frac{dI}{dS} = \frac{dI}{dS_0}\cos\alpha = j\cos\alpha = j_n, \tag{2.3}$$

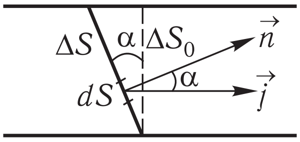

Fig. 2.1.

where $dS_0 = dS\cos\alpha$ is the element of transverse cross-section, and j_n is the projection of vector $\vec{j}$ onto the normal $\vec{n}$ to the cross-section ΔS.

From (2.3), it follows that the current strength through any surface S can be written in the form of the vector $\vec{j}$ flux,

$$I = \int_S j_n dS = \int_S \vec{j}d\vec{S}, \tag{2.4}$$

where $d\vec{S} = \vec{n}dS$. If the surface S is closed, the magnitude (strength) of the current flowing out from it equals

$$I = \oint_S j_n dS. \tag{2.5}$$

Then, according to the charge conservation law, the electric charge q inside the closed surface S has to decrease at the rate $-\partial q/\partial t$. Hence, we may write

$$-\frac{\partial q}{\partial t} = \oint j_n dS. \tag{2.6}$$

Taking into account that

$$q = \int_V \rho dV, \tag{2.7}$$

where V is the volume confined by the surface S, we obtain

$$-\frac{\partial}{\partial t}\int_V \rho dV = \oint j_n dS. \tag{2.8}$$

Applying Gauss's theorem (see Appendix B) to the right hand side of (2.8), we obtain the equality

$$\int_V \left(-\frac{\partial \rho}{\partial t}\right) dV = \int_V \operatorname{div} \vec{j}\, dV, \tag{2.9}$$

which gives rise to the equation

$$\frac{\partial \rho}{\partial t} + \operatorname{div} \vec{j} = 0 \tag{2.10}$$

called the **continuity equation.** This is a mathematical formulation for the law of conservation of electric charge.

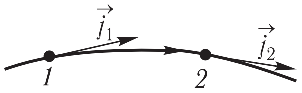

Fig. 2.2.

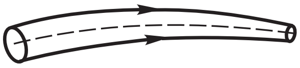

Fig. 2.3.

For illustrative purposes, the current is shown making use of **streamlines**. At every point of every stream line, the vector of current density $\vec{j}$ is directed tangentially (Fig. 2.2). A tube with the walls composed of streamlines is called the **tube of current** (Fig. 2.3). The strength of the current flowing through every cross-section of the tube of current is identical.

Linear, quasilinear, and bulk currents are distinguished. The current $I = j\Delta S$ is called **linear**, if the dimensions of any of its cross-sections are small enough in comparison with the distance from the cross-section to the observation point. It is clear that this condition is satisfied if the observation points are removed far enough from the current.

The current in a conductor is called **quasilinear** if the linearity condition are not obeyed, but the **current axis** can be so selected that, at every points belonging to any cross-section perpendicular to this axis, the values of such electric parameters as the current density $\vec{j}$, the potential

φ, the electric field strength $\vec{E}$, and so forth were identical. Then, the current can be considered to flow along the axis, which is called the **current path** in this case.

If neither linear nor quasilinear conditions are satisfied, the current is called the **bulk current**.

In what follows, we use the concept of **current element**. In the case of linear or quasilinear current, the current element is the quantity $Id\vec{l}$, where $d\vec{l}$ is the length element of conductor or current axis, the direction of which coincides with that of vector $\vec{j}$. In case of bulk current, the current element looks like $\vec{j}dV$. A formal transition from the linear current element, $Id\vec{l}$, to the bulk one, $\vec{j}dV$, is carried out as follows:

$$Id\vec{l} = j\Delta S d\vec{l} = \vec{j}\Delta S dl = \vec{j}dV.$$

2.2. Stationary electric field. Ohm's and Joule–Lenz laws

Below, we assume that the directed motion of electric charges occurs under the action of electric field. In the case of stationary current, the spatial distribution of charges does not change in time, i.e. $\partial\rho/\partial t = 0$; otherwise, the electric field would change and, hence, the current would not be constant. Therefore, if the current is constant, the spatial distribution of charges remains invariable in spite of their motion; the same as in the case of fixed charges. Therefore, the electric field of stationary current also does not change in time and is called **stationary**. Whence it follows that the stationary electric field, as well as the electrostatic one, is a potential field, i.e. the relation

$$\vec{E} = -\nabla\varphi \tag{2.11}$$

must be obeyed. On the basis of (2.10) and taking into account that, in the case of stationary electric field, $\partial\rho/\partial t = 0$, we obtain the equation

$$\operatorname{div}\vec{j} = 0, \tag{2.12}$$

which proves that the stationary current has no sources, at which the streamlines could start or end . In other words, the streamlines of stationary current are closed.

The main law for the stationary current is **Ohm's law**. In the case of quasilinear current in the conductor section, it reads

$$I = \frac{\varphi_1 - \varphi_2}{R}, \tag{2.13}$$

where φ_1 and φ_2 are the potentials at the beginning and the end, respectively, of conductor section oriented along the current direction, and R is the electrical resistance of the section.

The quantity

$$U = \varphi_1 - \varphi_2 = \int_1^2 \vec{E} d\vec{l}, \tag{2.14}$$

where $d\vec{l}$ is the section length element, is called the **electric voltage** across the section. Therefore, Ohm's law (2.13) is written in the form

$$I = \frac{U}{R}. \tag{2.15}$$

Ohm's law (2.15) is expressed in the integral form. In order to change to the differential form, let us suppose that the section of homogeneous conductor has the length l and the cross-section ΔS. Then,

$$R = \rho \frac{l}{\Delta S}, \tag{2.16}$$

where ρ is the resistivity (specific resistance) of conductor. In the case of homogeneous conductor, on the basis of (2.14), we may write

$$U = \varphi_1 - \varphi_2 = El. \tag{2.17}$$

After substituting (2.16) and (2.17) into (2.15), we obtain the relation

$$\frac{l}{\Delta S} = \frac{E}{\rho}$$

or

$$j = \sigma E, \tag{2.18}$$

where

$$\sigma = \frac{1}{\rho} \tag{2.19}$$

is the conductivity (specific conductance) of conductor. The vectors $\vec{j}$ and $\vec{E}$ are directed identically; therefore, instead of (2.18), we may write

$$\vec{j} = \sigma \vec{E}. \tag{2.20}$$

This is Ohm's law in the differential form.

If a current runs through the conductor, the heat generated in the conductor section per unit time, according to the **Joule–Lenz law**, equals

$$\frac{\Delta Q}{\Delta t} = I^2 R, \tag{2.21}$$

or, taking into account 2.16), (2.18), and (2.19),

$$q = \sigma E^2, \tag{2.22}$$

where $q = \Delta Q/(\Delta S l \Delta t)$ is the heat generated per unit time in a unit of conductor volume. Since $\vec{j} = \sigma \vec{E}$, instead of (2.22), we may write

$$q = \vec{j}\vec{E} = \frac{1}{\sigma} j^2. \tag{2.23}$$

Relation (2.22) or (2.23) is the Joule–Lenz law in the differential form.

Since rot grad = 0 (see Appendix B), on the basis of (2.11), we obtain the equation

$$\mathrm{rot}\ \vec{E} = 0. \tag{2.24}$$

Additionally, from (2.12) and (2.20), it follows that

$$\mathrm{div}\ \vec{E} = 0. \tag{2.25}$$

Equations (2.24) and (2.25) compose a set of equations for the stationary electric field generated by the stationary current. Comparing them with equations (1.59) and (1.61) for the electrostatic field, we notice that the fields are potential (rot $\vec{E} = 0$) in both cases, but the electrostatic field has sources, namely, electric charges (div $\vec{E} = \rho/\varepsilon_0$), whereas the stationary electric field of stationary currents has no sources and its strength lines are closed (div $\vec{E} = 0$). The difference between the electrostatic field and the stationary electric field of stationary currents is explained by the fact that no energy is needed to maintain the electrostatic field, whereas a permanent energy support is required to maintain the stationary electric field and compensate the released Joule–Lenz heat. The work executed by nonelectrostatic forces has to be the source of this energy. Such forces are called **external**.

They arise, in particular, when conductors with different chemical compositions are in contact, i.e. when they are associated with chemical interaction forces (galvanic cells, accumulators), in the case when conductors with different temperatures are in contact (thermoelements), and when a conductor moves in a magnetic field.

If the charge q undergoes the action of external force $\vec{F}^{ext}$, the ratio

$$\vec{E}^{ext} = \frac{\vec{F}^{ext}}{q} \tag{2.26}$$

is called the **strength of external force field**

Taking the external force field into account, Ohm's and Joule–Lenz laws must be generalized and written as follows

$$\vec{j} = \sigma\left(\vec{E} + \vec{E}^{ext}\right), \tag{2.27}$$

$$q = \vec{j}\left(\vec{E} + \vec{E}^{ext}\right). \tag{2.28}$$

Relation (2.27) is called the **generalized Ohm's law** in the differential form.

Let us consider a closed circuit with a stationary current, which can be regarded as quasilinear, and the integral

$$\mathscr{E} = \oint \left(\vec{E} + \vec{E}^{ext}\right) dl. \tag{2.29}$$

Since of the stationary electric field is potential, i.e.

$$\oint \vec{E} dl = 0, \tag{2.30}$$

then,

$$\mathscr{E} = \oint \vec{E}^{ext} dl. \tag{2.31}$$

The value of integral (2.31) is equal to the work executed by external forces when a positive unit charge is moved along the closed path. This quantity is called the **electromotive force**.

Let us multiply (2.27) by $(1/\sigma)d\vec{l} = \rho d\vec{l}$ and integrate the obtained equation from cross-section 1 to cross-section 2 in the conductor,

$$\int_1^2 \rho\vec{j}d\vec{l} = \int_1^2 \vec{E}d\vec{l} + \int_1^2 \vec{E}^{ext}d\vec{l}. \tag{2.32}$$

Since

$$\rho\vec{j}d\vec{l} = \rho j dl = \rho\frac{I}{\Delta S}dl = I\rho\frac{dl}{\Delta S} = IdR,$$

the left hand side of (2.32) takes the form

$$\int_1^2 \rho\vec{j}d\vec{l} = I\int_1^2 dR = IR_{12},$$

where R_{12} is the resistance of circuit section between cross-sections 1 and 2. The first integral in the right hand side of (2.32), in accordance with (2.14), equals the potential difference $\varphi_1 - \varphi_2$, and the second one, in accordance with (2.31), is the electromotive force $\mathscr{E}_{12}$ that acts between cross-sections 1 and 2. As a result, formula (2.32) reads

$$IR_{12} = \varphi_1 - \varphi_2 + \mathscr{E}_{12}. \tag{2.33}$$

This relation is the integral form of the **generalized Ohm's law**.

Integrating in (2.32) along the closed path, we obtain Ohm's law for a closed circuit in the integral form,

$$IR = \mathscr{E}, \tag{2.34}$$

where R is the total resistance of the circuit.

2.3. Magnetic interaction of currents. Ampère law. Magnetic induction

In the stationary field, electric charges move so that their spatial distribution does not change in time, and conductors remain electrically neutral. However, experimental data testify that conductors with currents interact with one another. Two parallel conductors with identically directed currents attract, whereas oppositely directed currents repel each other. Hence, interaction of currents is not electrostatic. At the same time, a conductor with a current was found (H. Ørsted) to interact with a permanent magnet. Therefore, interaction between conductors and currents is called the **magnetic interaction** of currents.

Since the electric current is created by moving charges, it means that magnetic interaction of currents gives rise to the magnetic interaction of charges. Therefore, there exists not only electrostatic but also magnetic interaction between moving charged particles.

According to the field theory (the theory of short-range interaction), magnetic interaction of currents, i.e. moving charges, takes place via the **magnetic field**; namely, every current generates a magnetic field in the surrounding space, and this field acts on other currents (moving charges). The magnetic field, as well as electrostatic one, is a material object with its own energy, momentum, and so forth.

The stationary magnetic field does not change in time. Therefore, it is called **stationary.** On the other hand, since it can also be generated by a permanent magnet, it is called **magnetostatic** as well.

The magnetic field acts on the electric current. Therefore, in order to study the former, a "test current" is required—by analogy with the test electric charge using to study the electrostatic field (see section 1.3). The current element $Id\vec{l}$ is used as a test current (the consideration of the difficulties associated with its "isolation" is omitted).

The force characteristic of magnetic field, which is analogous to that of electrostatic field (the field strength $\vec{E}$), is the **magnetic field induction $\vec{B}$**. The latter is introduced as follows. If the current element Idl $Id\vec{l}$ is arranged at a certain point in the "external" magnetic field, the latter will act on the former with a certain force $d\vec{F}$. The experiment shows that the magnitude of this force depends on both the current strength I and the current element orientation. At some orientation, the force $d\vec{F}$ turns out equal to zero. In this case, the direction of current element $Id\vec{l}$ determines the direction of magnetic induction vector $\vec{B}$. On the other hand, experimental data also demonstrate that the force dF is proportional to the product $Idl\sin\alpha$, where α is the angle between the current element and vector $\vec{B}$ directions. Hence, experimental data testify that the quantity

$$B = \frac{dF}{Idl\sin\alpha} \tag{2.35}$$

calculated for any field point does not depend on the current strength I in the current element and the angle α. Therefore, this is a characteristic of examined magnetic field, which is called the **absolute value of** the **magnetic induction vector** $\vec{B}$.

From (2.35), the relation

$$dF = IBdl\sin\alpha \tag{2.36}$$

or, in the vector form,

$$d\vec{F} = I\left[d\vec{l}\,B\right] \tag{2.37}$$

follows. The order of multipliers in the vector product is determine according to experimental data. Formula (2.37) is called **Ampère's law**.

The total ponderomotive force acting in an external magnetic field with the induction $\vec{B}$ on the closed loop L with the linear current I equals

$$\vec{F} = I\oint_L \left[d\vec{l}\,B\right]. \tag{2.38}$$

Illustratively, the magnetic field is represented making use of **magnetic flux lines** (the force lines of magnetic field), the curves, at every point of which the vector of magnetic induction $\vec{B}$ is tangent to them. The number of induction lines that cross a unit area perpendicular to the vector $\vec{B}$ must equal $\left|\vec{B}\right|$. The number of induction lines that cross the surface S is called the **magnetic flux** through this surface and is denoted as Φ. The flux is calculated using the formula (see also 1.27)

$$\Phi = \int_S \vec{B}d\vec{S} = \int_S B_n dS. \tag{2.39}$$

If the magnetic field is uniform, i.e. $\vec{B} \neq \vec{B}(\vec{r})$, the magnetic flux through the area of surface S oriented perpendicularly to the magnetic field direction equals

$$\Phi = BS. \tag{2.40}$$

In the SI system, the unit of magnetic flux is called **weber** (1 Wb). It is defined on the basis of Faraday's laws of electromagnetic induction (see section 3.1) as a magnetic flux through the loop area if its uniform reduction to zero in a unit time induces the appearance of electromotive force of 1 V in the loop. On the basis of the magnetic flux unit, Wb, the unit of magnetic induction is defined, which is called **tesla** (1 T). This is the induction of the uniform magnetic field, in which the magnetic flux that passes through a unit area of the surface perpendicular to the vector $\vec{B}$ equals 1 Wb.

2.4. Biot–Savart law

Let us elucidate how the magnetic induction is calculated. Let the magnetic field be created by a closed loop of linear current I. The loop consists of current elements $I d\vec{l}$'s. Every element creates a magnetic field with the induction $d\vec{B}$ at the observation point, and the resulting of magnetic induction $\vec{B}$ is a vector sum of inductions $d\vec{B}$'s.

In turn, the magnetic induction $d\vec{B}$ created by a current element $I d\vec{l}$ is determined using the law discovered by the French physicists J.-B. Biot and F. Savart in 1820 and formulated in the generalized form by P. Laplace. The **Biot–Savart law** looks like

$$d\vec{B} = k_1 \frac{I\left[d\vec{l}\,\vec{r}\right]}{r^3}, \tag{2.41}$$

where k_1 is a coefficient (its value will be determined later), $\vec{r}$ the radius vector drawn from the current element $I d\vec{l}$ to the observation point A (Fig. 2.4). In this figure, S is the plane of vectors $d\vec{l}$ and $\vec{r}$, and the vector $d\vec{B}$ is perpendicular to the plane S. The absolute value of vector $d\vec{B}$ is determined as

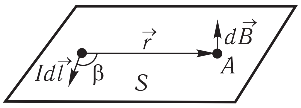

Fig. 2.4.

$$dB = k_1 \frac{I dl \sin\beta}{r^2}, \tag{2.42}$$

where β is the angle shown in Fig. 2.4.

Using formulas (2.41) and (2.42), it is easy in many cases to calculate the total magnetic induction $\vec{B}$. In the case of a ring current of the radius r, if the point observation is located at the loop center, $\beta = \pi/2$ for all current elements, and all elements create the induction $d\vec{B}$ perpendicular to the loop plane, with

$$dB = k_1 \frac{I dl}{r^2}.$$

Therefore, the total magnetic induction at the center of ring current equals

$$B = k_1 \frac{2\pi I}{r}. \tag{2.43}$$

The direction of vector $\vec{B}$ is governed by the direction of current in the loop according to the right-hand screw rule.

With the help of a little more complicated calculation, it is possible to demonstrate that, in case of a current running in a long straight wire, the magnetic field induction at the distance r from the wire is determined by the formula

$$B = k_1 \frac{2I}{r}. \tag{2.44}$$

The vector $\vec{B}$ itself is in the plane perpendicular to the current, it is tangent to the circle of the radius r lying in this plane, and its direction is related with the current direction also by the right-hand screw rule.

In order to determine the value of coefficient k_1, let us consider interaction of two long parallel linear currents I_1 and I_2 separated by the distance r. At the distance r, current I_1 creates a magnetic field with the induction (see formula (2.44))

$$B_1 = k_1 \frac{2I_1}{r},$$

and the vector $\vec{B}_1$ is perpendicular to the direction of current I_2. Therefore, on the basis of relation (2.36), the force acting on the section Δl_2 of current I_2 can be written as

$$\Delta F_2 = I_2 B_1 \Delta l_2 = k_1 \frac{2I_1 I_2}{r} \Delta l_2.$$

The unit length of current I_2 undergoes the action of force

$$\frac{\Delta F_2}{\Delta l_2} = k_1 \frac{2I_1 I_2}{r}.$$

It is easy to be convinced that the same force acts on the unit length of current I_1, i.e. for both currents,

$$\frac{\Delta F}{\Delta l} = k_1 \frac{2I_1 I_2}{r}. \tag{2.45}$$

This formula is called **Ampère's law**.

With the help of formula (2.37), it can be found that two parallel (with the same direction) currents attract, whereas the antiparallel ones (with the opposite directions) repulse each other.

In the SI system, the current strength unit, **amper** (1 A), is defined as the strength of a current that, while running in two straight and infinitely long parallel conductors with negligible circular cross section placed 1 m apart in vacuum, produces an interaction force between them equal to 2×10^{-7} N per every meter of length. Then, provided that $I_1 = I_2 = 1$ A, $r = 1$ m, and $\Delta F/\Delta l = 2 \times 10^{-7}$ N/m in formula (2.45), we obtain

$$\frac{\Delta F}{\Delta l} = 2 \times 10^{-7}\ \frac{\mathrm{N}}{\mathrm{m}} = k_1 \times 2\ \frac{\mathrm{A}^2}{\mathrm{m}}.$$

Whence it follows that

$$k_1 = 10^{-7}\ \frac{\mathrm{N}}{\mathrm{m}}.$$

In the SI system, the coefficient k_1 is written in the form (cf. formula 1.11)

$$k_1 = \frac{\mu_0}{4\pi}, \tag{2.46}$$

where the quantity $\mu_0 = 4\pi \times 10^{-7}\ \mathrm{N/A}^2$ is called **the magnetic constant** (cf. formula 1.12). Besides, in the SI system, $1\ \mathrm{N/A}^2 = 1$ H/m; 1 H (**henry**)is the unit of inductance. Therefore,

$$\mu_0 = 4\pi \times 10^{-7}\ \frac{\mathrm{H}}{\mathrm{m}}. \tag{2.47}$$

As a result, the Biot–Savart law (2.41) reads

$$d\vec{B} = \frac{\mu_0}{4\pi} \frac{I\left[d\vec{l}\,\vec{r}\right]}{r^3}. \tag{2.48}$$

The magnetic induction at the center of ring current equals

$$B = \mu_0 \frac{I}{2r}, \tag{2.49}$$

and the magnetic induction of straight current is written as

$$B = \mu_0 \frac{I}{2\pi r}. \tag{2.50}$$

In the case of bulk currents, the element of linear current $Id\vec{l}$ should be substituted by the element of bulk current $\vec{j}dV$ (see section 2.2). Then, instead of formula (2.48), we obtain

$$d\vec{B} = \frac{\mu_0}{4\pi}\frac{I\left[\vec{j}\,\vec{r}\right]}{r^3}dV. \tag{2.51}$$

Let us transform this formula to a form that is more convenient in applications. For this purpose, we designate the coordinate of current element $\vec{j}dV'$ as $\vec{r}\,'$ and the coordinate of observation point as $\vec{r}$ (Fig. 2.5). Then, instead of formula (2.51), we obtain

$$d\vec{B} = \frac{\mu_0}{4\pi}\frac{\left[\vec{j}(\vec{r}\,')(\vec{r}-\vec{r}\,')\right]}{|\vec{r}-\vec{r}\,'|^3}dV'. \tag{2.52}$$

On the basis of formula (2.52), the total magnetic induction $\vec{B}(\vec{r})$ is calculates in the form

$$\vec{B}(\vec{r}) = \frac{\mu_0}{4\pi}\int\frac{\left[\vec{j}(\vec{r}\,')(\vec{r}-\vec{r}\,')\right]}{|\vec{r}-\vec{r}\,'|^3}dV', \tag{2.53}$$

where $dV' = dx'dy'dz'$.

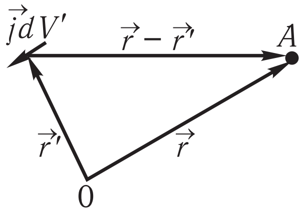

Fig. 2.5.

Let us consider two closed contours L_1 and L_2 with the linear currents I_1 and I_2, respectively, and distinguish the current elements $I_1d\vec{l}_1$ and $I_2d\vec{l}_2$ in them. The element $I_1d\vec{l}_1$ of the first contour creates a magnetic field at the point, where the element $I_2d\vec{l}_2$ of the second contour is located. According to formula (2.48), the induction of this field equals

$$d\vec{B}_{12} = \frac{\mu_0}{4\pi}\frac{I_1\left[d\vec{l}_1\vec{r}_{12}\right]}{r_{12}^3}, \tag{2.54}$$

where $\vec{r}_{12}$ is the radius vector drawn from the element $I_1d\vec{l}_1$ to the element $I_2d\vec{l}_2$ (see Fig. 2.6). Then, on the basis of relation (2.37), the force with which the element $I_1d\vec{l}_1$ acts on the element $I_2d\vec{l}_2$ is written as

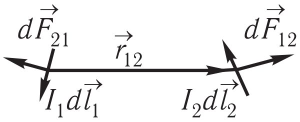

Fig. 2.6.

$$d\vec{F}_{12} = I_2\left[d\vec{l}_2 d\vec{B}_{12}\right] = \mu_0\frac{I_1I_2}{4\pi r_{12}^3}\left[d\vec{l}_2\left[d\vec{l}_1\vec{r}_{12}\right]\right]. \tag{2.55}$$

It is easy to be convinced that the element $I_2 d\vec{l}_2$ acts on the element $I_1 d\vec{l}_1$ with the force

$$d\vec{F}_{21} = \mu_0 \frac{I_1 I_2}{4\pi r_{21}^3} \left[d\vec{l}_1 \left[d\vec{l}_2 \vec{r}_{21} \right] \right], \tag{2.56}$$

where $\vec{r}_{21} = -\vec{r}_{12}$. The whole contour L_2 undergoes the action of the force

$$\vec{F}_{12} = \mu_0 \frac{I_1 I_2}{4\pi} \oint_{L_2} \frac{\left[d\vec{l}_2 \left[d\vec{l}_1 \vec{r}_{12} \right] \right]}{r_{12}^3}, \tag{2.57}$$

and the contour L_1, of the force

$$\vec{F}_{21} = \mu_0 \frac{I_1 I_2}{4\pi} \oint_{L_1} \frac{\left[d\vec{l}_1 \left[d\vec{l}_2 \vec{r}_{21} \right] \right]}{r_{21}^3}. \tag{2.58}$$

From Fig. 2.6, it follows that $d\vec{F}_{12} \neq d\vec{F}_{21}$. Hence, in the case of interaction between two current elements, Newton's third law is not obeyed. However, it can be proved that $\vec{F}_{12} = \vec{F}_{21}$, i.e. the third law of motion holds true in the case of interaction between the loops (see section 5.1 in [1]).

2.5. Vector potential of magnetic field

The calculation of magnetic induction can be facilitated by introducing an auxiliary function called the **vector potential** of magnetic field and denoted as $\vec{A}(\vec{r})$. Let us express the integrand in (2.53) in the form

$$\frac{\left[\vec{j}(\vec{r}\,')(\vec{r} - \vec{r}\,') \right]}{|\vec{r} - \vec{r}\,'|^3} = - \left[\vec{j}(\vec{r}\,') \nabla \frac{1}{|\vec{r} - \vec{r}\,'|} \right] = \left[\nabla \frac{1}{|\vec{r} - \vec{r}\,'|} \vec{j}(\vec{r}\,') \right], \tag{2.59}$$

where, while calculating the gradient of the function $|\vec{r} - \vec{r}\,'|^{-1}$, the differentiation is carry with respect to the coordinates of field point (the observation point), i.e. with respect to the coordinates (x, y, z) of the terminal point of vector $\vec{r}$ (see Fig. 2.5).

Using the following formula of vector analysis (see Appendix B),

$$\text{rot}\,(\psi \vec{a}) = \psi \text{ rot } \vec{a} + [\nabla \psi \; \vec{a}],$$

and assuming that

$$\psi = \frac{1}{|\vec{r}-\vec{r}'|}, \vec{a} = \vec{j}(\vec{r}'),$$

we obtain

$$\left[\nabla \frac{1}{|\vec{r}-\vec{r}'|}\vec{j}(\vec{r}')\right] = \text{rot}\frac{\vec{j}(\vec{r}')}{|\vec{r}-\vec{r}'|} - \frac{1}{|\vec{r}-\vec{r}'|}\text{rot}\ \vec{j}(\vec{r}'). \tag{2.60}$$

Since the vector $\vec{j}(\vec{r}')$ does not depend on the observation point coordinate, rot $\vec{j}(\vec{r}')$. Therefore, expression (2.59) takes the form

$$\frac{\left[\vec{j}(\vec{r}')(\vec{r}-\vec{r}')\right]}{|\vec{r}-\vec{r}'|^3} = \text{rot}\frac{\vec{j}(\vec{r}')}{|\vec{r}-\vec{r}'|},$$

and formula (2.53) looks like

$$\vec{B}(\vec{r}) = \frac{\mu_0}{4\pi}\int \text{rot}\frac{\vec{j}(\vec{r}')}{|\vec{r}-\vec{r}'|}dV'. \tag{2.61}$$

In (2.61), integration is carried out over the current coordinates, whereas the differentiation in the rot operator is fulfilled with respect to the field ones. Therefore, the order of integration and differentiation can changed, so that

$$\vec{B}(\vec{r}) = \text{rot}\left\{\frac{\mu_0}{4\pi}\int \frac{\vec{j}(\vec{r}')}{|\vec{r}-\vec{r}'|}dV'\right\}$$

or

$$\vec{B}(\vec{r}) = \text{rot}\ \vec{A}(\vec{r}), \tag{2.62}$$

where the quantity

$$\vec{A}(\vec{r}) = \frac{\mu_0}{4\pi}\int \frac{\vec{j}(\vec{r}')dV'}{|\vec{r}-\vec{r}'|} \tag{2.63}$$

is called the **vector potential** of magnetic field.

In the case of linear current, after making the transformation $\vec{j}(\vec{r}')dV' \rightarrow Id\vec{l}(\vec{r}')$ in formula (2.63), we obtain the vector potential in the form

$$\vec{A}(\vec{r}) = \frac{\mu_0 I}{4\pi}\oint \frac{d\vec{l}(\vec{r}')}{|\vec{r}-\vec{r}'|}. \tag{2.64}$$

The vector potential of magnetic field, $\vec{A}$, as well as the scalar potential of electrostatic field, φ, is an ambiguous function. Really, if it is summed up with the gradient of any scalar functions of coordinates, $f(\vec{r})$, i.e. in the case of substitution $\vec{A} \to \vec{A}' = \vec{A} + \text{grad } f$, the vector of magnetic induction does not change,

$$\vec{B}' = \text{rot } \vec{A}' = \text{rot } \vec{A} + \text{rot grad } f = \text{rot } \vec{A} = \vec{B}.$$

Here, we took into account that rot grad $\equiv 0$ (see Appendix B). This fact makes it possible to impose an additional condition on the vector potential. The most widely used is the condition

$$\text{div } \vec{A} = 0. \tag{2.65}$$

If div $\vec{A} \neq 0$, we can make such a substitution $\vec{A} \to \vec{A}' = \vec{A} + \text{grad } f$ that the vector $\vec{A}'$ would satisfy condition (2.65). The choice of function f is determined according to the equation

$$\text{div } \vec{A}' = \text{div } \vec{A} + \text{div grad } f = \text{div } \vec{A} + \Delta f = 0,$$

or, equivalently,

$$\Delta f = -\text{div } \vec{A},$$

which is similar to Poisson's equation (1.63). Therefore, its solution looks like

$$f(\vec{r}) = \frac{1}{4\pi} \int \frac{\text{div } \vec{A}(\vec{r}\,')}{|\vec{r} - \vec{r}\,'|} dV'.$$

Hence, condition (2.65) can always be satisfied. From (2.63), the formulas

$$\vec{A}_i(\vec{r}) = \frac{\mu_0}{4\pi} \int \frac{\vec{j}_i(\vec{r}\,')dV'}{|\vec{r} - \vec{r}\,'|}, \quad i = 1, 2, 3, \tag{2.66}$$

follow, the form of which is identical to that of formula (1.66) used to determine the scalar potential,

$$\varphi(\vec{r}) = \frac{1}{4\pi\varepsilon_0} \int \frac{\rho(\vec{r}\,')dV'}{|\vec{r} - \vec{r}\,'|}. \tag{2.67}$$

The latter is the general solution of Poisson's equation for the scalar potential,

$$\Delta\varphi = -\frac{\rho}{\varepsilon_0}. \tag{2.68}$$

Therefore, formulas (2.66) are the general solutions of the equations

$$\Delta A_i = -\mu_0 j_i, \quad i = 1, 2, 3,$$

and (2.63) is the general solution of Poisson's equation for the vector potential,

$$\Delta \vec{A} = -\mu_0 \vec{j}. \tag{2.69}$$

2.6. Magnetic field of distant currents. Magnetic moment of the current

Let a magnetic field be created by currents located in the space volume V' with the linear dimensions a. The coordinate origin is in the current section. Consider the vector potential

$$\vec{A}(\vec{r}) = \frac{\mu_0}{4\pi} \int \frac{\vec{j}(\vec{r}\,')dV'}{|\vec{r} - \vec{r}\,'|} \tag{2.70}$$

at distant points in space, provided that $a \ll r$. As was indicated in section 1.3, in this case,

$$\frac{1}{|\vec{r} - \vec{r}\,'|} \approx \frac{1}{r} + \frac{\vec{r}\vec{r}\,'}{r^3}. \tag{2.71}$$

Therefore, substituting (2.71) into (2.70), we obtain

$$\vec{A}(\vec{r}) = \frac{\mu_0}{4\pi r} \int \vec{j}(\vec{r}\,')dV' + \frac{\mu_0}{4\pi r^3} \int (\vec{r}\vec{r}\,')\,\vec{j}(\vec{r}\,')dV'. \tag{2.72}$$

In the case of constant currents,

$$\int \vec{j}(\vec{r}\,')dV' = 0. \tag{2.73}$$

Really, taking into account that

$$\vec{j}dV' = \vec{j}dS'dl(\vec{r}\,') = jdS'd\vec{l}(\vec{r}\,') = dId\vec{l}(\vec{r}\,') \tag{2.74}$$

where dI is the current strength in the tube of current with the cross-section dS', for stationary currents with closed streamlines, it is possible to write

$$\int \vec{j}(\vec{r}\,')dV' = \int dI \oint d\vec{l} = 0,$$

because the integral of complete differential over the closed path equals zero. Hence, we obtain

$$\vec{A}(\vec{r}) = \frac{\mu_0}{4\pi r^3}\int \vec{j}(\vec{r}\,')\,(\vec{r}\vec{r}\,')\,dV'. \tag{2.75}$$

The integrand in (2.75) is transformed using the identity

$$\vec{j}\ (\vec{r}\vec{r}\,') = \frac{1}{2}\left\{\vec{j}\ (\vec{r}\vec{r}\,') - \vec{r}\,'\left(\vec{r}\vec{j}\right)\right\} + \frac{1}{2}\left\{\vec{j}\ (\vec{r}\vec{r}\,') + \vec{r}\,'\left(\vec{r}\vec{j}\right)\right\}. \tag{2.76}$$

The expression in the first brackets in (2.76) can be written as the double vector product,

$$\vec{j}\ (\vec{r}\vec{r}\,') - \vec{r}\,'\left(\vec{r}\vec{j}\right) = \left[\vec{r}\left[\vec{j}\vec{r}\,'\right]\right] = \left[\left[\vec{r}\,'\vec{j}\right]\vec{r}\right].$$

Therefore,

$$\vec{A}(\vec{r}) = \frac{\mu_0}{8\pi r^3}\int \left[\left[\vec{r}\,'\vec{j}\right]\vec{r}\right]dV' + \frac{\mu_0}{8\pi r^3}\int \left\{\vec{r}\,'\left(\vec{r}\vec{j}\right) + \vec{j}\ (\vec{r}\vec{r}\,')\right\}dV'. \tag{2.77}$$

With the help of transformation $\vec{j}dV' = dId\vec{l}$, the second integral in (2.77) is written as

$$\begin{aligned} &\int \left\{\vec{r}\,'\left(\vec{r}\vec{j}\right) + \vec{j}\ (\vec{r}\vec{r}\,')\right\}dV' = \int dI \oint \left\{\vec{r}\,'\left(\vec{r}d\vec{l}\right) + d\vec{l}\,(\vec{r}\vec{r}\,')\right\} \\ = &\int dI \oint \left\{\vec{r}\,'\,(\vec{r}d\vec{r}) + d\vec{r}\,(\vec{r}\vec{r}\,')\right\} = \int dI \oint d\left\{\vec{r}\,'\,(\vec{r}\vec{r}\,')\right\} = 0. \end{aligned}$$

Here, we took into account that $d\vec{l} = d\vec{r}$, because the variation of the charge coordinate $d\vec{r}$ during the charge motion along the streamline coincides with $d\vec{l}$.

As a result, instead of (2.77), we obtain the expression

$$\vec{A}(\vec{r}) = \frac{\mu_0}{8\pi r^3}\int \left[\left[\vec{r}\,'\vec{j}\right]\vec{r}\right]dV' = \frac{\mu_0}{8\pi r^3}\left[\frac{1}{2}\int \left[\vec{r}\,'\vec{j}\right]dV'\ \vec{r}\right]. \tag{2.78}$$

The quantity

$$\vec{d}_m = \frac{1}{2}\int \left[\vec{r}'\vec{j}\right] dV' \tag{2.79}$$

is called the **magnetic moment of current**. With its help, the vector potential of the magnetic field of distant currents (2.78) can be written in the form

$$\vec{A}(\vec{r}) = \frac{\mu_0 \left[\vec{d}_m\vec{r}\right]}{4\pi r^3}. \tag{2.80}$$

In the case of plane loop L with the linear current I (Fig. 2.7), making the substitution $\vec{j}dV' \to Id\vec{l}$ in (2.79), we obtain

$$\vec{d}_m = \frac{1}{2}I\oint_L \left[\vec{r}'d\vec{l}\right] = I\int_L d\vec{S} = I\vec{S}, \tag{2.81}$$

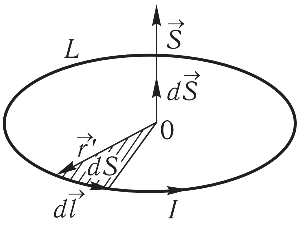

Fig. 2.7.

where $\left|d\vec{S}\right| = \frac{1}{2}\left|\left[\vec{r}'d\vec{l}\right]\right|$ is the area of triangle constructed on the vectors $\vec{r}'$ and $d\vec{l}$, and $\left|\vec{S}\right|$ is the area of loop L.

Now consider the magnetic moment of the current created by a particle with the mass m_0 and the electric charge e, which moves with the velocity $\vec{v}$. In this case, the charge density, in accordance with (1.7), equals $\rho(\vec{r}') = e\delta(\vec{r} - \vec{r}')$, where $\vec{r}$ is the particle's coordinate, and $\vec{r}'$ is an arbitrary coordinate. Then, $\vec{j}(\vec{r}') = \rho(\vec{r}')\vec{v} = e\vec{v}\delta(\vec{r} - \vec{r}')$. Hence,

$$\begin{aligned}\vec{d}_m &= \frac{1}{2}\int \left[\vec{r}'\vec{j}(\vec{r}')\right] dV' = \frac{1}{2}e\int \left[\vec{r}'\vec{v}\right]\delta(\vec{r} - \vec{r}')dV' \\ &= \frac{1}{2}e\left[\vec{r}\vec{v}\right] = \frac{e}{2m_0}\left[\vec{r}\vec{p}\right] = \frac{e}{2m_0}\vec{L},\end{aligned} \tag{2.82}$$

where $\vec{p} = m_0\vec{v}$ is the momentum and $\vec{L} = [\vec{r}\vec{p}]$ the angular momentum of the particle. Thus, the magnetic moment of the moving charged particle is proportional to its mechanical momentum (angular momentum). The proportionality coefficient $e/(2m_0)$ is called the **gyromagnetic ratio**.

Let us calculate the magnetic field induction far from currents. Using the formula $\vec{B} = \text{rot}\,\vec{A}$ and the vector potential in form (2.80), we obtain

$$\begin{aligned}\vec{B}(\vec{r}) &= \frac{\mu_0}{4\pi}\text{rot}\left\{\frac{1}{r^3}\left[\vec{d}_m\vec{r}\right]\right\} \\ &= \frac{\mu_0}{4\pi}\left\{\frac{1}{r^3}\text{rot}\left[\vec{d}_m\vec{r}\right] - \left[\left[\vec{d}_m\vec{r}\right]\nabla\frac{1}{r^3}\right]\right\}. \qquad (2.83)\end{aligned}$$

Here, both the rot and the gradient are calculated provided that the variables are the coordinates of the observation point, i.e. the terminal point of vector $\vec{r}$. In the first summand (see Appendix B),

$$\text{rot}\left[\vec{d}_m\vec{r}\right] = (\vec{r}\nabla)\,\vec{d}_m - \left(\vec{d}_m\nabla\right)\vec{r} + \vec{d}_m\text{div}\;\vec{r} - \vec{r}\text{div}\;\vec{d}_m.$$

Since $\vec{d}_m$ is independent of $\vec{r}$, $(\vec{r}\nabla)\,\vec{d}_m$ and div $\vec{d}_m = 0$.Therefore,

$$\text{rot}\left[\vec{d}_m\vec{r}\right] = -\left(\vec{d}_m\nabla\right)\vec{r} + \vec{d}_m\text{div}\;\vec{r} = -\vec{d}_m + 3\vec{d}_m = 2\vec{d}_m.. \qquad (2.84)$$

The second summand in the right hand side of (2.83) is transformed using the property of vector double product,

$$\begin{aligned}\left[\left[\vec{d}_m\vec{r}\right]\nabla\frac{1}{r^3}\right] &= -\frac{3}{r^5}\left[\left[\vec{d}_m\vec{r}\right]\vec{r}\right] = \frac{3}{r^5}\left\{\vec{d}_m r^2 - \vec{r}\left(\vec{d}_m\vec{r}\right)\right\} \\ &= \frac{3\vec{d}_m}{r^3} - \frac{3\vec{r}\left(\vec{d}_m\vec{r}\right)}{r^5}. \qquad (2.85)\end{aligned}$$

Substituting (2.84) and (2.85) into (2.83), we obtain

$$\vec{B}(\vec{r}) = \frac{\mu_0}{4\pi}\left\{-\frac{\vec{d}_m}{r^3} + 3\frac{\left(\vec{d}_m\vec{r}\right)\vec{r}}{r^5}\right\}. \qquad (2.86)$$

Hence, at large distances from currents, the magnetic induction associated with the second summands in expansions (2.71) and (2.72) decreases as r^{-3} as r grows. The magnetic induction associated with rejected summands (see (1.46)) decreases even more rapidly. Therefore, it is less substantial and, as a rule, is not taken into consideration.

2.7. Force on a current in a magnetic field

According to (2.38), the loop L with a linear current undergoes that action of the force

$$\vec{F} = I \oint \left[d\vec{l}(\vec{r})\vec{B}(\vec{r}) \right] \tag{2.87}$$

in the external magnetic field with the induction $\vec{B}(\vec{r})$. Here, $d\vec{l}(\vec{r})$ is the element to loop L. Changing to bulk currents with the help of transformation $Id\vec{l} \to \vec{j}dV$, the force $\vec{F}$ is obtained in the form

$$\vec{F} = \int \left[\vec{j}(\vec{r})\vec{B}(\vec{r}) \right] dV, \tag{2.88}$$

with integration is carried out over the whole current volume. Formula (2.88) demonstrate that every unit volume of currents undergoes the action of ponderomotive force

$$\vec{f}(\vec{r}) = \left[\vec{j}(\vec{r})\vec{B}(\vec{r}) \right]. \tag{2.89}$$

Besides, the unit current volume undergoes the action of the force moment $\vec{m}$ from the magnetic field, which in the mechanics is calculated as

$$\vec{m}(\vec{r}) = \left[\vec{r}\vec{f}(\vec{r}) \right]. \tag{2.90}$$

Hence, the force momentum acting on the whole system of currents equals

$$\vec{M} = \int \left[\vec{r}\vec{f}(\vec{r}) \right] dV = \int \left[\vec{r} \left[\vec{j}(\vec{r})\vec{B}(\vec{r}) \right] \right] dV. \tag{2.91}$$

Calculating the double vector product in (2.91), we obtain

$$\vec{M} = \int \vec{j}(\vec{r}) \left(\vec{r}\vec{B}(\vec{r}) \right) dV - \int \vec{B}(\vec{r}) \left(\vec{r}\vec{j}(\vec{r}) \right) dV. \tag{2.92}$$

Let the current section be small and the vector $\vec{B}$ be constant within its limits. Then,

$$\vec{M} = \int \vec{j}(\vec{r}) \left(\vec{r}\vec{B} \right) dV - \vec{B} \int \vec{r}\vec{j}(\vec{r}) dV. \tag{2.93}$$

Let us demonstrate that the second integral in (2.93) equals zero. Using the following formula of the vector analysis: $\operatorname{div}(\psi\vec{a}) = \psi\operatorname{div}\vec{a} + \vec{a}\nabla\psi$, and taking into account that $\operatorname{div}\vec{j} = 0$ in the stationary case, it is easy to verify the relation

$$\vec{r}\vec{j} = \frac{1}{2}\left\{\operatorname{div}\left(r^2\vec{j}\right) - r^2\operatorname{div}\vec{j}\right\} = \frac{1}{2}\operatorname{div}\left(r^2\vec{j}\right).$$

Then, with the help of Gauss's theorems, we obtain

$$\int \vec{r}\vec{j}(\vec{r})dV = \frac{1}{2}\int \operatorname{div}\left(r^2\vec{j}\right)dV = \frac{1}{2}\oint_S r^2 j_n dS = 0,$$

because the surface S envelops the current volume V, so that $j_n = 0$. Hence, formula (2.93) takes the form

$$\vec{M} = \int \vec{j}(\vec{r})\left(\vec{r}\vec{B}\right)dV. \tag{2.94}$$

A similar integral was considered in (2.75). Making allowance for (2.80), it looks like

$$\int \vec{j}(\vec{r})\,(\vec{r}\,'\vec{r})\,dV' = \left[\vec{d}_m\vec{r}\right],$$

i.e. differs from (2.94) by the $\vec{B} \to \vec{r}$ substitution. Therefore, it is possible to write

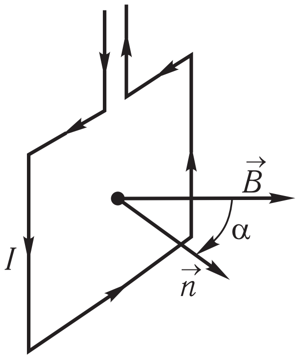

Fig. 2.8.

$$\vec{M} = \left[\vec{d}_m\vec{B}\right]. \tag{2.95}$$

A particular case of a current in the magnetic field is a loop of small area S with the linear current I and the magnetic moment $\vec{d}_m = IS\vec{n}$ (see (2.81)), where $\vec{n}$ is the unit vector normal to the loop area; its direction is coupled with the direction of current in the loop by the right-hand screw rule (Fig. 2.8). The torque created by the external magnetic field and acting on the loop, in accordance with (2.95), equals

$$\vec{M} = IS\left[\vec{n}\vec{B}\right]. \tag{2.96}$$

The absolute value of vector $\vec{M}$ is

$$\left|\vec{M}\right| = ISB\sin\alpha, \tag{2.97}$$

where α is the angle between the vectors $\vec{B}$ and $\vec{n}$ (Fig. 2.8). From (2.97), it follows the relation

$$B = \frac{\left|\vec{M}\right|}{IS\sin\alpha}, \tag{2.98}$$

which can be used for the determination of the absolute value of magnetic induction vector.

2.8. Equations for the stationary magnetic field in vacuum

Differential equations for the magnetic induction vector follow from (2.62) and (2.69), i.e. from the equations

$$\vec{B} = \text{rot } \vec{A}, \tag{2.99}$$

$$\Delta\vec{A} = -\mu_0 j. \tag{2.100}$$

Applying the divergence and rot operators to relation (2.99) in turn, we obtain

$$\text{div } \vec{B} = \text{div rot } \vec{A}, \tag{2.101}$$

$$\text{rot } \vec{B} = \text{rot rot } \vec{A}. \tag{2.102}$$

Since div rot $=$ 0 and rot rot $=$ grad div $-\ \Delta$ (see Appendix B), the following equations stem from (2.101) and (2.102):

$$\begin{aligned} \text{div } \vec{B} &= 0, \\ \text{rot } \vec{B} &= \text{grad div } \vec{A} - \Delta\vec{A}. \end{aligned}$$

Taking into consideration the conditions div $\vec{A} = 0$ (see (2.65)), equations (2.99) and (2.100) read

$$\text{div } \vec{B} = 0, \tag{2.103}$$

$$\text{rot } \vec{B} = \mu_0\vec{j}. \tag{2.104}$$

Equations (2.103) and (2.104) are differential equations for the magnetic induction vector in vacuum.

Let us determine integral relations corresponding to differential equation (2.103) and (2.104). First, consider a surface S confined by the contour L. For the flux of vector rot $\vec{B}$ through it, according to (2.104), we obtain

$$\int_S \text{rot } \vec{B} d\vec{S} = \mu_0 \int_S \vec{j} d\vec{S}. \tag{2.105}$$

Let us apply the Stokes theorem (see Appendix B) to the left hand side of this equation. Then,

$$\int_S \text{rot } \vec{B} d\vec{S} = \oint_L \vec{B} d\vec{l}.$$

At the same time, the integral in the right hand side of (2.105) is the total current I that passes through the surface S. Hence,

$$\oint_L \vec{B} d\vec{l} = \mu_0 I, \tag{2.106}$$

i.e. the circulation of vector $\vec{B}$ along a closed contour is proportional to the total current passing through the surface confined by this contour. This statement is called the **law of total current**. It is an integral relation corresponding to differential equation (2.104).

In order to find an integral relation corresponding to equation (2.103), let us integrate the latter over the volume V confined by the closed surface S,

$$\int_V \text{div } \vec{B} dV = 0.$$

Applying Gauss's theorem (see Appendix B)

$$\int_V \text{div } \vec{B} dV = \oint_S \vec{B} d\vec{S},$$

we come to a conclusion that

$$\oint_S \vec{B} d\vec{S} = 0. \tag{2.107}$$

This is the integral equation that corresponds to equation (2.103).

Let us compare the form and the content of basic equations for electrostatic and magnetostatic fields. As **magnetostatic**, we call magnetic fields of stationary currents and permanent magnets. The magnetic fields of permanent magnets will not be considered; therefore, under the term magnetostatic field, we will understand the magnetic field generated by stationary currents.

The equations of electrostatic field in the differential and integral forms look like (cf. (1.66) and (1.67))

$$\operatorname{rot} \vec{E} = 0, \qquad \oint_L \vec{E} d\vec{l} = 0, \tag{2.108}$$

$$\operatorname{div} \vec{E} = \frac{\rho}{\varepsilon_0}, \qquad \oint_S \vec{E} d\vec{S} = \frac{q}{\varepsilon_0}. \tag{2.109}$$

Equations (2.108) testify that the electrostatic field is irrotational. It means that the strength lines of electrostatic field are open. They either begin and end at electric charges or go to the infinity. The linear integral

$$\int_1^2 \vec{E} d\vec{l}$$

does not depend on the integration path. Its value is determined only by the coordinates of the initial and final points in the curve. Therefore, the concept of scalar potential can be introduced, so that

$$\int_1^2 \vec{E} d\vec{l} = \varphi_1 - \varphi_2.$$

Hence, the electrostatic field is a potential one.

From equations (2.109), it follows that the electrostatic field has sources, electric charges that generate the electrostatic field.

The equations of magnetostatic field are written as follows:

$$\operatorname{rot} \vec{B} = \mu_0 \vec{j}, \qquad \oint_L \vec{B} d\vec{l} = \mu_0 I, \tag{2.110}$$

$$\operatorname{div} \vec{B} = 0, \qquad \oint_S \vec{B} d\vec{S} = 0. \tag{2.111}$$

Equations (2.110) mean that the magnetostatic field is rotational, its force lines are closed, and it does not correspond to any scalar potential. From equations (2.111), it follows the source-free character of the magnetic field, i.e. the absence of magnetic charges. Such fields are called **solenoidal**.

2.9. Boundary conditions for magnetic induction vector

Differential equations (2.110) and (2.111) for the magnetic field have to be supplemented with boundary conditions describing the behavior of vector $\vec{B}$ across the discontinuity surface. The discontinuity surfaces for vector $\vec{B}$ are surfaces, in which the **surface current** runs. In other words, we have a current in a layer, the thickness of which tends to zero. This current is characterized by the **surface current density** $\vec{i}$. The absolute value of vector $\vec{i}$ equals the electric charge transferred per unit time across a line of unit length oriented perpendicularly to the current direction.

Consider a layer of the thickness dh (Fig. 2.9), which separates space 1 from space 2, and choose the directions of the normal to the layer $\vec{n}$ and a unit vector $\vec{\tau}$ along the layer. Let the electric current I with the bulk density $\vec{j}$ run in the layer. Provided that the layer thickness dh tends to zero ($dh \to 0$), whereas the current strength I remains constant, we obtain a surface current with the density $\vec{i}$.

In order to establish boundary conditions for the vector $\vec{B}$, let us first use differential equation (2.111), div $\vec{B} = 0$. In section 1.5, it was demonstrate that the electrostatic equation div $\vec{E} = \rho/\varepsilon_0$, where ρ is the bulk density of electric charge, gives rise to the boundary condition $E_{2n} - E_{1n} = \sigma/\varepsilon_0$ for a surface with the surface charge density σ. In our case, the magnetic charges are absent. Therefore, from the equation div $\vec{B} = 0$, we obtain the boundary condition

$$B_{2n} - B_{1n} = 0, \tag{2.112}$$

i.e. the normal component of vector $\vec{B}$ does not change when crossing the surface with the surface current.

To establish the behavior of the $\vec{B}$-vector component tangential to the surface, let us consider an element of the normal layer cross-section (hatched in Fig. 2.9) with the area $dS = dtdh$ and apply equation (2.110) to it,

$$\oint \vec{B} d\vec{l} = \oint B_l dl = \mu_0 I. \tag{2.113}$$

It is easy to understand that, at $dh \to 0$,

$$\oint B_l dl = B_{2\tau} dt - B_{1\tau} dt, \tag{2.114}$$

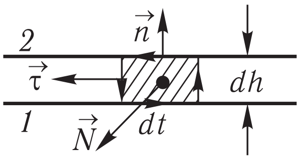

Fig. 2.9.

so that the current dI through the hatched area dS is

$$dI = i_N dt, \tag{2.115}$$

where i_N is the projection of the surface current density $\vec{i}$ on the unit vector $\vec{N}$ tangent to the surface current and perpendicular to the vectors $\vec{n}$ and $\vec{\tau}$ (see Fig. 2.9). Substituting (2.113) and (2.115) into (2.114), we obtain the boundary condition

$$B_{2\tau} - B_{1\tau} = \mu_0 i_N. \tag{2.116}$$

Hence, when crossing the surface with the surface current, the component of the magnetic induction vector tangent to the surface has a jump discontinuity of $\mu_0 i_N$.

Notice that, in the presence of surface currents, expression (2.53) and (2.63) for the magnetic field induction $\vec{B}$ and vector potential $\vec{A}$ should be generalized as follows:

$$\vec{B}(\vec{r}) = \frac{\mu_0}{4\pi} \int \frac{\left[\vec{j}(\vec{r}')(\vec{r} - \vec{r}')\right]}{|\vec{r} - \vec{r}'|^3} dV' + \frac{\mu_0}{4\pi} \int \frac{\left[\vec{i}(\vec{r}')(\vec{r} - \vec{r}')\right]}{|\vec{r} - \vec{r}'|^3} dS', \tag{2.117}$$

$$\vec{A}(\vec{r}) = \frac{\mu_0}{4\pi} \int \frac{\vec{j}(\vec{r}')}{|\vec{r} - \vec{r}'|} dV' + \frac{\mu_0}{4\pi} \int \frac{\vec{i}(\vec{r}')}{|\vec{r} - \vec{r}'|} dS'. \tag{2.118}$$

2.10. Relative character of electric and magnetic fields. Electromagnetic field

Fixed electric charges create an electrostatic field, whereas moving charges, both electrostatic and magnetic ones. However, the motion has a relative sense. Therefore, electric and magnetic fields must also have a relative character. Bearing all that in mind, let us consider the following example.

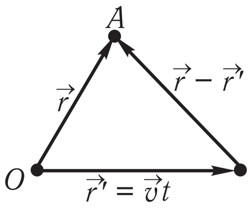

Fig. 2.10.

A point electric charge q moves at a constant velocity $\vec{v} \ll c$ in a fixed reference frame with the origin at the point O (Fig. 2.10). At the initial time moment, $t = 0$, it is located at O. Then, at the moment $t > 0$, the charge coordinate equals $\vec{r}' = \vec{v}t$, and, on the basis of expression (1.21), the strength of electric field created by the charge at the observation point A with the coordinate $\vec{r}$ equals

$$\vec{E}(\vec{r}) = \frac{q\,(\vec{r}-\vec{r}')}{4\pi\varepsilon_0\,|\vec{r}-\vec{r}'|^3} = \frac{q\,(\vec{r}-\vec{v}t)}{4\pi\varepsilon_0\,|\vec{r}-\vec{v}t|^3} \tag{2.119}$$

The moving charge also creates an electric current and, consequently, a magnetic field. The induction of the latter can be determined using (2.53). In this case, we write this formula in the form

$$\vec{B}(\vec{r}) = \frac{\mu_0}{4\pi}\int \frac{\left[\vec{j}(\vec{r}'')(\vec{r}-\vec{r}'')\right]}{|\vec{r}-\vec{r}''|^3}dV'', \tag{2.120}$$

where the integration coordinate is denoted as $\vec{r}''$. Then, according to (1.7), the charge density for the point charge with the coordinate $\vec{r}'$ equals $\rho(\vec{r}'') = q\delta(\vec{r}''-\vec{r}')$. Taking into account that the charge density ρ and the current density are connected by relations (see section 2.6)

$$\vec{j}(\vec{r}'') = \rho(\vec{r}'')\vec{v} = q\delta(\vec{r}''-\vec{r}')\vec{v},$$

we obtain

$$\begin{aligned}\vec{B}(\vec{r}) &= \frac{\mu_0}{4\pi}\int \frac{q\,[\vec{v}(\vec{r}-\vec{r}'')]\,\delta(\vec{r}''-\vec{r}')}{|\vec{r}-\vec{r}''|^3}dV'' \\ &= \frac{\mu_0 q}{4\pi}\frac{[\vec{v}(\vec{r}-\vec{r}')]}{|\vec{r}-\vec{r}'|^3} = \frac{\mu_0 q}{4\pi}\frac{[\vec{v}(\vec{r}-\vec{v}t)]}{|\vec{r}-\vec{v}t|^3} \end{aligned} \tag{2.121}$$

Comparing (2.119) and (2.121), we may write

$$\vec{B}(\vec{r}) = \varepsilon_0\mu_0\left[\vec{v}\vec{E}(\vec{r})\right]. \tag{2.122}$$

Formulas (2.119) and (2.121) determine the electric field strength and the magnetic field induction of a point charge that moves at the constant velocity $\vec{v}$. They demonstrate that a moving charge generates both electric and magnetic fields in the fixed reference frame.

In another inertial reference frame so moving at the velocity $\vec{v}$ with respect to the fixed reference frame that the coordinate origins of both systems coincide at the time moment $t = 0$, the charge q is at rest, and the corresponding field strength equals

$$\vec{E}(\vec{r}) = \frac{q\vec{r}}{4\pi\varepsilon_0 r^3},$$

but the magnetic field, as follows from (2.122), is absent ($\vec{v} = 0$). This example shows that 1) electric and magnetic fields have a relative character, because they depend on the choice of reference frame; 2) there exist a unique object, the electromagnetic field, as an interconnecting combination of electric and magnetic fields, which can reveal itself in some cases as either a purely electrostatic ($\vec{v} = 0$) or a purely magnetic ($\vec{v} \neq 0$) field.

2.11. Energy of magnetic field generated by current

As was already indicated, while calculating the energy of magnetic field created by the current, the electromagnetic induction law (see section 2.10) has to be taken into account. Let us consider this issue in more details.

The energy of magnetic field equals the work carried out by current sources (external forces) (strengths)) when the current is excited, i.e. when the current grows from zero to a definite value. This work is executed against the self-induction voltage that arises in the circuit when the current changes and equals (see section 3.1)

$$\mathscr{E}_i = -\frac{d\Phi}{dt}. \tag{2.123}$$

Hence, the determination of the magnetic field energy requires the application of the theory of alternating electromagnetic field. Such calculations for the general case are made in section 3.5. However, for this section to be comprehensive, let us calculate the energy of the magnetic field in a particular case, basing on the fact that the magnetic field in a long enough solenoid with the current can be considered uniform and concentrated only within the solenoid.

Let a solenoid without the core, of the length l, and with the cross-section area $S = \pi d^2/4$ $(l \gg d)$ be composed of N wire rings, and a current in it grow from zero to I. Then, the magnetic field grows from zero to B, and there emerges an electromotive force in the solenoid equal to

$$\mathscr{E}_i = -N\frac{d\Phi}{dt} = -N\frac{d}{dt}(BS) = -NS\frac{dB}{dt}. \qquad (2.124)$$

In this case, the work done by external forces equals

$$A = -\int_0^B I\mathscr{E}_i dt = NS\int_0^B IdB. \qquad (2.125)$$

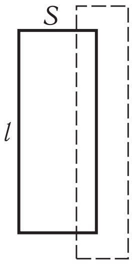

Fig. 2.11.

Applying the law of total current (2.106), selecting the integration path (see the dashed curve in Fig. 2.11), and neglecting the scattering of magnetic field in the space around the solenoid, we obtain the relation

$$Bl = \mu_0 IN, \qquad (2.126)$$

whence it follows that

$$I = \frac{Bl}{\mu_0 N}. \qquad (2.127)$$

Substituting (2.127) into (2.125), we obtain the energy of the magnetic field of solenoid in the form

$$W = A = \frac{1}{2\mu_0}B^2V, \qquad (2.128)$$

where $V = Sl$ is the solenoid volume. Since the solenoid volume is simultaneously the volume of uniform magnetic field, the density of magnetic field energy is

$$w = \frac{W}{V} = \frac{1}{2\mu_0}B^2. \qquad (2.129)$$

Generalizing the result obtained, we arrive at the conclusion that, in the general case, the energy of magnetic field equals

$$W = \frac{1}{2\mu_0}\int B^2 dV. \qquad (2.130)$$

Determining the magnetic induction B from (2.126) and substituting it to expression (2.128), we obtain the energy of magnetic field in the solenoid as the energy of current,

$$W = \frac{1}{2}LI^2, \tag{2.131}$$

where

$$L = \frac{1}{2}\mu_0 \left(\frac{N}{l}\right)^2 V \tag{2.132}$$

is the solenoid inductance.

3

ELECTROMAGNETIC FIELD IN VACUUM

3.1. Law of electromagnetic induction

Until now, we considered only electrostatic and magnetostatic fields generated by electric charges, the density of which did not change in time (fixed charges and stationary currents). If the distribution of those charges changes in time ($\rho = \rho(\vec{r}, t)$ and $\vec{j} = \vec{j}(\vec{r}, t)$), the electric and magnetic created by them also become dependent on the time, i.e. $\vec{E} = \vec{E}(\vec{r}, t)$ and $\vec{B} = \vec{B}(\vec{r}, t)$. In this case, there is no inertial reference frame in which only one field would be observed. In every inertial reference frame, there exists a combination of electric and magnetic fields, which is called the **electromagnetic field**. This means that an unbroken interrelation between the both fields, which is determined using Faraday's law of electromagnetic induction and Maxwell's law (hypothesis). According to those laws, the alternating magnetic field generates an electric field, and the alternating electric field generates a magnetic field.

The **law of electromagnetic induction** consists in the following. If the magnetic flux that passes through a closed conductor loop changes, there arises **an induction electromotive force** in the latter,

$$\mathscr{E}_i = -\frac{d\Phi}{dt}, \tag{3.1}$$

which stimulates an induction current in the conductor. The law of electromagnetic induction (3.1) was discovered by M. Faraday experimentally, but it can be deduced theoretically.

Let a conductor loop L move in a magnetic field at the velocity $\vec{u}$. Then, the charge carriers also have the velocity of directed motion $\vec{u}$, and, in accordance with formula (2.150) in section 3.1 of [10], they are subjected to the action of Lorentz force

$$\vec{f} = e\left[\vec{u}\vec{B}\right]. \tag{3.2}$$

Lorentz force (3.2) should be consider as and external force (see section 2.2). Therefore, the strength of equivalent electric field equals

$$\vec{E} = \frac{\vec{f}}{e} = \left[\vec{u}\vec{B}\right], \tag{3.3}$$

and the electromotive force, in accordance with (2.31), is determined as

$$\mathscr{E}_i = \oint_L \vec{E}d\vec{l} = -\oint_L \left[\vec{u}\vec{B}\right] d\vec{l}. \tag{3.4}$$

If $d\vec{\xi}$ is the shift of the loop element $d\vec{l}$ for the time interval dt, then, $\vec{u} = d\vec{\xi}/dt$; hence,

$$\mathscr{E}_i = \oint_L \left[\frac{d\vec{\xi}}{dt}\vec{B}\right] d\vec{l} = -\oint_L \vec{B}\left[\frac{d\vec{\xi}}{dt}d\vec{l}\right]. \tag{3.5}$$

The integral

$$\oint_L \vec{B}\left[d\vec{\xi}d\vec{l}\right] = \delta\Phi$$

is the variation of magnetic flux through the loop area associated with both the loop motion and its deformation, and $\left[d\vec{\xi}d\vec{l}\right]$ equals the area swept at that by the loop element $d\vec{l}$. Therefore, instead of (3.5), we may write

$$\mathscr{E}_i = -\frac{\delta\Phi}{dt}.$$

If the magnetic field also varies in time, i.e. $\vec{B} = \vec{B}(\vec{r}, t)$, we must write

$$\mathscr{E}_i = -\frac{d\Phi}{dt}, \tag{3.6}$$

where

$$\Phi = \int_S \vec{B} d\vec{S}$$

is the magnetic flux which passes through the loop area S; it can change in time owing to variations in both the magnetic field and the loop area.

If the loop is fixed case but $\vec{B} = \vec{B}(\vec{r}, t)$, we obtain

$$\mathscr{E}_i = -\int_S \frac{\partial \vec{B}}{\partial t} d\vec{S}, \tag{3.7}$$

where the derivative $\partial \vec{B}/\partial t$ is the rate of magnetic induction variation at the fixed spatial point.

On the basis of (3.4) and (3.7), we may write

$$\oint_L \vec{E} d\vec{l} = -\int_S \frac{\partial \vec{B}}{\partial t} d\vec{S}. \tag{3.8}$$

Then, using the Stokes theorem, we obtain

$$\int_S \text{rot } \vec{E} d\vec{S} = -\int_S \frac{\partial \vec{B}}{\partial t} d\vec{S}, \tag{3.9}$$

i.e.

$$\text{rot } \vec{E} = -\frac{\partial \vec{B}}{\partial t}. \tag{3.10}$$

Equation (3.10) is the law of electromagnetic induction in the differential form; the corresponding integral form is derived from relation (3.1).

From equation (3.10), it follows that the electric field induced by the alternating (ac) magnetic field has a vortex character, unlike the electrostatic field. Its force lines are closed, and, therefore, the concept of scalar potential is not applicable now. Besides, from equation (3.10), it also follows that the essence of the law of electromagnetic induction consists in the generation of vortex electric field by the ac magnetic field. Provided that there is an electric circuit, the induced electric field generates an induction current in it. If the circuit is absent, the current does not arise, but the vortex electric field does exist. Therefore, of key importance is not the motion of the circuit in the magnetic field, which gives rise to the action of Lorentz force on charge carriers (so that the Lorentz force plays an auxiliary role), but the variation of the magnetic field in the given reference frame and, thus, the magnetic flux that passes through the loop area. In the reference frame connected with the moving circuit, the latter remains fixed and the Lorentz force is absent; nevertheless, an induction electromotive force emerges as a result of magnetic flux variation.

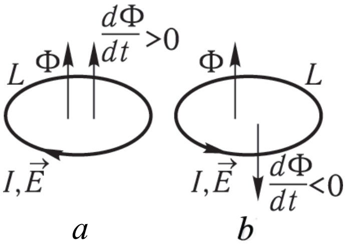

Fig. 3.1.

The sign "minus" in (3.1) and (3.6) demonstrates that the direction of the vector of induced electric field strength $\vec{E}$, i.e. the direction of induction current, together with the direction of magnetic induction flux growth $d\Phi/dt$ forms a left-handed coordinate system. The direction of flux Φ and the direction of its variation $d\Phi/dt$ are shown in Fig. 3.1 by arrows. In the case illustrated in Fig. 3.1a, the magnetic flux Φ grows ($d\Phi/dt > 0$), and the induction current I impedes this growth by its magnetic field; whereas in the case shown in Fig. 3.1b, where the magnetic flux decreases, the magnetic field of the current I supports the magnetic flux Φ. This conclusion corresponds to the known **Lenz's rule**: if the magnetic flux that passes through a closed conducting contour changes, an induction current emerges, being so directed that its magnetic field tends to compensate the variation of magnetic flux through the contour.

3.2. Displacement current. Maxwell's hypothesis

In the case of ac electromagnetic field, $\rho = \rho(\vec{r}, t)$, i.e. $\partial\rho/\partial t \neq 0$. Therefore, from the continuity equation

$$\frac{\partial\rho}{\partial t} + \operatorname{div}\vec{j} = 0, \tag{3.11}$$

it follows that $\operatorname{div}\vec{j} \neq 0$. This means that the equation of stationary magnetic field

$$\operatorname{rot}\vec{B} = \mu_0\vec{j} \tag{3.12}$$

is inapplicable to the ac electromagnetic field, because this equation brings about the relation $\operatorname{div}\vec{j} = \frac{1}{\mu_0}\operatorname{div}\operatorname{rot}\vec{B} = 0$.

This inconsistency can be removed by supplementing the conductivity current density $\vec{j}$ in (3.12) with the summand $\vec{j}_{dis}$, which is called the **displacement current density**, provided that

$$\mathrm{div}\left(\vec{j}+\vec{j}_{dis}\right)=0. \tag{3.13}$$

In this case, equation (3.12) should be rewritten in the form

$$\mathrm{rot}\ \vec{B}=\mu_0\left(\vec{j}+\vec{j}_{dis}\right). \tag{3.14}$$

From equations (3.13), (3.11), and div $\vec{E}=\rho/\varepsilon_0$ (see (1.66)), it follows that

$$\mathrm{div}\ \vec{j}_{dis}=-\mathrm{div}\ \vec{j}=\frac{\partial\rho}{\partial t}=\varepsilon_0\frac{\partial}{\partial t}\mathrm{div}\ \vec{E}. \tag{3.15}$$

Since

$$\frac{\partial}{\partial t}\mathrm{div}\ \vec{E}=\mathrm{div}\frac{\partial\vec{E}}{\partial t},$$

we may write

$$\mathrm{div}\ \vec{j}_{dis}=\varepsilon_0\mathrm{div}\frac{\partial\vec{E}}{\partial t}. \tag{3.16}$$

The simplest way is to adopt that

$$\vec{j}_{dis}=\varepsilon_0\frac{\partial\vec{E}}{\partial t}. \tag{3.17}$$

Using the electric induction vector (1.30b), we may write

$$\vec{j}_{dis}=\frac{\partial\vec{D}}{\partial t}. \tag{3.18}$$

Taking into account (3.17), equation (3.14) reads

$$\mathrm{rot}\ \vec{B}=\mu_0\vec{j}+\varepsilon_0\mu_0\frac{\partial\vec{E}}{\partial t}. \tag{3.19}$$

Hence, we arrive to the conclusion that the magnetic field can be generated not only by the conductivity current, but also the displacement one, i.e. by the ac electric field.

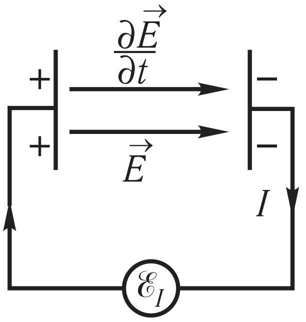

Fig. 3.2.

As an example, let us consider the electric circuit shown in Fig. 3.2 and consisting of a source of electromotive force $\mathscr{E}_i$, conductors, and two capacitor plates. The electromotive force $\mathscr{E}_i = -d\Phi/dt$ cannot be independent of time, because, otherwise, $d\Phi/dt = \text{const}$ and $\Phi = \text{const}\times t + \Phi_0$. As a result, the magnetic flux would have grown infinitely. Therefore, the electromotive force $\mathscr{E}_i$ must vary, and the current in the circuit must be alternating, as well as the electric field between the plates. The directions of vectors $\vec{E}$ and $\partial\vec{E}/\partial t$ are shown at the moment when the current increases.

The electric field between the plates is created by the charges of both plates. Therefore (see section 1.2), its strength equals $E = \sigma/\varepsilon_0 = q/(\varepsilon_0 S)$, where q is the charge of the plate, and S its area. Then, it is possible to write

$$\vec{j}_{dis} = \varepsilon_0 \frac{\partial E}{\partial t} = \frac{1}{S}\frac{\partial q}{\partial t} = \frac{I}{S} = j, \tag{3.20}$$

i.e. the displacement current density between the plates equals the conductivity current density in the plates. From Fig. 3.2, it also follows that the direction of displacement current coincides with that of conductivity current.

Hence, the conductivity current in conductors transforms into the equivalent displacement current between the plates, so that the streamline remain closed, which corresponds to condition (3.13). At the same time, it should be noted that the conductivity current and the displacement current in vacuum are different physical phenomena despite the similarity of their terms. The only property that unites them is their ability to identically excite a magnetic field. This fact was understood for the first time by Maxwell, who formulated the hypothesis about the magnetic field generation by the alternating electric field. However, the displacement current in vacuum is not accompanied by any charge motion and Joule heat release. It corresponds only to the variation of the electric field strength.

3.3. Equations of alternating electromagnetic field in vacuum

The system of equations for the alternating electromagnetic field consists of equations that enable the vectors of electric field strength, $\vec{E} = \vec{E}(\vec{r}, t)$, and magnetic induction, $\vec{B} = \vec{B}(\vec{r}, t)$, to be found knowing the distributions of charges, $\rho = \rho(\vec{r}, t)$, and currents, $\vec{j} = \vec{j}(\vec{r}, t)$, in the space. This system of equations has to be complete, i.e. all properties of the field must follow from it; in particular, the equations for electrostatic and magnetostatic fields must be obtained as the limiting cases. According to the general concept of theoretical physics, they should stem from the generalization of fundamental phenomenological, i.e. found experimentally, laws and be supplemented with boundary conditions determining the behavior of vectors $\vec{E}$ and $\vec{B}$ in the presence of surface charges and surface currents. This system of equations, was first formulated in the 1860s by J.C. Maxwell, who generalized the laws of electric and magnetic phenomena. Therefore, they are called **Maxwell's equations**. In the modern form, they were derived by the German physicist H. Hertz and the English physicist O. Heaviside.

The system of Maxwell's equations consists of four equations, which can be expressed in the differential and integral forms. The differential form will be considered first.

The first Maxwell's equation for the alternating electromagnetic field in vacuum is the already obtained equation (3.19),

$$\operatorname{rot} \vec{B} = \mu_0 \vec{j} + \varepsilon_0 \mu_0 \frac{\partial \vec{E}}{\partial t}. \tag{3.21}$$

The second Maxwell's equation is equation (3.10),

$$\operatorname{rot} \vec{E} = -\frac{\partial \vec{B}}{\partial t}. \tag{3.22}$$

The other two equations follow from (3.21) and (3.22), provided some additional assumptions.

Since the divergence of a rot equals zero, it follows from (3.22) that

$$\operatorname{div} \frac{\partial \vec{B}}{\partial t} = \frac{\partial}{\partial t} \operatorname{div} \vec{B} = 0,$$

hence, $\operatorname{div}\vec{B} = f(\vec{r})$, and, at every point in the space, $\operatorname{div}\vec{B}$ must have a time-independent arbitrary value. This means that the function $f(\vec{r})$ plays the role of initial conditions at integration. Therefore, putting it equal to zero over the whole space, we obtain the third Maxwell's equation,

$$\operatorname{div}\vec{B} = 0. \tag{3.23}$$

Taking the divergence of the both sides of equation (3.21), we find

$$\operatorname{div}\vec{j} + \varepsilon_0 \frac{\partial}{\partial t}\operatorname{div}\vec{E} = 0. \tag{3.24}$$

Applying continuity equation (3.11), equation (3.24) reads

$$\frac{\partial}{\partial t}\left(-\rho + \varepsilon_0 \operatorname{div}\vec{E}\right) = 0, \tag{3.25}$$

whence it follows that

$$-\rho + \varepsilon_0 \operatorname{div}\vec{E} = \psi(\vec{r}). \tag{3.26}$$

Repeating the arguments used while substantiating equation (3.23), we may assume that $\psi(\vec{r}) = 0$. Then, on the basis of (3.26), we obtain the fourth Maxwell's equation

$$\operatorname{div}\vec{E} = \frac{\rho}{\varepsilon_0}. \tag{3.27}$$

Equations (3.21)–3.23, and (3.27) compose the system of Maxwell's equations in the differential form for the alternating electromagnetic field in vacuum.

Now, let us establish the integral form for Maxwell's equations. Calculating the flux of vectors in the both sides of equation (3.21) through a certain surface S, we obtain

$$\int_S \operatorname{rot}\vec{B}\, d\vec{S} = \mu_0 \int_S \vec{j}\, d\vec{S} + \varepsilon_0\mu_0 \int_S \frac{\partial \vec{E}}{\partial t}\, d\vec{S}. \tag{3.28}$$

Using the Stokes theorem, we write

$$\oint_L \operatorname{rot}\vec{B}\, d\vec{S} = \oint_L \vec{B}\, d\vec{l}, \tag{3.29}$$

where L is the contour that bounds the surface S (or the contour that the surface S leans on). In the right hand side of (3.28),

$$\int_S \vec{j}\, d\vec{S} = I \tag{3.30}$$

and

$$\varepsilon_0 \int_S \frac{\partial \vec{E}}{\partial t}\, d\vec{S} = \int_S \vec{j}_{dis}\, d\vec{S} = I_{dis}, \tag{3.31}$$

where I is the conductivity current strength, and I_{dis} the strength of displacement current through the surface S. Then instead of (3.28), we obtain the equation

$$\oint_L \vec{B}\, d\vec{l} = \mu_0 I + \mu_0 I_{dis}, \tag{3.32}$$

which is the integral form of Maxwell's equation (3.21). It establishes the relation between the rotor of vector $\vec{B}$ along the contour L and the total current through the surface S bounded by this contour (the **law of total current**). This relation is a consequence of the generalization of Biot–Savart law (see section 2.4).

For the fluxes of vectors in equation (3.22) through the surface S, we have the relation

$$\int_S \operatorname{rot} \vec{E}\, d\vec{S} = -\frac{\partial}{\partial t} \int_S \vec{B}\, d\vec{S}.$$

Since

$$\oint_L \operatorname{rot} \vec{E}\, d\vec{S} = \oint_L \vec{E}\, d\vec{l},$$

the integral form of Maxwell's equation (3.22) looks like

$$\oint_L \vec{E}\, d\vec{l} = -\frac{\partial}{\partial t} \int \vec{B}\, d\vec{S}, \tag{3.33}$$

which corresponds to Faraday's law of electromagnetic induction.

On the basis of equation (3.23) and Gauss's theorems, it is possible to write

$$\int_V \operatorname{div} \vec{B}\, dV = \oint_S \vec{B}\, d\vec{S} = 0.$$

Therefore, the equation

$$\oint_S \vec{B}\, d\vec{S} = 0 \tag{3.34}$$

is the integral form of Maxwell's equation (3.23), which declares the absence of magnetic charges.

Integrating equation (3.27) over the volume V and applying Gauss's theorem, it is easy to be convinced that

$$\oint_S \vec{E}\, d\vec{S} = \frac{q}{\varepsilon_0} = \frac{1}{\varepsilon_0}\int \rho\, dV. \tag{3.35}$$

Equation (3.35) is the integral form of equation (3.27), which is a consequence of Coulomb's law.

To summarize, we present the combined system of Maxwell's equations in the differential and integral forms:

$$\operatorname{rot}\vec{B} = \mu_0\vec{j} + \varepsilon_0\mu_0\frac{\partial\vec{E}}{\partial t}, \qquad \oint_L \vec{B}\, d\vec{l} = \mu_0\left(I + I_{dis}\right); \tag{3.36}$$

$$\operatorname{rot}\vec{E} = -\frac{\partial\vec{B}}{\partial t}, \qquad \oint_L \vec{E}\, d\vec{l} = -\frac{\partial}{\partial t}\int_S \vec{B}\, d\vec{S}; \tag{3.37}$$

$$\operatorname{div}\vec{B} = 0, \qquad \oint_S \vec{B}\, d\vec{S} = 0; \tag{3.38}$$

$$\operatorname{div}\vec{E} = \frac{\rho}{\varepsilon_0}, \qquad \oint_S \vec{E}\, d\vec{S} = \frac{1}{\varepsilon_0}\int_V \rho\, dV. \tag{3.39}$$

3.4. Boundary conditions for Maxwell's equations

Let us establish boundary conditions for Maxwell's equations (3.36)–(3.39) in the framework of the arguments used while finding boundary conditions for electrostatic (see section 1.5) and magnetostatic (see section 2.8) equations.

The equation of alternating electromagnetic field (3.36) differs from magnetostatic equation (2.110) by the summand $\varepsilon_0\mu_0\partial\vec{E}/\partial t$. However, repeating our speculations presented in section 2.9 when substantiating condition (2.116), it is easy to be convinced that the flux of vector $\vec{E}$ and, hence, the vector $\varepsilon_0\mu_0\partial\vec{E}/\partial t$ through the hatched area (see Fig. 2.9) also tends to zero if the contour height vanishes. Therefore, condition (2.116)

$$B_{2\tau} - B_{1\tau} = \mu_0 i_N \tag{3.40}$$

is applicable in the case of alternating electromagnetic field as well. It will be recalled that that $B_{2\tau}$ and $B_{1\tau}$ are the tangent components of vector $\vec{B}$ on both sides of the surface current, $\vec{i}$ is the vector of surface current density (its magnitude equals the current strength passing through a unit-length interval oriented perpendicularly to the surface current direction), and i_N is the projection of vector $\vec{i}$ on the unit vector $\vec{N} = [\vec{n}\,\vec{\tau}]$, where $\vec{n}$ is the normal to the surface current, and $\vec{\tau}$ the unit vector tangential to the surface current (see Fig. 2.9).

Equations (3.38) coincide with equations (2.111). Therefore, condition (2.112),

$$B_{2n} - B_{1n} = 0, \tag{3.41}$$

remains valid.

Analogously, from the similarity of equations (3.39) and (1.66), it follows that boundary condition (1.73),

$$E_{2n} - E_{1n} = \frac{\sigma}{\varepsilon_0}, \tag{3.42}$$

where σ is the surface electric charge density, is applicable to the alternating electromagnetic field.

Maxwell's equation (3.37) differs from equation (1.67) by the summand $-\partial\vec{B}/\partial t$. However, using Fig. 1.7 and the arguments applied to substantiate condition (3.40), we come to the conclusion that condition (1.75),

$$E_{2\tau} - E_{1\tau} = 0, \tag{3.43}$$

also remains valid.

The system of Maxwell's equations (3.36)–(3.39) supplemented with the equation

$$\vec{j} = \sigma\left(\vec{E} + \vec{E}^{ext}\right) \tag{3.44}$$

and boundary conditions (3.40)–(3.43) is complete. This means that, for the given charge, $\rho(\vec{r}, t)$, and current, $\vec{j}(\vec{r}, t)$, distributions and taking into account boundary, (3.40)–(3.43), and initial (at $t = 0$) conditions for the vectors $\vec{E}$ and $\vec{B}$ at infinity, one can unambiguously determine, by integrating those equations, the distributions of electric and magnetic fields in the space at an arbitrary time moment.

3.5. Energy conservation law for electromagnetic field. Poynting vector

A number of important consequences of the general character follow from Maxwell's equations. In particular, they concern the existence of energy and momentum for the electromagnetic field and their conservation in a system consisting of electromagnetic field and charged.

Let us consider a closed system consisting of electromagnetic field, charges, and currents in the volume V. The field acts on a unit volume of charges with the force

$$\vec{f} = \rho\vec{E} + \rho\left[\vec{v}\,\vec{B}\right] = \rho\vec{E} + \left[\vec{j}\,\vec{B}\right]. \tag{3.45}$$

Therefore, the work done by the field on the charges in the volume V per unit time is determined by the formula

$$\frac{dA}{dt} = \int_V \vec{f}\,\vec{v}\,dV = \int_V \rho\vec{v}\vec{E}\,dV + \int_V \rho\vec{v}\left[\vec{v}\,\vec{B}\right]\,dV = \int_V \vec{j}\vec{E}\,dV, \tag{3.46}$$

where we took into account that $\rho\vec{v} = \vec{j}$ and $\vec{v}\left[\vec{v}\vec{B}\right] = \left[\vec{v}\vec{v}\right]\vec{B} = 0$. Calculating the scalar product of equation (3.36) and vector $\vec{E}$, as well as the scalar product of equation (3.37) and vector $\vec{B}$, and subtracting the results, we obtain

$$\vec{E}\mathrm{rot}\,\vec{B} - \vec{B}\mathrm{rot}\,\vec{E} = \mu_0\vec{j}\vec{E} + \varepsilon_0\mu_0\vec{E}\frac{\partial\vec{E}}{\partial t} + \vec{B}\frac{\partial\vec{B}}{\partial t}, \tag{3.47}$$

whence it follows that

$$\vec{j}\vec{E} = \frac{1}{\mu_0}\left(\vec{E}\mathrm{rot}\,\vec{B} - \vec{B}\mathrm{rot}\,\vec{E}\right) - \frac{\partial}{\partial t}\left(\frac{1}{2}\varepsilon_0 E^2 + \frac{1}{2\mu_0}B^2\right). \tag{3.48}$$

With the help of the vector analysis formula (see Appendix B)

$$\mathrm{div}\,\vec{a}\vec{b} = \vec{b}\mathrm{rot}\,\vec{a} - \vec{a}\mathrm{rot}\,\vec{b},$$

we may write, instead of (3.48),

$$\vec{j}\vec{E} = -\frac{1}{\mu_0}\mathrm{div}\left[\vec{E}\vec{B}\right] - \frac{\partial}{\partial t}\left(\frac{1}{2}\varepsilon_0 E^2 + \frac{1}{2\mu_0}B^2\right). \tag{3.49}$$

Then, relationship (3.46) reads

$$\begin{aligned}\frac{dA}{dt} &= -\frac{1}{\mu_0}\int_V \operatorname{div}\left[\vec{E}\vec{B}\right] dV \\ &\quad -\frac{\partial}{\partial t}\int_V \left(\frac{1}{2}\varepsilon_0 E^2 + \frac{1}{2\mu_0}B^2\right) dV.\end{aligned} \tag{3.50}$$

Let us analyze the result obtained. The first summand in the right hand side of (3.50), according to Gauss's theorem, can be rewritten as

$$-\frac{1}{\mu_0}\int_V \operatorname{div}\left[\vec{E}\vec{B}\right] dV = -\frac{1}{\mu_0}\oint_S \left[\vec{E}\vec{B}\right] d\vec{S}, \tag{3.51}$$

where the closed surface S bounds the volume V. Then, from (3.50), it follows the relation

$$\begin{aligned}\frac{dA}{dt} &= -\frac{1}{\mu_0}\oint_S \left[\vec{E}\vec{B}\right] d\vec{S} \\ &\quad -\frac{\partial}{\partial t}\int_V \left(\frac{1}{2}\varepsilon_0 E^2 + \frac{1}{2\mu_0}B^2\right) dV.\end{aligned} \tag{3.52}$$

Let the surface S envelops the entire electromagnetic field. Then, the field should be absent at any of its point, so that the first integral in the right hand side of relation (3.52) equals zero. In this case,

$$\frac{dA}{dt} = -\frac{\partial}{\partial t}\int_V \left(\frac{1}{2}\varepsilon_0 E^2 + \frac{1}{2\mu_0}B^2\right) dV. \tag{3.53}$$

In the closed system, the field work has to be equal to the reduction of its energy. Therefore, the quantity

$$W = \int_V \left(\frac{1}{2}\varepsilon_0 E^2 + \frac{1}{2\mu_0}B^2\right) dV \tag{3.54}$$

has the meaning of the **electromagnetic field energy** in the volume V, and the quantity

$$w = \frac{1}{2}\varepsilon_0 E^2 + \frac{1}{2\mu_0}B^2 \tag{3.55}$$

the meaning of the **electromagnetic field energy density**. According to (1.89), the summand

$$w_e = \frac{1}{2}\varepsilon_0 E^2 \tag{3.56}$$

has the meaning of the **electric field energy density**. Therefore, the summand

$$w_m = \frac{1}{2\mu_0} B^2 \tag{3.57}$$

should be considered as the **magnetic field energy density in vacuum**.

Suppose that the closed surface S in (3.52) does not envelop the entire field. Then, relation (3.52) has the form

$$-\frac{\partial W}{\partial t} = \frac{dA}{dt} + \frac{1}{\mu_0} \oint_S \left[\vec{E}\vec{B}\right] d\vec{S} \tag{3.58}$$

This formula means that the electromagnetic field energy in the closed surface S diminishes because the field does the work and the field energy is lost through the surface S, with

$$-\left(\frac{\partial W}{\partial t}\right)_{outflow} = \frac{1}{\mu_0} \oint_S \left[\vec{E}\vec{B}\right] d\vec{S} = \oint_S \vec{\Pi}\, d\vec{S}. \tag{3.59}$$

The vector

$$\vec{\Pi} = \frac{1}{\mu_0} \left[\vec{E}\vec{B}\right] \tag{3.60}$$

is called the **Poynting vector**. By magnitude, it equals the energy following out per unit time through the unit area of the surface in the direction perpendicular to the vectors $\vec{E}$ and $\vec{B}$. Hence, the relation (3.58) has the meaning of the energy conservation law for a system with the electromagnetic field and electric charges in vacuum.

3.6. Momentum conservation law for electromagnetic field in vacuum

Besides the energy, the electromagnetic field also has a momentum. To prove this, let us consider a system of electromagnetic field and electric charges. The force acting on a unit volume of charges, in accordance with (3.45), equals

$$\vec{f} = \rho\vec{E} + \left[\vec{j}\vec{B}\right]. \tag{3.61}$$

From (3.39) and (3.36), it follows that

$$\begin{aligned} \rho &= \varepsilon_0 \mathrm{div}\, \varepsilon_0 \mathrm{div}\, \vec{E}, && (3.62) \\ \vec{j} &= \frac{1}{\mu_0} \mathrm{rot}\, \vec{B} - \varepsilon_0 \frac{\partial \vec{E}}{\partial t}. && (3.63) \end{aligned}$$

Therefore,

$$\vec{f} = \varepsilon_0 \vec{E} \mathrm{div}\, \vec{E} + \frac{1}{\mu_0} \left[\mathrm{rot}\, \vec{B}\, \vec{B} \right] - \varepsilon_0 \left[\frac{\partial \vec{E}}{\partial t} \vec{B} \right]. \qquad (3.64)$$

The last summand in (3.64) can be transformed as follows:

$$\begin{aligned} -\varepsilon_0 \left[\frac{\partial \vec{E}}{\partial t} \vec{B} \right] &= -\varepsilon_0 \frac{\partial}{\partial t} \left[\vec{E} \vec{B} \right] + \varepsilon_0 \left[\vec{E} \frac{\partial \vec{B}}{\partial t} \right] \\ &= -\varepsilon_0 \frac{\partial}{\partial t} \left[\vec{E} \vec{B} \right] + \varepsilon_0 \left[\mathrm{rot}\, \vec{E}\, \vec{E} \right] && (3.65) \end{aligned}$$

Then, instead of (3.64), we obtain

$$\vec{f} = \varepsilon_0 \vec{E} \mathrm{div}\, \vec{E} + \frac{1}{\mu_0} \left[\mathrm{rot}\, \vec{B}\, \vec{B} \right] - \varepsilon_0 \frac{\partial}{\partial t} \left[\vec{E} \vec{B} \right] + \varepsilon_0 \left[\mathrm{rot}\, \vec{E}\, \vec{E} \right]. \qquad (3.66)$$

Added the expression equal zero,

$$\frac{1}{\mu_0} \vec{B} \mathrm{div}\, \vec{B} + \varepsilon_0 \left[\mathrm{rot}\, \vec{E}\, \vec{E} \right] - \varepsilon_0 \left[\mathrm{rot}\, \vec{E}\, \vec{E} \right] = 0,$$

to (3.36), we obtain the formula

$$\begin{aligned} \vec{f} &= \varepsilon_0 \vec{E} \mathrm{div}\, \vec{E} + \frac{1}{\mu_0} \vec{B} \mathrm{div}\, \vec{B} + \varepsilon_0 \left[\mathrm{rot}\, \vec{E}\, \vec{E} \right] \\ &\quad + \frac{1}{\mu_0} \left[\mathrm{rot}\, \vec{B}\, \vec{B} \right] - \varepsilon_0 \frac{\partial}{\partial t} \left[\vec{E} \vec{B} \right], && (3.67) \end{aligned}$$

which can be rewritten in the form

$$\begin{aligned} \vec{f} + \varepsilon_0 \frac{\partial}{\partial t} \left[\vec{E} \vec{B} \right] &= \varepsilon_0 \vec{E} \mathrm{div}\, \vec{E} + \frac{1}{\mu_0} \vec{B} \mathrm{div}\, \vec{B} \\ &\quad + \varepsilon_0 \left[\mathrm{rot}\, \vec{E}\, \vec{E} \right] + \frac{1}{\mu_0} \left[\mathrm{rot}\, \vec{B}\, \vec{B} \right]. && (3.68) \end{aligned}$$

With the help of straightforward but rather cumbersome calculations, we can show that relation (3.68) can be written in the form

$$\vec{f} + \varepsilon_0 \frac{\partial}{\partial t} \left[\vec{E} \vec{B} \right] = \sum_{i,j=1}^{3} \frac{\partial T_{ij}}{\partial x_i} \vec{m}_j, \tag{3.69}$$

where $\vec{m}_j$ are the unit vectors in the Cartesian coordinate systems, $T_{ij} = T_{ji}$ is the symmetric tensor

$$T_{ij} = \varepsilon_0 E_i E_j + \frac{1}{\mu_0} B_i B_j - \delta_{ij} w, \tag{3.70}$$

which is called the **Maxwell stress tensor**, and w is the density of electromagnetic field energy (3.55). Integrating (3.69) over the volume V, we obtain

$$\int_V \vec{f}\, dV + \varepsilon_0 \frac{\partial}{\partial t} \int_V \left[\vec{E} \vec{B} \right] dV = \int_V \sum_{i,j=1}^{3} \frac{\partial T_{ij}}{\partial x_i} \vec{m}_j dV. \tag{3.71}$$

The right hand side of (3.71) can be transformed with the help of Gauss's theorems as follows:

$$\begin{aligned} \int_V \sum_{i,j} \frac{\partial T_{ij}}{\partial x_i} \vec{m}_j dV &= \sum_j \vec{m}_j \int_V \sum_i \frac{\partial T_{ij}}{\partial x_i} dV \\ = \quad \sum_j \vec{m}_j \int_V \operatorname{div} \vec{T}_j dV &= \oint_S \sum_{i,j} T_{ij} n_i \vec{m}_j dS, \end{aligned} \tag{3.72}$$

where n_i is the projection of normal $\vec{n}$ to the surface element $d\vec{S}$ on the axis x_i. Then, (3.71) takes the form

$$\int_V \vec{f}\, dV + \varepsilon_0 \frac{\partial}{\partial t} \int_V \left[\vec{E} \vec{B} \right] dV = \oint_S \sum_{i,j} T_{ij} n_i \vec{m}_j dS. \tag{3.73}$$

The integral

$$\vec{F} = \int_V \vec{f}\, dV$$

is the force acting on the electric charge in the volume V. Using the equation of motion, it can be written as

$$\vec{F} = \int_V \vec{f}\, dV = \frac{d\vec{P}_{part}}{dt}, \tag{3.74}$$

where $\vec{P}_{part}$ is the total momentum of charged particles in the volume V. On the basis of (3.60) and (3.74), instead of (3.73), we obtain

$$\frac{d}{dt}\left\{\vec{P}_{part} + \varepsilon_0\mu_0 \int_V \vec{\Pi}\, dV\right\} = \oint_S \sum_{i,j} T_{ij} n_i \vec{m}_j dS. \tag{3.75}$$

If the surface S envelops the entire electromagnetic field, the right hand side in (3.75) equals zero. Then,

$$\frac{d}{dt}\left\{\vec{P}_{part} + \varepsilon_0\mu_0 \int_V \vec{\Pi}\, dV\right\} = 0. \tag{3.76}$$

The meaning of this relation is the momentum conservation law for a closed system consisting of electric charges and electromagnetic field. The quantity

$$\vec{P}_{field} = \varepsilon_0\mu_0 \int_V \vec{\Pi}\, dV \tag{3.77}$$

is the **momentum of electromagnetic field**, and, respectively, the vector

$$\vec{g} = \varepsilon_0\mu_0\vec{\Pi} = \varepsilon_0 \left[\vec{E}\vec{B}\right] \tag{3.78}$$

is the **momentum density of electromagnetic field**.

If the surface S does not envelops the whole field,

$$\frac{d}{dt}\left\{\vec{P}_{part} + \vec{P}_{field}\right\} = \oint_S \sum_{i,j} T_{ij} n_i \vec{m}_j dS. \tag{3.79}$$

Then, the right hand side in (3.79) equals the momentum flux through the surface S, and relation (3.79) is the momentum conservation law.

3.7. Potentials of alternating electromagnetic field in vacuum. D'Alembert equations

In sections 1.3 and 2.5, in order to simplify the calculations of the electrostatic field strength $\vec{E}$ and the stationary magnetic field induction $\vec{B}$, the scalar electrostatic, $\varphi(\vec{r})$, and vector, $\vec{A}(\vec{r})$, potentials were introduced, which satisfy the equations $\vec{E} = -\nabla\varphi$ and $\vec{B} = \operatorname{rot}\vec{A}$, respectively. Similar auxiliary functions can also be introduced for the alternating electromagnetic field.

On the basis of Maxwell's equation (3.38), taking into account that $\operatorname{div}\operatorname{rot} = 0$, we may assume that

$$\vec{B}(\vec{r},t) = \operatorname{rot}\vec{A}(\vec{r},t), \tag{3.80}$$

where $\vec{A}(\vec{r},t)$ is the vector potential of alternating electromagnetic field. Then, Maxwell's equation (3.37) can be written in the form

$$\operatorname{rot}\vec{E} = -\frac{\partial\vec{A}}{\partial t},$$

i.e.

$$\operatorname{rot}\left(\vec{E} + \frac{\partial\vec{A}}{\partial t}\right) = 0. \tag{3.81}$$

Since $\operatorname{rot}\operatorname{grad} = 0$, the vector $\vec{E} + \partial\vec{A}/\partial t$ can be written as

$$\vec{E} + \frac{\partial\vec{A}}{\partial t} = -\nabla\varphi,$$

where $\varphi(\vec{r},t)$ is the scalar potential of alternating electromagnetic field. Hence,

$$\vec{E}(\vec{r},t) = -\nabla\varphi(\vec{r},t) - \frac{\partial\vec{A}(\vec{r},t)}{\partial t}. \tag{3.82}$$

The potentials $\varphi(\vec{r},t)$ and $\vec{A}(\vec{r},t)$ are called the **electromagnetic potentials**. If they are known, formulas (3.80) and (3.82) can be used to determine the field vectors $\vec{E}(\vec{r},t)$ and $\vec{B}(\vec{r},t)$. From this viewpoint, let us consider the main properties of electromagnetic potentials and derive equations used to calculate those properties.

Notice that the electromagnetic potentials, like the electrostatic, $\varphi(\vec{r}\,)$, and vector, $\vec{A}(\vec{r}\,)$, potentials, are ambiguous functions. If we sum up the vector potential $\vec{A}(\vec{r},t)$ with the vector $\nabla\psi(\vec{r},t)$, where $\psi(\vec{r},t)$ is an arbitrary scalar function of coordinates and time, and subtract the derivative $\partial\psi(\vec{r},t)/\partial t$ from the scalar potential $\varphi(\vec{r},t)$, i.e. if we make the transformations

$$\vec{A} \quad\to\quad \vec{A}' = \vec{A} + \nabla\psi, \tag{3.83}$$

$$\varphi \quad\to\quad \varphi' = \varphi - \frac{\partial\psi}{\partial t}, \tag{3.84}$$

the field vectors $\vec{E}$ and $\vec{B}$ will not change, because

$$\begin{aligned}\vec{B}' &= \operatorname{rot}\vec{A}' = \operatorname{rot}\vec{A} + \operatorname{rot}\operatorname{grad}\psi = \operatorname{rot}\vec{A} = \vec{B},\\ \vec{E}' &= -\nabla\varphi' - \frac{\partial\vec{A}'}{\partial t} = -\nabla\varphi + \nabla\frac{\partial\psi}{\partial t} - \frac{\partial\vec{A}}{\partial t} - \frac{\partial}{\partial t}\nabla\psi = \vec{E},\end{aligned}$$

since

$$\nabla\left(\partial\psi/\partial t\right) = \partial\nabla\psi/\partial t.$$

Hence, the electromagnetic field is invariant with respect to the potential transformations (3.83) and (3.84), which are called the gauge transformations, and the corresponding field invariance is called the gauge invariance.

In the case of electrostatic and magnetostatic fields, the equations for potentials $\varphi(\vec{r}\,)$ and $\vec{A}(\vec{r}\,)$ are Poisson's equations (1.68) and (2.69), respectively. Let us establish the form of equations for the electromagnetic potentials $\varphi(\vec{r},t)$ and $\vec{A}(\vec{r},t)$. From equation (3.80), it follows that

$$\operatorname{rot}\vec{B} = \operatorname{rot}\operatorname{rot}\vec{A} = \operatorname{grad}\operatorname{div}\vec{A} - \Delta\vec{A}. \tag{3.85}$$

On the basis of (3.36) and (3.82), we obtain

$$\operatorname{rot}\vec{B} = \mu_0\vec{j} + \varepsilon_0\mu_0\frac{\partial\vec{E}}{\partial t} = \mu_0\vec{j} - \varepsilon_0\mu_0\frac{\partial}{\partial t}\nabla\varphi - \varepsilon_0\mu_0\frac{\partial^2\vec{A}}{\partial t^2}. \tag{3.86}$$

From (3.85) and (3.86), we find the equation

$$\Delta\vec{A} - \varepsilon_0\mu_0\frac{\partial^2\vec{A}}{\partial t^2} = -\mu_0\vec{j} + \operatorname{grad}\left(\operatorname{div}\vec{A} + \varepsilon_0\mu_0\frac{\partial\varphi}{\partial t}\right) \tag{3.87}$$

and, with the help of (3.39) and (3.82), the equation

$$\Delta\varphi = -\frac{\rho}{\varepsilon_0} - \frac{\partial}{\partial t}\mathrm{div}\,\vec{A}. \tag{3.88}$$

The gauge invariance of electromagnetic field, i.e. the ambiguity of electromagnetic potentials, enables an additional condition to be imposed on them. We select it in the form of **Lorentz condition**

$$\mathrm{div}\,\vec{A} + \varepsilon_0\mu_0\frac{\partial\varphi}{\partial t} = 0. \tag{3.89}$$

Then, instead of equations (3.87) and (3.88), we obtain the equations for the electromagnetic potentials in the form

$$\Delta\vec{A} - \varepsilon_0\mu_0\frac{\partial^2\vec{A}}{\partial t^2} = -\mu_0\vec{j}, \tag{3.90}$$

$$\Delta\varphi - \varepsilon_0\mu_0\frac{\partial^2\varphi}{\partial t^2} = -\frac{\rho}{\varepsilon_0}. \tag{3.91}$$

It is easy to prove that the Lorentz gauge transformation of potentials (3.89) leaves the electromagnetic field invariant (see section 11 in [2]).

Equation for potentials (3.90) and (3.91) are called the **d'Alembert equations**. They can be written with the help of the d'Alembert operator

$$\Box = \Delta - \varepsilon_0\mu_0\frac{\partial^2}{\partial t^2} \tag{3.92}$$

in the form

$$\Box\vec{A} = -\mu_0\vec{j}, \tag{3.93}$$

$$\Box\varphi = -\frac{\rho}{\varepsilon_0}. \tag{3.94}$$

The application of d'Alembert equations simplifies the calculation of the fields $\vec{E}$ and $\vec{B}$, because their solution is much simpler than the solution of Maxwell's equations (see section 4).

Note that, according to the correspondence principle, in the case when the potentials are independent of time, d'Alembert equations (3.90) and (3.91) give rise to Poisson's equations for electrostatic and magnetostatic fields (1.68) and (2.69), respectively. Note also that, in the quantum field theory, the Coulomb gauge transformation of potentials

$$\mathrm{div}\,\vec{A} = 0 \tag{3.95}$$

rather than Lorentz's condition (3.89) is used. Then, from equations (3.87) and (3.88), the equations for the electromagnetic potentials are obtained in the form

$$\Delta\vec{A} - \varepsilon_0\mu_0\frac{\partial^2\vec{A}}{\partial t^2} - \varepsilon_0\mu_0\nabla\frac{\partial\varphi}{\partial t} = -\mu_0\vec{j}, \tag{3.96}$$

$$\Delta\varphi = -\frac{\rho}{\varepsilon_0}. \tag{3.97}$$

4

ELECTROMAGNETIC WAVES

4.1. Wave equations. Electromagnetic waves

One of important consequences following from Maxwell's equations (3.36)–(3.39) is the conclusion about the existence of electromagnetic waves. From this viewpoint, let us consider an alternating electromagnetic field at those points in the space, where charges and conductivity currents are absent. Under such conditions, Maxwell's equations take the form

$$\operatorname{rot}\vec{B} = \varepsilon_0\mu_0\frac{\partial\vec{E}}{\partial t}, \tag{4.1}$$

$$\operatorname{rot}\vec{E} = -\frac{\partial\vec{B}}{\partial t}, \tag{4.2}$$

$$\operatorname{div}\vec{B} = 0, \tag{4.3}$$

$$\operatorname{div}\vec{E} = 0. \tag{4.4}$$

Equations (4.1) and (4.2) lead to the relation

$$\operatorname{rot}\operatorname{rot}\vec{E} - \frac{\partial}{\partial t}\operatorname{rot}\vec{B} = -\varepsilon_0\mu_0\frac{\partial^2\vec{E}}{\partial t^2}.$$

Therefore, taking into account that $\operatorname{rot}\operatorname{rot} = \operatorname{grad}\operatorname{div} - \Delta$ (see Appendix B) and equation (4.4), we obtain

$$\Delta\vec{E} - \varepsilon_0\mu_0\frac{\partial^2\vec{E}}{\partial t^2} = 0. \tag{4.5}$$

Analogously, using equations (4.1)–(4.3), we obtain

$$\Delta\vec{B} - \varepsilon_0\mu_0\frac{\partial^2\vec{B}}{\partial t^2} = 0. \tag{4.6}$$

Two more equations, similar to (4.5) and (4.6), are obtained from d'Alembert equations (3.90) and (3.91), provided that $\rho = 0$ and $\vec{j} = 0$,

$$\Delta\vec{A} - \varepsilon_0\mu_0\frac{\partial^2\vec{A}}{\partial t^2} = 0, \tag{4.7}$$

$$\Delta\varphi - \varepsilon_0\mu_0\frac{\partial^2\varphi}{\partial t^2} = 0. \tag{4.8}$$

All four equations (4.5)–(4.8) were obtained under the condition $\rho = 0$ and $\vec{j} = 0$ for the alternating electromagnetic field, and all the four have the identical form. Therefore, we may analyze only one of them, e.g., equation (4.5). An equation of such a form is called the **wave equation**. It describes the propagation of a periodic alternating electric field at the velocity

$$c = \frac{1}{\sqrt{\varepsilon_0\mu_0}}. \tag{4.9}$$

To prove this statement, let us rewrite equation (4.5), taking into account (4.9), in the form

$$\Delta\vec{E} - \frac{1}{c^2}\frac{\partial^2\vec{E}}{\partial t^2} = 0 \tag{4.10}$$

and be convinced that any function of the form

$$\vec{E}(\vec{r},t) = \vec{E}\left(t \pm \frac{1}{c}\vec{k}_0\vec{r}\right), \tag{4.11}$$

where $\vec{k}_0$ is a constant unit vector, satisfies equation (4.10). Really, introducing the notation

$$\varphi = t \pm \frac{1}{c}\vec{k}_0\vec{r}, \tag{4.12}$$

we obtain

$$\Delta\vec{E} = \nabla^2\vec{E} = \left(\frac{\partial}{\partial\vec{r}}\right)^2, \quad \vec{E} = \frac{\partial^2\vec{E}}{\partial t^2}\left(\frac{\partial\varphi}{\partial\vec{r}}\right)^2 = \frac{1}{c^2}\frac{\partial^2\vec{E}}{\partial\varphi^2},$$

$$\frac{1}{c^2}\frac{\partial^2 \vec{E}}{\partial t^2} = \frac{1}{c^2}\frac{\partial^2 \vec{E}}{\partial \varphi^2}\left(\frac{\partial \varphi}{\partial t}\right)^2 = \frac{1}{c^2}\frac{\partial^2 \vec{E}}{\partial \varphi^2}.$$

Hence, equation (4.10) is satisfied by functions (4.11).

If we remove the sign "minus" from (4.11), then, for

$$\Delta t = \frac{1}{c}\Delta\left(\vec{k}_0\vec{r}\right), \tag{4.13}$$

the relation

$$\vec{E}\left(t - \frac{1}{c}\vec{k}_0\vec{r}\right) = \vec{E}\left(t + \Delta t - \frac{\vec{k}_0\vec{r} + \Delta\left(\vec{k}_0\vec{r}\right)}{c}\right) \tag{4.14}$$

is obtained, whence it follows that the value of function $\vec{E}\left(t - \frac{1}{c}\vec{k}_0\vec{r}\right)$ taken at the time moment t at a point with a certain $\vec{k}_0\vec{r}$-value coincides with the value taken at the next time moment $t + \Delta t$ at the point, for which $\vec{k}_0\vec{r}$ changes to $\vec{k}_0\vec{r} + \Delta\left(\vec{k}_0\vec{r}\right)$. Hence, this value of function $\vec{E}$ shifts along the vector $\vec{k}_0$ at the velocity $v = c$.

In the case of sign "plus" in (4.11),

$$\vec{E}\left(t + \frac{1}{c}\vec{k}_0\vec{r}\right) = \vec{E}\left(t + \Delta t + \frac{\vec{k}_0\vec{r} - \Delta\left(\vec{k}_0\vec{r}\right)}{c}\right), \tag{4.15}$$

i.e. some value of the function $\vec{E}\left(t + \frac{1}{c}\vec{k}_0\vec{r}\right)$ coincides with the function value at the next time moment $t + \Delta t$ if $\vec{k}_0\vec{r}$ is replaced by $\vec{k}_0\vec{r} - \Delta\left(\vec{k}_0\vec{r}\right)$. In other words, we observe the propagation of a certain value of the field $\vec{E}$ oppositely to the vector $\vec{k}_0$.

Since equations (4.5)–(4.8) are identical by their form, similar conclusions can be automatically extended on the magnetic induction $\vec{B}$ and the electromagnetic potentials $\vec{A}$ and φ.

The propagation of alternating and interdependent electric and magnetic fields is called the **electromagnetic wave**. Hence, in the space free of electric charges and currents, there may exist an electromagnetic field in the form of electromagnetic waves, which are real (observable) physical objects. The waves of potentials, which follow from equations (4.7) and (4.8), have and abstract meaning. At the same time, one should not forget that, although electromagnetic waves can propagate being not associated with charges, this is electric charges that are their sources (see section 5).

There are electromagnetic waves of various types. We will confine the consideration to plane and spherical ones.

4.2. Plane electromagnetic waves

The solution of equation (4.10) is sought in the form

$$\vec{E}(\vec{r},t) = \vec{E}(\vec{r}\,)f(t). \tag{4.16}$$

Substituting of product (4.16) into (4.10) and dividing the result by (4.16), we obtain

$$\frac{1}{\vec{E}(\vec{r}\,)}\Delta\vec{E}(\vec{r}\,) - \frac{1}{c^2}\frac{1}{f(t)}\frac{\partial^2 f(t)}{\partial t^2} = 0. \tag{4.17}$$

The first summand depends on the coordinate $\vec{r}$, and the second one on the time t. Therefore, each of them must be equal to a constant, and their difference to zero. Those constants are designates as follows:

$$\frac{1}{\vec{E}(\vec{r}\,)}\Delta\vec{E}(\vec{r}\,) \quad = \quad -\vec{k}^2, \tag{4.18}$$

$$\frac{1}{f(t)}\frac{\partial^2 f(t)}{\partial t^2} \quad = \quad -\omega^2. \tag{4.19}$$

Then,

$$-\vec{k}^2 + \frac{\omega^2}{c^2} = 0, \tag{4.20}$$

i.e.

$$\omega = kc. \tag{4.21}$$

It is easy to be convinced that (4.18) and (4.19) have the following partial solutions:

$$\vec{E}_1(\vec{r}\,) \quad = \quad \vec{a}(\vec{k})\mathrm{e}^{i\vec{k}\vec{r}}, \tag{4.22}$$

$$\vec{E}_2(\vec{r}\,) \quad = \quad \vec{a}(\vec{k})\mathrm{e}^{-i\vec{k}\vec{r}}, \tag{4.23}$$

$$f_1(t\,) \quad = \quad f_0(\omega)\mathrm{e}^{i\omega t}, \tag{4.24}$$

$$f_2(t\,) \quad = \quad f_0(\omega)\mathrm{e}{-}^{i\omega t}, \tag{4.25}$$

where $\vec{a}(\vec{k})$ and $f_0(\omega)$ are arbitrary complex coefficients. Substituting them into (4.16), we obtain four particular solutions of wave equation (4.10). Two of them look like

$$\vec{E}_1(\vec{r},t) = \vec{E}_{10}(\vec{k})\mathrm{e}^{i(\vec{k}\vec{r}-\omega t)}, \tag{4.26}$$

$$\vec{E}_2(\vec{r},t) = \vec{E}_{20}(\vec{k})\mathrm{e}^{i(\vec{k}\vec{r}+\omega t)}, \tag{4.27}$$

where $\vec{E}_{10}(\vec{k}) = \vec{a}(\vec{k})f_0^*(\omega)$ and $\vec{E}_{20}(\vec{k}) = \vec{a}(\vec{k})f_0(\omega)$. The other two are the functions $\vec{E}_1^*(\vec{r},t)$ and $\vec{E}_2^*(\vec{r},t)$, the real parts of which coincide with those of functions $\vec{E}_1(\vec{r},t)$ and $\vec{E}_2(\vec{r},t)$. Therefore, the latter may be not considered because only the real parts of complex functions have the physical sense.

If the phases of complex amplitudes of functions (4.26) and (4.27) are designates as α_1 and α_2, their real parts are written as follows:

$$\vec{\mathscr{E}}_1(\vec{r},t) = \vec{\mathscr{E}}_{01}(\vec{k})\cos\left(\vec{k}\vec{r}-\omega t+\alpha_1\right), \tag{4.28}$$

$$\vec{\mathscr{E}}_2(\vec{r},t) = \vec{\mathscr{E}}_{02}(\vec{k})\cos\left(\vec{k}\vec{r}+\omega t+\alpha_1\right). \tag{4.29}$$

Equations (4.26)–(4.29) describe of the waves of electric field strength of electromagnetic wave in the complex and real forms. The quantity

$$\varphi = \vec{k}\vec{r} \pm \omega t + \alpha \tag{4.30}$$

is called the wave phase (do not confuse with the scalar potential φ!). The set of space points with the same phase value at the given time moment form the wave front (the wave surface). Its shape is determined from the equation $\varphi = \text{const}$, or, equivalently, from the equation

$$\vec{k}\vec{r} = \text{const} \tag{4.31}$$

for the fixed time moment. Equation (4.31) describes a plane oriented perpendicularly to the vector $\vec{k}$. Therefore, (4.26)–(4.29) are equations of plane monochromatic harmonic waves with the frequency ω and the **wave vector**

$$\vec{k} = k\vec{k}_0 = \frac{\omega}{c}\vec{k}_0 = \frac{2\pi}{\lambda}\vec{k}_0, \tag{4.32}$$

where $\vec{k}_0$ is the dimensionless unit vector oriented along the vector $\vec{k}$.

Equation (4.26) describes plane waves propagating along the wave vector $\vec{k}$. Really, in this case, the wave phase looks like

$$\varphi = \vec{k}\vec{r} - \omega t + \alpha.$$

Therefore, under the condition $\varphi = \text{const}$, its time derivative equals

$$\frac{d\varphi}{dt} = \vec{k}\frac{d\vec{r}}{dt} - \omega = 0, \tag{4.33}$$

and the speed of propagation of a certain phase value is determined by the formula

$$\vec{v} = \frac{d\vec{r}}{dt} = \frac{\omega}{\vec{k}} = \frac{\omega\vec{k}}{k^2} = c\vec{k}_0. \tag{4.34}$$

Hence, its direction coincides with that of the wave vector.

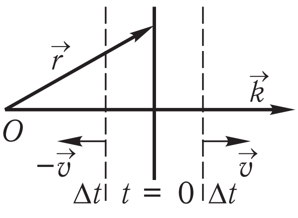

Fig. 4.1.

In the same way, it is easy to prove that plane wave (4.27) propagates oppositely to the vector $\vec{k}$. In Fig. 4.1, the position of the plane fronts of both waves at the time moment $t = 0$ is schematically shown by the vertical solid line, and the positions of the fronts in the time interval Δt are marked by the dashed lines.

While solving equation (4.10), the $\vec{k}$-vector direction was chosen arbitrary. Since all directions in the empty space are equivalent, there is no principal difference between waves (4.26) and (4.27). Therefore, in what follows, we will consider plane waves propagating along the vector $\vec{k}$, and their equations will be written in complex form (4.26) or real form (4.28).

The general solution of equation (4.10) is constructed as a superposition of plane monochromatic harmonic waves (4.26) in the form

$$\vec{E}(\vec{r}, t) = \int \vec{E}_0(\vec{k})\mathrm{e}^{i(\vec{k}\vec{r} - \omega t)} d^3k, \;\; d^3k = dk_x dk_y dk_z. \tag{4.35}$$

It can also be considered as an expansion of an arbitrary field of vector $\vec{E}(\vec{r}, t)$ in **plane monochromatic harmonic waves** (in **harmonics**).

The particular solutions of wave equations (4.6)–(4.8) can be written similarly to (4.26) as

$$\begin{aligned} \vec{B}(\vec{r},t) &= \vec{B}_0(\vec{k})\mathrm{e}^{i(\vec{k}\vec{r}-\omega t)}, & (4.36)\\ \vec{A}(\vec{r},t) &= \vec{A}_0(\vec{k})\mathrm{e}^{i(\vec{k}\vec{r}-\omega t)}, & (4.37)\\ \varphi(\vec{r},t) &= \varphi_0(\vec{k})\mathrm{e}^{i(\vec{k}\vec{r}-\omega t)}. & (4.38) \end{aligned}$$

Maxwell's equation (4.1) and (4.2) illustrate the interrelation between the electric and magnetic fields in the electromagnetic wave. Therefore, the plane monochromatic harmonic wave is described using two equations,

$$\begin{aligned} \vec{E}(\vec{r},t) &= \vec{E}_0(\vec{k})\mathrm{e}^{i(\vec{k}\vec{r}-\omega t)}, & (4.39)\\ \vec{B}(\vec{r},t) &= \vec{B}_0(\vec{k})\mathrm{e}^{i(\vec{k}\vec{r}-\omega t)}. & (4.40) \end{aligned}$$

Functions (4.37) and (4.38) describe abstract plane harmonic waves of electromagnetic potentials.

Let us consider the main properties of plane monochromatic electromagnetic waves in vacuum. Substituting (4.39) and (4.40) into Maxwell's equations (4.3) and (4.4), we obtain the relations

$$\begin{aligned} \operatorname{div}\vec{E} &= \nabla\vec{E} = i\vec{k}\vec{E} = 0, & (4.41)\\ \operatorname{div}\vec{B} &= \nabla\vec{B} = i\vec{k}\vec{B} = 0, & (4.42) \end{aligned}$$

whence it follows that $\vec{E} \perp \vec{k}$ and $\vec{B} \perp \vec{k}$, i.e. electromagnetic waves are transverse.

On the basis of (4.2), (4.39), and (4.40), we obtain

$$\begin{aligned} \operatorname{rot}\vec{E} &= \left[\nabla\,\vec{E}\right] = i\left[\vec{k}\,\vec{E}\right]\\ &= -\frac{\partial\vec{B}}{\partial t} = i\omega\vec{B}, \end{aligned}$$

i.e.

$$\left[\vec{k}\,\vec{E}\right] = \omega\vec{B}. \qquad (4.43)$$

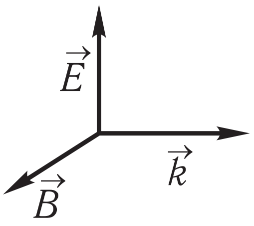

Fig. 4.2.

From relation (4.43), it follows that the vectors $\vec{E}$, $\vec{B}$, and $\vec{k}$ form a right-handed system (Fig. 4.2). From equations (4.39) and (4.40), it also follows that the vectors $\vec{E}$ and $\vec{B}$ change in phase.

Since $\vec{E} \perp \vec{k}$, equation (4.43) brings about the relation $kE = \omega B$. On its basis, taking into account that $k = \omega/c = \omega\sqrt{\varepsilon_0\mu_0}$, we obtain the relationship between the absolute values of vectors $\vec{E}$ and $\vec{B}$ for the electromagnetic wave in vacuum in the form

$$\sqrt{\varepsilon_0}E = \frac{1}{\sqrt{\mu_0}}B \tag{4.44}$$

or

$$E = cB. \tag{4.45}$$

Relation (4.44) makes it possible to write the energy density of electromagnetic wave (3.55) as

$$w = \frac{1}{2}\varepsilon_0 E^2 + \frac{1}{2\mu_0}B^2 = \varepsilon_0 E^2 = \frac{1}{\mu_0}B^2. \tag{4.46}$$

In the case of plane monochromatic harmonic waves, $\vec{E} \perp \vec{B}$, and the absolute value of Poynting vector (3.60) can be written as

$$\Pi = \frac{1}{\mu_0}EB = \sqrt{\frac{\varepsilon_0}{\mu_0}}E^2 = \frac{c}{\mu_0}B^2 = cw. \tag{4.47}$$

4.3. Spherical electromagnetic waves

The solutions of wave equations (4.5) and (4.6) in the form of plane electromagnetic waves (4.39) and (4.40) were obtained in the absence of electromagnetic field symmetry. Let us consider an important case when the electromagnetic field has the central symmetry. It is easy to understand that such a field will be formed if the source is a point charge located at the coordinate origin. Then, the latter is the center of symmetry for the alternating electromagnetic field. In this case, it is expedient to use the spherical coordinate system, in which the Laplace operator looks like

$$\Delta = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}, \tag{4.48}$$

where r is the absolute value of the radius vector drawn to the observation point, and θ and φ are the polar and azimuthal angles, respectively.

If the electromagnetic field is central symmetric, the field vectors $\vec{E}$ and $\vec{B}$, as well as the potentials $\vec{A}$ and φ, do not depend on the angles θ and φ. Then, wave equation (4.10) reads

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial \vec{E}}{\partial r}\right)-\frac{1}{c^2}\frac{\partial^2 \vec{E}}{\partial t^2}=0. \tag{4.49}$$

The solution of equation (4.49) is sought in the form

$$\vec{E}(\vec{r},t)=\frac{\vec{v}(\vec{r},t)}{r}. \tag{4.50}$$

Substitution (4.50) into (4.49), we obtain the equation for the vector $\vec{v}(\vec{r},t)$,

$$\frac{d^2\vec{v}(\vec{r},t)}{dr^2}-\frac{1}{c^2}\frac{d^2\vec{v}(\vec{r},t)}{dt^2}=0. \tag{4.51}$$

Equation (4.51) is similar to equation (4.10), the solution of which was found in section 4.2. Therefore, the particular solution can now be written as

$$\vec{v}(\vec{r},t)=\vec{E}_0(k)\mathrm{e}^{i(kr-\omega t)}, \tag{4.52}$$

where $\omega = kc$. Then, equation (4.50) takes the form

$$\vec{E}(\vec{r},t)=\frac{\vec{E}_0(k)}{r}\mathrm{e}^{i(kr-\omega t)}. \tag{4.53}$$

Function (4.53) is the equation for the spherical monochromatic harmonic wave propagating from the coordinate origin.

From (4.53), it follows that the amplitude of spherical wave is reciprocal to the distance reckoned from the coordinate origin, and the wave vector $\vec{k}$ is directed along the radius vector $\vec{r}$, since $\vec{k}\vec{r}=kr$.

Particular solutions of equations (4.6)–(4.8) in the form of spherical waves can be written similarly to (4.53). Therefore, the spherical monochromatic harmonic electromagnetic wave looks like

$$\vec{E}(\vec{r},t) = \frac{\vec{E}_0(k)}{r}\mathrm{e}^{i(kr-\omega t)},. \tag{4.54}$$

$$\vec{B}(\vec{r},t) = \frac{\vec{B}_0(k)}{r}\mathrm{e}^{i(kr-\omega t)} \tag{4.55}$$

Substituting (4.54) and (4.55) into Maxwell's equations (4.3) and (4.4), after simple calculations, we obtain the formulas $\vec{E}_0(k)\vec{r} = 0$ and $\vec{B}_0(k)\vec{r} = 0$, which testify that spherical electromagnetic waves are transverse.

4.4. Electromagnetic wave polarization

The electromagnetic wave acts on electric charges owing to its electric and magnetic fields. The electric field acts on the charge q with the force $\vec{F}_e = q\vec{E}$, and the magnetic one with the force $\vec{F}_m = q\left[\vec{v}\vec{B}\right]$. Hence, $F_m/F_e \sim vB/E$ and, taking into account that $E = cB$, we obtain

$$\frac{F_m}{F_e} \sim \frac{v}{c}.$$

In most cases, $v \ll c$; therefore, $F_m \ll F_e$. Hence, the action of electromagnetic wave on electric charges is mainly governed by its electric field. From this viewpoint, the orientation of the vector of electric field strength of the wave, or the **wave polarization**, is of large importance.

In the case of arbitrary electromagnetic field, according to (4.35),

$$\vec{E}(\vec{r},t) = \int \vec{E}_0(\vec{k})\mathrm{e}^{i(\vec{k}\vec{r}-\omega t)}d^3k, \tag{4.56}$$

i.e. the electric field strength can be expanded in plane harmonic monochromatic waves (harmonics), which look like (4.39),

$$\vec{E}(\vec{k},\vec{r},t) = \vec{E}_0(\vec{k})\mathrm{e}^{i(\vec{k}\vec{r}-\omega t)}, \tag{4.57}$$

where, owing to the transverse character of electromagnetic waves, the vector $\vec{E}_0(\vec{k})$ lies in the plane normal to the wave vector $\vec{k}$. Choosing two mutually perpendicular unit vectors $\vec{e}_1(\vec{k})$ and $\vec{e}_2(\vec{k})$ lying in this plane, the amplitude $\vec{E}_0(\vec{k})$ can written as an expansion in those directions,

$$\begin{aligned} \vec{E}_0(\vec{k}) &= \vec{E}_{01}(\vec{k}) + \vec{E}_{02}(\vec{k}) \\ &= \vec{e}_1(\vec{k})E_{01}(\vec{k}) + \vec{e}_2(\vec{k})E_{02}(\vec{k}), \end{aligned} \tag{4.58}$$

and superposition (4.56) in the form

$$\vec{E}(\vec{r},t) = \int \sum_{i=1}^{2} \vec{e}_i(\vec{k})E_{0i}(\vec{k})\mathrm{e}^{i(\vec{k}\vec{r}-\omega t)}d^3k. \tag{4.59}$$

The unit vectors $\vec{e}_i(\vec{k})$ $(i = 1, 2)$ are called the **polarization vectors**.

Let the electromagnetic wave propagate along the axis z of coordinate system, and the polarization vectors $\vec{e}_1(\vec{k})$ and $\vec{e}_2(\vec{k})$ are directed along the axis x and y, respectively. Suppose first that the electromagnetic wave looks like harmonic (4.57), i.e.

$$\vec{\mathscr{E}} = \vec{\mathscr{E}}_0 \mathrm{e}^{i(kz-\omega t)} \tag{4.60}$$

or

$$\vec{\mathscr{E}} = \vec{\mathscr{E}}_0 \cos(kz - \omega t + \alpha) \tag{4.61}$$

in the real form. In this case, the direction of vector $\vec{\mathscr{E}}$ oscillations in the plane xy remains constant in time. Such an electromagnetic wave is called **plane-polarized** or **linearly polarized**, and the plane of vectors $\vec{k}$ and $\vec{\mathscr{E}}$, the **polarization plane**.

According to (4.58), wave (4.61) can be written as a sum of two waves

$$\vec{\mathscr{E}} = \vec{\mathscr{E}}_{0x} \cos(kz - \omega t + \alpha) + \vec{\mathscr{E}}_{0y} \cos(kz - \omega t + \alpha), \tag{4.62}$$

which have identical initial phases but are polarized perpendicularly to each other along the axes x and y, respectively, with the condition $\vec{\mathscr{E}}_{0x} + \vec{\mathscr{E}}_{0y} = \vec{\mathscr{E}}_0$ being satisfied. If those two waves have different initial phases,

$$\vec{\mathscr{E}}_x = \vec{\mathscr{E}}_{0x} \cos(kz - \omega t + \alpha_1), \tag{4.63}$$

$$\vec{\mathscr{E}}_y = \vec{\mathscr{E}}_{0y} \cos(kz - \omega t + \alpha_2), \tag{4.64}$$

then, two mutually perpendicular electric harmonic oscillations with the same frequency but different amplitudes and initial phases are summed up in the plane xy. In this case, as is known from the theory of oscillations, the terminal point of the resulting vector of electric field strength $\vec{\mathscr{E}}$ describes a path in the plane xy as the time varies. The path equation in the $(\mathscr{E}_x, \mathscr{E}_y)$-coordinates looks like

$$\frac{\mathscr{E}_x^2}{\mathscr{E}_{0x}^2} + \frac{\mathscr{E}_y^2}{\mathscr{E}_{0y}^2} - \frac{2\mathscr{E}_x\mathscr{E}_y}{\mathscr{E}_{0x}\mathscr{E}_{0y}} \cos(\alpha_2 - \alpha_1) = \sin(\alpha_2 - \alpha_1). \tag{4.65}$$

In the general case, this is the equation describing an ellipse in the $(\mathscr{E}_x, \mathscr{E}_y)$-plane.

If $\Delta\alpha = \alpha_2 - \alpha_1 = 2\pi m$, where $m \in Z$, equation (4.65) takes the form

$$\mathscr{E}_y = \frac{\mathscr{E}_{0y}}{\mathscr{E}_{0x}} \mathscr{E}_x. \tag{4.66}$$

This equation is the equation of a straight line in the $(\mathscr{E}_x, \mathscr{E}_y)$-coordinates (Fig. 4.3). Hence, the electromagnetic wave is plane-polarized.

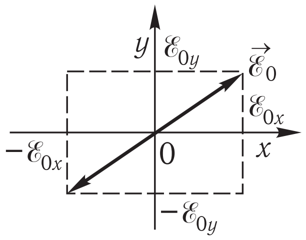

Fig. 4.3.

If $\Delta\alpha = \alpha_2 - \alpha_1 = (2m+1)\pi$, where $m \in Z$, equation (4.65) gives rise to

$$\mathscr{E}_y = -\frac{\mathscr{E}_{0y}}{\mathscr{E}_{0x}}\mathscr{E}_x. \tag{4.67}$$

This is also the equation of a straight line, i.e. this electromagnetic wave it is plane-polarized as well, but with a different orientation of vector $\vec{\mathscr{E}}$ (Fig. 4.4).

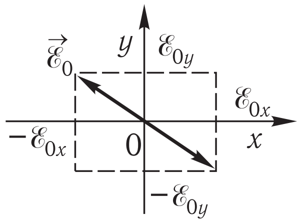

Fig. 4.4.

Let $\Delta\alpha = \left(2m \pm \frac{1}{2}\right)\pi$, where $m \in Z$. Substituting this value phase difference into (4.65), we obtain the equation for the ellipse

$$\frac{\mathscr{E}_x^2}{\mathscr{E}_{0x}^2} + \frac{\mathscr{E}_y^2}{\mathscr{E}_{0y}^2} = 1, \tag{4.68}$$

i.e. the terminal point of vector $\vec{E}$ of electromagnetic wave describes an ellipse in the plane xy as the time changes (Fig. 4.5). The terminal point of vector $\vec{\mathscr{E}}$ in Fig. 4.5 moves counterclockwise in the case $\Delta\alpha = \left(2m + \frac{1}{2}\right)\pi$, $m \in Z$, and clockwise if $\Delta\alpha = \left(2m - \frac{1}{2}\right)\pi$, $m \in Z$. To prove this, let $z = 0$, $\alpha_1 = 0$, and $\alpha_2 = \Delta\alpha$ in the equations for harmonics (4.63) and (4.64). Then, provided that $\Delta\alpha = \left(2m + \frac{1}{2}\right)\pi$, we obtain

$$\mathscr{E}_x = \mathscr{E}_{0x}\cos\omega t, \tag{4.69}$$

$$\mathscr{E}_y = \mathscr{E}_{0y}\sin\omega t. \tag{4.70}$$

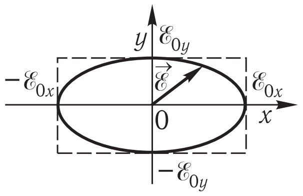

Fig. 4.5.

It is easy to see that, in this case, the point with the coordinates $(\mathscr{E}_x, \mathscr{E}_y)$ moves counterclockwise in time. Similarly, under the condition $\Delta\alpha = \left(2m - \frac{1}{2}\right)\pi$, we obtain

$$\mathscr{E}_x = \mathscr{E}_{0x}\cos\omega t, \tag{4.71}$$

$$\mathscr{E}_y = -\mathscr{E}_{0y}\sin\omega t. \tag{4.72}$$

and the point with the coordinates $(\mathscr{E}_x, \mathscr{E}_y)$ moves clockwise in time.

The described electromagnetic waves are called **elliptically polarized**. In this case, if $\mathscr{E}_{0x} = \mathscr{E}_{0y}$, the ellipse degenerates into a circle (see Fig. 4.5). The corresponding wave is called **circularly polarized**; left- of right-hand circularly polarized, depending on the direction of vector $\vec{\mathscr{E}}$ rotation. Notice that, if the phases α_1 and α_2 are equal $(\Delta\alpha = \alpha_2 - \alpha_1 = 0)$, the ellipse degenerates into a straight line described by equation (4.66), i.e. the electromagnetic wave becomes linearly polarized.

In the case of linear, elliptic, or circular polarization, the phase difference between components (4.63) and (4.64) does not change in time. Then, electromagnetic waves are called **completely polarized**. As a whole, these are elliptically polarized waves, but linearly and circularly polarized ones are their partial cases. However, if there is no mutual correlation between the phases of the waves polarized along the axes x and y, i.e. the phase difference $\Delta\alpha$ changes chaotically, the direction of vector $\vec{E}$ in the plane xy changes also chaotically. Such electromagnetic waves are called **non-polarized**.

5

EMISSION OF ELECTROMAGNETIC WAVES

5.1. Retarded potentials

As was mentioned above, the free electromagnetic field can exist in vacuum in the form of electromagnetic waves; however, its sources are electric charges. Therefore, it is of interest to establish the relationship between the parameters describing the system of charges and those describing electromagnetic waves. The latter include the vectors $\vec{E}$ and $\vec{B}$, which are determined by means of electromagnetic potentials as

$$\vec{E} = -\nabla\varphi - \frac{\partial \vec{A}}{\partial t}, \tag{5.1}$$

$$\vec{B} = \operatorname{rot}\vec{A}. \tag{5.2}$$

In turn, the potentials satisfy d'Alembert equation (see section 3.7)

$$\Delta\varphi - \frac{1}{c^2}\frac{\partial^2\varphi}{\partial t^2} = -\frac{\rho}{\varepsilon_0}, \tag{5.3}$$

$$\Delta\vec{A} - \frac{1}{c^2}\frac{\partial^2\vec{A}}{\partial t^2} = -\mu_0\vec{j}. \tag{5.4}$$

Equations (5.3) and (5.4) have the identical form. Therefore, if a solution is found for one of them, the solution for the other can be written by analogy.

First, let the electric charge be a point charge, $\delta q = \rho(t)\delta V$, located at the coordinate origin. Then the alternating electromagnetic field generated by this charge is spherically symmetric. As a result, in the spherical coordinate system (see Appendix C), the potential φ does not depend on the polar and azimuthal angles, i.e. we may write $\varphi = \varphi(r,t)$. Out of the coordinate origin, $\rho = 0$. Therefore, in this space region, equation (5.3) looks like

$$\Delta\varphi - \frac{1}{c^2}\frac{\partial^2\varphi}{\partial t^2} = 0. \tag{5.5}$$

Bearing in mind the form of Laplace operator Δ (see (4.48)) and taking into account that $\varphi = \varphi(r,t)$, we obtain the relation

$$\Delta\varphi = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\varphi}{\partial r}\right) = \frac{1}{r}\frac{\partial^2}{\partial r^2}(r\varphi). \tag{5.6}$$

Then, equation (5.5) can be written as

$$\frac{\partial^2 f}{\partial r^2} - \frac{1}{c^2}\frac{\partial^2 f}{\partial t^2} = 0, \tag{5.7}$$

where

$$f(r,t) = r\varphi(r,t). \tag{5.8}$$

Let us introduce new variables,

$$\xi = t - \frac{r}{c}, \tag{5.9}$$

$$\eta = t + \frac{r}{c}. \tag{5.10}$$

Then, taking into account that

$$\begin{aligned}\frac{\partial f}{\partial r} &= \frac{\partial f}{\partial \xi}\frac{\partial \xi}{\partial r} + \frac{\partial f}{\partial \eta}\frac{\partial \eta}{\partial r} = -\frac{1}{c}\frac{\partial f}{\partial \xi} + \frac{1}{c}\frac{\partial f}{\partial \eta},\\ \frac{\partial f}{\partial t} &= \frac{\partial f}{\partial \xi}\frac{\partial \xi}{\partial t} + \frac{\partial f}{\partial \eta}\frac{\partial \eta}{\partial t} = \frac{\partial f}{\partial \xi} + \frac{\partial f}{\partial \eta},\end{aligned}$$

we obtain, instead of (5.7), the equation

$$\frac{\partial^2 f}{\partial \xi \partial \eta} = 0. \tag{5.11}$$

Integrating it over η, we obtain a new equation

$$\frac{\partial f}{\partial \xi} = \psi(\xi),$$

and integrating the latter, we obtain the function f in the

$$f = \int \psi(\xi) d\xi + \psi_2(\eta) = \psi_1(\xi) + \psi_2(\eta). \tag{5.12}$$

From (5.8) and (5.12), it follows that

$$\begin{aligned}\varphi(r,t) &= \frac{f(r,t)}{r} = \frac{1}{r}\psi_1(\xi) + \frac{1}{r}\psi_2(\eta)\\ &= \frac{1}{r}\psi_1\left(t - \frac{r}{c}\right) + \frac{1}{r}\psi_2\left(t + \frac{r}{c}\right).\end{aligned} \tag{5.13}$$

The term

$$\varphi_1(r,t) = \frac{1}{r}\psi_1\left(t - \frac{r}{c}\right) \tag{5.14}$$

describes, as was shown in section 4.3, diverging spherical waves propagating at the velocity c from the coordinate origin, whereas the term

$$\varphi_2(r,t) = \frac{1}{r}\psi_2\left(t + \frac{r}{c}\right) \tag{5.15}$$

describes spherical waves converging to the coordinate origin.

Let us consider first solution (5.14) and establish the form of the function $\psi_1\left(t - \frac{r}{c}\right)$. If $c = \infty$, equation (5.3) is Poisson's equation

$$\Delta\varphi = -\frac{\rho}{\varepsilon_0}, \tag{5.16}$$

the solution of which looks like (see section 1.5)

$$\varphi = \frac{1}{4\pi\varepsilon_0}\int\frac{\rho dV}{r}. \tag{5.17}$$

In our case, the field is created by the point charge $\delta q = \rho(t)\delta V$. Therefore, instead of (5.17), we must write

$$\delta\varphi_1 = \frac{\rho(t)\delta V}{4\pi\varepsilon_0 r}. \tag{5.18}$$

Comparing (5.18) with (5.14), where $c = \infty$, we obtain the relationship

$$\psi_1 = \frac{\rho(t)\delta V}{4\pi\varepsilon_0}, \tag{5.19}$$

whence it follows that

$$\psi_1\left(t - \frac{r}{c}\right) = \frac{\rho\left(t - \frac{r}{c}\right)\delta V}{4\pi\varepsilon_0}, \tag{5.20}$$

hence,

$$\delta\varphi_1(r,t) = \frac{\rho\left(t - \frac{r}{c}\right)\delta V}{4\pi\varepsilon_0 r}. \tag{5.21}$$

If the charge δq is not located at the coordinate origin, but at the point with the coordinate $\vec{r}'$, and the observation point has the radius vector $\vec{r}$ (see Fig. 1.5), then,

$$\delta q = \rho\left(\vec{r}', t - \frac{|\vec{r} - \vec{r}'|}{c}\right)\delta V',$$

and, instead of (5.21), we must write

$$\delta\varphi_1(|\vec{r} - \vec{r}'|, t) = \frac{\rho\left(\vec{r}', t - \frac{|\vec{r} - \vec{r}'|}{c}\right)\delta V'}{4\pi\varepsilon_0 |\vec{r} - \vec{r}'|}. \tag{5.22}$$

At last, if the electric charge is not a point charge, but is distributed in a certain volume V', on the basis of (5.22), we may write the solution of d'Alembert equation (5.3) as

$$\varphi_1(\vec{r},t) = \frac{1}{4\pi\varepsilon_0} \int \frac{\rho\left(\vec{r}\,',t - \dfrac{|\vec{r}-\vec{r}\,'|}{c}\right) dV'}{|\vec{r}-\vec{r}\,'|}, \tag{5.23}$$

where $dV' = d^3r' = dx'dy'dz'$. In the same way, using the second summand in the right hand side of relation (5.13), we obtain another solution of equation (5.3) in the form

$$\varphi_2(\vec{r},t) = \frac{1}{4\pi\varepsilon_0} \int \frac{\rho\left(\vec{r}\,',t + \dfrac{|\vec{r}-\vec{r}\,'|}{c}\right) dV'}{|\vec{r}-\vec{r}\,'|}, \tag{5.24}$$

The value of potential $\varphi_1(\vec{r},t)$ at the moment t is determined by the distribution of charges at the previous moment $t - |\vec{r}-\vec{r}\,'|/c$. Therefore, the potential $\varphi_1(\vec{r},t)$ is called **retarded**. It corresponds to the causality principle, because the cause (the distribution of charges) precedes the consequence (the value of potential). The potential $\varphi_1(\vec{r},t)$ at the moment t is determined by the distribution of charges at the next moment $t + |\vec{r}-\vec{r'}|/c$. Therefore, it is called **advanced**. In this case, the consequence precedes the cause, which contradicts the causality principle. Below, we confine the consideration to retarded potentials and adopt that the solution (decision) of d'Alembert equation (5.3) looks like

$$\varphi(\vec{r},t) = \frac{1}{4\pi\varepsilon_0} \int \frac{\rho\left(\vec{r}\,',t - \dfrac{|\vec{r}-\vec{r}\,'|}{c}\right) dV'}{|\vec{r}-\vec{r}\,'|}, \tag{5.25}$$

Note that solution (5.24) formally appeared as a result of equation (5.3) invariance with respect to time inversion $t \to -t$, which means that the cause becomes a consequence, and the consequence becomes a reason. However, in our case, this situation contradicts the formulation of the problem, according to which electromagnetic waves are a consequence, whereas the variation of electric charge distribution in time is a cause.

The solution of d'Alembert equation (5.4) is written similarly to (5.25) in the form of retarded vector potential

$$\vec{A}(\vec{r},t) = \frac{\mu_0}{4\pi} \int \frac{\vec{j}\left(\vec{r}\,',t - \dfrac{|\vec{r}-\vec{r}\,'|}{c}\right) dV'}{|\vec{r}-\vec{r}\,'|}. \tag{5.26}$$

5.2. Electromagnetic field at large distances from the system of charges

The form of alternating electromagnetic field substantially depends on the distance from the system of radiating charges. The simplest and the most important case is the field at large distances from a radiator. At short distances, the field is very complicated and will not be considered.

To establish the criterion of a large distance from the charges, let us assume that the coordinate origin is located within the limits of volume V containing the charges. Then the absolute value of the radius vector of charges, r', does not exceed the linear dimensions of the volume V'. Therefore, the observation point is located at a large distance from the charges if the condition $r' \ll r$ is satisfied, where r is the distance reckoned from the coordinate origin to the observation point at which the potentials φ and $\vec{A}$ are calculated.

The quantity $|\vec{r} - \vec{r}'|$ that appears in (5.25) and (5.26) can be written in the form

$$|\vec{r} - \vec{r}'| = \sqrt{(\vec{r} - \vec{r}')^2} = \sqrt{r^2 - 2\vec{r}\vec{r}' + r'^2} = r\sqrt{1 - \frac{2\vec{r}\vec{r}'}{r^2} + \frac{r'^2}{r^2}}. \quad (5.27)$$

If $r' \ll r$, i.e. $r'/r \ll 1$, the last summand in (5.27) can be neglected. Then,

$$|\vec{r} - \vec{r}'| \approx r\left(1 - \frac{2\vec{r}\vec{r}'}{r^2}\right)^{\frac{1}{2}} \approx r\left(1 - \frac{\vec{r}\vec{r}'}{r^2}\right) = r = \vec{n}\vec{r}', \quad (5.28)$$

where $\vec{n} = \vec{r}/r$ is the unit vector directed along the vector $\vec{r}$. In transformation (5.28), we took into account that, if $x \ll 1$, $(1+x)^n \approx 1 + nx$. Using formula (5.28) in the numerators of integrands in (5.25) and (5.26), and neglecting the vector $\vec{r}'$ in the denominators, we obtain the electromagnetic potentials in the form

$$\begin{aligned} \varphi(\vec{r}, t) &= \frac{1}{4\pi\varepsilon_0 r}\int \rho\left(\vec{r}', t - \frac{r - \vec{r}'\vec{n}}{c}\right) dV' \\ &= \frac{1}{4\pi\varepsilon_0}\int \rho\left(\vec{r}', t_1\right) dV', \end{aligned} \quad (5.29)$$

$$\begin{aligned}\vec{A}(\vec{r},t) &= \frac{\mu_0}{4\pi r}\int \vec{j}\left(\vec{r}',t-\frac{r-\vec{r}'\vec{n}}{c}\right)dV' \\ &= \frac{\mu_0}{4\pi\varepsilon_0}\int \vec{j}\,(\vec{r}',t_1)\,dV' \end{aligned} \tag{5.30}$$

where

$$t_1 = t - \frac{r-\vec{r}'\vec{n}}{c}. \tag{5.31}$$

With the help of formulas (5.1), (5.2), (5.29), and (5.30), let us calculate the magnetic field induction $\vec{B}$ and the electric field strength $\vec{E}$ of emitted electromagnetic radiation. In the formula $\vec{B} = \operatorname{rot}\vec{A}$, the vector $\vec{B}$ is calculated at the observation point (the terminal point of vector $\vec{r}$), i.e. differentiation is carried out with respect to the coordinates of the terminal point of vector $\vec{r}$ (see Appendix B). Therefore,

$$\begin{aligned}\vec{B}(\vec{r},t) &= \operatorname{rot}\vec{A} = \frac{\mu_0}{4\pi}\int \operatorname{rot}\frac{\vec{j}(\vec{r}',t_1)}{r}dV' \\ &= \frac{\mu_0}{4\pi}\int\left\{\frac{1}{r}\operatorname{rot}\vec{j}(\vec{r}',t_1)+\left[\nabla\frac{1}{r}\vec{j}(\vec{r}',t_1)\right]\right\}dV' \\ &= \frac{\mu_0}{4\pi}\int\left\{\frac{1}{r}\operatorname{rot}\vec{j}(\vec{r}',t_1)-\left[\frac{\vec{r}}{r^3}\vec{j}(\vec{r}',t_1)\right]\right\}dV'. \end{aligned} \tag{5.32}$$

The first summand is proportional to $1/r$, and the second to $1/r^2$. However, at large r-values, the second summand can be neglected, and as a result, instead of (5.32), we obtain

$$\vec{B}(\vec{r},t) = \frac{\mu_0}{4\pi r}\int \operatorname{rot}\vec{j}(\vec{r}',t_1)dV'. \tag{5.33}$$

Let us calculate $\operatorname{rot}\vec{j}$ taking into account (5.31):

$$\begin{aligned}\operatorname{rot}\vec{j}(\vec{r}',t_1) &= \left[\nabla\vec{j}(\vec{r}',t_1)\right] = \left[\nabla r\frac{\partial\vec{j}}{\partial t_1}\frac{\partial t_1}{\partial r}\right] \\ &= \left[\frac{\vec{r}}{r}\frac{\partial\vec{j}}{\partial t_1}\left(-\frac{1}{c}\right)\right] = \frac{1}{c}\left[\dot{\vec{j}}\vec{n}\right]. \end{aligned} \tag{5.34}$$

Hence,

$$\begin{aligned}\vec{B}(\vec{r},t) &= \frac{\mu_0}{4\pi rc}\int\left[\dot{\vec{j}}\vec{n}\right]dV' = \frac{1}{c}\left[\frac{\mu_0}{4\pi}\int\frac{\dot{\vec{j}}dV'}{r}\vec{n}\right]\\ &= \frac{1}{c}\left[\dot{\vec{A}}(\vec{r},t)\vec{n}\right]. \end{aligned} \tag{5.35}$$

The vector $\vec{E}(\vec{r},t)$ is calculated by formula (5.1),

$$\vec{E} = -\nabla\varphi - \dot{\vec{A}}. \tag{5.36}$$

Using (5.29), we obtain

$$\begin{aligned}\nabla\varphi &= \frac{1}{4\pi\varepsilon_0}\int\nabla\frac{\rho(\vec{r}\,',t)}{r}dV' = \frac{1}{4\pi\varepsilon_0}\int\left\{-\frac{\vec{r}}{r^3}\rho - \frac{1}{rc}\dot{\rho}\vec{n}\right\}dV'\\ &\approx -\frac{\vec{n}}{4\pi\varepsilon_0 c}\int\frac{\dot{\rho}(\vec{r}\,',t)}{r}dV'. \end{aligned} \tag{5.37}$$

In this case, the continuity equation

$$\frac{\partial\rho}{\partial t} = -\operatorname{div}\vec{j}$$

can be written in the form

$$\dot{\rho} = -\nabla\vec{j} = \frac{1}{c}\dot{\vec{j}}\vec{n}. \tag{5.38}$$

Then,

$$\begin{aligned}\nabla\varphi &= -\frac{\vec{n}}{4\pi\varepsilon_0 c^2}\int\frac{\dot{\vec{j}}\vec{n}}{r}dV'\\ &= -\frac{\vec{n}}{\varepsilon_0\mu_0 c^2}\left\{\frac{\mu_0}{4\pi}\int\frac{\dot{\vec{j}}}{r}dV'\cdot\vec{n}\right\} = -\vec{n}\left(\dot{\vec{A}}\vec{n}\right). \end{aligned} \tag{5.39}$$

Returning to (5.36) and taking (5.35) into account, we obtain

$$\vec{E}(\vec{r},t) = \vec{n}\left(\dot{\vec{A}}\vec{n}\right) - \dot{\vec{A}} = \left[\left[\dot{\vec{A}}\vec{n}\right]\vec{n}\right] = c\left[\vec{B}(\vec{r},t)\vec{n}\right]. \tag{5.40}$$

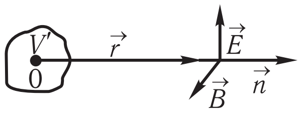

Fig. 5.1.

The relative orientation of vectors $\vec{E}$, $\vec{n}$, and $\vec{B}$ following from (5.40) is depicted in Fig. 5.1. Hence, their relative arrangement and interconnection are the same as in the case of plane electromagnetic wave (see section 4.2). It is so because a small section of spherical wave front can be regarded as a plane at large distances from the radiator in the so-called wave zone.

5.3. Electric dipole radiation

According to (5.35) and (5.40), the vectors $\vec{E}$ and $\vec{B}$ for electromagnetic radiation can be determined from vector potential (5.30)

$$\vec{A}(\vec{r},t) = \frac{\mu_0}{4\pi r}\int \vec{j}\left(\vec{r}', t - \frac{r - \vec{r}'\vec{n}}{c}\right) dV'. \tag{5.41}$$

If $r' \ll r$, the integrand in (5.41) can be expanded in a power series of $\vec{r}'\vec{n}/c$,

$$\begin{aligned}\vec{j}\left(\vec{r}', t - \frac{r}{c} + \frac{\vec{r}'\vec{n}}{c}\right) &\approx \vec{j}\left(\vec{r}', t - \frac{r}{c}\right) + \frac{\partial}{\partial t}\vec{j}\left(\vec{r}', t - \frac{r}{c}\right)\frac{\vec{r}'\vec{n}}{c} \\ &\equiv \vec{j}\left(\vec{r}', t - \frac{r}{c}\right) + \dot{\vec{j}}\left(\vec{r}', t - \frac{r}{c}\right)\frac{\vec{r}'\vec{n}}{c},\end{aligned}$$

where $\dot{\vec{j}} = \partial\vec{j}/\partial t$. Then, instead of (5.41), we obtain

$$\begin{aligned}\vec{A}(\vec{r},t) &= \frac{\mu_0}{4\pi r}\int \vec{j}\left(\vec{r}', t - \frac{r}{c}\right) dV' \\ &+ \frac{\mu_0}{4\pi r c}\frac{\partial}{\partial t}\int (\vec{r}'\vec{n})\,\vec{j}\left(\vec{r}', t - \frac{r}{c}\right) dV'\end{aligned} \tag{5.42}$$

Notice that expansion (5.42) is valid if the currents $\vec{j}$ do not change substantially within the time interval $\vec{r}'\vec{n}/c$; otherwise, the derivative of $\vec{j}$ is large, and the second summand cannot be taken as small. The time $\vec{r}'\vec{n}/c$ is the time of field propagation within the radiating source limits. Therefore, if electric charges the motion of which creates the current have the velocity v, their displacement for the time interval $\vec{r}'\vec{n}/c$ equals $v\left(\vec{r}'\vec{n}\right)/c = vL/c$, where L is the linear size of radiator $\left(L \sim V^{1/3}\right)$. For the currents and, hence, currents to change slightly within the time interval $\vec{r}'\vec{n}/c$, the condition $vL/c \ll L$ or

$$v \ll c \tag{5.43}$$

has to be satisfied. In what follows, we assume that this condition is obeyed and expansion (5.42) is applicable.

Let us consider the first (dominating) summand in (5.42),

$$\vec{A}(\vec{r}, t) = \frac{\mu_0}{4\pi r} \int \vec{j}\left(\vec{r}', t - \frac{r}{c}\right) dV'. \tag{5.44}$$

Here, the integral differs from zero, because, owing to the current nonstationarity, the conductivity current lines are not closed (see sections 2.6 and 3.2). However, it can be transformed to the form

$$\int \vec{j}\left(\vec{r}', t - \frac{r}{c}\right) dV' = -\int \vec{r}' \operatorname{div} \vec{j}\left(\vec{r}', t - \frac{r}{c}\right) dV'. \tag{5.45}$$

Really, according to the formula of vector analysis (see Appendix B),

$$\psi \operatorname{div} \vec{a} = \operatorname{div}\, (\psi \vec{a}) - \nabla \psi \vec{a},$$

we may write

$$\int x' \operatorname{div} \vec{j} dV' = \int \operatorname{div}\, \left(x'\vec{j}\right) dV' - \int \nabla x' \vec{j}\left(\vec{r}', t - \frac{r}{c}\right) dV''.$$

Since

$$\nabla x' = \vec{i}',$$

where $\vec{i}'$ is the unit vector of Cartesian coordinate systems, we have

$$\int x' \operatorname{div} \vec{j} dV' = \oint_S x' \vec{j}\, d\vec{S} - \int j_{x'} dV'. \tag{5.46}$$

The closed surface S envelops the whole volume of currents; therefore, $\vec{j} = 0$ on it, and, hence,

$$\int x' \operatorname{div} \vec{j} dV' = -\int j_{x'} dV'. \tag{5.47}$$

Thus, formula (5.47) gives rise to (5.45).

The continuity equation can be written in the form

$$\operatorname{div} \vec{j}\left(\vec{r}\,', t - \frac{r}{c}\right) = -\dot{\rho}\left(\vec{r}\,', t - \frac{r}{c}\right), \tag{5.48}$$

where $\dot{\rho} = \partial \rho / \partial t$. Therefore, relation (5.45) reads

$$\int \vec{j}\left(\vec{r}\,', t - \frac{r}{c}\right) dV' = \int \vec{r}\,' \dot{\rho}\left(\vec{r}\,', t - \frac{r}{c}\right) dV' = \dot{\vec{d}}\left(t - \frac{r}{c}\right). \tag{5.49}$$

Here, $\dot{\vec{d}}\left(t - \frac{r}{c}\right)$ is the electric dipole moment of radiating system at the time moment $t - \frac{r}{c}$. On the basis of (5.44) and (5.49), we obtain

$$\vec{A}_1(\vec{r}, t) = \frac{\mu_0}{4\pi r} \dot{\vec{d}}\left(t - \frac{r}{c}\right). \tag{5.50}$$

Substituting (5.50) into (5.35), we have

$$\vec{B}_{dip}(\vec{r}, t) = \frac{\mu_0}{4\pi r c}\left[\ddot{\vec{d}}\left(t - \frac{r}{c}\right) \vec{n}\right]. \tag{5.51}$$

Then, in accordance with (5.40),

$$\vec{E}_{dip}(\vec{r}, t) = c\left[\vec{B}(\vec{r}, t) \vec{n}\right] = \frac{\mu_0}{4\pi r}\left[\left[\ddot{\vec{d}}\left(t - \frac{r}{c}\right) \vec{n}\right] \vec{n}\right]. \tag{5.52}$$

Formulas (5.51) and (5.52) are used to determine the vectors of an electromagnetic radiation field generated by an alternating electric dipole moment of radiating source, which is called the **electric dipole**.

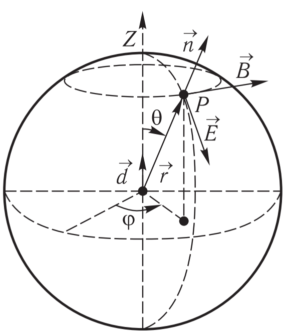

Fig. 5.2.

Let an alternating electric dipole be oriented along the axis z of spherical coordinate system, and the observation point has the coordinates (r, θ, φ) (see Fig. 5.2). Then, from (5.51), it follows that the vector $\vec{B}$ must be perpendicular to the plane of vectors $\vec{d}$ and $\vec{n}$, i.e. be tangent to the "parallel" at which the point P is located. The direction of vector $\vec{E}$ is determined according to relation (5.52): it should be tangent to the "meridian" of point P. The directions of vectors $\vec{E}$ and $\vec{B}$ at a certain time moment in the case when the source of radiation is at the coordinate origin and its electric dipole moment is oriented along the axis Z are shown in Fig. 5.2.

The absolute values of vectors $\vec{E}$ and $\vec{B}$ are determined by the formulas

$$B_{dip}(\vec{r},t) = \frac{\mu_0 \ddot{d}\left(t-\dfrac{r}{c}\right)}{4\pi r c}\sin\theta, \tag{5.53}$$

$$E_{dip}(\vec{r},t) = \frac{\mu_0 \ddot{d}\left(t-\dfrac{r}{c}\right)}{4\pi r}\sin\theta, \tag{5.54}$$

whereas the absolute value of Poynting vector (3.60), i.e. the energy flux through a unit area perpendicular to $\vec{r}$, by the formula

$$\Pi_{dip}(\vec{r},t) = \frac{1}{\mu_0}\left|\left[\vec{E}_{dip}\vec{B}_{dip}\right]\right| = \frac{\mu_0\left\{\ddot{d}\left(t-\dfrac{r}{c}\right)\right\}}{16\pi^2 r^2 c}\sin^2\theta. \tag{5.55}$$

From (5.55), it follows that the intensity of electric dipole radiation is maximum if $\theta = \pi/2$. At the same time, radiation is absent along the direction of dipole moment ($\theta = 0$) and opposite to it ($\theta = \pi$).

The instant radiation power of electric dipole, i.e. the energy emitted by the dipole per unit time equals

$$\begin{aligned} I_{dip}(t) &= \oint \Pi_{dip} dS = \frac{\mu_0\left\{\ddot{d}\left(t-\dfrac{r}{c}\right)\right\}^2}{16\pi^2 c}\int_0^{\pi}\sin^3\theta d\theta\int_0^{2\pi}d\varphi \\ &= \frac{\mu_0\left\{\ddot{d}\left(t-\dfrac{r}{c}\right)\right\}^2}{6\pi c}. \end{aligned} \tag{5.56}$$

Of special importance is the case when the electric dipole moment oscillates harmonically, i.e. when $d = d_0\cos\omega t$. Then, relation (5.56) takes the form

$$I_{dip}(t) = \frac{\mu_0 d_0^2\omega^4}{6\pi c}\cos^2\omega\left(t-\frac{r}{c}\right), \tag{5.57}$$

and the average power of radiation by the electric dipole equals

$$\langle I_{dip}(t)\rangle = \frac{\mu_0 d_0^2\omega^4}{12\pi c}. \tag{5.58}$$

If the source of electromagnetic radiation is a single moving particle with the electric charge e, its electric dipole moment, in accordance with (1.8) and (1.49), equals $\vec{d} = e\vec{r}$. Then, on the basis of relationship (5.56), we obtain

$$I_{dip}(t) = \frac{\mu_0 e^2 \left(\ddot{\vec{r}}\right)^2}{6\pi c}, \tag{5.59}$$

i.e. the charged particle radiates energy if it moves with the acceleration. Notice that this conclusion concerns particles which motion can be described with the help of classical mechanics. However, it is inapplicable to the particles, the behavior of which is described by quantum-mechanical laws. The latter particles include, in particular, electrons in atoms and molecules.

B. Electromagnetic field in the medium

6

EQUATIONS FOR THE ELECTROMAGNETIC FIELD IN THE MEDIUM

6.1. Maxwell–Lorentz equations

As was already mentioned, the macroscopic electromagnetic field in vacuum can be described using the vectors of electric field strength $\vec{E}$ and magnetic induction $\vec{B}$, which satisfy Maxwell's equation (3.36)–(3.39),

$$\operatorname{rot}\vec{B} = \mu_0\vec{j} + \varepsilon_0\mu_0\frac{\partial\vec{E}}{\partial t}, \tag{6.1}$$

$$\operatorname{rot}\vec{E} = -\frac{\partial\vec{B}}{\partial t}, \tag{6.2}$$

$$\operatorname{div}\vec{B} = 0, \tag{6.3}$$

$$\operatorname{div}\vec{E} = \frac{\rho}{\varepsilon_0}. \tag{6.4}$$

Those equations should be supplemented with boundary conditions (3.40)–(3.43) and the equation $\vec{j} = \sigma\left(\vec{E} + \vec{E}^{ext}\right)$. The field vectors $\vec{E}$ and $\vec{B}$ in equations (6.1)–(6.4) have a macroscopic meaning: when acting on electric charges, they have the same value within the physically infinitesimal volume containing a large number of particles (see section 1.1) and change only when the coordinate varies at the macroscopic scale, the latter exceeding the atomic size very much.

Let us proceed to the consideration of electromagnetic field in a material medium (in a substance). The latter, being electrically neutral, consists of charged particles: atomic nuclei, electrons, and so forth. "External" charges and currents are considered as field sources, but the observation point is always located inside the medium, among the atoms of the substance. Then, the field vectors must considerably depend on the observation point position in the microscopic scale (e.g., near atomic nuclei, in the interatomic space, and so forth) and on the time momentum (owing to the motion of nuclei and electrons). In this case, actual $\vec{E}$ and $\vec{B}$ vectors strongly vary if either the observation point location or/and the time moment changes. Below, those actual instant values of the field vectors will be associated with the **microscopic fields** and denoted as $\vec{E}_{micro}$ and $\vec{B}_{micro}$.

The field inside charged particles—atomic nuclei, nucleons, electrons—will not be considered (this issue is analyzed in the theory of elementary particles). Therefore, the observation point is assumed to be outside the particles, i.e. in the vacuum. Then, it is reasonable to suppose that equations (6.1)–(6.4) are also applicable for the microfield, but they should be rewritten in the form

$$\begin{aligned}
\operatorname{rot}\vec{B}_{micro} &= \mu_0\vec{j}_{micro} + \varepsilon_0\mu_0\frac{\partial\vec{E}_{micro}}{\partial t}, && (6.5)\\
\operatorname{rot}\vec{E}_{micro} &= -\frac{\partial\vec{B}_{micro}}{\partial t}, && (6.6)\\
\operatorname{div}\vec{B}_{micro} &= 0, && (6.7)\\
\operatorname{div}\vec{E}_{micro} &= \frac{\rho_{micro}}{\varepsilon_0}. && (6.8)
\end{aligned}$$

The system of equations (6.5)–(6.8) was applied by the Dutch physicist H. Lorentz to describe the microfield in the substance in the framework of electron theory developed by him. Therefore, they are called the **Maxwell–Lorentz** (in some sources, **Lorentz–Maxwell**) **equations**.

At the same time, directly measured are fields that act on a macroscopic charge or current rather than a separate particle. In other words, those are **macroscopic fields** denoted as $\vec{E}_{macro} = \vec{E}$ and $\vec{B}_{macro} = \vec{B}$. Besides, it is the macrofields $\vec{E}$ and $\vec{B}$ that are usually used in applied physical aspects. Taking those speculations into account, it is easy to understand that the vectors $\vec{E}$ and $\vec{B}$ have the meaning of averaged microfields $\vec{E}_{micro}$ and $\vec{B}_{micro}$. Let us consider this issue in more details.

6.2. Averaged Maxwell–Lorentz equations

The differential equations of electromagnetic field must couple the values of problem quantities that they have at a given point in the space and at a given time moment. Therefore, the Maxwell–Lorentz equations should be averaged over a small volume δV regarded as infinitesimal at the macroscopic scale; it must be a point. On the other hand, the volume δV has to contain a large enough number of particles for the concepts of average values to be applicable. Such a volume is called the **physically infinitesimal volume** (see section 1.1). The time interval τ over which the averaging is carried out must be much shorter than the observation time Δt, but much longer than the period of thermal vibrations of atoms in molecules, the periods of intra-atomic motions, and so forth. In this case, the average value of a certain quantity $\mathscr{F}$ can be written down in the form

$$\langle \mathscr{F} \rangle = \frac{1}{\tau} \int_0^\tau \left\{ \frac{1}{\delta V} \int_{\delta V} \mathscr{F}(\vec{r}, \vec{r}', t, t') dV' \right\} dt'. \tag{6.9}$$

However, it is impossible to use this formula for direct calculations of average quantities, because the integrand is unknown. At the same time, this difficulty can be overcome using macroscopic distributions of charges and currents (see below), and macroscopic characteristics of the substance.

In order to averaging the Maxwell–Lorentz equations in accordance with (6.9), let us first rewrite them as follows:

$$\left\langle \mathrm{rot}\, \vec{B}_{micro} \right\rangle = \mu_0 \left\langle \vec{j}_{micro} \right\rangle + \varepsilon_0 \mu_0 \left\langle \frac{\partial \vec{E}_{micro}}{\partial t} \right\rangle, \tag{6.10}$$

$$\left\langle \mathrm{rot}\, \vec{E}_{micro} \right\rangle = -\left\langle \frac{\partial \vec{B}_{micro}}{\partial t} \right\rangle, \tag{6.11}$$

$$\left\langle \mathrm{div}\, \vec{B}_{micro} \right\rangle = 0, \tag{6.12}$$

$$\left\langle \mathrm{div}\, \vec{E}_{micro} \right\rangle = \frac{1}{\varepsilon_0} \langle \rho_{micro} \rangle . \tag{6.13}$$

Pay attention that that averaging in (6.9)–(6.13) is carried out over the microscopic coordinates $\vec{r}'$ and time t', whereas the operations div, rot, and $\partial/\partial t$ are calculated with respect to the macroscopic coordinates

$\vec{r}$ and time t. Therefore, their order can be changed. As a result,

$$\begin{aligned}
\operatorname{rot}\left\langle \vec{B}_{micro}\right\rangle &= \mu_0\left\langle \vec{j}_{micro}\right\rangle + \varepsilon_0\mu_0\frac{\partial}{\partial t}\left\langle \vec{E}_{micro}\right\rangle, &(6.14)\\
\operatorname{rot}\left\langle \vec{E}_{micro}\right\rangle &= -\frac{\partial}{\partial t}\left\langle \vec{B}_{micro}\right\rangle, &(6.15)\\
\operatorname{div}\left\langle \vec{B}_{micro}\right\rangle &= 0, &(6.16)\\
\operatorname{div}\left\langle \vec{E}_{micro}\right\rangle &= \frac{1}{\varepsilon_0}\left\langle \rho_{micro}\right\rangle. &(6.17)
\end{aligned}$$

The quantities $\left\langle \vec{E}_{micro}\right\rangle$ and $\left\langle \vec{B}_{micro}\right\rangle$ are the averaged microfield vectors, i.e. the electric strength and magnetic induction, respectively, of the macroscopic electromagnetic field in the medium,

$$\begin{aligned}
\vec{E}(\vec{r},t) &= \left\langle \vec{E}_{micro}\right\rangle, &(6.18)\\
\vec{B}(\vec{r},t) &= \left\langle \vec{B}_{micro}\right\rangle. &(6.19)
\end{aligned}$$

As a result, the system of equations (6.14)–(6.17) takes the form

$$\begin{aligned}
\operatorname{rot}\vec{B} &= \mu_0\left\langle \vec{j}_{micro}\right\rangle + \varepsilon_0\mu_0\frac{\partial\vec{E}}{\partial t}, &(6.20)\\
\operatorname{rot}\vec{E} &= -\frac{\partial\vec{B}}{\partial t}, &(6.21)\\
\operatorname{div}\vec{B} &= 0, &(6.22)\\
\operatorname{div}\vec{E} &= \frac{1}{\varepsilon_0}\left\langle \rho_{micro}\right\rangle, &(6.23)
\end{aligned}$$

and the task consists in calculating $\left\langle \vec{j}_{micro}\right\rangle$ and $\langle \rho_{micro}\rangle$.

6.3. Polarization and electric induction vectors

While calculating $\langle \rho_{micro}\rangle$, we will distinguish between free and bound electric charges in the medium. **Free charges** are charges that can move over macroscopic distances. These are free electrons in metals, ions in electrolytes and gases, and charges deposited on the insulator and violating its electroneutrality. **Bound charges** are electric charges in atoms and molecules of insulator which can move only over microscopic distances under the action of electric field; they remain in “their” atoms and molecules and are responsible for insulator polarization. Hence, it is possible to write

$$\langle \rho_{micro} \rangle = \langle \rho_{free} \rangle + \langle \rho_{bound} \rangle \,. \tag{6.24}$$

In what follows, we will use the notation

$$\langle \rho_{free} \rangle = \rho \tag{6.25}$$

If the insulator is electrically neutral and the external electric field is absent, then $\langle \rho_{bound} \rangle = 0$ as a rule. However, under the action of external field, the charges shift and the average density of bound charges can become different from zero.

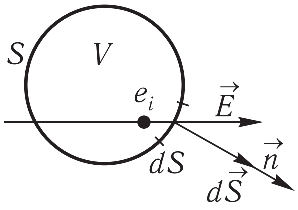

Fig. 6.1.

Consider an insulator in the electric field and some volume V in it confined by the surface S (Fig. 6.1). Let the insulator consist of charged particles of several types (electrons, nuclei, ions, and so forth) with charges e_i and the concentrations n_{0i}. If the displacement of the charges of the i-th type under the field action is designated as $\delta\vec{r}_i$, the amount of charge leaving the volume V through the surface element dS (see Fig. 6.1) equals

$$dq = \sum_i e_i n_{0i} \delta\vec{r}_i d\vec{S}.$$

As a result, the volume V acquires the bound charge

$$dq_{bound} = -dq = -\sum_i e_i n_{0i} \delta\vec{r}_i d\vec{S} = -\vec{P} d\vec{S}, \tag{6.26}$$

where

$$\vec{P} = \sum_i e_i n_{0i} \delta\vec{r}_i \tag{6.27}$$

is the electric dipole moment per unit insulator volume, which is called the **polarization vector**.

Besides charges leaving the volume V under the field action, it is necessary to take into account charges that, shifting, enter this volume. Then, the resulting bound charge acquired by the volume confined by the closed surface S is determines as

$$q_{bound} = -\oint_S \vec{P}\, d\vec{S} = -\int_V \operatorname{div} \vec{P}\, dV. \tag{6.28}$$

Expressing the bound charge in the form

$$q_{bound} = \int_V \langle \rho_{bound} \rangle \, dV, \tag{6.29}$$

we obtain the relation

$$\langle \rho_{bound} \rangle = -\operatorname{div} \vec{P} \tag{6.30}$$

On the basis of (6.24), (6.25), and (6.30), equation (6.23) takes the form

$$\operatorname{div} \left(\varepsilon_0 \vec{E} + \vec{P} \right) = \rho. \tag{6.31}$$

The vector

$$\vec{D} = \varepsilon_0 \vec{E} + \vec{P} \tag{6.32}$$

is called the **electric induction vector** (the obsolete term is the electric displacement vector). Then, instead of (6.31), we may write the equation

$$\operatorname{div} \vec{D} = \rho, \tag{6.33}$$

which is a consequence of the averaging of Maxwell–Lorentz equation (6.8).

6.4. Magnetization and magnetic field strength vectors

While calculating $\left\langle \vec{\jmath}_{micro} \right\rangle$, it is necessary to bear in mind that microscopic currents $\vec{\jmath}_{micro}$ result from the motion of free charges, the displacement of bound charges, and the closed currents in atoms and molecules are caused by the motion of electrons in them.

The directed motion of free charges is the **conductivity current.** Its density will be designated $\vec{j}_{cond}$.

The displacement of bound charges is associated with polarization of medium. Therefore, the corresponding microcurrent is called the **polarization current**. The density of polarization current equals

$$\left\langle \vec{j}_{pol} \right\rangle = \sum_i e_i n_{0i} \vec{v}_i = \frac{\partial}{\partial t} \sum_i e_i n_{0i} \delta \vec{r}_i, \tag{6.34}$$

where the averaging operation $\langle \ldots \rangle$ is associated with the application of macroscopic quantity n_{0i}. Taking into account (6.27), we obtain

$$\left\langle \vec{j}_{pol} \right\rangle = \frac{\partial \vec{P}}{\partial t}. \tag{6.35}$$

The closed currents in atoms (molecules) are responsible for the medium magnetization and therefore are called the **magnetization currents**. To determine their density, let us consider the closed currents in atoms (molecules) to be linear with the current strength I_0 and the contour area ΔS_0 (Fig. 6.2a).

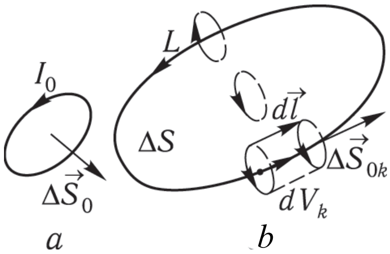

Fig. 6.2.

In the medium, let us select a macroscopic section ΔS confined by the contour L (Fig. 6.2a) and find the strength of a current associated with intra-atomic currents that passes through the contour. From Fig. 6.2b, is follows that those currents give a contribution to the current concerned if the contour of current I_0 intersects the section ΔS only once, i.e. if the contour L intersects the area ΔS_0 of intra-atomic currents. Let us select an element $d\vec{l}$ in the contour and consider an infinitesimal volume $dV_k = d\vec{l} \Delta \vec{S}_{0k}$ of a cylinder formed by the generatrix $d\vec{l}$ and two bases, which are two identically oriented areas of current contours ΔS_{0k}. Taking into account different orientations of current contours on the element $d\vec{l}$, the resulting current passing through the section ΔS on the element $d\vec{l}$ of its contour can be written in the form

$$dI_{mag} = \sum_k I_{0k} n_{0k} dV_k = \sum_k I_{0k} n_{0k} d\vec{l} \Delta \vec{S}_{0k}. \tag{6.36}$$

The vector

$$\vec{p}_{mk} = I_{0k} \Delta \vec{S}_{0k}$$

is the magnetic moment of closed current I_{0k} with the vector contour section ΔS_{0k} (see section 2.6), and the vector

$$\vec{P}_m = \sum_k \vec{p}_{mk} n_{0k} \tag{6.37}$$

is called the **magnetization vector** or the **magnetization**. Its magnitude equals the magnetic moment per unit volume of the medium. Taking into account (6.37), relation (6.36) can be written as

$$dI_{mag} = \vec{P}_m d\vec{l}. \tag{6.38}$$

In the non-ferromagnetic medium, the magnetic moments $\vec{p}_{mk}$ in sum (6.37) are oriented chaotically in the absence of external magnetic field. As a result, the magnetization vector $\vec{P}_m$ equals zero. In ferromagnets, the vector $\vec{P}_m$ can differ from zero even in the absence of external magnetic field, but ferromagnetism is a quantum-mechanical effect. Therefore, it will not be consider here.

On the basis of (6.38), the total magnetization current reads

$$I_{mag} = \oint_L \vec{P}_m d\vec{l} = \int_{\Delta S} \operatorname{rot} \vec{P}_m \, d\vec{S}. \tag{6.39}$$

Therefore, taking into account that

$$I_{mag} = \int_{\Delta S} \left\langle \vec{j}_{mag} \right\rangle d\vec{S}, \tag{6.40}$$

we obtain the average density of magnetization current in the form

$$\left\langle \vec{j}_{mag} \right\rangle = \operatorname{rot} \vec{P}_m. \tag{6.41}$$

Summing up the averaged microcurrents, we may write

$$\left\langle \vec{j}_{micro} \right\rangle = \left\langle \vec{j}_{cond} \right\rangle + \left\langle \vec{j}_{pol} \right\rangle + \left\langle \vec{j}_{mag} \right\rangle = \vec{j} + \frac{\partial \vec{P}}{\partial t} + \operatorname{rot} \vec{P}_m. \tag{6.42}$$

Substituting (6.42) into (6.20), we obtain the equation

$$\operatorname{rot}\left(\frac{\vec{B}}{\mu_0} - \vec{P}_m\right) = \vec{j} + \frac{\partial}{\partial t}\left(\varepsilon_0\vec{E} + \vec{P}\right). \tag{6.43}$$

The vector

$$\vec{H} = \frac{\vec{B}}{\mu_0} - \vec{P}_m \tag{6.44}$$

is called the **vector of magnetic field strength**. Then, taking into account (6.44) and (6.32), equation (6.43) is written as

$$\operatorname{rot}\vec{H} = \vec{j} + \frac{\partial\vec{D}}{\partial t}. \tag{6.45}$$

Equation (6.45) is a consequence of averaging of Maxwell–Lorentz equation (6.5).

6.5. Maxwell's equations for the electromagnetic field in the medium

After averaging Maxwell–Lorentz equations (6.5)–(6.8) for the microfield, the differential equations for the electromagnetic field in a medium, in accordance with (6.45), (6.33), (6.21), and (6.22), look like

$$\begin{aligned} \operatorname{rot}\vec{H} &= \vec{j} + \frac{\partial\vec{D}}{\partial t}, && (6.46)\\ \operatorname{rot}\vec{E} &= -\frac{\partial\vec{B}}{\partial t}, && (6.47)\\ \operatorname{div}\vec{B} &= 0, && (6.48)\\ \operatorname{div}\vec{D} &= \rho. && (6.49) \end{aligned}$$

They are the equations of classical macroscopic electrodynamics for the electromagnetic field in the medium and are called the **Maxwell equations in the medium**. Unlike Maxwell's equations (3.36)–(3.39), they are applicable to the macroscopic field in both the vacuum (then, $\vec{D} = \varepsilon_0\vec{E}$ and $\vec{H} = \vec{B}/\mu_0$) and the medium.

In the integral form, Maxwell's equations read

$$\oint \vec{H}\,d\vec{l} = \int \vec{j}\,d\vec{S} + \int \frac{\partial \vec{D}}{\partial t} d\vec{S}, \tag{6.50}$$

$$\oint \vec{E}\,d\vec{l} = -\int \frac{\partial \vec{B}}{\partial t} d\vec{S}, \tag{6.51}$$

$$\oint \vec{B}\,d\vec{S} = 0, \tag{6.52}$$

$$\oint \vec{D}\,d\vec{S} = \int \rho\,dV. \tag{6.53}$$

In equations (6.46)–(6.53), two "magnetic" vectors, $\vec{H}$ and $\vec{B}$, are used, which are coupled by relation (6.44), and two "electric" ones, $\vec{E}$ and $\vec{D}$, interrelated by (6.32). The force characteristic of electric field, which determines its action on an electric charge, is the strength $\vec{E}$, and the force characteristic of magnetic field, which determines its action on an electric current, is the magnetic induction $\vec{B}$. Therefore, despite of a terminological mismatch, which emerged historically, an analog of the vector of electric field strength $\vec{E}$ is the vector of magnetic induction $\vec{B}$, and an analog of electric induction $\vec{D}$ is the magnetic field strength $\vec{H}$.

In some sense, the vectors $\vec{D}$ and $\vec{H}$ are auxiliary to the vectors $\vec{E}$ and $\vec{B}$. Namely, in accordance with (6.32) and (6.44), $\vec{D} = \varepsilon_0 \vec{E}$ and $\vec{H} = \vec{B}/\mu_0$ in vacuum, i.e. the difference between the vectors $\vec{D}$ and $\vec{E}$ and between the vectors $\vec{H}$ and $\vec{B}$ is practically absent (in the Gauss system of units, $\varepsilon_0 = \mu_0 = 1$; therefore, $\vec{D} = \vec{E}$ and $\vec{H} = \vec{B}$ in vacuum). However, in a medium, $\vec{D} = \varepsilon_0 \vec{E} + \vec{P}$ and $\vec{H} = \vec{B}/\mu_0 + \vec{P}_m$. Therefore, for the system of Maxwell's equations (6.46)–(6.53) to be complete, we must know the dependences $\vec{P} = \vec{P}(\vec{E})$ and $\vec{P}_m = \vec{P}_m(\vec{B})$. In macroscopic electrodynamics, those dependences are determined on the basis of experimental data and theoretically substantiated in the framework of electron theory and quantum mechanics.

From experimental data, it follows that, in the case of weak electric fields in an isotropic medium, the polarization vector $\vec{P}$ is proportional to the field strength $\vec{E}$. This dependence is written in the form

$$\vec{P} = \varepsilon_0 \kappa \vec{E}, \tag{6.54}$$

where the coefficient κ is called the **dielectric susceptibility** of the medium. Then, dependence (6.32) takes the form

$$\vec{D} = \varepsilon_0 \vec{E} + \vec{P} = \varepsilon_0 (1 + \kappa) \vec{E} = \varepsilon_0 \varepsilon \vec{E}. \tag{6.55}$$

The quantity

$$\sigma = 1 + \kappa \tag{6.56}$$

is called the **dielectric permittivity** of the medium. In the case of anisotropic medium, for instance, in crystals,

$$\begin{aligned} P_i &= \sum_{j=1}^{3} \varepsilon_0 \kappa_{ij} E_j, && (6.57) \\ D_i &= \varepsilon_0 \sum_{j=1}^{3} \varepsilon_{ij} E_j, && (6.58) \\ \varepsilon_{ij} &= 1 + \kappa_{ij}, && (6.59) \end{aligned}$$

where κ_{ij} is the **tensor** (see Appendix D) **of dielectric susceptibility**, and ε_{ij} the **tensor of dielectric permittivity**.

As was already indicated, formulas (6.54)–(6.59) are applicable for weak electric fields that satisfy the condition

$$E \ll E_{at} \sim 10^{11} \text{ V/m},$$

where E_{at} is the strength of electric field created by the atomic nucleus in the atom. This condition is usually obeyed, but it can be violated in the case of laser radiation.

In most cases, the dependence $\vec{P}_m(\vec{H})$ rather than $\vec{P}_m(\vec{B})$ is used. If the magnetic field in an isotropic non-ferromagnetic medium is weak,

$$\begin{aligned} \vec{P}_m &= \kappa_m \vec{H}, && (6.60) \\ \vec{B} &= \mu_0 \vec{H} + \mu_0 \vec{P}_m = \mu_0 (1 + \kappa_m) \vec{H} = \mu_0 \mu \vec{H}. && (6.61) \end{aligned}$$

The coefficient κ_m is called the **magnetic susceptibility** and the quantity $\mu = 1 + \kappa_m$ the **magnetic permeability** of the medium. In an anisotropic medium,

$$P_{mi} = \mu_0 \sum_{j=1}^{3} \kappa_{mij} H_j, \tag{6.62}$$

$$B_i = \mu_0 \sum_{j=1}^{3} \mu_{ij} H_j, \tag{6.63}$$

$$\mu_{ij} = 1 + \kappa_{mij}, \tag{6.64}$$

where κ_{mij} and μ_{ij} are the tensors of magnetic susceptibility and magnetic permeability, respectively.

The relations

$$\vec{D} = \varepsilon \vec{E}, \tag{6.65}$$

$$\vec{B} = \mu\mu_0 \vec{H} \tag{6.66}$$

and

$$\vec{P} = \varepsilon_0 \kappa \vec{E}, \tag{6.67}$$

$$\vec{P}_m = \kappa_m \vec{H} \tag{6.68}$$

are called the **constitutive equations**; they supplement Maxwell equations (6.46)–(6.53). The constitutive equations also include the equation mentioned above

$$\vec{j} = \sigma \left(\vec{E} + \vec{E}^{ext} \right). \tag{6.69}$$

In the general case, the function $P(E)$ can be expanded in a power series of E,

$$P(E) = P(0) + \left(\frac{\partial P}{\partial E} \right)_0 E + \left(\frac{\partial^2 P}{\partial E^2} \right)_0 E^2 + \ldots. \tag{6.70}$$

In ordinary insulators, $P(0) = 0$. Then, if the field is weak, expansion (6.70) can be made up to the second summand, so that

$$P(E) = \left(\frac{\partial P}{\partial E} \right)_0 E = \varepsilon_0 \kappa E. \tag{6.71}$$

However, if the condition $E \ll E_{at}$ is not satisfied, it is necessary to hold some nonlinear summands as well,

$$P(E) = \varepsilon_0 \kappa E + \alpha E^2 + \beta E^3 + \dots . \tag{6.72}$$

The nonlinear terms in (6.72) are responsible for the nonlinear effects of interaction between radiation and the medium, generation of radiation with the frequencies 2ω, 3ω, $\dots$, frequency summation, etc. However, those effects can be observed only in the case if the condition $E \ll E_{at}$ is violated. Therefore, they will not be considered here.

6.6. Boundary conditions for the vectors of electromagnetic field in the medium

Besides constitutive equations (6.65)–(6.69), the system of differential Maxwell's equations (6.46)–(6.49) has to be supplemented with boundary conditions for the vectors $\vec{E}$, $\vec{D}$, $\vec{B}$, and $\vec{H}$ at the places where they are discontinuous and, therefore, the equations in the differential form are inapplicable (the differential form assumes that the infinitesimal change of variable corresponds to an infinitesimal change of function). At the same time, Maxwell's equations in the integral form, which generalize experimental facts, are applicable in all cases. Therefore, they are used to determine boundary conditions for the differential equations.

In sections 1.5, 2.9, and 3.4, we noticed that discontinuities for electromagnetic field vectors exist in the presence of surface charges and currents. In the medium, such charges and currents can arise at the interface between two different media. When crossing this interface, the field vectors can change in a jump-like manner. In the noted sections, we also considered how the boundary conditions can be found making use of Maxwell's equations in the integral form. Here, we will apply this procedure, without detailed calculations, to the electromagnetic field in a medium.

Using Fig. 1.6, Maxwell's equation (6.53), and speculations given in section 1.5, it is easy to be convinced that the boundary condition for the normal component of electromagnetic induction vector $\vec{D}$ across the charged interface between two media looks like

$$D_{2n} - D_{1n} = \sigma, \tag{6.73}$$

where $D_n = \vec{D}\vec{n}$ is the projection of vector $\vec{D}$ on the normal $\vec{n}$ to the interface directed from medium 1 to medium 2 (see Fig. 1.6), and σ is the surface density of free charges at the interface. In the medium, $\vec{D} = \varepsilon_0\vec{E} + \vec{P}$; therefore, condition (6.73) can be written as

$$E_{2n} - E_{1n} = \frac{\sigma}{\varepsilon_0} - \left(\frac{P_{2n}}{\varepsilon_0} - \frac{P_{1n}}{\sigma_0}\right). \tag{6.74}$$

According to (6.30), $\operatorname{div}\vec{P} = -\left\langle\rho_{bound}\right\rangle$ By its form, this equation is similar to Maxwell equation (6.49), which gives rise to the condition (6.73). Therefore, by analogy, it is possible to write

$$P_{2n} - P_{1n} = -\left\langle\sigma_{bound}\right\rangle. \tag{6.75}$$

Then, instead of (6.74), we obtain

$$E_{2n} - E_{1n} = \frac{\sigma}{\varepsilon_0} + \frac{\sigma^{(2)}_{bound} - \sigma^{(1)}_{bound}}{\varepsilon_0}, \tag{6.76}$$

where $\sigma^{(1)}_{bound}$ and $\sigma^{(2)}_{bound}$ are the average values of bound charge density on the surfaces of both media that create the interface.

The boundary condition for the tangential component of vector $\vec{E}$ follows from Maxwell's equation (6.51). On the basis of speculations presented in section 1.5 and taking into account that (see Fig. 1.7) the flux of magnetic induction vector $\vec{B}$ through the contour L tends to zero if the contour height does, we obtain

$$E_{2\tau} - E_{1\tau} = 0. \tag{6.77}$$

Now, let us determine boundary conditions for the vectors $\vec{B}$ and $\vec{H}$. Equation (6.53) was used to obtain boundary condition (6.73). Therefore, on the basis of equation (6.52) and taking into account the absence of magnetic charges, it is possible to write

$$B_{2n} - B_{1n} = 0. \tag{6.78}$$

In order to establish the boundary condition for the tangential components of vector $\vec{H}$, let us use Maxwell's equation (6.50) and apply the reasoning given in section 2.9. If the contour height in Fig. 2.9 tends to zero, the flux of vector $\vec{D}$ through it also tends to zero. Therefore,

$$H_{2\tau} - H_{1\tau} = i_N, \tag{6.79}$$

where i_N is the projection of surface current density on the normal to the contour.

Hence, the following boundary conditions are satisfied across the interface between two media:

$$\begin{aligned} D_{2n} - D_{1n} &= \sigma, & (6.80)\\ E_{2\tau} - E_{1\tau} &= 0, & (6.81)\\ B_{2n} - B_{1n} &= 0, & (6.82)\\ H_{2\tau} - H_{1\tau} &= i_N. & (6.83) \end{aligned}$$

6.7. Energy conservation law for the electromagnetic field in the medium

As was noticed in section 3.5, the electromagnetic field in vacuum has an energy and a momentum. Let us consider the issue of electromagnetic field energy in a medium consisting of conductors and insulators. The medium is assumed to be isotropic, non-ferromagnetic, non-absorbing, and motionless.

The electromagnetic field acts on the electric charge in a unit volume with the Lorentz force

$$\vec{f} = \rho\vec{E} + \rho\left[\vec{v}\vec{B}\right].$$

The corresponding work done in a unit volume per unit time equals

$$\frac{dA}{dt} = \rho\vec{v}\vec{E} + \rho\vec{v}\left[\vec{v}\vec{B}\right].$$

Taking into account that $\rho\vec{v} = \vec{j}$ and $\vec{v}\left[\vec{v}\vec{B}\right] = 0$, we obtain

$$\frac{dA}{dt} = \vec{j}\vec{E}. \tag{6.84}$$

Then, the work done by the electromagnetic field on the charges in the volume V per unit time is determined by the formula

$$\frac{dA}{dt} = \int \vec{j}\vec{E}\, dV. \tag{6.85}$$

Maxwell's equation (6.46) leads to

$$\vec{j} = \operatorname{rot} \vec{H} - \frac{\partial \vec{D}}{\partial t}.$$

Therefore, instead of (6.85), we may write

$$\frac{dA}{dt} = \int \left\{ \vec{E} \operatorname{rot} \vec{H} - \vec{E}\frac{\partial \vec{D}}{\partial t} \right\} dV. \tag{6.86}$$

On the basis of equation (6.47), we obtain the expression equal to zero,

$$\vec{H} \operatorname{rot} \vec{E} + \vec{H}\frac{\partial \vec{B}}{\partial t} = 0.$$

Taking it into account, we obtain

$$\frac{dA}{dt} = \int \left\{ \vec{E} \operatorname{rot} \vec{H} - \vec{H} \operatorname{rot} \vec{E} \right\} dV - \int \left\{ \vec{E}\frac{\partial \vec{D}}{\partial t} + \vec{H}\frac{\partial \vec{B}}{\partial t} \right\} dV. \tag{6.87}$$

If the relationship between the vectors $\vec{D}$ and $\vec{E}$ is determined by formula (6.65) and between the vectors $\vec{B}$ and $\vec{H}$ by formula (6.66), it is possible to write

$$\vec{E}\frac{\partial \vec{D}}{\partial t} + \vec{H}\frac{\partial \vec{B}}{\partial t} = \frac{\partial w}{\partial t}, \tag{6.88}$$

$$w = \frac{1}{2}\left(\vec{E}\vec{D} + \vec{H}\vec{B}\right) = \frac{1}{2}\varepsilon\varepsilon_0 E^2 + \frac{1}{2}\mu\mu_0 H^2, \tag{6.89}$$

$$\vec{E} \operatorname{rot} \vec{H} - \vec{H} \operatorname{rot} \vec{E} = -\operatorname{div}\left[\vec{E}\vec{H}\right] = -\frac{1}{\mu\mu_0}\operatorname{div}\left[\vec{E}\vec{B}\right] = -\operatorname{div}\vec{\Pi}. \tag{6.90}$$

Then, formula (6.87) is written as

$$\frac{dA}{dt} = -\oint \operatorname{div}\vec{\Pi}\, dV - \frac{\partial}{\partial t}\int w\, dV. \tag{6.91}$$

Applying Gauss's theorem to the first term, we obtain

$$\frac{dA}{dt} = -\oint \vec{\Pi}\, d\vec{S} - \frac{\partial}{\partial t}\int w\, dV. \tag{6.92}$$

If the system consisting of the electromagnetic field and charges is closed, i.e. if it does not interact with the environment, the fields on the surface that surrounds it equal zero, $\vec{E} = \vec{H} = 0$, so that the vector $\vec{\Pi} = \left[\vec{E}\vec{H}\right]$ (see (6.90)) also equals zero. Hence,

$$\frac{dA}{dt} = -\frac{\partial}{\partial t}\int w\, dV = -\frac{dW}{dt}. \tag{6.93}$$

This means that the work in the system is done owing to the reduction of the quantity

$$W = \int w\, dV, \tag{6.94}$$

which therefore should be adopted as the **energy of electromagnetic field in** the **medium**. Then, the quantity

$$w = \frac{1}{2}\vec{E}\vec{D} + \frac{1}{2}\vec{H}\vec{B} \tag{6.95}$$

is the **energy density of electromagnetic field in the medium**. The first summand in (6.95), in accordance with (1.89), is the **energy density of electric field** component and the second one the **energy density of magnetic field** component of the total electromagnetic field.

If the system is not closed, but does not contain conductors, the work done by the field, in accordance with (6.85), equals zero ($\vec{j} = 0$). Hence, from (6.92) and (6.94), it follows that

$$\oint_S \vec{\Pi}\, d\vec{S} = \frac{\partial W}{\partial t}. \tag{6.96}$$

In other words, the decrease of the field energy W occurs owing to the field energy loss through the surface S. Therefore, the vector

$$\vec{\Pi} = \left[\vec{E}\vec{H}\right] \tag{6.97}$$

has the meaning of Poynting vector (see section 3.5), i.e. a vector, the absolute value of which is equal to the energy of electromagnetic field transferred per unit time through a section of unit area oriented perpendicularly to the vector $\vec{\Pi}$ direction, the direction of energy propagation.

The momentum of electromagnetic field in the medium is not defined unambiguously. Therefore, it will not be considered.

6.8. Potentials of electromagnetic field in the medium

The electromagnetic potentials in a medium are determined in the same way as in vacuum (see section 3.7), i.e. on the basis of Maxwell's equations (6.46)–(6.49). If the electromagnetic field is weal, $\vec{D} = \varepsilon\varepsilon_0\vec{E}$ and $\vec{B} = \mu\mu_0\vec{H}$ (see section 6.6). Then, Maxwell's equation (6.46)–(6.49) for the electromagnetic field in a medium can be written in the form

$$\begin{aligned} \operatorname{rot}\vec{B} &= \mu\mu_0\vec{j} + \varepsilon\varepsilon_0\mu\mu_0\frac{\partial\vec{E}}{\partial t}, && (6.98)\\ \operatorname{rot}\vec{E} &= -\frac{\partial\vec{B}}{\partial t}, && (6.99)\\ \operatorname{div}\vec{B} &= 0, && (6.100)\\ \operatorname{div}\vec{E} &= \frac{\rho}{\varepsilon\varepsilon_0}. && (6.101) \end{aligned}$$

Those equations differ from Maxwell's equations for the electromagnetic field in vacuum (3.36)–(3.39) by the substitutions

$$\begin{gathered} \varepsilon_0 \to \varepsilon\varepsilon_0, \quad \mu_0 \to \mu\mu_0, \\ c = \frac{1}{\sqrt{\varepsilon_0\mu_0}} \to \frac{1}{\sqrt{\varepsilon\varepsilon_0\mu\mu_0}} = \frac{c}{\sqrt{\varepsilon\mu}}. \end{gathered} \tag{6.102}$$

Therefore, the definitions of electromagnetic potentials φ and $\vec{A}$, the formulation of d'Alembert equations that those potentials satisfy, and their solutions in the medium can be done using the corresponding formulas in sections 3.7 and 5.1, in which the indicated substitutions should be made.

The scalar, $\varphi(\vec{r},t)$, and vector, $\vec{A}(\vec{r},t)$, potentials are defined using formulas (3.80) and (3.82),

$$\begin{aligned}\vec{E}(\vec{r},t) &= -\nabla\varphi(\vec{r},t) - \frac{\partial \vec{A}(\vec{r},t)}{\partial t}, &(6.103)\\ \vec{B}(\vec{r},t) &= \operatorname{rot}\vec{A}(\vec{r},t). &(6.104)\end{aligned}$$

On the basis of equations (3.90), (3.91), and (6.102), d'Alembert equations for the electromagnetic potentials in a medium take the form

$$\begin{aligned}\Delta\varphi - \frac{\varepsilon\mu}{c^2}\frac{\partial^2\varphi}{\partial t^2} &= -\frac{\rho}{\varepsilon\varepsilon_0}, &(6.105)\\ \Delta\vec{A} - \frac{\varepsilon\mu}{c^2}\frac{\partial^2\vec{A}}{\partial t^2} &= -\mu\mu_0\vec{j}. &(6.106)\end{aligned}$$

The solutions of d'Alembert equations in vacuum are given in section 5.1 by formulas (5.25) and (5.26). Therefore, let us write the solutions of d'Alembert equations (6.105) and (6.106) as

$$\begin{aligned}\varphi(\vec{r},t) &= \frac{1}{4\pi\varepsilon\varepsilon_0}\int\frac{\rho\left(\vec{r}\,',t-\frac{|\vec{r}-\vec{r}\,'|}{c}\sqrt{\varepsilon\mu}\right)}{|\vec{r}-\vec{r}\,'|}dV', &(6.107)\\ \vec{A}(\vec{r},t) &= \frac{\mu\mu_0}{4\pi}\int\frac{\vec{j}\left(\vec{r}\,',t-\frac{|\vec{r}-\vec{r}\,'|}{c}\sqrt{\varepsilon\mu}\right)}{|\vec{r}-\vec{r}\,'|}dV' &(6.108)\end{aligned}$$

From (6.102) and (6.105)–(6.108), it follows that the electromagnetic perturbation propagates in the medium at the velocity

$$v = \frac{c}{\sqrt{\varepsilon\mu}}. \qquad (6.109)$$

Notice that the electromagnetic field in the medium, as well as in vacuum, is invariant with respect to the gauge transformation of potentials

$$\vec{A}\to\vec{A}+\nabla\psi, \quad \varphi\to\varphi-\frac{\partial\psi}{\partial t}. \qquad (6.110)$$

The corresponding proof is identical to that given in section 3.7.

7

ELECTROMAGNETIC WAVES IN THE MEDIUM

7.1. Plane electromagnetic waves in dielectrics

In section 4, a possibility for electromagnetic waves to exist in vacuum was demonstrated and their main properties were determined. Now, let us consider the features of electromagnetic waves in a medium.

Let the space be filled with a uniform isotropic insulator ($\varepsilon = \text{const}$, $\mu = \text{const}$, $\sigma = 0$). We also suppose that free charges are absent ($\rho = 0$), and the relationships between the vectors $\vec{E}$ and $\vec{D}$ and between the vectors $\vec{H}$ and $\vec{B}$ are $\vec{D} = \varepsilon\varepsilon_0\vec{E}$ and $\vec{B} = \mu\mu_0\vec{H}$, respectively. Then the system of Maxwell's equations (6.46)–(6.49) has the form

$$\begin{aligned} \operatorname{rot}\vec{H} &= \varepsilon\varepsilon_0\frac{\partial\vec{E}}{\partial t}, && (7.1)\\ \operatorname{rot}\vec{E} &= -\mu\mu_0\frac{\partial\vec{H}}{\partial t}, && (7.2)\\ \operatorname{div}\vec{H} &= 0, && (7.3)\\ \operatorname{div}\vec{E} &= 0. && (7.4) \end{aligned}$$

Applying the rot-operator to equation (7.2), we obtain

$$\operatorname{rot}\operatorname{rot}\vec{E} = \operatorname{grad}\operatorname{div}\vec{E} - \Delta\vec{E} = -\mu\mu_0\frac{\partial}{\partial t}\operatorname{rot}\vec{H}.$$

Making allowance for (7.1) and (7.4), we obtain the equation

$$\Delta\vec{E} - \varepsilon\varepsilon_0\mu\mu_0\frac{\partial^2\vec{E}}{\partial t^2} = 0. \tag{7.5}$$

Analogously, applying the rot-operator to equation (7.1) and making allowance for (7.2) and (7.3), we obtain

$$\Delta\vec{H} - \varepsilon\varepsilon_0\mu\mu_0\frac{\partial^2\vec{H}}{\partial t^2} = 0. \tag{7.6}$$

Equations (7.5) and (7.6) are wave equations (see section 4.1). They describe the propagation of electromagnetic waves in a dielectric medium at the velocity

$$v = \frac{1}{\sqrt{\varepsilon\varepsilon_0\mu\mu_0}} = \frac{c}{\sqrt{\varepsilon\mu}}, \tag{7.7}$$

where $c = 3\times 10^8$ m/s is the velocity of electromagnetic wave propagation in vacuum (see section 4.1).

In sections 4.2 and 4.3, the solutions of wave equations in the form of plane and spherical electromagnetic waves were obtained. The solutions of equations (7.5) and (7.6) can be found in the same way. In the case of plane electromagnetic waves, the solutions of equations (7.5) and (7.6) in the complex form look like (see (4.39) and (4.40))

$$\begin{aligned} \vec{E}(\vec{r},t) &= \vec{E}_0(\vec{k})\mathrm{e}^{i(\vec{k}\vec{r}-\omega t)}, && (7.8)\\ \vec{H}(\vec{r},t) &= \vec{H}_0(\vec{k})\mathrm{e}^{i(\vec{k}\vec{r}-\omega t)}, && (7.9) \end{aligned}$$

where $\vec{k}$ is the wave vector, the direction of which in the isotropic insulator coincides with that of wave propagation and the absolute value (the **wave number**) is coupled with the frequency ω by the relation (cf. (4.34))

$$k = \frac{\omega}{v} = \frac{\omega}{c}\sqrt{\varepsilon\mu}. \tag{7.10}$$

According to section 4.4, the velocity

$$v = \frac{\omega}{k} = \frac{c}{\sqrt{\varepsilon\mu}} \tag{7.11}$$

is the phase velocity, i.e. the propagation velocity for a certain wave phase.

Substituting (7.8) and (7.9) into Maxwell's equation (7.1)–(7.4), we obtain the following relations:

$$\begin{aligned} \left[\vec{k}\vec{H}\right] &= -\varepsilon\varepsilon_0\omega\vec{E}, && (7.12)\\ \left[\vec{k}\vec{E}\right] &= \mu\mu_0\varepsilon\vec{H}, && (7.13)\\ \vec{k}\vec{H} &= 0, && (7.14)\\ \vec{k}\vec{E} &= 0, && (7.15) \end{aligned}$$

whence it follows that plane electromagnetic waves are transverse in an isotropic insulator, i.e. $\vec{E} \perp \vec{k}$, $\vec{H} \perp \vec{k}$, and the vectors $\vec{E}$ and $\vec{H}$ are also mutually perpendicular ($\vec{E} \perp \vec{H}$). Besides, from (7.12) and (7.13), we obtain the relationship between the absolute values of vectors $\vec{E}$ and $\vec{H}$ (cf. (4.44)),

$$\sqrt{\varepsilon\varepsilon_0}E = \sqrt{\mu\mu_0}H. \tag{7.16}$$

The energy density in the electromagnetic wave in the medium is determined by the formula (6.95),

$$w = \frac{1}{2}\varepsilon\varepsilon_0 E^2 + \frac{1}{2}\mu\mu_0 H^2,$$

which taking into account (7.16) can be rewritten as

$$w = \varepsilon\varepsilon_0 E^2 = \mu\mu_0 H^2. \tag{7.17}$$

The electric and magnetic field strength (7.8) and (7.9) are complex-valued quantities, whereas the energy density (7.17) is a square-law function of the field. Therefore, substituting (7.8) and (7.9) into (7.17), me must take their real parts

$$\begin{aligned} \mathrm{Re}\left\{\vec{E}_0(\vec{k})\mathrm{e}^{i(\vec{k}\vec{r}-\omega t)}\right\} &= \vec{E}_0(\vec{k})\cos\left(\vec{k}\vec{r}-\omega t\right), \\ \mathrm{Re}\left\{\vec{H}_0(\vec{k})\mathrm{e}^{i(\vec{k}\vec{r}-\omega t)}\right\} &= \vec{H}_0(\vec{k})\cos\left(\vec{k}\vec{r}-\omega t\right). \end{aligned} \tag{7.18}$$

Finally, we obtain

$$w = \varepsilon\varepsilon_0 E_0^2 \cos^2\left(\vec{k}\vec{r}-\omega t\right) = \mu\mu_0 H_0^2 \cos^2\left(\vec{k}\vec{r}-\omega t\right). \tag{7.19}$$

As a rule, we are interested in the average value of energy density rather than the instant one, i.e. in

$$\langle w\rangle = \frac{1}{2}\varepsilon\varepsilon_0 E_0^2 = \frac{1}{2}\mu\mu_0 H_0^2. \tag{7.20}$$

The propagation of a plane monochromatic electromagnetic wave in a dielectric medium is associated with the transfer of energy, which id determined using Poynting vector (6.97),

$$\vec{\Pi} = \left[\vec{E}\vec{H}\right],$$

the absolute value of which equals the energy transmitted per unit time through a unit area of the surface oriented perpendicularly to the direction of wave propagation,

$$\Pi = EH = \sqrt{\frac{\varepsilon\varepsilon_0}{\mu\mu_0}}E^2 = \sqrt{\frac{\varepsilon\varepsilon_0}{\mu\mu_0}}E_0^2 \cos^2\left(\vec{k}\vec{r} - \omega t\right) \tag{7.21}$$

The average value of the absolute value of Poynting vector (7.21), i.e. the density of energy flux equals

$$\langle\Pi\rangle = \frac{1}{2}\sqrt{\frac{\varepsilon\varepsilon_0}{\mu\mu_0}}E^2 = \sqrt{\frac{\varepsilon\varepsilon_0}{\mu\mu_0}}\frac{\langle w\rangle}{\varepsilon\varepsilon_0} = v\,\langle w\rangle\,. \tag{7.22}$$

7.2. Reflection and refraction of electromagnetic waves

In section 6.6, the boundary conditions for the electromagnetic field vectors at the medium interface were formulated. Let us apply them to the interface between two dielectric media with different dielectric and magnetic permittivities. The interface is assumed to be plane and oriented perpendicularly to the axis z (Fig. 7.1).

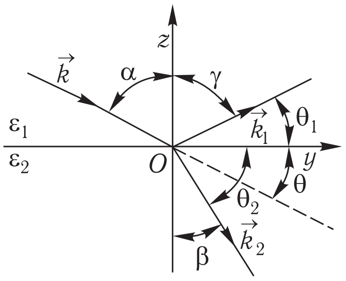

Fig. 7.1.

Suppose that a plane monochromatic electromagnetic wave with

$$\vec{E} = \vec{E}_0 \mathrm{e}^{i(\vec{k}\vec{r}-\omega t)}, \quad \vec{H} = \vec{H}_0 \mathrm{e}^{i(\vec{k}\vec{r}-\omega t)} \tag{7.23}$$

falls on the interface between media 1 and 2. The wave vector $\vec{k}$ of the wave is in the plane yz (the figure plane in Fig. 7.1), which is called the **plane of incidence**. At the medium interface, the incident electromagnetic wave becomes partially reflected into medium 1 and partially transmitted (refracted) into medium 2. The reflected wave can be written in the form

$$\vec{E}_1 = \vec{E}_{10} \mathrm{e}^{i(\vec{k}_1\vec{r}-\omega_1 t)}, \quad \vec{H}_1 = \vec{H}_{10} \mathrm{e}^{i(\vec{k}_1\vec{r}-\omega_1 t)}, \tag{7.24}$$

and the reflected one in the form

$$\vec{E}_2 = \vec{E}_{20} e^{i(\vec{k}_2 \vec{r} - \omega_2 t)}, \quad \vec{H}_2 = \vec{H}_{20} e^{i(\vec{k}_2 \vec{r} - \omega_2 t)}. \tag{7.25}$$

Let us use boundary conditions (6.81) and (6.83), i.e. the equality of tangential components of vectors $\vec{E}$ and $\vec{H}$. In our case, those conditions look like

$$E_\tau + E_{1\tau} = E_{2\tau}, \tag{7.26}$$

$$H_\tau + H_{1\tau} = H_{2\tau}. \tag{7.27}$$

Taking into account (7.23)–(7.25), conditions (7.26) and (7.27) take the form

$$E_{0\tau} e^{i(\vec{k}\vec{r} - \omega t)} + E_{10\tau} e^{i(\vec{k}_1 \vec{r} - \omega_1 t)} = E_{20\tau} e^{i(\vec{k}_2 \vec{r} - \omega_2 t)}, \tag{7.28}$$

$$H_{0\tau} e^{i(\vec{k}\vec{r} - \omega t)} + H_{10\tau} e^{i(\vec{k}_1 \vec{r} - \omega_1 t)} = H_{20\tau} e^{i(\vec{k}_2 \vec{r} - \omega_2 t)}. \tag{7.29}$$

Those conditions must be satisfied at any interface point at any time moment. Therefore, the exponents in the exponential functions in formulas (7.28) and (7.29) must coincide at $z = 0$. Then,

$$\omega = \omega_1 = \omega_2, \tag{7.30}$$

$$\vec{k}\vec{r} = \vec{k}_1 \vec{r} = \vec{k}_2 \vec{r}. \tag{7.31}$$

From relation (7.30), it follows that the frequency of reflected and refracted electromagnetic waves does not change. Relation (7.31) at $z = 0$ looks like

$$k_x x + k_y y = k_{1x} x + k_{1y} y = k_{2x} x + k_{2y} y. \tag{7.32}$$

This relation must be satisfied at arbitrary values of coordinates x and y. Therefore, the corresponding coefficients at x and y must be equal,

$$k_x = k_{1x} = k_{2x}, \tag{7.33}$$

$$k_y = k_{1y} = k_{2y}. \tag{7.34}$$

Since $k_x = 0$ (see Fig. 7.1), $k_{1x} = k_{2x} = 0$ as well, i.e. the vectors $\vec{k}_1$ and $\vec{k}_2$ also lie in the incidence plane xz. Whence, the **first law of reflection and refraction** follows:

- the wave vectors (the propagation directions) of the incident, reflected, and refracted waves, as well as the normal to the medium interface, lie in the same plane.

Formula (7.34), taking into account Fig. 7.1, can be rewritten in the form

$$k\cos\theta = k_1\cos\theta_1 = k_2\cos\theta_2. \tag{7.35}$$

Then, on the basis (7.10), we have

$$\begin{aligned} k &= k_1 = \frac{\omega}{c}\sqrt{\varepsilon_1\mu_1}, \\ k_2 &= \frac{\omega}{c}\sqrt{\varepsilon_2\mu_2}. \end{aligned}$$

Moreover, is follows from Fig. 7.1 that $\theta = \frac{\pi}{2} - \alpha$, $\theta_1 = \frac{\pi}{2} - \gamma$, and $\theta_2 = \frac{\pi}{2} - \beta$. Therefore, relation (7.35) can be written as

$$\frac{\omega}{c}\sqrt{\varepsilon_1\mu_1}\sin\alpha = \frac{\omega}{c}\sqrt{\varepsilon_1\mu_1}\sin\gamma = \frac{\omega}{c}\sqrt{\varepsilon_2\mu_2}\sin\beta. \tag{7.36}$$

Hence,

$$\alpha = \gamma, \tag{7.37}$$

$$\frac{\sin\alpha}{\sin\gamma} = \sqrt{\frac{\varepsilon_2\mu_2}{\varepsilon_1\mu_1}} = n_{12}, \tag{7.38}$$

i.e. (the **second law of reflection**)

- the angle of incidence equals the angle of reflection,

and (the **second law of refraction**)

- for the given two media, the ratio between the sines of the angle of incidence and the angle of refraction is a constant.

This constant is called the **relative refractive index** n_{12} of medium 2 with respect to medium 1.

For insulators, which are para- and diamagnets, the magnetic permeability μ only slightly differs from unity. Then,

$$n_{12} = \sqrt{\frac{\varepsilon_2}{\varepsilon_1}}. \tag{7.39}$$

If medium 1 is vacuum ($\varepsilon_1 = 1$), i.e. if the electromagnetic wave falls from vacuum on the surface of dielectric medium with the dielectric permittivity ε_2, the quantity

$$n_2 = \sqrt{\varepsilon_2} \tag{7.40}$$

is called the **absolute refractive index of the dielectric medium**. Hence, the relative refractive index n_{12} can be written in the form

$$n_{12} = \frac{n_2}{n_1}, \tag{7.41}$$

where n_1 is the absolute refractive index of medium 1 (for vacuum, $n_1 = 1$) and n_2 the absolute refractive index of medium 2.

7.3. Fresnel formulas

The average absolute values of Poynting vector for the incident, reflected, and refracted waves (see section 7.2) under the condition $\mu = 1$, in accordance with (7.22), equal

$$\langle \Pi \rangle = v_1 \langle w \rangle = \frac{1}{2}\varepsilon_0 c \sqrt{\varepsilon_1} E_0^2, \tag{7.42}$$

$$\langle \Pi_1 \rangle = v_1 \langle w \rangle = \frac{1}{2}\varepsilon_0 c \sqrt{\varepsilon_1} E_{10}^2, \tag{7.43}$$

$$\langle \Pi_2 \rangle = v_2 \langle w \rangle = \frac{1}{2}\varepsilon_0 c \sqrt{\varepsilon_2} E_{20}^2. \tag{7.44}$$

The quantity

$$R = \frac{\langle \Pi_1 \rangle}{\langle \Pi \rangle} = \frac{E_{10}^2}{E_0^2}, \tag{7.45}$$

which evaluates what fraction of the incident wave energy is contained in the reflected wave is called the **reflection coefficient** (or the **reflectance**). At the same time, the quantity (see Fig. 7.1)

$$D = \frac{\langle \Pi_2 \rangle \cos\beta}{\langle \Pi \rangle \cos\alpha} = \sqrt{\frac{\varepsilon_2}{\varepsilon_1}} \frac{E_{20}^2}{E_0^2} \frac{\cos\beta}{\cos\alpha} = n_{12} \frac{E_{20}^2}{E_0^2} \frac{\cos\beta}{\cos\alpha} \tag{7.46}$$

is called the **transmission coefficient** (or the **transmittance)**. It equals the fraction of the incident wave energy that is contained in the refracted wave. The coefficient $\cos\beta/\cos\alpha$ makes allowance for the change in the wave cross-section after the refraction. If the dielectric meduim does not absorb the wave energy, the energy conservation law brings about the relation

$$R + D = 1. \tag{7.47}$$

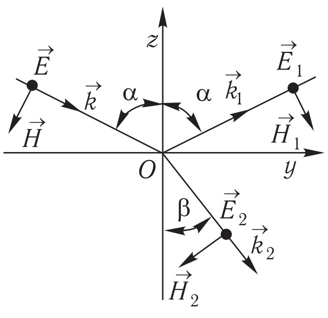

Fig. 7.2.

In order to determine the coefficients R and D, we must express the amplitudes E_{10} and E_{20} in terms of the incident wave amplitude E_0. Let the incident, reflected, and refracted waves be linearly polarized with the corresponding vectors of electric field strength being oriented perpendicularly to the plane of incidence, i.e. directed along the axis x (see Fig. 7.2). Then, the vectors of magnetic field strength must lie in the plane of incidence (see Fig. 7.2, where the vectors of electric field strength are directed towards the reader and are represented by dots). In this geometry, boundary conditions (7.28) and (7.29) can be written in the form of their projections on the axes x and y, respectively, taking into account the equality of exponential functions (see section 7.2), as follows:

$$E_0^{\perp} + E_{10}^{\perp} = E_{20}^{\perp}, \tag{7.48}$$

$$-H_0^{\parallel}\cos\alpha + H_{10}^{\parallel}\cos\alpha = -H_{20}^{\parallel}\cos\beta. \tag{7.49}$$

The superscript $\perp$ means that the electric field strength is directed perpendicularly to the plane of incidence, and the superscript $\parallel$, that the magnetic field strength lies in the plane of incidence. On the basis of (7.16), it is possible to write

$$H_0^{\parallel} = \sqrt{\frac{\varepsilon_1\varepsilon_0}{\mu_0}}E_0^{\perp}, \tag{7.50}$$

$$H_{10}^{\parallel} = \sqrt{\frac{\varepsilon_1\varepsilon_0}{\mu_0}}E_{10}^{\perp}, \tag{7.51}$$

$$H_{20}^{\parallel} = \sqrt{\frac{\varepsilon_2\varepsilon_0}{\mu_0}}E_{20}^{\perp}. \tag{7.52}$$

Substituting (7.50)–(7.52) into (7.49), we obtain the system of two equations (7.48) and (7.49) with the unknown E_{10} and E_{20},

$$\begin{aligned} E_0^{\perp} + E_{10}^{\perp} &= E_{20}^{\perp}, && (7.53)\\ \sqrt{\varepsilon_1} E_0^{\perp} \cos\alpha - \sqrt{\varepsilon_1} E_{10}^{\perp} \cos\alpha &= \sqrt{\varepsilon_2} E_{20}^{\perp} \cos\beta. && (7.54) \end{aligned}$$

The solutions of this system of equations look like

$$E_{10}^{\perp} = \frac{\sin(\beta - \alpha)}{\sin(\beta + \alpha)} E_0^{\perp}, \qquad (7.55)$$

$$E_{20}^{\perp} = \frac{2 \sin\beta \cos\alpha}{\sin(\beta + \alpha)} E_0^{\perp}. \qquad (7.56)$$

Now, let us consider the case when the vector of electric field strength lies in the plane of incidence, and the vector of magnetic field strength is directed along the axis x (Fig. 7.3). In this geometry, boundary conditions (7.28) and (7.29) can be written as

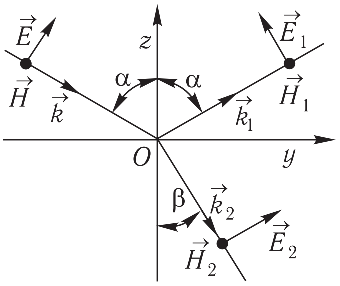

Fig. 7.3.

$$\begin{aligned} H_0^{\perp} + H_{10}^{\perp} &= H_{20}^{\perp}, && (7.57)\\ E_0^{\parallel} \cos\alpha - E_{10}^{\parallel} \cos\alpha &= E_{20}^{\parallel} \cos\beta. && (7.58) \end{aligned}$$

or, taking into account (7.16), as

$$\begin{aligned} \sqrt{\varepsilon_1} E_0^{\parallel} + \sqrt{\varepsilon_1} E_{10}^{\parallel} &= \sqrt{\varepsilon_2} E_{20}^{\parallel}, && (7.59)\\ E_0^{\parallel} \cos\alpha - E_{10}^{\parallel} \cos\alpha &= E_{20}^{\parallel} \cos\beta. && (7.60) \end{aligned}$$

By solving the system of equations (7.59) and (7.60) with the unknown $E_{10}^{\parallel}$ and $E_{20}^{\parallel}$, we obtain

$$\begin{aligned} E_{10}^{\parallel} &= \frac{\tan(\alpha - \beta)}{\tan(\alpha + \beta)} E_0^{\parallel}, && (7.61)\\ E_{20}^{\parallel} &= \frac{2 \cos\alpha \sin\beta}{\sin\alpha \cos\alpha + \sin\beta \cos\beta} E_0^{\parallel} = \frac{2 \cos\alpha \sin\beta}{\frac{1}{2}(\sin 2\alpha + \sin 2\beta)} E_0^{\parallel} \\ &= \frac{2 \cos\alpha \sin\beta}{\sin(\alpha + \beta) \cos(\alpha - \beta)} E_0^{\parallel}. && (7.62) \end{aligned}$$

Formulas (7.55), (7.56) and (7.61), (7.62), which couple the amplitudes of electric field strength in the reflected and refracted waves with the amplitude of electric field strength in the incident wave are called the **Fresnel formulas**.

With the help of Fresnel formulas (7.55), (7.56), (7.61), and (7.62), as well as formulas (7.45) and (7.46), we can find the reflection, R, and transmission, D, coefficients for linearly polarizable waves,

$$\begin{aligned}
R_{\perp} &= \left(\frac{E_{10}^{\perp}}{E_0^{\perp}}\right)^2 = \frac{\sin^2(\beta-\alpha)}{\sin^2(\beta+\alpha)}, && (7.63)\\
R_{\parallel} &= \left(\frac{E_{10}^{\parallel}}{E_0^{\parallel}}\right)^2 = \frac{\tan^2(\beta-\alpha)}{\tan^2(\beta+\alpha)}, && (7.64)\\
D_{\perp} &= \frac{\sin\alpha\cos\beta}{\sin\beta\cos\alpha}\left(\frac{E_{20}^{\perp}}{E_0^{\perp}}\right)^2 = \frac{\sin 2\alpha\sin 2\beta}{\sin^2(\beta+\alpha)}, && (7.65)\\
D_{\parallel} &= \frac{\sin\alpha\cos\beta}{\sin\beta\cos\alpha}\left(\frac{E_{20}^{\parallel}}{E_0^{\parallel}}\right)^2 = \frac{\sin 2\alpha\sin 2\beta}{\sin^2(\alpha+\beta)\cos^2(\alpha-\beta)}. && (7.66)
\end{aligned}$$

It is easy to be convinced that, in accordance with

$$\begin{aligned}
R_{\perp}+D_{\perp} &= 1,\\
R_{\parallel}+D_{\parallel} &= 1.
\end{aligned}$$

Let the vector $\vec{E}$ of the incident wave form the angle φ with the plane of incidence. Then, $E_0^{\perp} = E_0\sin\varphi$ and $E_0^{\parallel} = E_0\cos\varphi$. Therefore, on the basis of (7.45), it is possible to write

$$\begin{aligned}
\left(E_{10}^{\perp}\right)^2 &= R_{\perp}\left(E_0^{\perp}\right)^2 = R_{\perp}E_0^2\sin^2\varphi,\\
\left(E_{10}^{\parallel}\right)^2 &= R_{\parallel}\left(E_0^{\parallel}\right)^2 = R_{\parallel}E_0^2\cos^2\varphi.
\end{aligned}$$

In this case, reflection coefficient (7.45) takes the form

$$R_{\varphi} = \frac{E_{10}^2}{E_0^2} = \frac{\left(E_{10}^{\perp}\right)^2 + \left(E_{10}^{\parallel}\right)^2}{E_0^2} = R_{\perp}\sin^2\varphi + R_{\parallel}\cos^2\varphi. \qquad (7.67)$$

If the incident wave is not polarized, all directions of vector $\vec{E}$ in the plane perpendicular to the wave vector $\vec{k}$, i.e. all values of angle φ, are equally probable (see section 4.4). Hence, $\langle\cos^2\varphi\rangle = \langle\sin^2\varphi\rangle = 1/2$. Therefore, the coefficient of non-polarized electromagnetic wave reflection from the interface between two insulators equals

$$R = \frac{1}{2}\left(R_\perp + R_\parallel\right) = \frac{1}{2}\left\{\frac{\sin^2(\beta-\alpha)}{\sin^2(\beta+\alpha)} + \frac{\tan^2(\beta-\alpha)}{\tan^2(\beta+\alpha)}\right\}. \tag{7.68}$$

In order to analyze the results obtained, let us consider some specific cases. If the electromagnetic wave falls normally on the interface between two insulators, $\alpha = \beta = 0$. However, when substituting those values of incidence and refraction angles into formulas (7.63)–(7.66), all reflection and transmission coefficients become uncertain like $0/0$. To eliminate this uncertainty, we should take into account that, at $\alpha \ll 1$, we have $\sin\alpha \approx \alpha$ and $\tan\alpha \approx \alpha$. Therefore, under the condition $\alpha \to 0$ and $\beta \to 0$, formula (7.63) takes the form

$$R_\perp \approx \frac{(\beta-\alpha)^2}{(\beta+\alpha)^2} = \frac{\left(1-\dfrac{\alpha}{\beta}\right)^2}{\left(1+\dfrac{\alpha}{\beta}\right)^2} = \frac{\left(1-n_{12}\right)^2}{\left(1+n_{12}\right)^2}. \tag{7.69}$$

Here, we took into account that, in the case concerned,

$$n_{12} = \frac{\sin\alpha}{\sin\beta} \approx \frac{\alpha}{\beta}. \tag{7.70}$$

In the same way, we can find that, for the normal incidence,

$$R_\parallel = R_\perp = \frac{\left(1-n_{12}\right)^2}{\left(1+n_{12}\right)^2}, \tag{7.71}$$

$$D_\parallel = D_\perp = \frac{4n_{12}}{\left(1+n_{12}\right)^2}. \tag{7.72}$$

There is an important case of non-polarized electromagnetic wave incidence on the interface separating two insulators, when

$$\alpha + \beta = \frac{\pi}{2}. \tag{7.73}$$

In this case, $\tan(\alpha+\beta)=\infty$. Therefore, in accordance with (7.64), $R_\parallel = 0$, i.e. the reflected wave is linearly polarized, with the vector $\vec{E}$ being perpendicular to the plane of incidence (the **Brewster law**). The angle of incidence α_B which satisfies condition (7.73) is called the **Brewster angle**. It can be determined from the condition

$$n_{12} = \frac{\sin\alpha_B}{\sin\beta} = \frac{\sin\alpha_B}{\sin\left(\dfrac{\pi}{2}-\alpha_B\right)} = \tan\alpha_B, \tag{7.74}$$

so that

$$\alpha_B = \arctan n_{12}. \tag{7.75}$$

It is easy to be convinced that, under condition (7.73), the propagation directions of reflected and refracted waves are mutually perpendicular, i.e. $\vec{k}_1 \perp \vec{k}_2$.

Now, let us consider the case when the electromagnetic wave propagates from an optically denser medium to a less dense one ($\varepsilon_1 > \varepsilon_2$). Then,

$$n_{12} = \frac{\sin\alpha}{\sin\beta} = \sqrt{\frac{\varepsilon_1}{\varepsilon_2}} < 1, \tag{7.76}$$

and, hence, $\beta > \alpha$. In this case, the maximum value of angle β equals $\dfrac{\pi}{2}$. It corresponds to the angle of incidence

$$\alpha_0 = \arcsin n_{12} = \arcsin\frac{1}{n_{21}}. \tag{7.77}$$

If $\alpha > \alpha_0$, the refracted wave is absent. This phenomenon is called the **total internal reflection**, and the angle α_0 the **critical angle** of total internal reflection. Under the condition $\alpha < \alpha_0$,

$$\sin\beta = \frac{\sin\alpha}{n_{12}} > 1.$$

Therefore,

$$\cos\beta = \pm\sqrt{1-\left(\frac{\sin\alpha}{n_{12}}\right)^2} = \pm i\sqrt{\frac{\sin^2\alpha}{n_{12}^2}-1}, \tag{7.78}$$

and the phase of refracted wave (7.25) becomes complex-valued,

$$\begin{aligned}\vec{k}_2\vec{r}-\omega t &= k_{2y}y+k_{2z}z-\omega t=k_{2y}y+k_2\cos\beta\, z-\omega t\\ &= k_{2y}y\pm ik_2z\sqrt{\frac{\sin^2\alpha}{n_{12}^2}-1}-\omega t,\quad z<0.\end{aligned}\tag{7.79}$$

Substitution (7.79) into (7.25), we obtain the electric field strength in medium 2 in the form

$$\vec{E}_2=\vec{A}(z)\mathrm{e}^{i(k_{2y}y-\omega t)},\tag{7.80}$$

where

$$\vec{A}(z) = \vec{E}_0\mathrm{e}^{\pm\delta z},\quad z<0,\tag{7.81}$$

$$\delta = k_2\sqrt{\frac{\sin^2\alpha}{n_{12}^2}-1}.\tag{7.82}$$

The condition that the electric field strength must have a finite value at $z\to-\infty$ dictates the availability of sign “plus” in (7.81). Then,

$$\vec{A}(z)=\vec{E}_0\mathrm{e}^{\delta z},\quad z<0.\tag{7.83}$$

Hence, under the condition $\alpha>\alpha_0$, the electric field strength in medium 2 decreases exponentially as the distance from the interface grows. The amplitude of $\vec{A}(z)$ becomes smaller by a factor of e at the distance

$$|\Delta z|=\frac{1}{\delta}=\frac{1}{k_2\sqrt{\dfrac{\sin^2\alpha}{n_{12}^2}-1}}=\frac{c}{\sqrt{\varepsilon_2}\omega\sqrt{\dfrac{\sin^2\alpha}{n_{12}^2}-1}}.\tag{7.84}$$

At optical frequencies $\omega\sim10^{15}$ s^{-1}, we have $|\Delta z|\sim10^{-7}$ m. Hence, at the total internal reflection, the electromagnetic wave penetrates medium 2, but to a very short distance.

7.4. Electromagnetic waves in conductors

While considering the propagation of electromagnetic waves in conductors, we will assume the latter to be uniform, isotropic, electrically neutral

($\rho = 0$) and characterized by the specific conductance σ. Besides, the relationships $\vec{D} = \varepsilon\varepsilon_0\vec{E}$, $\vec{B} = \mu\mu_0\vec{H}$, and $\vec{j} = \sigma\vec{E}$ are supposed to be true. In this case, the system of Maxwell's equations (6.46)–(6.49) takes the form

$$\begin{aligned} \operatorname{rot}\vec{H} &= \sigma\vec{E} + \varepsilon\varepsilon_0\frac{\partial\vec{E}}{\partial t}, & (7.85)\\ \operatorname{rot}\vec{E} &= -\mu\mu_0\frac{\partial\vec{H}}{\partial t}, & (7.86)\\ \operatorname{div}\vec{H} &= 0, & (7.87)\\ \operatorname{div}\vec{E} &= 0. & (7.88) \end{aligned}$$

Applying the rot-operator to equation (7.86) and taking into account equations (7.85) and (7.87), we obtain

$$\Delta\vec{E} - \frac{\varepsilon\mu}{c^2}\frac{\partial^2\vec{E}}{\partial t^2} = \mu\mu_0\sigma\frac{\partial\vec{E}}{\partial t}. \tag{7.89}$$

Analogously, applying the rot-operator to equation (7.85), we obtain

$$\Delta\vec{H} - \frac{\varepsilon\mu}{c^2}\frac{\partial^2\vec{H}}{\partial t^2} = \mu\mu_0\sigma\frac{\partial\vec{H}}{\partial t}. \tag{7.90}$$

The solutions of equations (7.89) and (7.90) are sought in the form of plane monochromatic waves,

$$\begin{aligned} \vec{E} &= \vec{E}_0\mathrm{e}^{i(\vec{k}\vec{r}-\omega t)},\\ \vec{H} &= \vec{H}_0\mathrm{e}^{i(\vec{k}\vec{r}-\omega t)}. \end{aligned} \tag{7.91}$$

Substituting the solutions obtained into equations (7.89) and (7.90), we obtain the relation

$$k^2 = \frac{\varepsilon\mu}{c^2} + i\omega\mu\mu_0\sigma = \frac{\omega^2}{c^2}\mu\left(\varepsilon + i\frac{\sigma}{\varepsilon_0\mu}\right). \tag{7.92}$$

Introducing the notation

$$\tilde{\varepsilon} = \varepsilon + i\frac{\sigma}{\varepsilon_0\omega}, \tag{7.93}$$

relation (7.92) can be written as

$$k = \frac{\omega}{c}\sqrt{\tilde{\varepsilon}\mu}, \tag{7.94}$$

which looks similar to relation (7.10) obtained for insulators. Therefore, the quantity $\tilde{\varepsilon}$ should be considered as a complex dielectric permittivity of conductors.

For insulators, the quantity $n = \sqrt{\varepsilon\mu}$ is the refractive index. Then, the quantity

$$\tilde{n} = \sqrt{\tilde{\varepsilon}\mu} = \sqrt{\mu\left(\varepsilon + i\frac{\sigma}{\varepsilon_0\omega}\right)} \tag{7.95}$$

can be regarded as the complex refractive index of conductors. Expressing it in the form

$$\tilde{n} = n + in', \tag{7.96}$$

where n is the real and n' the imaginary part of complex refractive index $\tilde{n}$, we obtain the equation

$$n + in' = \sqrt{\mu\left(\varepsilon + i\frac{\sigma}{\varepsilon_0\omega}\right)}, \tag{7.97}$$

which can be used to determine n and n'. Adopting that $\mu = 1$, we write

$$(n + in')^2 = \varepsilon + i\frac{\sigma}{\varepsilon_0\omega}. \tag{7.98}$$

Equating the real and imaginary trerms in (7.98), we obtain a system of two equations

$$\begin{aligned} n^2 - n'^2 &= \varepsilon, \\ 2nn' &= \frac{\sigma}{\varepsilon_0\omega}, \end{aligned} \tag{7.99}$$

whence it follows that

$$n = \sqrt{\frac{\varepsilon}{2}\left\{\sqrt{1 + \left(\frac{\sigma}{\varepsilon\varepsilon_0\omega}\right)^2} + 1\right\}^2}, \tag{7.100}$$

$$n' = \sqrt{\frac{\varepsilon}{2}\left\{\sqrt{1 + \left(\frac{\sigma}{\varepsilon\varepsilon_0\omega}\right)^2} - 1\right\}^2}. \tag{7.101}$$

If $\sigma = 0$, equations (7.100) and (7.101) give rise to the relations for insulators: $n = \sqrt{\varepsilon}$ and $n' = 0$.

To establish the physical meaning of the real and imaginary parts of complex refractive index, let us rewrite relation (7.94) in the form

$$k = \frac{\omega}{c}\sqrt{\tilde{\varepsilon}} = \frac{\omega}{c}\tilde{n} = \frac{\omega}{c}(n + in') \tag{7.102}$$

and substitute it into (7.91) assuming that the electromagnetic wave propagates along the axis z,

$$\begin{aligned} \vec{E} &= \vec{E}_0 \mathrm{e}^{i(kz-\omega t)} \\ &= \vec{E}_0 \exp\left(-\frac{\omega}{c}n'z\right) \exp\left[i\left(\frac{\omega}{c}nz - \omega t\right)\right], \qquad (7.103) \\ \vec{H} &= \vec{H}_0 \mathrm{e}^{i(kz-\omega t)} \\ &= \vec{H}_0 \exp\left(-\frac{\omega}{c}n'z\right) \exp\left[i\left(\frac{\omega}{c}nz - \omega t\right)\right]. \qquad (7.104) \end{aligned}$$

Hence, the strengths of electric and magnetic fields generated by the electromagnetic wave in a conductor decrease with the distance following the exponential law. This can be explained by the fact that the electric field of the wave excites an electric current in the conductor. As a result, the warmth is released, so that the wave energy and, consequently, the electric and magnetic field strengths decrease as the wave penetrates the conductor.

The rate of electromagnetic wave damping, in accordance with (7.103) and (7.104), is determined using the imaginary part n' of the complex refractive index, which is called the **extinction coefficient**. The filed strengths decrease by a factor of e at the distance

$$|\Delta z| = \frac{c}{\omega n'}. \tag{7.105}$$

The real part n of the complex refractive index determines the wave phase. Therefore, it is called the **refractive index**. If all free charges in a conductor had become bound, the refractive index of the resulting insulator would have had just this value.

Substituting (7.91) into Maxwell's equations (7.85)–(7.88), we obtain the relations

$$\left[\vec{k}\vec{H}\right] = -\omega\left(\varepsilon\varepsilon_0 + i\frac{\sigma}{\omega}\right)\vec{E}, \tag{7.106}$$

$$\left[\vec{k}\vec{E}\right] = \mu\mu_0\omega\vec{H}, \tag{7.107}$$

$$\vec{k}\vec{H} = 0, \tag{7.108}$$

$$\vec{k}\vec{E} = 0, \tag{7.109}$$

whence it follows that the electromagnetic wave remains transverse in the conductor, and the vectors $\vec{E}$ and $\vec{E}$ mutually perpendicular. Besides, the following relation is valid:

$$H = \frac{k}{\mu\mu_0\omega}E, \tag{7.110}$$

whence, in the case $\mu = 1$, we obtain

$$H_0 = \frac{k}{\mu_0\omega}E_0 = \frac{\dfrac{\omega}{c}\tilde{n}}{\mu_0\omega}E_0 = \frac{n + in'}{\mu_0 c}E_0. \tag{7.111}$$

Expressing the complex quantity $n + in'$ in the exponential form

$$n + in' = \sqrt{n^2 + n'^2}\mathrm{e}^{i\varphi}, \quad \varphi = \arctan\frac{n'}{n}, \tag{7.112}$$

we obtain

$$H_0 = \frac{\sqrt{n^2 + n'^2}}{\mu_0 c}E_0\mathrm{e}^{i\varphi}. \tag{7.113}$$

Substituting (7.113) into (7.104), we find that

$$H = \frac{\sqrt{n^2 + n'^2}}{\mu_0 c}E_0 \exp\left(-\frac{\omega}{c}n'z\right)\exp\left[i\left(\frac{\omega}{c}nz - \omega t + \varphi\right)\right]. \tag{7.114}$$

Comparing (7.114) with (7.103), one can see that the electric and magnetic field strengths of electromagnetic wave in the conductor have a phase shift, which, in accordance with (7.112), (7.100), and (7.101), equals

$$\varphi = \arctan\frac{n'}{n} = \arctan\sqrt{\frac{\sqrt{1 + \left(\dfrac{\sigma}{\varepsilon\varepsilon_0\omega}\right)^2} - 1}{\sqrt{1 + \left(\dfrac{\sigma}{\varepsilon\varepsilon_0\omega}\right)^2} + 1}}. \tag{7.115}$$

At the normal incidence of electromagnetic wave on the conductor, the reflection coefficient, according to (7.71), is determined as

$$R = \left|\frac{1-\tilde{n}}{1-\tilde{n}}\right|^2 = \frac{|1-n-in'|^2}{|1+n+in'|^2} = \frac{(1-n)^2+n'^2}{(1+n)^2+n'^2}. \tag{7.116}$$

It is easy to be convinced that in a medium with a high enough conductivity, $\sigma/(\varepsilon\varepsilon_0\omega) \gg 1$. Then, in accordance with (7.100) and (7.101),

$$n = n' = \sqrt{\frac{\sigma}{2\varepsilon_0\omega}},$$

and phase shift (7.115) equals $\pi/4$.

7.5. Skin effect

The properties of alternating electromagnetic field substantially depend on its frequency. As a result, low-frequency fields have certain peculiarities. In many cases, they are similar to stationary fields, i.e. the fields of stationary currents. Therefore, they are called **quasi-stationary**.

The main attribute of quasi-stationary electromagnetic field is almost identical values of field phases at all field points. Therefore, the quasi-stationary character of the field is determined not only by the frequency, but also by the dimensions of corresponding space region. If the linear sizes of the region equal L and λ is the wavelength, the variation of the field phase over the distance L is determined by the formula

$$\Delta\varphi = 2\pi\frac{L}{\lambda}.$$

For the $\Delta\varphi$-value to be neglected ($\Delta\varphi \ll 2\pi$), the following condition must be satisfied:

$$L \ll \lambda. \tag{7.117}$$

Dividing inequality (7.117) by the velocity of field propagation, we obtain the **criterion of quasi-stationary state** in the form

$$\tau \ll T, \tag{7.118}$$

where $T = \lambda/c$ is the period of field oscillations and $\tau = L/c$ is the system delay time. Neglecting the phase change or the delay time means that the electromagnetic perturbation propagates in the system almost immediately ($c \to \infty$).

From (7.117), it follows that the frequency ν of quasi-stationary electromagnetic field has to satisfy the condition

$$\nu \ll \frac{c}{L}, \tag{7.119}$$

i.e. the value of quasi-stationary field frequencies depends on the system size. In particular, if $L = 3 \times 10^5$ m, $\nu \ll 3 \times 10^3$ Hz. The frequency of industrial alternating current (50 Hz) satisfies this criterion. For $L = 1$ m, $\nu \ll 3 \times 10^8$ Hz; this is the range of radio frequencies.

At low frequencies, the displacement current

$$\vec{j}_{disp} = \frac{\partial \vec{D}}{\partial t} = \varepsilon\varepsilon_0 \frac{\partial \vec{E}}{\partial t} = \varepsilon\varepsilon_0 \omega \vec{E}$$

is small and can be neglected in the case of quasi-stationary electromagnetic field. Then, the relation $\vec{j}_{disp} \ll \vec{j}$, where $\vec{j} = \sigma\vec{E}$ is the density of conduction current, has to be obeyed in Maxwell's equation (6.46). Whence one more criterion for the quasi-stationary state of electromagnetic field follows:

$$\omega = 2\pi\nu \ll \frac{\sigma}{\varepsilon\varepsilon_0}. \tag{7.120}$$

For metals, $\sigma \sim 10^7\ \Omega^{-1}\times\mathrm{m}^{-1}$. Therefore, displacement currents can be neglected in them at frequencies $\nu \ll 10^{18}$ Hz. From the comparison of (7.119) and (7.120), it follows that, in the case of macroscopic systems with metal conductors, the condition of quasi-stationary state (7.119) is crucial.

Taking into account the determinations and remarks made above, the system of Maxwell's equations (6.46)–(6.49) for the quasi-stationary electromagnetic field in the differential and integral forms is written as follows:

$$\operatorname{rot}\vec{H} = \vec{j}, \qquad \oint_L \vec{H}\,d\vec{l} = I; \tag{7.121}$$

$$\operatorname{rot}\vec{E} = -\frac{\partial \vec{B}}{\partial t}, \qquad \oint_L \vec{E}\,d\vec{l} = -\frac{\partial \Phi}{\partial t}; \tag{7.122}$$

$$\operatorname{div}\vec{B} = 0, \qquad \oint_S \vec{B}\,d\vec{S} = 0; \tag{7.123}$$

$$\operatorname{div}\vec{D} = \rho, \qquad \oint_S \vec{D}\,d\vec{S} = 0. \tag{7.124}$$

The corresponding constitutive (coupling) equations are

$$\vec{D} = \varepsilon\varepsilon_0\vec{E}, \vec{B} = \mu\mu_0\vec{H}, \vec{j} = \sigma\left(\vec{E} + \vec{E}^{ext}\right). \tag{7.125}$$

Then, the continuity equation looks like

$$\operatorname{div}\vec{j} + \frac{\partial \rho}{\partial t} = \operatorname{div}\vec{j} + \frac{\partial \vec{D}}{\partial t} = \operatorname{div}\left(\vec{j} + \vec{j}_{disp}\right) \approx \operatorname{div}\vec{j} = 0, \tag{7.126}$$

and coincides by form with the continuity equation for the stationary field of stationary current.

At the same time, it should be noted that in a good many cases, it in impossible to neglect the displacement current it is impossible, e.g., id there is a capacitor in the alternating current circuit. In this case, on the basis of (6.27), (6.32) and (6.50), we may write

$$\vec{j}_{disp} = \frac{\partial \vec{D}}{\partial t} = \varepsilon_0\frac{\partial \vec{E}}{\partial t} + \frac{\partial \vec{P}}{\partial t} = \varepsilon_0\frac{\partial \vec{E}}{\partial t} + \sum_i e_i\frac{\partial \vec{r}_i}{\partial t}, \tag{7.127}$$

i.e. the displacement current in the capacitor's insulator consists of the "pure" displacement current $\varepsilon_0\partial\vec{E}/\partial t$ and the term $\sum_i e_i\partial\vec{r}_i/\partial t$. The latter makes allowance for the motion of bound charges in the molecules of insulator. It is this component that is responsible for the heat release when the displacement current runs through polar insulators (cf. section 3.2).

The stationary current in the conductor is uniformly distributed over its cross-section, i.e. the current density is identical at every cross-section point. However, in the case of alternating current, this uniform distribution becomes violated: the current density near the conductor axis is lower in comparison with that near the surface. This phenomenon is called the **skin effect.** It can be explained with the help of the electromagnetic induction law.

A section of conductor with the alternating current I at the time moment when the current grows and is directed upwards is exhibited in Fig. 7.4. The dashed line represents the induction line of the growing magnetic field of the current. Then, according to the law of electromagnetic induction, there must be an induction current I_{ind} in the conductor, the closed line and the direction of which are shown in the figure by the solid curve (the circle). The directions of induction and conduction currents coincide near the conductor surface, but are opposite near the conductor axis. Therefore, the density of current near the surface turns out larger than that near the conductor axis.

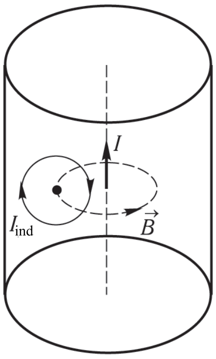

Fig. 7.4.

The same conclusion can be obtained in the framework of theoretical calculation as well. To make it easier, let a conductor with the specific conductance σ fill the half-space $z > 0$, and the current-inducing alternating electric field be directed along the axis x (Fig. 7.5). We suppose the conductor to be electrically neutral ($\rho = 0$) and non-ferromagnetic ($\mu = 1$). The system of Maxwell's equations for the quasi-stationary electromagnetic field (7.121)–(7.125), taking into account that $\vec{j} = \sigma\vec{E}$, $\vec{D} = \varepsilon\varepsilon_0\vec{E}$, and $\vec{B} = \mu_0\vec{H}$, looks like

$$
\begin{aligned}
\operatorname{rot}\vec{H} &= \sigma\vec{E}, && (7.128)\\
\operatorname{rot}\vec{E} &= -\mu_0\frac{\partial\vec{H}}{\partial t}, && (7.129)\\
\operatorname{div}\vec{H} &= 0, && (7.130)\\
\operatorname{div}\vec{E} &= 0. && (7.131)
\end{aligned}
$$

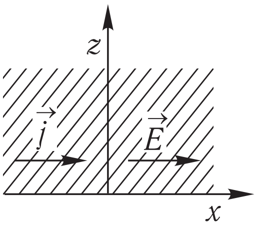

Fig. 7.5.

Applying the rot-operator to equation (7.129) and taking into account that $\operatorname{rot}\operatorname{rot} = \operatorname{grad}\operatorname{div} - \Delta$, as well as equations (7.128) and (7.131), we obtain

$$\Delta\vec{E} = \mu_0\sigma\frac{\partial\vec{E}}{\partial t}. \tag{7.132}$$

Applying the rot-operator to equation (7.128) and taking into account equations (7.129) and (7.130), we obtain

$$\Delta\vec{H} = \mu_0\sigma\frac{\partial\vec{H}}{\partial t}. \tag{7.133}$$

Equations (7.47) and (7.48) are identical by form. Therefore, it is sufficient to find the solution of equation (7.47), which will be sought in the form

$$\vec{E} = \vec{E}_0\mathrm{e}^{-i\omega t},$$

where $\vec{E}_0$ is a function of coordinates. From Fig. 7.5 and on the basis of symmetry arguments, it follows that the amplitude $\vec{E}_0$ is independent of the coordinates x and y, but it may depend on the coordinate z. Therefore, we assume that

$$\vec{E}(z,t) = \vec{E}_0(z)\mathrm{e}^{-i\omega t}. \tag{7.134}$$

Substituting (7.134) into (7.135), we obtain an equation for the amplitude $\vec{E}_0(z)$,

$$\frac{d^2\vec{E}_0(z)}{dz^2} = -i\mu_0\omega\sigma\vec{E}_0(z). \tag{7.135}$$

Let us rewrite it in the form

$$\frac{d^2\vec{E}_0(z)}{dz^2} = -k^2\vec{E}_0(z), \tag{7.136}$$

where

$$k^2 = -i\mu_0\omega\sigma. \tag{7.137}$$

The general solution of equation (7.136) is written as

$$\vec{E}_0(z) = \vec{a}e^{ikz} + \vec{b}e^{-ikz}, \quad k = \sqrt{i\mu_0\omega\sigma}. \tag{7.138}$$

Here, $\vec{a}$ and $\vec{b}$ are integration constants. Since

$$\sqrt{i} = \mathrm{e}^{i\pi/4} = \frac{i+1}{\sqrt{2}},$$

we may write

$$\begin{aligned} k &= \sqrt{i\mu_0\omega\sigma} = k' + ik', \\ k' &= \sqrt{\frac{\mu_0\omega\sigma}{2}}. \end{aligned} \tag{7.139}$$

Then, instead of (7.138), we obtain

$$\vec{E}_0(z) = \vec{a}e^{-k'z}\mathrm{e}^{ik'z} + \vec{b}e^{k'z}\mathrm{e}^{-ik'z}. \tag{7.140}$$

If $z \to \infty$, the second term in (7.140) grows infinitely; therefore, we must put $\vec{b} \equiv 0$. Hence, solution (7.140) takes the form

$$\vec{E}_0(z) = \vec{a}e^{-k'z}\mathrm{e}^{ik'z}, \tag{7.141}$$

and the electric field strength (7.134) looks like

$$\vec{E}(z,t) = \vec{a}e^{-k'z}\mathrm{e}^{i(k'z-\omega t)}. \tag{7.142}$$

Extracting the real part of (7.142) and multiplying it by the specific conductance σ, we obtain the electric current density

$$\vec{j}(z,t) = \vec{j}_0(z)\cos(k'z - \omega t), \tag{7.143}$$

where

$$\vec{j}_0(z) = \vec{a}\sigma e^{-k'z}. \tag{7.144}$$

From (7.143) and (7.144), it follows that the current density decreases following the exponential law towards the conductor depth (along the axis z) and becomes smaller by a factor of e at the depth

$$\Delta z = \frac{1}{k'} = \sqrt{\frac{2}{\mu_0\omega\sigma}}, \tag{7.145}$$

which is called the **skin depth** and where the main part of the current is concentrated. The skin depth substantially depends on the frequency. For example, for copper ($\sigma = 6 \times 10^7\ \Omega^{-1}\times\mathrm{m}^{-1}$), $\Delta z = 9$ mm at a frequency of 50 Hz and 0.2 mm at a frequency of 10^5 Hz.

Going deeper into the conductor, the current phase (see (7.144)) also changes, but this variation has a linear character. The results obtained are also applicable to a cylindrical conductor.

The skin effect reveals itself as if the transverse cross-section of the conductor becomes smaller, so that its electrical resistance increases. This phenomenon is especially appreciable at high frequencies. To reduce the influence of skin effect in radio equipment, special methods are applied. For instance, the surface of conductors is covered with a layer of metal possessing a high conductivity, e.g., silver. In the case of stationary current ($\omega = 0$), in accordance with (7.145), $\Delta z = \infty$, which means that the current density does not change as the conductor depth increases.

8

DISPERSION OF ELECTROMAGNETIC WAVES

8.1. Frequency dispersion of electromagnetic waves

Until now, we considered the dielectric permittivity ε to be constant for a definite medium. However, experimental data testify to its dependence on some parameters of electromagnetic field, in particular, the frequency ω. For example, the dielectric permittivity of water in the electrostatic field and in low-frequency fields equals 81 (the **static dielectric permittivity**). At the same time, $\varepsilon = n^2 = 1.77$ at optical frequencies (the **high-frequency dielectric permittivity**). The dependence of dielectric permittivity ε or the refractive index $n = \sqrt{\varepsilon}$ (we assume $\mu = 1$) on the frequency is called the **frequency dispersion**. The explanation and the calculation of the dependence $\varepsilon = \varepsilon(\omega)$ can be done only in the framework of microscopic theory that takes into account the interaction of electromagnetic waves with atoms and molecules in the medium.

The propagation of electromagnetic wave in the medium is accompanied by the polarization of the latter, which is a consequence of electron (electron polarization) and ion (ionic polarization) shifts and dipole molecule rotations (orientation polarization) in the medium under the action of the wave electric field. However, owing to the inertial properties of particles, the response of the medium to the electromagnetic wave action substantially depends on the wave frequency. As an example, let us consider the polarization associated with the shifts of electrons.

At low frequencies, the electrons move (forced oscillations) in phase with the variations of the wave electric field. However, at higher frequencies, owing to the inertia, their shifts begin to lag with respect to the field changes. If the frequency is too high, electrons have no time to trace those changes, and their response to the electromagnetic field action tends to zero. Hence, the amplitude of polarization vector, the dielectric susceptibility, and the dielectric permittivity depend on the field frequency. Similar reasoning also concerns the polarization associated with the shifts of ions and the rotations of dipole molecules. However, owing to different inertial properties, the ranges of low and high frequencies are also different in those cases.

The time delay in the medium response (polarization) to the field action means that the polarization and electric induction vectors at a definite time moment t are determined by the electric strength values at previous time moments t', i.e.

$$\vec{D}(\vec{r},t) = \varepsilon_0 \int_{-\infty}^{t} \varepsilon(t,t')\vec{E}(\vec{r},t')dt'. \tag{8.1}$$

If the properties of the medium do not change in time, $\varepsilon(t,t') = \varepsilon(t-t')$, so that

$$\vec{D}(\vec{r},t) = \varepsilon_0 \int_{-\infty}^{t} \varepsilon(t-t')\vec{E}(\vec{r},t')dt'. \tag{8.2}$$

Expanding the functions $\vec{E}(\vec{r},t')$ and $\vec{D}(\vec{r},t)$ into Fourier integrals, we obtain

$$\vec{D}(\vec{r},t) \quad = \quad \int_{-\infty}^{\infty} \vec{D}(\vec{r},\omega)\mathrm{e}^{-i\omega t}d\omega, \tag{8.3}$$

$$\vec{E}(\vec{r},t) \quad = \quad \int_{-\infty}^{\infty} \vec{E}(\vec{r},\omega)\mathrm{e}^{-i\omega t}d\omega, \tag{8.4}$$

The Fourier coefficient $\vec{D}(\vec{r},\omega)$ is determined as

$$\begin{aligned}\vec{D}(\vec{r},\omega) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \vec{D}(\vec{r},t)\mathrm{e}^{i\omega t}dt \\ &= \frac{\varepsilon_0}{2\pi}\int_{-\infty}^{\infty} dt\,\mathrm{e}^{i\omega t}\int_{-\infty}^{t} \varepsilon(t-t')\vec{E}(\vec{r},t')dt'. \end{aligned} \tag{8.5}$$

Substitution expansion (8.4) into (8.5) and introducing a new variable $\tau = t - t'$, we obtain

$$\vec{D}(\vec{r},\omega) = \frac{\varepsilon_0}{2\pi}\int_{-\infty}^{\infty} dt\, \mathrm{e}^{i\omega t}\int_{0}^{\infty} d\tau\, \varepsilon(\tau)\int_{-\infty}^{\infty} \vec{E}(\vec{r},\omega')\mathrm{e}^{-i\omega' t}\mathrm{e}^{i\omega'\tau}d\omega'. \quad (8.6)$$

Using the Fourier transform of delta-function (see Appendix A)

$$\delta(\omega-\omega') = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{i(\omega-\omega')t}dt, \quad (8.7)$$

we obtain (8.6) in a form

$$\begin{aligned} \vec{D}(\vec{r},\omega) &= \varepsilon_0\varepsilon(\omega)\vec{E}(\vec{r},\omega), && (8.8)\\ \varepsilon(\omega) &= \int_0^{\infty} \varepsilon(\tau)\mathrm{e}^{i\omega\tau}d\tau. && (8.9) \end{aligned}$$

Hence, the dielectric permittivity depends not only on the properties of the medium, but also on the electromagnetic wave frequency.

8.2. Classical electron theory of dispersion in gases

Let us consider a gas consisting of nonpolar molecules with a single optical electron, provided that the interaction between the molecules can be neglected. The optical electron is assumed to be a harmonic oscillator with the characteristic frequency of oscillations ω_0 and the damping coefficient γ.

The field of plane monochromatic electromagnetic wave with the frequency ω acts on the electron with the force

$$\vec{F} = -e\vec{E} = -e\vec{E}_0\mathrm{e}^{i(\vec{k}\vec{r}-\omega t)}. \quad (8.10)$$

Since the electron is located within the molecule, $\vec{k}\vec{r} \sim 2\pi a/\lambda$, where a is the molecule (atom) size. Taking into account that $a \sim 10^{-10}$ m and $\lambda \sim 5\times 10^{-7}$ m, we obtain $\vec{k}\vec{r} \sim 10^{-3}$; hence,

$$\mathrm{e}^{i\vec{k}\vec{r}} \approx 1.$$

Therefore, the force $\vec{F}$ that acts on the electron can be written as

$$\vec{F} = -e\vec{E}_0 \mathrm{e}^{-i\omega t}. \tag{8.11}$$

Then, the equations of motion for the electron looks like

$$m_0\ddot{\vec{r}} = -m_0\omega_0^2\vec{r} - m_0\gamma\dot{\vec{r}} - e\vec{E}_0\mathrm{e}^{-i\omega t}, \tag{8.12}$$

where $m_0\omega_0^2\vec{r}$ is the quasi-elastic force, and $m_0\gamma\dot{\vec{r}}$ is the radiation friction force. Rewriting equation (8.12) in the form

$$\ddot{\vec{r}} + \gamma\dot{\vec{r}} + \omega_0^2\vec{r} = -\frac{e}{m_0}\vec{E}_0\mathrm{e}^{-i\omega t}, \tag{8.13}$$

we obtain a linear nonhomogeneous differential equation of the second order. Its general solution is the sum of the general solution of homogeneous equation

$$\ddot{\vec{r}} + \gamma\dot{\vec{r}} + \omega_0^2\vec{r} = 0 \tag{8.14}$$

and the particular solution of equation (8.13).

Equation (8.14) is the equation describing damped oscillations. Therefore, its solution can be neglected. The particular solution of equation (8.13) is sought in the form

$$\vec{r} = \vec{r}_0\mathrm{e}^{-i\omega t}. \tag{8.15}$$

Substituting (8.15) into (8.13) and determining $\vec{r}_0$, we obtain the electron shift

$$\vec{r} = \frac{-\dfrac{e}{m_0}\vec{E}}{\omega_0^2 - \omega^2 - i\gamma\omega} \tag{8.16}$$

and the vector of gas polarization

$$\vec{P} = -Ne\vec{r} = \frac{-\dfrac{e^2N}{m_0}\vec{E}}{\omega_0^2 - \omega^2 - i\gamma\omega}, \tag{8.17}$$

where N is the concentration of molecules. Formula (8.17) enables the complex dielectric permittivity to be obtained in the form

$$\tilde{\varepsilon} = 1 + \tilde{\kappa} = 1 + \frac{\vec{P}}{\varepsilon_0\vec{E}} = 1 + \frac{\omega_p^2}{\omega_0^2 - \omega^2 - i\gamma\omega}. \tag{8.18}$$

Here, we used the notation for the plasma frequency ω_p (see section 8.5),

$$\omega_p = \sqrt{\frac{e^2 N}{\varepsilon_0 m_0}}. \tag{8.19}$$

Since the dielectric permittivity of gases differs inconsiderably from unity, the complex refractive index can be written as

$$\tilde{n} = \sqrt{\tilde{\varepsilon}} \approx 1 + \frac{\frac{1}{2}\omega_p^2}{\omega_0^2 - \omega^2 - i\gamma\omega} = n + in'. \tag{8.20}$$

Then,

$$n = 1 + \frac{\frac{1}{2}\omega_p^2\left(\omega_0^2 - \omega^2\right)}{\left(\omega_0^2 - \omega^2\right)^2 + \gamma^2\omega^2}, \tag{8.21}$$

$$n' = \frac{\frac{1}{2}\omega_p^2\omega\gamma}{\left(\omega_0^2 - \omega^2\right)^2 + \gamma^2\omega^2}. \tag{8.22}$$

It will be recalled (see section 7.4) that n is the refractive index, which determines the wave phase, and n' is the extinction coefficient, which describes the damping of electromagnetic wave, the vectors $\vec{E}$ and $\vec{H}$ of which look like (7.103) and (7.104). Then, the average absolute value of Poynting vector, i.e. the density of energy flux, takes the form

$$\langle\Pi\rangle = \Pi_0 \mathrm{e}^{-z/\delta}, \tag{8.23}$$

where the quantity

$$\frac{1}{\delta} = \frac{2\omega n'}{c} \tag{8.24}$$

is called the **absorption coefficient**. It equals the distance at which the wave energy decreases by a factor of e.

The frequency dependences of the refractive index n (the dispersion curve) and the extinction coefficient n' (the absorption curve), calculated according to formulas (8.21) and (8.22), are schematically shown in Fig. 8.1. At frequencies ω far from the characteristic frequency ω_0, the refractive index grows together with the frequency. Such a dependence of the refractive index on the frequency is called the **normal dispersion**. Near the characteristic frequency ω_0, the refractive index decreases as the frequency grows. This phenomenon is called the **anomalous dispersion**.

The extinction coefficient is substantial only at frequencies close to ω_0, forming the **absorption band**. The corresponding curve looks like a peak with a maximum at $\omega = \omega_0$ and a half-width at half-maximum that, as follows from (8.22), equals γ, Therefore, the quantity 2γ is considered as the **absorption band width**.

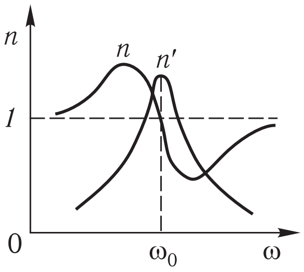

Fig. 8.1.

If the molecule has a few optical electrons with different characteristic frequencies ω_{0i} giving different contributions to polarization, formula (8.21) should be generalized to the form

$$n = 1 + \frac{1}{2} \sum_i \frac{f_i \omega_{0i}^2 \left(\omega_{0i}^2 - \omega^2\right)}{\left(\omega_{0i}^2 - \omega^2\right)^2 + \gamma_i^2 \omega^2}, \tag{8.25}$$

where the factors f_i correspond to the weights of indicated contributions.

8.3. Dispersion in dense non-polar dielectrics. Lorentz–Lorenz formula

The electron theory of dispersion developed in section 8.2 is applicable to non-polar rarefied gases, i.e. in the case when it is possible to consider that a gas molecule is subjected to the action of only the external electric field generating by the propagating electromagnetic wave. However, for compressed gases, liquids, and solids, it is necessary to take also into account the electric field created by other polarized molecules. Hence, the strength of electric field that acts on the molecule is equal to the sum of strengths of the external (electromagnetic wave) field and the field of polarized medium.

In [10] (section 7.1), it was shown that the microscopic strength value of electric field acting on a molecule in a non-polar isotropic dielectric is described by the formula

$$\vec{E}' = \vec{E} + \frac{\vec{P}}{3\varepsilon_0}, \tag{8.26}$$

where $\vec{E}$ is the strength of external macroscopic field, $\vec{E}'$ the actual field strength, and $\vec{P}$ the polarization vector. If $\vec{E}$ is the strength of electric field created by the electromagnetic wave, the equation of motion for the optical electron in the molecule can be written as

$$m_0\ddot{\vec{r}} + m_0\omega_0^2\vec{r} = -e\vec{E}'. \tag{8.27}$$

Here, the electron oscillation damping was neglected ($\gamma = 0$). Making allowance for (8.26), this equation can be rewritten as

$$\ddot{\vec{r}} + \omega_0^2\vec{r} + \frac{e}{3\varepsilon_0 m_0}\vec{P} = -\frac{e}{m_0}\vec{E}. \tag{8.28}$$

Multiplying the result by $-eN$ and taking into account that $-eN\vec{r} = \vec{P}$, we obtain

$$\ddot{\vec{P}} + \left(\omega_0^2 - \frac{e^2N}{3\varepsilon_0 m_0}\right)\vec{P} = \frac{e^2N}{m_0}\vec{E}. \tag{8.29}$$

Since $\vec{E} = \vec{E}_0\mathrm{e}^{-i\omega t}$ (see section 8.2), the particular solution of equation (8.29) should be sought in the form $\vec{P} = \vec{P}_0\mathrm{e}^{-i\omega t}$. As a result, we obtain

$$\vec{P} = \frac{\dfrac{e^2N}{m_0}\vec{E}}{\omega_0^2 - \omega^2 - \dfrac{e^2N}{3\varepsilon_0 m_0}}. \tag{8.30}$$

The polarization vector is coupled with the refractive index by the relation

$$\vec{P} = \varepsilon_0\kappa\vec{E} = \varepsilon_0(\varepsilon - 1)\vec{E} = \varepsilon_0(n^2 - 1)\vec{E}.$$

Therefore,

$$n^2 - 1 = \frac{\dfrac{e^2N}{\varepsilon_0 m_0}}{\omega_0^2 - \omega^2 - \dfrac{e^2N}{3\varepsilon_0 m_0}}. \tag{8.31}$$

It is easy to be convinced that relation (8.31) can be transformed as follows:

$$\frac{n^2 - 1}{n^2 + 2} = \frac{\dfrac{e^2N}{3\varepsilon_0 m_0}}{\omega_0^2 - \omega^2}. \tag{8.32}$$

This formula was discovered independently by the Danish physicist L. Lorenz (in 1869) and the Dutch physicist H. Lorentz (in 1878). Therefore, it is called the **Lorentz–Lorenz formula**.

The number of molecules in a unit volume equals $N = \rho N_A/M$, where ρ is the density, N_A the Avogadro constant; and M the molar mass. Then, the Lorentz–Lorenz formula can be rewritten in the form

$$\frac{n^2-1}{n^2+2}\frac{1}{\rho} = \frac{e^2 N_A}{3\varepsilon_0 m_0 M}\frac{1}{\omega_0^2-\omega^2}. \tag{8.33}$$

From (8.33), it follows that, for the definite medium, if the frequency ω is constant, the quantity

$$r = \frac{n^2-1}{n^2+2}\frac{1}{\rho} \tag{8.34}$$

is also constant, and experimental data confirm this conclusion. The quantity r is called the **specific refraction**. For example, when going from the gaseous state into the liquid one, the density of oxygen becomes 800 times as large, but the specific refraction remains invariable to an accuracy of 2–3%. If the air pressure varies from 1 (98.0665 kPa) to 200 at (19613 kPa), the corresponding specific refraction remains constant to within 0.1%.

8.4. Dispersion in metals

While studying the dispersion of electromagnetic waves in metals, both free electrons and electrons bound in ions have to be taken into consideration. The bound electrons are characterized by the characteristic frequencies ω_{0i}, and $\omega_0 = 0$ for free electrons. Designating the concentration of free electrons as N and the number of bound electrons in the i-th state in ions as N_i, the polarization vector, on the basis of (8.17), can be written in the form

$$\vec{P} = \sum_i \frac{\frac{e^2 N_i}{m_0}\vec{E}}{\omega_{0i}^2-\omega^2-i\gamma_i\omega} - \frac{\frac{e^2 N}{m_0}\vec{E}}{\omega^2+i\gamma\omega}. \tag{8.35}$$

Then, on the basis of (8.35), putting $\mu = 1$, the complex dielectric permittivity of metals reads

$$\tilde{\varepsilon} = (n+in')^2 = 1 + \sum_i \frac{\omega_{pi}^2}{\omega_{0i}^2-\omega^2-i\gamma_i\omega} - \frac{\omega_p^2}{\omega^2+i\gamma\omega}, \tag{8.36}$$

where

$$\omega_{pi}^2 = \frac{e^2 N_i}{\varepsilon_0 m_0}, \omega_p^2 = \frac{e^2 N}{\varepsilon_0 m_0}.$$

Extracting the real and imaginary parts of (8.36), we obtain

$$n^2 - n'^2 = 1 + \sum_i \frac{\omega_{pi}^2 \left(\omega_{0i}^2 - \omega^2\right)}{\left(\omega_{0i}^2 - \omega^2\right)^2 + \gamma_i^2 \omega^2} - \frac{\omega_p^2}{\omega^2 + \gamma^2}, \tag{8.37}$$

$$2nn' = \sum_i \frac{\omega_{pi}^2 \gamma_i \omega}{\left(\omega_{0i}^2 - \omega^2\right)^2 + \gamma_i^2 \omega^2} + \frac{\omega_p^2 \gamma}{\omega\left(\omega^2 + \gamma^2\right)}. \tag{8.38}$$

Formulas (8.37) and (8.38) agree with experimental data in a wide frequency range.

The characteristic frequencies of bound electrons ω_{0i} belong to the ultra-violet spectral interval. Provided $\omega \ll \omega_{0i}$, relations (8.37) and (8.38) look like

$$n^2 - n'^2 = 1 - \frac{\omega_p^2}{\omega^2 + \gamma^2}, \tag{8.39}$$

$$2nn' = \frac{\omega_p^2 \gamma}{\omega\left(\omega^2 + \gamma^2\right)}, \tag{8.40}$$

which is in agreement with experimental data for the infra-red range.

8.5. Dispersion of electromagnetic waves in plasma

Plasma is a gas consisting of ions and electrons. Hence, plasma is similar to metal. The only difference is that metal ions are located at the crystal lattice sites, whereas plasma ions can move rather freely. Besides, the concentrations of ions and electrons in plasma are usually lower than in metal. At the same time, this similarity enable relations (8.37) and (8.38), with corresponding amendments, to be used in the case of plasma consideration.

Under the action of the electric field of electromagnetic wave in plasma, electrons and ions shift. However, since ions are much heavier than electrons, their shift is shorter than that of electrons and, hence, can be neglected. As a result, only the shift of electrons should be taken into account. Moreover, in the case of rarefied plasma, the frequency of collisions between plasma particles is rather low. This means that the wave energy loss is also low, and we may put $\gamma = 0$. Bearing those remarks in mind, relations (8.37) and (8.38) for the plasma case can be written as

$$n^2 = 1 - \frac{\omega_p^2}{\omega^2}, \tag{8.41}$$

$$n'^2 = 0, \tag{8.42}$$

where the parameter $\omega_p = \sqrt{\dfrac{e^2 N}{\varepsilon_0 m_0}}$ is called the **plasma frequency**, and N is the concentration of free electrons.

From (8.41), it follows that, at $\omega < \omega_p$, the refractive index

$$n = i\sqrt{\frac{\omega_p^2}{\omega^2} - 1} = i\,|n|$$

is imaginary. Then, from (7.71), it follows that the coefficient of electromagnetic wave reflection from plasma at the normal incidence equal unity,

$$R = \left|\frac{1-n}{1+n}\right|^2 = 1. \tag{8.43}$$

This fact can explain, in particular, the reflection of long radio waves from the ionosphere.

If $\omega > \omega_p$, it follows from (8.41) that $n < 1$, which means that the phase velocity of electromagnetic waves in plasma exceeds the light velocity in vacuum, $v_{ph} = c/n > c$. However, this circumstance does not contradict the relativity principle, because the concept of phase speed is abstract: it is not associated with the energy (signal) transfer. The group velocity

$$v_{gr} = \frac{d\omega}{dk} = \left(\frac{dk}{d\omega}\right)^{-1},$$

which determines the velocity of electromagnetic signal propagation, does not exceed the light velocity in vacuum c. Really, taking into account that

$$k = \frac{\omega n}{c} = \frac{\omega}{c}\sqrt{1 - \frac{\omega_p^2}{\omega^2}} \tag{8.44}$$

and carrying out come simple calculations, we obtain

$$\frac{dk}{d\omega} = \frac{1}{c\sqrt{1 - \dfrac{\omega_p^2}{\omega^2}}} = \frac{1}{cn}, \tag{8.45}$$

so that $v_{gr} = cn < c$.

What is the meaning of plasma frequency $\omega_p = \left\{e^2 N/(\varepsilon_0 m_0)\right\}^{1/2}$? In section 7.4, it was noticed that the conductivity of the medium can be taken into account with the help of complex dielectric permittivity

$$\tilde{\varepsilon} = \varepsilon + i\frac{\delta}{\varepsilon_0 \omega}, \tag{8.46}$$

where ε is the dielectric permittivity of non-conducting subsystem in the medium, and σ the specific conductance of conducting subsystem. In the case of plasma, ε is the dielectric permittivity of ionic system, which equals unity if the shift of ions is neglected. The specific conductance σ can be found with the help of relationship

$$\sigma = \frac{\vec{j}}{\vec{E}} = \frac{-eN\dot{\vec{r}}}{\vec{E}}, \tag{8.47}$$

where N is the concentration of electrons, and $\vec{r}$ is their shift under the action of electromagnetic wave. The latter can be determined using formula (8.16). For electrons in plasma, $\omega_0 = 0$. Therefore,

$$\vec{r} = \frac{\dfrac{e}{m_0}\vec{E}}{\omega^2 + i\gamma\omega}. \tag{8.48}$$

Since $\vec{E} = \vec{E}_0 \mathrm{e}^{-i\omega t}$, using (8.47) and (8.48), we obtain

$$\sigma = \frac{\dfrac{e^2 N}{m_0}}{\gamma - i\omega} = \frac{\varepsilon_0 \omega_p^2}{\gamma - i\omega}. \tag{8.49}$$

Substituting (8.49) into (8.46), we have

$$\tilde{\varepsilon} = 1 - \frac{\omega_p^2}{\omega(\omega + i\gamma)}. \tag{8.50}$$

Hence, the vector of electric induction that makes allowance for the electric conductivity in the medium looks like

$$\vec{D} = \varepsilon_0 \tilde{\varepsilon} \vec{E}. \tag{8.51}$$

In the case of rarefied plasma, $\gamma \to 0$. Therefore (cf. (8.41)),

$$\tilde{\varepsilon} = 1 - \frac{\omega_p^2}{\omega^2}, \tag{8.52}$$

$$\vec{D} = \varepsilon_0 \left(1 - \frac{\omega_p^2}{\omega^2}\right) \vec{E}. \tag{8.53}$$

If the vectors $\vec{E}$ and $\vec{D}$ oscillate harmonically in time,

$$\begin{aligned} \frac{\partial \vec{D}}{\partial t} &= \varepsilon_0 \frac{\partial \vec{E}}{\partial t} - \varepsilon_0 \frac{\omega_p^2}{\omega^2} \frac{\partial \vec{E}}{\partial t} = \varepsilon_0 \frac{\partial \vec{E}}{\partial t} + i\varepsilon_0 \frac{\omega_p^2}{\omega^2} \vec{E} \\ &= \varepsilon_0 \frac{\partial \vec{E}}{\partial t} + \sigma \vec{E} = \vec{j}_{disp} + \vec{j}. \end{aligned} \tag{8.54}$$

Then, the system of Maxwell's equations (6.46)–(6.49) for plasma, taking into account the electric neutrality of plasma ($\rho = 0$) and the relationship $\vec{B} = \mu_0 \vec{H}$, should be written in the form

$$\begin{array}{ll} \operatorname{rot} \vec{H} = \dfrac{\partial \vec{D}}{\partial t}, & \operatorname{rot} \vec{E} = -\mu_0 \dfrac{\partial \vec{H}}{\partial t}, \\ \operatorname{div} \vec{H} = 0, & \operatorname{div} \vec{D} = 0. \end{array} \tag{8.55}$$

Substituting the vectors $\vec{E}$, $\vec{D}$, and $\vec{H}$ in the form

$$\vec{E} = \vec{E}_0 \mathrm{e}^{i(\vec{k}\vec{r} - \omega t)}, \vec{D} = \vec{D}_0 \mathrm{e}^{i(\vec{k}\vec{r} - \omega t)}, \vec{H} = \vec{H}_0 \mathrm{e}^{i(\vec{k}\vec{r} - \omega t)}$$

into the system of equations (8.55), we obtain

$$\left[\vec{k}\vec{H}\right] = -\varepsilon_0 \omega \left(1 - \frac{\omega_p^2}{\omega^2}\right) \vec{E}, \tag{8.56}$$

$$\left[\vec{k}\vec{E}\right] = \mu_0 \omega \vec{H}, \tag{8.57}$$

$$\vec{k}\vec{H} = 0, \tag{8.58}$$

$$\vec{k}\vec{D} = 0. \tag{8.59}$$

This system has a solution

$$\vec{H} = \vec{D} = 0, \quad \vec{E} \neq 0. \tag{8.60}$$

However, it follows then from (8.57) that $\vec{E} \parallel \vec{k}$, i.e. there are longitudinal oscillations of electric vector $\vec{E}$. Moreover, in this case, in accordance with (8.56), $\omega = \omega_0$. Hence, the plasma frequency ω_p is the frequency of longitudinal electric field oscillations.

Plasma oscillations with the frequency ω_p arise if any factor perturbs the electric neutrality of plasma, and negative and positive charges in plasma (electrons and ions) become separated. Then, under the influence of Coulomb forces, electrons will start to move in order to restore the electric neutrality, whereas ions, which are much heavier, can be considered motionless. As a result, electron oscillations with respect to ions arise.

C. Principles of relativity theory

9

FUNDAMENTALS OF THE SPECIAL THEORY OF RELATIVITY

9.1. Experimental basis of the special theory of relativity

In the last quarter of 19th century, despite of success attained in classical electrodynamics, a principal problem remained unresolved. It was the problem of light propagation in moving media. Its importance consisted in its close relation to the properties of space and time.

Until the end of 19th century, the Newtonian concept of space and time dominated in physics. According to it, space and time are absolute entities, independent of each other, as well as of material objects, which are contained in them. Space was considered as a container for the matter, and time was adopted to run identically at every point in the Universe. Mathematically, this concept was implemented in the Galilean transformations for coordinates and time.

While studying physical phenomena, a reference frame, i.e. a coordinate system attached to a reference object, and a clock are always used. In most cases, inertial reference frames (IRFs) are used in which Newton's first law of motion (the law of inertia) is obeyed. In such reference frames, any object either remains at rest or continues to move at a constant velocity and in the same direction if other objects do not act upon it or their actions are mutually compensated.

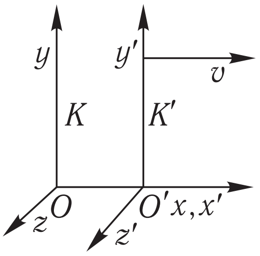

Fig. 9.1.

Let us consider two IRFs, K and K' (Fig. 9.1). Let the frame K be fixed and the frame K' move with respect to K at the velocity v along the axis. At the initial moment, $t = t' = 0$, the coordinate origins O and O' of both IRFs coincide. Then, according to the laws of classical mechanics, the transformations of coordinates and time for a material point look like

$$\begin{aligned} x' &= x - vt, \\ y' &= y, \quad z' = z; \quad t' = t. \end{aligned} \tag{9.1}$$

Formulas (9.1) are called the **Galilean transformations**. They bring about the velocity addition law in classical mechanics,

$$u'_x = u_x - v, \quad u'_y = u_y, \quad u'_z = u_z, \tag{9.2}$$

where $u'_x = dx'/dt'$ and $u_x = dx/dt$ are the velocities of material point in the reference frames K' and K, respectively.

Newton's second law

$$m\frac{d^2\vec{r}}{dt^2} = \vec{F} \tag{9.3}$$

is invariant with respect to the Galilean transformations. Really, the mass m is a constant quantity in classical mechanics, and the acceleration, in accordance (9.1), also has the same value in every IRF,

$$\frac{d^2\vec{r}'}{dt'^2} = \frac{d^2\vec{r}}{dt^2}$$

The force $\vec{F}$ acting on the material point is determined as

$$\vec{F} = -\frac{\partial V}{\partial \vec{r}}.$$

Here, V is the potential energy of point interaction with other material points. It depends on relative distances, which, in accordance with (9.1), are identical in various IRFs. Therefore,

$$\frac{\partial V'}{\partial \vec{r}'} = \frac{\partial V}{\partial \vec{r}}. \tag{9.4}$$

Hence, Newton's laws of motion are invariant with respect to the Galilean transformations. This means that mechanical phenomena occur identically in ever IRF and any mechanical experiments in those systems cannot reveal their motion with respect to other IRFs. This conclusion is called the **Galilean relativity principle**.

This principle was formulated in the 17th century, and it concerned only mechanical phenomena. After the creation of classical macroscopic electrodynamics (the 19th century), an issue emerged concerning the invariance of Maxwell's equations, which electrodynamics is based on, with respect to the Galilean transformations. However, Maxwell's equations include the light speed (see (3.36)–(3.39)), since $\varepsilon_0\mu_0 = 1/c^2$. Therefore, if the velocity addition law (9.2) following from the Galilean transformations had been applicable in this case, the light speed would have had different values in different IRFs. As a result, Maxwell's equations would have been non-invariant with respect to the Galilean transformations, and electromagnetic phenomena would have run differently in different IRFs.

After the electromagnetic nature of light was established, light was considered as a propagation of electromagnetic waves in the **aether**, a special medium that fills the whole Universe, with the speed $c = 3 \times 10^8$ m/s being the speed of light with respect to the aether. In this case, the speed of light with respect to moving objects must depend on the aether behavior with respect to them. The following assumptions are possible.

- The aether is completely stationary. Then, the motion of objects can be regarded as absolute (the **Lorentz hypothesis**).
- The aether is partially dragged by moving objects (the **Fresnel–Fizeau hypothesis**).
- The aether is completely dragged by a moving object. Therefore, the relative velocity of light equals c irrespective of the object's motion state. As a result, the observation of electromagnetic phenomena in a moving medium does not allow one to detect the object's motion (the **Hertz hypothesis**).

The validity or invalidity of each of those assumptions has to be elucidated on the basis of experimental data. However, it was found that there were experimental proofs that testified both to and against each of them. The main of them are the starlight aberration phenomenon and the Fizeau and Michelson experiments.

The phenomenon of starlight aberration was discovered by the English astronomer Bradley in 1728. It consists in that the visible positions of stars perform ellipses on the celestial sphere for a year. For stars in the ecliptic plane, the ellipses degenerate into straight lines, whereas for stars observed perpendicularly to the ecliptic plane, into circles.

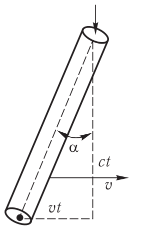

Fig. 9.2.

The starlight aberration can be is easily explained if we adopt the hypothesis of stationary aether and take into account the finiteness of light speed and the annual motion of the Earth around the Sun. Let a star be in the zenith. Then, the light emitted by the star will propagate normally to the Earth motion velocity $\vec{v}$ (Fig. 9.2). For the time interval t required for the light to pass the length of telescope tube, ct, the latter shifts by the distance vt. Therefore, for the light to get into the observer's eye, the telescope tube should be tilted towards direction of velocity v by the angle α determined by the condition

$$\tan\alpha = \frac{vt}{ct} = \frac{v}{c}. \tag{9.5}$$

In half a year, the direction of velocity vector $\vec{v}$ changes by 180°, and the telescope becomes tilted by the same angle α, but in the opposite direction. If we measure the angle 2α, we will be able to calculate the speed of light. According to Bradley, $c \approx 3.08 \times 10^8$ m/s. Hence, the starlight aberration proves that the aether is stationary.

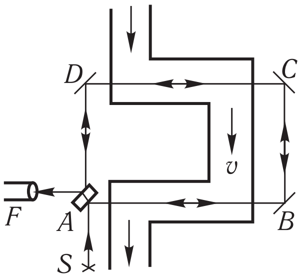

Fig. 9.3.

The **Fizeau experiment** (1851) consisted in measuring the phase velocity of light in a moving medium (water). The experimental setup is shown on Fig. 9.3. Here, v is the velocity of water motion in a bent pipe; S the light source; A the translucent plate reflecting some fraction of incident light and transmitting the other; B, C, and D are mirrors; and F the telescope. The light wave from source S becomes split by plate A into two coherent waves. One of them, reflected, propagates counterclockwise along the path $ABCDAF$, and gets into the objective of telescope F. The second wave passes through plate A, propagates clockwise along the path $ADCBAF$, and also gets into the telescope objective. As a result of interference between two coherent light waves, a system of interference bands is observed in the field of telescope F . Comparing the interference patterns registered at stationary and moving water, it is possible to calculate the phase velocity of light in moving water.

If the reference frame K' is coupled with moving water, then, according to the velocity addition law (9.2), the speed of light in the stationary (laboratory) reference frame K must equal

$$u = \frac{c}{n} + v.$$

However, the experiment shows that

$$u = \frac{c}{n} + v\left(1 - \frac{1}{n^2}\right). \tag{9.6}$$

In the framework of aether theories, this result testifies to a partial drag of aether by moving water, and the quantity

$$\alpha = 1 - \frac{1}{n^2}, \tag{9.7}$$

where n is the refractive index of water should be taken as the **drag coefficient**. Hence, the Fizeau experiment confirms assumption 2, but contradicts assumptions 1 and 3.

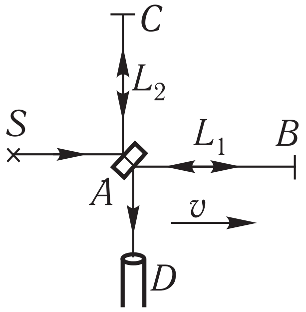

Fig. 9.4.

Now, let us consider the Michelson experiments, which were carried out, with improvements, between 1881 and 1929 in order to detect "the aether wind", i.e. the Earth's motion with respect to the aether, the latter being considered as stationary. For this purpose, Michelson designed an interferometer, the schematic diagram of which is exhibited in Fig. 9.4.

The light wave emitted by source S falls on translucent plate A, where it becomes split into two coherent waves. The reflected wave propagates to mirror C, reflects from it, and transmitting through plate A gets into telescope D. The second wave transmits through plate A, reflects first from mirror B and afterwards from plate A, and also gets into the telescope, where two waves interfere. The arms AB and AC of interferometer are mutually perpendicular with the lengths L_1 and L_2, respectively.

Let the Earth move with respect to the stationary aether at the speed v directed along arm L_1. The time for light to travel from plate A to mirror B and backwards equals

$$t_1 = \frac{L_1}{c - v} + \frac{L_1}{c + v} = \frac{2L_1}{c}\frac{1}{1 - \dfrac{v^2}{c^2}}, \tag{9.8}$$

where, in accordance with (9.2), $c - v$ is the speed of light directed along the Earth velocity v, and $c + v$ opposite to it.

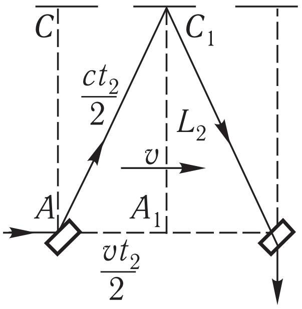

Fig. 9.5.

When calculating the time t_2, for which light travels along arm L_2 to mirror C and backwards, we must take into account that the interferometer moves. For the time interval $t_2/2$, light travels the path AC_1 (Fig. 9.5) of the length $ct_2/2$, and the interferometer shifts by the distance $AA_1 = vt_2/2$. Then,

$$\left(\frac{1}{2}ct_2\right)^2 = L_2^2 + \left(\frac{1}{2}vt_2\right)^2 .$$

Hence,

$$t_2 = \frac{2L_2}{c}\frac{1}{\sqrt{1 - \dfrac{v^2}{c^2}}}. \tag{9.9}$$

The time difference $t_1 - t_2$, which determines the interference pattern, equals

$$t_1 - t_2 = \frac{2}{c}\left\{\frac{L_1}{1 - \dfrac{v^2}{c^2}} - \frac{L_2}{\sqrt{1 - \dfrac{v^2}{c^2}}}\right\}. \tag{9.10}$$

If we rotate the interferometer by 90°, arms L_1 and L_2 interchange their positions, and the time difference must equal

$$t_1' - t_2' = \frac{2}{c}\left\{\frac{L_1}{\sqrt{1 - \dfrac{v^2}{c^2}}} - \frac{L_2}{1 - \dfrac{v^2}{c^2}}\right\}. \tag{9.11}$$

Hence, owing to the interferometer rotation, the time difference must change by

$$\Delta t = (t_1' - t_2') - (t_1 - t_2) = \frac{2\,(L_1 + L_2)}{c}\left(\frac{1}{\sqrt{1 - \dfrac{v^2}{c^2}}} - \frac{1}{1 - \dfrac{v^2}{c^2}}\right), \tag{9.12}$$

where v is the velocity of Earth motion around the Sun ($v \approx 30$ km/s) and, together with the Sun, around the Galaxy center ($v \approx 200$ km/s), and so on. However, the condition $v \ll c$ remains valid, so that (9.12) can be transformed as

$$\Delta t = \frac{(L_1 + L_2)\, v^2}{c^3} \tag{9.13}$$

The change in the time difference Δt means the change in the optical path difference by $c\Delta t$ and gives rise to the shift of interference pattern by N bands, where

$$N = \frac{c\Delta t}{\lambda} = \frac{L_1 + L_2}{\lambda} \frac{v^2}{c^2}. \tag{9.14}$$

In the last experiments carried out following the Michelson method, the magnitude of $L_1 + L_2$, owing to multiple reflections, reached 50 m. Therefore, for $v = 30$ km/s and $\lambda = 5 \times 10^{-7}$ m, the quantity $N \approx 1$, i.e. the interference pattern should have been shifted by one band. However, neither the first nor the later experiments revealed any shift in the interference patterns, despite that the installation enabled the shifts of small fractions of the band width to be registered. In the framework of aether theories, the negative result of Michelson experiments testifies to a complete drag of aether by moving objects, but contradicts the starlight aberration phenomenon and the Fizeau experiment.

In order to explain the negative result of Michelson experiments, FitzGerald and Lorentz put forward a hypothesis that the length of objects contracts by $\sqrt{1 - v^2/c^2}$ along the direction of motion. Then, before the interferometer was rotated, $L_1 = L_{10}\sqrt{1 - v^2/c^2}$ and $L_2 = L_{20}$, where L_{10} and L_{20} are the lengths of arms at rest; and, instead of (9.10), we obtain the formula

$$t_1 - t_2 = \frac{2}{c} \frac{L_{10} - L_{20}}{\sqrt{1 - \dfrac{v^2}{c^2}}}. \tag{9.15}$$

It is easy to be convinced that the rotation of interferometer by 90° does not change the time difference and, consequently, the interference pattern.

At the same time, as a result of Earth's daily rotation and its annual motion around the Sun, the speed v has to change in time. Therefore, if $L_{10} \neq L_{20}$, the time difference $t_1 - t_2$ and the interference pattern must also change. However, long-term observations (the Kennedy–Thorndike experiment) proved that there is no shift of interference bands in this case as well. Hence, the FitzGerald–Lorentz hypothesis about the size contraction in moving objects, which follows from the Newtonian concept of space, time, and aether, turned out incapable.

In addition, let us recall the Ritz hypothesis that the speed of light is constant with respect to the source, so that it equals the sum of the light speed with respect to the motionless source in the given IRF and the velocity of the source itself (the **ballistic hypothesis**). In this case, the speed of light with respect to the interferometer in the Michelson experiment should not depend on the Earth motion and, consequently, the interference pattern should not change if the interferometer is rotated. However, observations of binary stars resolutely deny the ballistic hypothesis. Without calculations, we would like to point out that should the Ritz hypothesis be valid, the observable motion would have a complicated and nonperiodical character contradicting the results of observations.

To summarize, notice that, before the special theory of relativity (STR) was developed, all attempts to explain the available body of experimental facts making use of Newtonian concept of space and time turned out inefficient.

9.2. Postulates of special relativity. Relativity of simultaneity

A new doctrine of space and time, which was called the **special** (partial) **theory of relativity** (STR), was formulated by A. Einstein in 1905. It is based on two postulates, which are verified by the agreement between the consequences that follow from it and the whole body of experimental data.

Einstein's first postulate, the **principle of relativity**, extends the Galilean principle of relativity onto all physical phenomena. Its formulation is as follows:

- *All physical phenomena of the same type run identically in all inertial reference frames, provided identical initial conditions.*

Einstein's second postulate is called the **principle of light speed constancy**:

- *The speed of light in vacuum* $c = 3 \times 10^8$ *m/s is identical in all IRFs and does not depend on the propagation direction.*

The first postulate declares the absence of exclusive (absolute) reference frame and the absolute motion. This means, in particular, that the STR completely denies the existence of ether as a carrier of electromagnetic (light) waves. The latter are not considered in the STR as the ether's state, but as a material object that does not require a special medium. From the first postulate, it follows that physical laws have to be formulated identically in all IRFs. Mathematically, those laws are expressed in the form of equations. Therefore, the form of equations must not vary when changing from one IRF to another. In other words, the equations must be covariant with respect to the transformations of coordinates and time associated with the IRF change. The transformations themselves (the Lorentz transformation) are established below (see section 9.4).

Einstein's second postulate explains the negative result of Michelson experiment: the speed of light is the same along both interferometer's arms, therefore its rotation does not change the interference pattern. Besides, the second postulate can be formulated as the statement that the speed of light in vacuum, which has the same value in every IRF, is the limit for the signal (energy, interaction) propagation velocity. However, its invariance in all IRFs contradicts, at first glance, the common sense.

Taking all that into account, let us consider the following thought experiment. There are two reference frames, K and K' (Fig. 9.1). Suppose that, when the coordinate origins O and O' coincide, a light flash occurs, and a light wave starts to propagate in both reference frames. Since the speed of light propagation in all directions is identical, the front of light wave in the stationary reference frame K must look like a sphere with the center at point O. However, for the same reasons, the wave front in system K' also must look like a sphere, but with the center at moving point O'. Just this situation contradicts our intuition. From the viewpoint of "common sense", this is a paradox. It can be explained by the relativity of simultaneity.

Events (physical phenomena) are called **simultaneous** if they occur at the same time momentum. Let us consider another thought experiment. Let observer 1 and a source of light flashes be located at the center of a train car (Einstein's train) that moves uniformly and rectilinearly at the relativistic speed $v \leq c$ (this is admissible in thought experiments). The doors at the both car's ends are located equidistantly with respect to the car's center and can be opened if illuminated with a light flash signal. Besides, there is another observer (observer 2) standing on a motionless railway platform.

When the car moves past observer 2, the lamp in the car emits a flash, and the light signal starts to propagate towards the front and back doors. Since the distances from the lamp to both doors are equivalent, and the speed of light does not depend on the direction, observer 1 will see that the both doors will open simultaneously, i.e. these are two simultaneous events. However, observer 2 on the platform will see a different picture. In his stationary reference frame, the speeds of light propagating forward and backward with respect to the direction of car motion are also identical, but the back door "approaches" the light, whereas the front one "runs away" from it. Therefore, for observer 2, the back door will be opened earlier than the front one. Hence, two events, the door openings, will be non-simultaneous for him.

Let us come back to the first thought experiment. The **wave front** is a set of points, which are reached simultaneously by light excitation. Therefore, as a result of the relativity of simultaneity, the wave front in reference frame K is composed by one set of points, and, in frame K' by the other. The relativity of simultaneity brings us to a basic conclusion on the dependence of time rate on the IRF choice. In other words, the time rate is different in different IRFs.

9.3. Reference frame in the special theory of relativity. Synchronization of clocks

In classical mechanics, the reference frame is composed, as was already said in section 9.1, of a reference object, an attached coordinate system, and a clock. The reference object is required to mark the coordinate origin, and the clock can be only one, because the speed of signal (interaction) propagation is considered infinitely high in classical mechanics. In this case, the both doors of the car in the thought experiment with "Einstein's train" must be opened with a flash of lamp simultaneously for both observer 1 and 2.

In the STR, the reference frame also has a reference object—it can be a material object with the non-zero rest mass (see section 10.2)—and a coordinate system with rigid axes; but, owing to the finite speed of signal transmission (the speed of light), there must be a clock at every point where the event time is registered. Every clock in the given IRF must be at rest, because its transportation from one point to another is associated with an acceleration that affects the clock rate. Besides, the clocks must be identical. In particular, as such, there can be atoms of the same kind, because their characteristic periods of oscillations precisely coincide, and the radiation wavelength can be taken as the length standard.

The application of atomic oscillations as a time standard and the radiation wavelength as a length one provides their complete identity in every IRF. For this purpose, it is enough to take the same substance in all IRFs.

Besides the identity of all clocks in the given IRF, the start of time counting must be agreed; i.e. the clocks must be synchronized. Synchronization can be carried out with the help of light signals as follows. The clock at the coordinate origin will be considered as reference; i.e. all other clocks are synchronized with it. Let the distance from one of them to the reference clock equals r. For its measurement, only the length standard is required. At some time moment t_0, a light signal is sent from the reference clock. Fo the signal to travel over the distance r, the time r/c is required. Therefore, at the moment when the signal arrives at the synchronized clock, the latter has to show the time $t = t_0 + r/c$. In such a way, it is possible to synchronize all clocks in the given IRF.

9.4. Lorentz transformations

Among the important concepts of STR, there is a concept of event. The **event** is an arbitrary physical phenomenon that occurs at a definite point in the space and at a definite time moment (a light flash, a shot, an elementary particle decay, and so forth). The point coordinates and the time moment are called the **event coordinates**. In the stationary reference frame K, these are (x, y, z, t), and in the moving frame K', these are $(x'.y', z', t')$. The event coordinates in the frame K are connected with the event coordinates in the frame K' by the formulas of coordinate transformation. This is not the Galilean transformation (9.1), because the latter does not agree with the STR postulates. Therefore, we have to find a proper coordinate transformation, which would agree with them.

From Einstein's postulates, it follows that space and time in every IRF are homogeneous, i.e. all points of in the space and all time moments are equivalent. Therefore, the relationships between the event coordinates in the frames K and K' must be linear and look like

$$x\prime = a_1 x + a_2 y + a_3 z + a_4 t + a_0,$$

where a_i are coefficients independent of coordinates x, y, z and time t. If the transformation formulas had been non-linear and the coefficients a_i had been dependent on the event coordinates, and it would have meant that coordinates of different space regions and different time moments would have been transformed according to different laws, which contradicts the homogeneity of space and time.

Let the coordinate axes of the reference frames K and K' be parallel, and the axes x and x' coincide (Fig. 9.1). The frame K' moves with respect to the frame K along the axis x at the velocity v. At the moment $t = t' = 0$, the coordinate origins O and O' coincide. Since the axes y, y' and z, z' are parallel and the frame K' moves only along the axis x, we may write

$$y' = y, z' = z, \tag{9.16}$$

i.e. the coordinates y, z are not coupled with the coordinates x', t'. Hence, we may write

$$\begin{aligned} x' &= \alpha_1 x + \alpha_2 t + \alpha_0, & (9.17)\\ t' &= \beta_1 x + \beta_2 t + \beta_0. & (9.18) \end{aligned}$$

At the moment $t = t' = 0$, $x' = x = 0$; therefore, $\alpha_0 = \beta_0 = 0$, and instead of (9.17) and (9.18), we obtain

$$\begin{aligned} x' &= \alpha_1 x + \alpha_2 t, & (9.19)\\ t' &= \beta_1 x + \beta_2 t. & (9.20) \end{aligned}$$

For point O', $x' = 0$ and $x = vt$. Then, from (9.19), it follows that $\alpha_2 = -\alpha_1 v$. Hence,

$$\begin{aligned} x' &= \alpha_1 (x - vt), & (9.21)\\ t' &= \beta_1 x + \beta_2 t. & (9.22) \end{aligned}$$

Suppose that, at the moment $t = t' = 0$, when points O and O' coincide, the lamp flashes. Then, the spherical fronts of light waves start to propagate in the reference frames K and K'. Their equations look like

$$\begin{aligned} x^2 + y^2 + z^2 &= c^2 t^2, & (9.23)\\ x'^2 + y'^2 + z'^2 &= c^2 t'^2. & (9.24) \end{aligned}$$

Substituting x' and t' from (9.21) and (9.22) into (9.24) and taking into account (9.16), we obtain the equation

$$\alpha_1^2(x - vt)^2 + y^2 + z^2 = c^2(\beta_1 x + \beta_2 t)^2, \tag{9.25}$$

which must be equivalent to (9.23). Comparing those two equations, we obtain a system of equations for the unknowns α_1, β_1, and β_2,

$$\begin{aligned} \alpha_1^2 - c^2\beta_1 &= 1, && (9.26)\\ \alpha_1^2 v + c^2\beta_1\beta_2 &= 0, && (9.27)\\ -\alpha_1^2 v + c^2\beta_2^2 &= c^2. && (9.28) \end{aligned}$$

Its solution is

$$\begin{aligned} \alpha_1 &= \beta_2 = \frac{1}{\sqrt{1 - \dfrac{v^2}{c^2}}}, && (9.29)\\ \beta_1 &= -\frac{v}{c^2}\frac{1}{\sqrt{1 - \dfrac{v^2}{c^2}}}. && (9.30) \end{aligned}$$

Substituting coefficients (9.29) and (9.30) into (9.21) and (9.22) and taking into account (9.16), we obtain the following formulas for the transformation of event coordnates:

$$\begin{aligned} x' &= \frac{x - vt}{\sqrt{1 - \dfrac{v^2}{c^2}}}, && (9.31)\\ y' &= y, && (9.32)\\ z' &= z, && (9.33)\\ t' &= \frac{t - \dfrac{vx}{c^2}}{\sqrt{1 - \dfrac{v^2}{c^2}}}. && (9.34) \end{aligned}$$

Formulas for inverse transformations can be found by solving the system of equations (9.31)–(9.34) to determine the unknowns x, y, z, and t. However, it is easier to use a different way. As the motion has a relative character, we may assume that it is the reference frame K that moves with respect to the frame K' at the velocity $-v$. Therefore, on the basis of (9.31)–(9.34), we may write

$$x = \frac{x' + vt'}{\sqrt{1 - \frac{v^2}{c^2}}}, \tag{9.35}$$

$$y = y', \tag{9.36}$$

$$z = z', \tag{9.37}$$

$$t = \frac{t' + \frac{vx'}{c^2}}{\sqrt{1 - \frac{v^2}{c^2}}}. \tag{9.38}$$

Formulas (9.31)–(9.38) describing the transformation of event coordinates are called the **Lorentz transformations**. They were derived by H. Lorentz in 1904, but in the framework of classical concepts for space and time. A. Einstein, who obtained those transformations in the framework of STR on the basis of postulates formulated by himself, gave them physical meaning.

It is easy to be convinced that, if $v \ll c$, the Lorentz transformation transforms into Galilean transformation (9.1). This fact agrees with the **correspondence principle,** according to which a new theory describing the physical reality in more details should take into account the previous theory as the limiting case.

The choice of directions for the coordinate axes, which is shown in Fig. 9.1, does not make the scope of Lorentz transformations applicability narrower, because only the relative velocity of reference frames $\vec{v}$ has a physical meaning. For any other mutual orientation of axes and any other choice of coordinate origins O and O', it is always possible to apply known geometrical transformations (translations and rotations of coordinate systems) in order to reduce the specific problem to that considered above.

In section 9.2, it was noticed that according to the principle of relativity, the laws of the nature must be covariant with respect to the transformations of coordinates and time if the inertial reference frame changes. Now, we may assert that such transformations are the Lorentz transformations. Hence, the physical laws must be expressed in the form that is covariant with respect to the Lorentz transformations, i.e. this form must not change at transformations (9.31)–(9.38).

9.5. Time dilation in moving inertial reference frames. Length contraction of moving objects

The Lorentz transformations concern the spatial and temporal coordinates of events. At the same time, physical objects and phenomena are characterized by spatial and temporal intervals. Therefore, let us find the transformation formulas for them.

Let two events occur in the moving reference frame K' at the same point with the coordinate x' but at the time moments t'_1 and t'_2 measured by the clock attached to the frame. The time measured by a clock attached to the given reference frame is called the **proper time**. Therefore, the difference $\Delta t' = t'_2 - t'_1$ is the interval of proper time between two events. Let us determine the time interval Δt between the same events, but in the stationary reference frame K, where they occur at the moments t_1 and t_2. Since the coordinate x' does not change, we may use the Lorentz transformation (9.38), according to which

$$\Delta t = t_2 - t_1 = \frac{t'_2 + \dfrac{vx'}{c^2}}{\sqrt{1 - \dfrac{v^2}{c^2}}} - \frac{t'_1 + \dfrac{vx'}{c^2}}{\sqrt{1 - \dfrac{v^2}{c^2}}} = \frac{\tau}{\sqrt{1 - \dfrac{v^2}{c^2}}}. \tag{9.39}$$

From (9.39), it follows that the interval of proper time, i.e. the time interval measured using the clock attached to the point where the events occur, is minimum.

The time measurement using a clock is based on periodic processes running in it, such as pendulum oscillations in mechanical clocks, elastic vibrations of piezocrystal in quartz ones, intratomic and intramolecular oscillations in atomic clocks. The time interval between two events is evaluated as a number of oscillations (vibrations) between them. The reduction of time interval means the reduction in the number of oscillations, i.e. the growth of oscillation period or the dilation of clock rate (the time rate). Hence, relation (9.39) means the dilation of time rate in moving reference frames.

Let two events occur in the frame K at the point with the coordinate x at the time moments t_1 and t_2. Then, the proper time interval between the events equals $\tau = t_2 - t_1$, whereas in the frame K', in which the frame K moves at the velocity $-v$, the time interval, in accordance with (9.34), equals

$$\Delta t' = t'_2 - t'_1 = \frac{t_2 - \dfrac{vx}{c^2}}{\sqrt{1 - \dfrac{v^2}{c^2}}} - \frac{t_1 - \dfrac{vx}{c^2}}{\sqrt{1 - \dfrac{v^2}{c^2}}} = \frac{\tau}{\sqrt{1 - \dfrac{v^2}{c^2}}}, \tag{9.40}$$

i.e. we obtain formula (9.39) once more, which testifies to the time dilation in moving reference frames.

There is some experimental evidence that the time rate changes in moving reference frames. Let us consider one of them. Among elementary particles, there are muons. Under laboratory conditions, it was found that they are unstable, and the average characteristic time of their existence τ equals 2.2×10^6 s. For this time interval, muons, even when moving at the light speed, can pass the path length $h = 660$ m. At the same time, it was found that muons are generated in the upper atmosphere, at the heights $H = 5 \div 6$ km, in nuclear reactions occurring under the influence of cosmic rays, but can be registered at sea level. In other words, muons pass a way of 6000 m in length. This fact can be explained if we take into account that the average time of their existence in the reference frame K (the Earth) is $\Delta t = H/c = 2 \times 10^{-5}$ s, i.e. $\Delta t/\tau \geq 10$. According to (9.39), this means that the velocity of muons $v \geq 0.995c$.

Now, let a rod be immovably fixed along the axis x' in the moving IRF K'. The coordinates of rod's ends are x'_1 and x'_2, i.e. the rod length if $l_0 = x'_2 - x'_1$. The length of the rod l_0 in the reference frame, where it is at rest (in the **proper reference frame**), is called the **proper length**. Let us determine the rod length in the reference frame K, where it moves at the velocity v. For this purpose, it is necessary to register the coordinates x_1 and x_2 of the moving rod's ends simultaneously, at the same moment t. Then, its length can be determined as $l = x_2 - x_1$.

Since the coordinates x_1 and x_2 are registered simultaneously, in accordance with (9.31), we obtain

$$l_0 = x'_2 - x'_1 = \frac{x_2 - vt}{\sqrt{1 - \dfrac{v^2}{c^2}}} - \frac{x_1 - vt}{\sqrt{1 - \dfrac{v^2}{c^2}}} = \frac{l}{\sqrt{1 - \dfrac{v^2}{c^2}}}, \tag{9.41}$$

so that

$$l = l_0\sqrt{1 - \frac{v^2}{c^2}}. \tag{9.42}$$

Since $l_0 < l$, we come to the conclusion that the dimensions of moving objects become shorter in the direction of their motion (the **length contraction**).

As a result of motion relativity, the length contraction of moving objects also has a relative character. The reference frame K moves at the velocity $-v$ in the frame K'. Let a rod with the proper length $l_0 = x_2 - x_1$ be immovably fixed along the axis x in the frame K. In order to determine its length in the frame K', where it moves, we must simultaneously register the coordinates of its ends x'_1 and x'_2 at the moment t'. Then, in accordance with (9.35),

$$l_0 = x_2 - x_1 = \frac{x'_2 + vt}{\sqrt{1 - \dfrac{v^2}{c^2}}} - \frac{x'_1 + vt}{\sqrt{1 - \dfrac{v^2}{c^2}}} = \frac{l}{\sqrt{1 - \dfrac{v^2}{c^2}}}, \tag{9.43}$$

i.e. we obtain relation (9.42) again.

The contraction of the dimensions of moving objects is a kinematic effect associated only with the relative motion rather than deformations induced by internal stresses. Therefore, the issue concerning the “true’ length of the rod has no meaning: in every IRF, the rod length determined following the method described above is valid. If the rod is at rest, its true length equals its proper length l_0, and if it moves at the velocity v, the true length equals $l_0\sqrt{1 - v^2/c^2}$.

The experimental observation of length contraction for moving objects is rather complicated. The essence is associated with the registration on a film of photons that simultaneously reach it from different object regions and, therefore, is macroscopic. However, for the length contraction by 1% to take place, the object has to move at the velocity $v \approx 42000$ m/s. Such velocities under laboratory conditions are accessible only to elementary particles. There are also other difficulties, which are omitted.

At the same time, the change of time intervals in moving IRFs, which also follows from the Lorentz transformations, has experimental proofs. Therefore, there is no ground to doubt in the validity of the conclusion concerning the length contraction of moving objects. In particular, from this viewpoint, we can explain why muons can reach the Earth’s surface. For this purpose, let as attach the reference frame K' to a muon. In this reference frame, the frame K attached to the Earth moves at the muon's velocity v. As a result, the distance from the muon to the Earth, H, becomes $\sqrt{1-v^2/c^2}$ times shorter, and equals $H\sqrt{1-v^2/c^2}$ in the reference frame of muon. The muon passes this distance for the time interval that, at the velocity $v = 0.995c$, is equal to 2×10^{-6} s, i.e. it does not exceed the characteristic time of muon existence.

9.6. Velocity addition law in the special theory of relativity

The law of velocity addition (9.2) in classical mechanics does not agree with Einstein's postulates. Therefore, to find the law of velocity addition in the STR, let us use the Lorentz transformations. The components of the velocity vector of material point are defined as the derivatives of its coordinates in the given reference frame with respect to the time in this reference frame, i.e., in the reference frame K, as

$$u_x = \frac{dx}{dt}, u_y = \frac{dy}{dt}, u_z = \frac{dz}{dt}, \tag{9.44}$$

and, in the reference frame K', as

$$u'_x = \frac{dx'}{dt'}, u'_y = \frac{dy'}{dt'}, u'_z = \frac{dz'}{dt'}. \tag{9.45}$$

Using Lorentz transformations (9.31)–(9.33), we obtain

$$u'_x = \frac{dx'}{dt'} = \frac{dx'}{dt}\frac{dt}{dt'} = \frac{u_x - v}{\sqrt{1-\dfrac{v^2}{c^2}}}\frac{dt}{dt'}, \tag{9.46}$$

$$u'_y = \frac{dy'}{dt'} = \frac{dy}{dt'} = \frac{dy}{dt}\frac{dt}{dt'} = u_y\frac{dt}{dt'}, \tag{9.47}$$

$$u'_z = \frac{dz'}{dt'} = \frac{dz}{dt'} = \frac{dz}{dt}\frac{dt}{dt'} = u_z\frac{dt}{dt'}. \tag{9.48}$$

In the right hand sides of relations (9.46)–(9.48), there must be quantities inherent to the frame K. Therefore, the derivative dt/dt' is determined using (9.34),

$$\frac{dt}{dt'} = \frac{1}{\dfrac{dt'}{dt}} = \frac{\sqrt{1-\dfrac{v^2}{c^2}}}{1-\dfrac{vu_x}{c^2}}. \tag{9.49}$$

Substituting (9.49) into (9.46)–(9.48), we obtain the law of velocity addition for the transition from the reference frame K to the reference frame K',

$$u'_x = \frac{u_x - v}{1-\dfrac{vu_x}{c^2}}, \tag{9.50}$$

$$u'_y = u_y\frac{\sqrt{1-\dfrac{v^2}{c^2}}}{1-\dfrac{vu_x}{c^2}}, \tag{9.51}$$

$$u'_z = u_z\frac{\sqrt{1-\dfrac{v^2}{c^2}}}{1-\dfrac{vu_x}{c^2}}. \tag{9.52}$$

The formulas for the inverse transition from the frame K' to the frame K can be obtained by changing v to $-v$ in (9.50)–(9.52) and swapping u_x, u_y, and u_z with u'_x, u'_y, and u'_z, respectively,

$$u_x = \frac{u'_x + v}{1+\dfrac{vu'_x}{c^2}}, \tag{9.53}$$

$$u_y = u'_y\frac{\sqrt{1-\dfrac{v^2}{c^2}}}{1+\dfrac{vu'_x}{c^2}}, \tag{9.54}$$

$$u_z = u'_z\frac{\sqrt{1-\dfrac{v^2}{c^2}}}{1+\dfrac{vu'_x}{c^2}}. \tag{9.55}$$

The relativistic law of velocity addition enables us to find a method for the calculation of relative velocities in the STR. Suppose that two point move at the velocities v_1 and v_2 along the axis x in the stationary reference frame K. Another, moving reference frame K' is attached to point 2, which moves at the velocity v_2. In this case, $u_{1x} = v_1$, $v - v_2$, and the velocity u'_{1x} of point 1 in the frame K' is its relative velocity with respect to point 2, i.e. $u'_{1x} = v_{rel}$. Applying formula (9.50), we obtain

$$v_{rel} = u'_{1x} = \frac{u_{1x} - v}{1 - \dfrac{u_{1x}v}{c^2}} = \frac{v_1 - v_2}{1 - \dfrac{v_1 v_2}{c^2}}. \tag{9.56}$$

In general, we have to write

$$v_{rel} = \frac{v_1 + v_2}{1 \pm \dfrac{v_1 v_2}{c^2}}, \tag{9.57}$$

where the sign "plus" and "minus" correspond to oppositely and equally, respectively, directed velocities v_1 and v_2.

While analyzing the formulas of velocity transformation, notice first of all that, at low velocities, when, formally, $c \to \infty$, we obtain Galilean transformation (9.2) and the relative velocity of classical mechanics, $v_{rel} = v_1 + v_2$.

Suppose that a beam of light propagates at the velocity $u'_x = c$ along the axis x' in the reference frame K'. Let us determine its velocity in the frame K. According to (9.53) and the second Einstein's postulate, we obtain

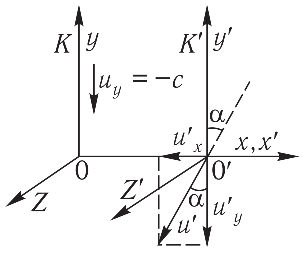

Fig. 9.6.

$$u_x = \frac{c + v}{1 + \dfrac{v}{c}} = c.$$

If the light propagates at the velocity c in the reference frame K ($u_x = c$), its velocity in the reference frame K' also equals c,

$$u'_x = \frac{c - v}{1 - \dfrac{v}{c}} = c.$$

Using the relativistic law of velocity addition, we can explain the starlight aberration and the results of Fizeau experiment (see section 9.1).

In order to explain the annual aberration of starlight, let us consider a stationary star and an attached reference frame K with the axis y directed to the star (Fig. 9.6). Then, in the reference frame K, $u_x = u_z = 0$ and $u_y = -c$. The reference frame K' is attached to the Earth moving round the Sun at the velocity v. In the Earth's reference frame,

$$\begin{aligned} u'_x &= \frac{u_x - v}{1 - \dfrac{u_x v}{c^2}} = -v, \\ u'_y &= u_y \frac{\sqrt{1 - \dfrac{v^2}{c^2}}}{1 - \dfrac{u_x v}{c^2}} = -c\sqrt{1 - \frac{v^2}{c^2}}, \\ u'_z &= u_z \frac{\sqrt{1 - \dfrac{v^2}{c^2}}}{1 - \dfrac{u_x v}{c^2}} = 0. \end{aligned}$$

Hence, in the Earth's reference frame, the direction of light velocity u' forms the angle α with the direction to the star (see Fig. 9.6), with

$$\tan\alpha = \frac{u'_x}{u'_y} = \frac{\dfrac{v}{c}}{\sqrt{1 - \dfrac{v^2}{c^2}}}.$$

Taking into account that $v/c \ll 1$, we obtain, in accordance with (9.5),

$$\tan\alpha = \frac{v}{c}.$$

To explain the Fizeau experiment, suppose that the stationary reference frame K is the laboratory one, and the reference frame K' moves with water at the velocity v. The speed of light in still water $u'_x = c/n$, where n is the refractive index of water. Therefore, the speed of light in moving water in the laboratory reference frame, in accordance with (9.53), equals

$$u_x = \frac{u'_x + v}{1 + \dfrac{u'_x v}{c^2}} = \frac{\dfrac{c}{n} + v}{1 + \dfrac{v}{cn}} = \left(\frac{c}{n} + v\right)\left(1 + \frac{v}{cn}\right)^{-1}.$$

Expanding the multiplier $\left(1+\frac{v}{cn}\right)^{-1}$ in a series according to the binomial formula and confining the expansion to the term linear in $\frac{v}{cn}$, we obtain

$$\left(1+\frac{v}{cn}\right)^{-1}=1-\frac{v}{cn}. \tag{9.58}$$

Hence,

$$u_x=\left(\frac{c}{n}+v\right)\left(1-\frac{v}{cn}\right)\approx\frac{c}{n}+v-\frac{v}{n^2}=\frac{c}{n}+v\left(1-\frac{1}{n^2}\right), \tag{9.59}$$

i.e. we obtain formula (9.6).

9.7. Space-time interval

Let two events with the coordinates (x_1, y_1, z_1, t_1) and (x_2, y_2, z_2, t_2) occur in the reference frame K. Then, the quantity

$$S_{12}=\sqrt{(x_2-x_1)^2+(y_2-y_1)^2+(z_2-z_1)^2-c\,(t_2-t_1)^2} \tag{9.60}$$

is called the **space-time interval** between the events. It is a generalization of the space interval

$$l_{12}=\sqrt{(x_2-x_1)^2+(y_2-y_1)^2+(z_2-z_1)^2} \tag{9.61}$$

and the time one

$$t_{12}=t_2-t_1. \tag{9.62}$$

In the reference frame K', the space-time interval between the same events looks like

$$S'_{12}=\sqrt{(x'_2-x'_1)^2+(y'_2-y'_1)^2+(z'_2-z'_1)^2-c\,(t'_2-t'_1)^2} \tag{9.63}$$

In classical mechanics, the space, l_{12}, and time, t_{12}, intervals are separately invariant with respect to the Galilean transformations. In the STR, they are not invariant with respect to the Lorentz transformations, but the space-time interval S_{12} (below, the interval) is. Really, with the help of Lorentz transformations (9.31)–(9.34), instead of (9.63), we obtain the expression

$$\begin{aligned} S'_{12} &= \left\{ \frac{(x_2 - vt_2 - x_1 + vt_1)^2}{1 - \dfrac{v^2}{c^2}} + (y_2 - y_1)^2 \right. \\ &\quad \left. + (z_1 - z_2)^2 - c^2 \frac{\left(t_2 - \dfrac{vx_2}{c^2} - t_1 + \dfrac{vx_1}{c^2}\right)^2}{1 - \dfrac{v^2}{c^2}} \right\}, \end{aligned}$$

whence, after simple calculations, we obtain

$$S'_{12} = S_{12}. \tag{9.64}$$

Using (9.61) and (9.62), the square of interval (9.60) can be written as

$$S_{12}^2 = l_{12}^2 - c^2 t_{12}^2. \tag{9.65}$$

If events occur at infinitely close points in an infinitesimal time interval, the square of interval takes the form

$$dS_{12}^2 = dx^2 + dy^2 + dz^2 - c^2 dt^2. \tag{9.66}$$

Depending on the summands in the right hand side of relation (9.65), the square of the interval can be negative, positive, and equal to zero. Let us analyze those cases.

If we have

$$S_{12}^2 < 0, \tag{9.67}$$

the interval is called **time-like** and, in accordance with (9.65),

$$l_{12}^2 - c^2 t_{12}^2 < 0. \tag{9.68}$$

In this case, we can choose a frame K', in which both events occur at the same point ($l'_{12} = 0$). To prove this statement, we have to demonstrate that the interval remains time-like in the frame K' provided that $l'_{12} = 0$. Really, according to the interval invariance,

$$S_{12}^2 = l_{12}^2 - c^2 t_{12}^2 = S'^2_{12} = l'^2_{12} - c^2 t'^2_{12}. \tag{9.69}$$

Therefore, at $l'_{12} = 0$,

$$S'^2_{12} = -c^2 t'^2_{12} < 0.$$

Let us determine the velocity v of frame K', in which the events occur at the same point (the compatible events). The **event** is a physical phenomenon that occurs at a definite point in space; this point should be regarded as a material point that moves in the frame K at the velocity u. For the events attached to a material point in the frame K' to be compatible, it is enough that the frame K' should be the proper (accompanying) reference frame for the material point, i.e. the condition $v = u$ should be satisfied.

In the case of time-like interval between events, the latter can be causally linked. The causal relationship between events implies that one of them must precede the other in every IRF, i.e. if $t_{12} = t_2 - t_1 > 0$, then, $t'_{12} = t'_2 - t'_1 > 0$ as well. In order to prove that the time-like interval ensures this requirement, let us write a relation that follows from (9.34),

$$\Delta t' = \frac{\delta t - v\Delta x/c^2}{\sqrt{1 - \dfrac{v^2}{c^2}}} = \left(1 - \frac{v\Delta x}{c^2 \Delta t}\right) \frac{\Delta t}{\sqrt{1 - \dfrac{v^2}{c^2}}}. \tag{9.70}$$

Whence it follows that

$$t'_2 - t'_1 = \left(1 - \frac{v\Delta x}{c^2 \Delta t}\right) \frac{t_2 - t_1}{\sqrt{1 - \dfrac{v^2}{c^2}}}.$$

In the case of time-like interval, $l_{12} < ct_{12} = c\Delta t$; hence, $\Delta x < c\Delta t$, i.e. $\Delta x < (c\Delta t) < 1$. Taking also into account that $v/c < 1$, we obtain the relation

$$\frac{v\Delta x}{c^2 \Delta t} < 1.$$

Hence, if $t_2 - t_1 > 0$, we have $t'_2 - t'_1 > 0$, i.e. the sequence of events keeps its order in every IRF.

An opportunity of causal relationship between events in the case of time-like interval also follows immediately from relation (9.68). Having rewritten it in the form $l_{12} < ct_{12}$, we come to a conclusion that, for the time interval t_{12} between the events, the light signal travels the distance l_{12} between the points where they occur. Hence, the first event causes the second one.

If we have

$$S_{12}^2 > 0, \tag{9.71}$$

the interval is called **space-like**. From (9.69), it follows that, in the case $t'_{12} = 0$, condition (9.71) holds true,

$$l_{12}^2 - c^2 t_{12}^2 = l'^2_{12} > 0. \tag{9.72}$$

Therefore, we can choose a frame K', where the events occur simultaneously. According to (9.70), such a reference frame has to move at the velocity

$$v = \frac{c^2 \Delta t}{\Delta x}$$

and satisfy the condition $v < c$. Really, since in this case, $l_{12} < ct_{12}$, i.e. $\Delta x > c\Delta t$ or $c\Delta t/\Delta x < 1$, we obtain

$$v = c\frac{c\Delta t}{\Delta x} < c.$$

Provided that $c\Delta t/\Delta x < 1$, the quantity $1 - \dfrac{v}{c}\dfrac{\Delta x}{c\Delta t}$ in (9.70) can be negative, However, then, if $\Delta t > 0$, we obtain $\Delta t' < 0$, i.e. the sequence of events in time can be inversed when changing to the other IRF. Therefore, the events coupled by a space-like interval cannot form the causal link. This conclusion also follows from the condition $l_{12} < ct_{12}$, because the light signal cannot travel the distance l_{12} for the time interval t_{12}.

In the case

$$S_{12}^2 = 0,$$

we obtain the equation $l_{12}^2 = c^2 t_{12}^2$, which describes events associated with the consecutive passage of light wave through different points in space, i.e. the reception and emission of light signal. In this case, some authors call the interval S_{12} **light-like**.

The results obtained in this section can be given a geometrical interpretation, which, in turn, enables one to simplify calculations in a number of cases.

Let event 1 occurs at the coordinate origin, i.e. at the point $x_1 = y_1 = z_1 = 0$ and the time moment $t_1 = 0$. The coordinates of event 2 are denoted as (x, y, z, t). Then, provided that $S_{12}^2 = 0$, we obtain the equation

$$x^2 + y^2 + z^2 - c^2 t^2 = 0, \tag{9.73}$$

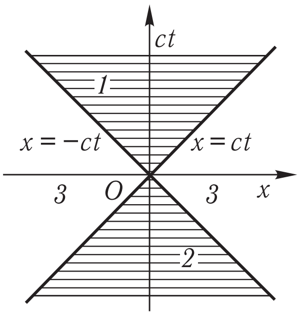

Fig. 9.7.

which describes the propagation of the spherical front of light wave from the coordinate origin O. From the geometrical point of view, equation (9.73) is the equation of a conical surface with the vertex at the point $x = y = z = ct = 0$ in the four-dimensional space with the coordinates (x, y, z, ct). Its cross-section in the plane xz is shown in Fig. 9.7. The region inside the conical surface (it is hatched in Fig. 9.7) is called the **light cone**. The surface of light cone divides the space and time into three regions marked as 1, 2, and 3 in Fig. 9.7.

Regions 1 and 2 correspond to the condition $x^2 + y^2 + z^2 < c^2t^2$, i.e. $S_{12}^2 < 0$, so that the events in this regions can be causally linked. In region 1, $t > 0$, i.e. event 2 occurs after event 1. Therefore, the points (events) in region 1 are called the **absolute future**, because they can be a consequences of the event at point O. For points in region 2, $t < 0$. Therefore, events of this region occur before the event at point O. In this connection, region 2 is called the **absolute past**, because events from this region can cause an event at point O.

Hence, an observer at point O can know about events in region 2 and affect events in section 1. However, he cannot know about events in region 3 and, respectively, affect them.

9.8. Four-dimensional Minkowski space-time

The principle of relativity demands that, in the case of Lorentz transformations, no physical law should change its form. It is possible if all the summands in the corresponding equation are tensors of the same rank, which change identically at Lorentz transformations. At the same time Lorentz transformations (9.31)–(9.38) are non-symmetric with respect to spatial coordinates and time, i.e. the forms of their transformations are different. To make the Lorentz transformations symmetric and formulate the transformation rules for tensor components, we add the coordinate x_4 associated with the time to three spatial coordinates designates as $x_1 = x$, $x_2 = y$, and $x_3 = z$.

To determine the form of x_4, let us use space-time interval (9.66)

$$dS^2 = dx^2 + dy^2 + dz^2 - c^2dt^2.$$

Expressing it in the symmetric form

$$dS^2 = \sum_{i=1}^{4} dx_i^2, \quad x_1 = x,\ x_2 = y,\ x_3 = z,$$

we see that the fourth coordinate looks like $x_4 = ict$. It is imaginary, but, as a result of its formal introduction, this fact does not affect the physical conclusions drawn from calculations where this coordinate appears.

With the help of coordinates x_1, x_2, x_3, x_4, Lorentz transformations (9.31)–(9.34) can be written as follows ($\beta = v/c$):

$$\begin{aligned}
x'_1 &= \frac{1}{\sqrt{1-\beta^2}}\,x_1 + 0\times x_2 + 0\times x_3 + \frac{i\beta}{\sqrt{1-\beta^2}}\,x_4,\\
x'_2 &= 0\times x_1 + 1\times x_2 + 0\times x_3 + 0\times x_4,\\
x'_3 &= 0\times x_1 + 0\times x_2 + 1\times x_3 + 0\times x_4,\\
x'_4 &= \frac{-i\beta}{\sqrt{1-\beta^2}}\,x_1 + 0\times x_2 + 0\times x_3 + \frac{1}{\sqrt{1-\beta^2}}\,x_4.
\end{aligned} \tag{9.74}$$

Transformations (9.74) can be written in a compact form as

$$x'_i = \sum_{j=1}^{4} a_{ij}x_j, \quad i = 1, 2, 3, 4; \tag{9.75}$$

where the transformation coefficients a_{ij} compose the matrix

$$a_{ij} = \begin{pmatrix} \frac{1}{\sqrt{1-\beta^2}} & 0 & 0 & \frac{i\beta}{\sqrt{1-\beta^2}} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \frac{-i\beta}{\sqrt{1-\beta^2}} & & & \frac{1}{\sqrt{1-\beta^2}} \end{pmatrix}. \tag{9.76}$$

In the symmetric notation, the inverse Lorentz transformations (9.35)–(9.38) look like

$$x_i = \sum_{j=1}^{4} a_{ji}x'_j, \quad i = 1, 2, 3, 4. \tag{9.77}$$

It is easy to be convinced that the coefficients a_{ij} satisfy the relations

$$a_{ij} = a^*_{ji}, \tag{9.78}$$

$$\sum_j a_{ji}a_{jk} = \sum_j a^*_{ij}a_{jk} = \delta_{ik} = \begin{cases} 1 \text{ if } i = k \\ 0 \text{ if } i \neq k \end{cases} . \tag{9.79}$$

Transformations that satisfy condition (9.79) are called **orthogonal**.

Transformation of coordinates x_1, x_2, x_3, x_4 (9.75) and (9.77) are similar by their form to the transformations of coordinates in a real three-dimensional space corresponding to the rotation of coordinate system and looking like

$$x'_i = \sum_{j=1}^{3} a_{ij}x_j, \quad i = 1, 2, 3. \tag{9.80}$$

In particular, for the rotation of reference frame around the axis $x_3 = z$ by the angle φ, the matrix of coefficients a_{ij} looks like

$$a_{ij} = \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} . \tag{9.81}$$

Therefore, the coordinate x_4 in the Lorentz transformations can formally be regarded as the fourth coordinate in the abstract four-dimensional space (the 4-space introduced by the German physicist Minkowski), whereas the Lorentz transformations themselves as a result of the reference frame rotation in the 4-space.

9.9. Scalars, vectors, and tensors in the 4-space

The application of four-dimensional Minkowski space-time enables the physical laws to be formulated in the relativistically covariant form. However, we must first define the concepts of scalar, vector, and tensor quantities in the four-dimensional space. For this purpose, let us recall their definitions in the real three-dimensional space.

By **a scalar** is meant a quantity that remains invariant at the rotation of reference frame (below, the coordinate transformation).

The coordinates $x_1 = x$, $x_2 = y$, and $x_3 = z$ of three-dimensional space can be considered as the components of radius vector $\vec{r}$, which are transformed according to law (9.80). Therefore, by **a vector** $\vec{A}$ is meant a set of three quantities (components) A_i $(i = 1, 2, 3)$, which, in the case of coordinate transformation, are transformed according to the law

$$A'_i = \sum_{j=1}^{3} \alpha_{ij} A_j, \quad (i = 1, 2, 3). \tag{9.82}$$

By **a tensor of the second rank** are meant nine components A_{ij} $(i, j = 1, 2, 3)$, which, in the case of coordinate transformation, are transformed according to the law

$$A'_{ij} = \sum_{k,l=1}^{3} \alpha_{ik} \alpha_{jl} A_{kl}, \quad (i, j = 1, 2, 3). \tag{9.83}$$

In the same way, we can define the tensors of the third and higher ranks.

Formally speaking, all physical quantities are tensors of the n-th rank with a number of components equal to 3^n. If $n = 0$, the quantity is a tensor of the zeroth rank consisting of one component, it is called the **scalar**. If $n = 1$, we have a tensor of the first rank consisting of three components, and it is called the **vector**. In the case $n = 2$, the quantity is a tensor of the second rank consisting of nine components. And so on.

In the 4-space, the coordinate transformations are Lorentz transformations (9.75). By **a scalar** (a 4-tensor of the zero rank) is meant a quantity invariant with respect to Lorentz transformations (9.75). By **a vector** in the 4-space (a 4-vector, a 4-tensor of the first rank), we mean a set of four ($4^n = 4^1 = 4$) quantities (components) A_i $(i = 1, 2, 3, 4)$, which, in the case of Lorentz transformations are transformed according to the law

$$A'_i = \sum_{j=1}^{4} \alpha_{ij} A_j, \quad (i = 1, 2, 3, 4), \tag{9.84}$$

i.e.

$$A'_1 = \frac{A_1 + i\beta A_4}{\sqrt{1-\beta^2}}, A'_2 = A_2, A'_3 = A_3, A'_4 = \frac{A_4 - i\beta A_1}{\sqrt{1-\beta^2}}. \tag{9.85}$$

The scalar product of two 4-vectors is the sum

$$\sum_{i=1}^{4} A_i B_i,$$

which is invariant with respect to the Lorentz transformations. Really, taking into account (9.79), we obtain

$$\sum_{i=1}^{4} A'_i B'_i = \sum_{i,j,k=1}^{4} a_{ij} a_{ik} A_j B_k = \sum_{j,k=1}^{4} \delta_{jk} A_j B_k = \sum_{j=1}^{4} A_j B_j. \tag{9.86}$$

Putting $B_i = A_i$ in (9.86), we come to the conclusion that the **square of 4-vector is invariant**.

By **a tensor of the second rank** in the 4-space (a 4-tensor of the second rank) is meant the set of sixteen components ($4^2 = 16$) A_{ij} ($i, j = 1, 2, 3, 4$), that, in the case of Lorentz transformations (9.75), are transformed according to the law

$$A'_{ij} = \sum_{k,l=1}^{4} \alpha_{ik} \alpha_{jl} A_{kl}. \tag{9.87}$$

The 4-tensors of the higher ranks are defined analogously.

The inverse transformations of the 4-vector and 4-tensor components look like

$$A_i = \sum_{j=1}^{4} \alpha_{ji} A'_j, \tag{9.88}$$

$$A_{ij} = \sum_{k,l=1}^{4} \alpha_{ki} \alpha_{lj} A'_{kl}. \tag{9.89}$$

10

RELATIVISTIC MECHANICS

10.1. 4-velocity and 4-acceleration

The laws of classical mechanics are covariant with respect to the Galilean transformations, but non-covariant with respect to the Lorentz ones. Hence, the former are inapplicable at high velocities of object motion.

This circumstance brings us to a necessity of establishing relativisticall covariant laws of mechanics, which transform into the laws of classical mechanics at $v \ll c$. This task can be fulfilled if we formulate those laws in terms of 4-vectors and 4-tensors in the four-dimensional Minkowski space. Therefore, it is necessary first to find the form for the 4-vectors of velocity and acceleration for a particle (a material point).

In general, the components of particle velocity are defined as the derivatives of particle coordinates with respect to the time. Therefore, the components u_i of 4-velocity has to be the time derivatives of coordinates x_i $(i = 1, 2, 3, 4)$ and transform in the case of Lorentz transformations according to law (9.75). However, the derivatives dx_i/dt do not satisfy this requirement, because dt is not invariant.

At the same time, the interval $d\tau$ of proper time, i.e. of the time measured by the clock in the reference frame K' attached to the particle, is invariant. In this system, events occur at the same point; therefore, the invariant square of space-time interval dS^2 looks like $dS^2 = -c^2 d\tau^2$; hence, $d\tau = idS/c = \text{const}$. The interval $d\tau$ of proper time is connected with the time interval dt in the frame K by the relationship $dS^2 = dl^2 - c^2dt^2 = -c^2d\tau^2$, whence it follows that

$$d\tau = dt\sqrt{1 - \frac{v^2}{c^2}}. \tag{10.1}$$

Hence, the 4-velocity of the particle is defined as

$$u_i = \frac{dx_i}{d\tau}, \quad (i = 1, 2, 3, 4), \tag{10.2}$$

and, making allowance for (10.1), as

$$u_1 = \frac{dx_1}{d\tau} = \frac{v_x}{\sqrt{1-\beta^2}}, \tag{10.3}$$

$$u_2 = \frac{dx_2}{d\tau} = \frac{v_y}{\sqrt{1-\beta^2}}, \tag{10.4}$$

$$u_3 = \frac{dx_3}{d\tau} = \frac{v_z}{\sqrt{1-\beta^2}}, \tag{10.5}$$

$$u_4 = \frac{dx_4}{d\tau} = \frac{ic}{\sqrt{1-\beta^2}}, \tag{10.6}$$

where $\beta = v/c$ and $v^2 = v_x^2 + v_y^2 + v_z^2$. These components can be written in the form

$$\{u_i\} = \left\{ \frac{\vec{v}}{\sqrt{1-\beta^2}}, \frac{ic}{\sqrt{1-\beta^2}} \right\}, \tag{10.7}$$

where the braces mean the set of four components. It is easy to demonstrate that

$$\sum_{i=1}^{4} u_i^2 = -c^2, \tag{10.8}$$

i.e., according to the conclusion about the invariance of 4-vector obtained in section 9.8, the square of 4-velocity is invariant.

The acceleration is defined as the derivative of velocity with respect to the time. Therefore, let us construct the 4-vector of acceleration similarly to the 4-vector of velocity in the form

$$w_i = \frac{du_i}{d\tau}, \quad (i = 1, 2, 3, 4). \tag{10.9}$$

Calculating w_1, we obtain

$$w_1 = \frac{du_1}{d\tau} = \frac{1}{\sqrt{1-\beta^2}} \frac{d}{dt} \frac{v_x}{\sqrt{1-\beta^2}} = \frac{\dot{v}_x}{1-\beta^2} + \frac{v_x \left(\vec{v}\dot{\vec{v}}\right)}{c^2 \left(1-\beta^2\right)^2}. \tag{10.10}$$

Here, we took into account that $\beta^2 = v^2/c^2 = \vec{v}^2/c^2$, and the dot means the differentiation with respect to the time.

The components w_2 and w_3 can be found by substituting v_x in (10.10) by v_y and v_z, respectively. For the component w_4, we obtain the expression

$$w_4 = \frac{du_4}{d\tau} = \frac{1}{\sqrt{1-\beta^2}} \frac{d}{dt} \frac{ic}{\sqrt{1-\beta^2}} = \frac{i}{c} \frac{\left(\vec{v}\dot{\vec{v}}\right)}{\left(1-\beta^2\right)^2}. \tag{10.11}$$

As a result, the 4-acceleration looks like

$$\{w_i\} = \left\{ \frac{\dot{\vec{v}}}{1-\beta^2} + \frac{\vec{v}\left(\vec{v}\dot{\vec{v}}\right)}{c^2\left(1-\beta^2\right)^2}, \frac{i}{c} \frac{\left(\vec{v}\dot{\vec{v}}\right)}{\left(1-\beta^2\right)^2} \right\}. \tag{10.12}$$

10.2. Relativistic equation of motion

In classical mechanics, the equation of motion for a material point looks like

$$\frac{d\vec{p}_{cl}}{dt} = \vec{F}, \tag{10.13}$$

where $\vec{p}_{cl} = m_0\vec{v}$ is the classical three-dimensional momentum and $\vec{F}$ the classical three-dimensional force. Therefore, on the basis of reasons given in section 10.1, the relativistic equation of motion for a particle should be written as

$$\frac{dp_i}{d\tau} = \mathscr{F}_i, \quad (i = 1, 2, 3, 4). \tag{10.14}$$

Here, p_i is the i-th components of the 4-vector of momentum (the **4-momentum**), and $\mathscr{F}_i$ the i-th component of the 4-vector of force (the **4-force**), which is also called the **Minkowski force**. Hence, we have to determine the forms of the 4-momentum and Minkowski 4-force.

The 4-momentum components have to be transformed according to law (9.75). Therefore,

$$p_i = m_0 u_i, \quad (i = 1, 2, 3, 4). \tag{10.15}$$

where m_0 is the invariant mass of the particle. This is the mass of the particle in the accompanying reference where it is at rest. Therefore, this parameter is called the **rest mass**. Taking into account (10.7), we obtain

$$\{p_i\} = \left\{ \frac{m_0\vec{v}}{\sqrt{1-\beta^2}}, \frac{im_0c}{\sqrt{1-\beta^2}} \right\}. \tag{10.16}$$

The first three components of 4-momentum,

$$\vec{p} = \frac{m_0\vec{v}}{\sqrt{1-\beta^2}}, \tag{10.17}$$

compose the **relativistic three-dimensional momentum**, in which the quantity

$$m = \frac{m_0}{\sqrt{1-\beta^2}} \tag{10.18}$$

is called the **relativistic mass** of the particle. Taking into account (10.18), the relativistic 3-momentum can be written in the conventional form $\vec{p} = m\vec{v}$, and the 4-momentum in the form

$$\{p_i\} = \left\{ \vec{p}, \frac{i m_0 c}{\sqrt{1-\beta^2}} \right\}. \tag{10.19}$$

Now, let us find the expression for the 4-force. Analogously to the first three components of 4-velocity (see (10.7)) and 4-momentum (see (10.16)), the first three components for the 4-force are written down as

$$\{\mathscr{F}_\alpha\} = \frac{\vec{F}}{\sqrt{1-\beta^2}}, \quad (\alpha = 1, 2, 3), \tag{10.20}$$

where $\vec{F}$ is the classical 3-force. To find the component $\mathscr{F}_4$, let us multiply (10.14) by u_i and sum up the result over the index i,

$$\sum_{i=1}^{4} u_i \frac{dp_i}{d\tau} = \sum_{i=1}^{4} \mathscr{F}_i u_i. \tag{10.21}$$

Taking into account that $p_i = m_0 u_i$, we obtain the relation

$$m_0 \sum_{i=1}^{4} u_i \frac{du_i}{d\tau} = \sum_{i=1}^{4} \mathscr{F}_i u_i.$$

On the other hand, by differentiating (10.8) with respect to the proper time τ, we obtain

$$2 \sum_{i=1}^{4} u_i \frac{du_i}{d\tau} = 0.$$

Therefore,

$$\sum_{i=1}^{4} \mathscr{F}_i u_i = 0$$

or

$$\frac{\vec{F}\vec{v}}{1-\beta^2} + \mathscr{F}_4 \frac{ic}{\sqrt{1-\beta^2}} = 0. \tag{10.22}$$

Whence it follows that

$$\mathscr{F}_4 = \frac{i}{c}\frac{\vec{F}\vec{v}}{\sqrt{1-\beta^2}}. \tag{10.23}$$

Hence, the 4-vector of Minkowski force looks like

$$\{\mathscr{F}_i\} = \left\{\frac{\vec{F}}{\sqrt{1-\beta^2}}, \frac{i}{c}\frac{\vec{F}\vec{v}}{\sqrt{1-\beta^2}}\right\}. \tag{10.24}$$

Let us come back to equations (10.14). Taking into account (10.16) and (10.24), they can be written in the form

$$\frac{d}{d\tau}\left\{\vec{p}, \frac{im_0c}{\sqrt{1-\beta^2}}\right\} = \left\{\frac{\vec{F}}{\sqrt{1-\beta^2}}, \frac{i}{c}\frac{\vec{F}\vec{v}}{\sqrt{1-\beta^2}}\right\}. \tag{10.25}$$

Whence, taking into account that $d\tau = dt\sqrt{1-\beta^2}$, we obtain two equations,

$$\frac{d\vec{p}}{dt} = \vec{F}, \tag{10.26}$$

$$\frac{d}{dt}\frac{m_0c^2}{\sqrt{1-\beta^2}} = \vec{F}\vec{v}. \tag{10.27}$$

Equations (10.26) is called the **relativistic three-dimensional equation of motion** for the particle, in which $\vec{p}$ is the relativistic 3-momentum.

In equation (10.27), the product $\vec{F}\vec{v}$ is the external-force power, i.e. the work done by the external force $\vec{F}$ per unit time. This work equals the variation of particle energy per unit time. Therefore, the quantity

$$E = \frac{m_0c^2}{\sqrt{1-\beta^2}} = mc^2 \tag{10.28}$$

is the energy of a particle that moves at the velocity v, and equation (10.27) is the energy conservation law. For a free particle, it looks like

$$\frac{d}{dt}\frac{m_0c^2}{\sqrt{1-\beta^2}} = \frac{dE}{dt} = 0.$$

Notice that formula (10.28) is applicable to particles with nonzero rest masses.

From (10.28), an important conclusion can be drawn: even if a particle dies not move ($v = 0$) in the given reference frame, its energy differs from zero and equals $E_0 = m_0 c^2$. This energy is called the **rest energy**.

The relativistic kinetic energy of the particle is defined as

$$T = E - E_0 = m_0 c^2 \left(\frac{1}{\sqrt{1-\beta^2}} - 1 \right). \tag{10.29}$$

If $v \ll c$ (the non-relativistic particle),

$$\frac{1}{\sqrt{1-\beta^2}} = \left(1 - \frac{v^2}{c^2}\right)^{-\frac{1}{2}} \approx 1 + \frac{v^2}{2c^2}.$$

Therefore,

$$T = \frac{m_0 v^2}{2}. \tag{10.30}$$

In the case of Lorentz transformations, the component of 4-force are transformed according to rule (9.84), i.e.

$$\mathscr{F}'_i = \sum_{j=1}^{4} a_{ij} \mathscr{F}_j. \tag{10.31}$$

Denoting the velocity of the particle in the reference frame K as $\vec{v}$, in the frame K' as $\vec{v}'$, and the velocity of frame K' in the frame K as $\vec{v}_0$, and taking into account (9.76) and (10.24), we obtain the relations

$$\begin{aligned}
\frac{F'_x}{\sqrt{1-\beta'^2}} &= \frac{F_x - \dfrac{\beta_0}{c}\left(\vec{F}\vec{v}\right)}{\sqrt{1-\beta_0^2}\sqrt{1-\beta^2}}, \\
\frac{F'_y}{\sqrt{1-\beta'^2}} &= \frac{F_y}{\sqrt{1-\beta^2}}, \\
\frac{F'_z}{\sqrt{1-\beta'^2}} &= \frac{F_z}{\sqrt{1-\beta^2}}, \\
\frac{\left(\vec{F}'\,\vec{v}'\right)}{\sqrt{1-\beta'^2}} &= \frac{\left(\vec{F}\vec{v}\right) - \beta_0 c F_x}{\sqrt{1-\beta_0^2}\sqrt{1-\beta^2}}.
\end{aligned}$$

The transformation of the 4-velocity component is reduced to the form

$$\frac{ic}{\sqrt{1-\beta'^2}} = \frac{ic - i\beta_0 v_x}{\sqrt{1-\beta_0^2}\sqrt{1-\beta^2}}. \tag{10.32}$$

Then, we finally obtain

$$\begin{aligned} F'_x &= \frac{F_x - \left(\vec{F}\vec{v}\right) v_0/c^2}{1 - \dfrac{v_x v_0}{c^2}}, \\ F'_y &= \frac{\sqrt{1-\beta_0^2}F_y}{1 - \dfrac{v_x v_0}{c^2}}, \\ F'_z &= \frac{\sqrt{1-\beta_0^2}F_z}{1 - \dfrac{v_x v_0}{c^2}}, \\ \vec{F}'\,\vec{v}' &= \frac{\left(\vec{F}\vec{v}\right) - v_0 F_x}{\sqrt{1-\beta_0^2}\sqrt{1-\beta^2}}. \end{aligned}$$

At velocities much lower than the speed of light, we obtain the classical relations

$$\vec{F}' = \vec{F}, \quad \vec{F}'\,\vec{v}' = \vec{F}\,\vec{v} - v_0 F_x = \vec{F}\left(\vec{v} - \vec{v}_0\right).$$

10.3. 4-vector of energy-momentum

The fourth component of 4-momentum (10.19) is connected with the particle energy E by the relation

$$p_4 = \frac{im_0 c}{\sqrt{1-\beta^2}} = \frac{i}{c}\frac{m_0 c^2}{\sqrt{1-\beta^2}} = \frac{i}{c}E. \tag{10.33}$$

Hence, the 4-momentum can be written in the form

$$\{p_i\} = \left\{\vec{p}, \frac{i}{c}E\right\}. \tag{10.34}$$

Therefore, it is called the **energy-momentum 4-vector**. At the Lorentz transformations, the components of energy-momentum 4-vector transform as follows:

$$p_1' = \frac{p_1 + i\beta p_4}{\sqrt{1-\beta^2}}, p_2' = p_2, p_3' = p_3, p_4' = \frac{-i\beta p_1 + p_4}{\sqrt{1-\beta^2}}. \tag{10.35}$$

For the corresponding three-dimensional quantities, we obtain

$$p_x' = \frac{p_x - \dfrac{1}{c}\beta E}{\sqrt{1-\beta^2}}, p_y' = p_y, p_z' = p_z, E' = \frac{E - c\beta p_x}{\sqrt{1-\beta^2}}. \tag{10.36}$$

If the particle is at rest in the frame K ($p_x = p_y = p_z = 0$ and $E = m_0c^2$), then, in the frame K',

$$\begin{aligned} p_x' &= -\frac{v}{c^2}\frac{m_0c^2}{\sqrt{1-\beta^2}} = -mv, \quad p_y' = 0, \\ p_z' &= 0, \quad E' = \frac{m_0c^2}{\sqrt{1-\beta^2}} = mc^2. \end{aligned} \tag{10.37}$$

These relations are associated with the fact that a particle that is at rest in the frame K moves at the velocity v in the frame K'.

The square of the energy-momentum 4-vector (10.34),

$$\sum_{i=1}^{4} p_i^2 = p^2 - \frac{E^2}{c^2}, \tag{10.38}$$

must be invariant under the Lorentz transformations. Really, taking into account (10.17) and (10.28), we obtain

$$\sum_{i=1}^{4} p_i^2 = -m_0^2c^2 = \text{const}. \tag{10.39}$$

Equating the right hand sides in (10.38) and (10.39), we obtain the relativistic relationship between the energy and momentum of free particle,

$$E^2 = p^2c^2 + m_0^2c^4, \tag{10.40}$$

i.e. the Hamilton function

$$E = c\sqrt{p^2 + m_0^2 c^2}. \tag{10.41}$$

Substituting energy (10.29) into formula (10.41) for the kinetic energy of the particle, we obtain the expression

$$T = c\sqrt{p^2 + m_0^2 c^2} - m_0 c^2,$$

whence it follows the relationship between the relativistic momentum of the particle and its kinetic energy,

$$p = \frac{1}{c}\sqrt{T\left(T + 2m_0 c^2\right)}. \tag{10.42}$$

If $v \ll c$ (a non-relativistic particle),

$$T = E - m_0 c^2 = m_0 c^2 \left(\frac{1}{\sqrt{1 - \dfrac{v^2}{c^2}}} - 1 \right) \approx m_0 c^2 \frac{v^2}{2c^2} = \frac{m_0 v^2}{2} \ll m_0 c^2.$$

Then, from (10.42), we find

$$p = \sqrt{2m_0 T}. \tag{10.43}$$

On the other hand, in the case $T \gg m_0 c^2$, from relation (10.42), we obtain

$$p = \frac{T}{c}. \tag{10.44}$$

Such a particle is called **ultra-relativistic**.

Formula (10.18) is used to determine the relativistic energy of a free particle moving at the velocity v. Let this particle be in a potential field. Then, the field acts on the particle with the force

$$\vec{F} = -\frac{dV}{d\vec{r}}, \tag{10.45}$$

where $V = V(\vec{r})$ is the potential energy of the particle. Substituting (10.45) into (10.27) and taking into account that $\vec{v} = d\vec{r}/dt$, we obtain the equality

$$\frac{d}{dt}\frac{m_0 c^2}{\sqrt{1-\beta^2}} = -\frac{dV}{dt},$$

the integral of which is the energy conservation law,

$$\frac{m_0c^2}{\sqrt{1-\beta^2}} + V = \text{const.} \tag{10.46}$$

The quantity

$$\mathscr{E} = \frac{m_0c^2}{\sqrt{1-\beta^2}} + V = E + V \tag{10.47}$$

should be regarded as the total particle energy in the potential field. In this case, the 4-vector of particle energy-momentum (10.34) reads

$$\{p_i\} = \left\{\vec{p}, \frac{i}{c}(\mathscr{E} - V)\right\}, \tag{10.48}$$

and its square looks like

$$p^2 - \frac{1}{c^2}(\mathscr{E} - V)^2 = -m_0c^2. \tag{10.49}$$

Hence, the Hamilton function for a particle in a potential field V equals

$$\mathscr{E} = c\sqrt{p^2 + m_0^2c^2} + V. \tag{10.50}$$

10.4. System of particles. Binding energy

Let us apply the results obtained in the previous section to study a system of particles. An example of such a system is the atomic nucleus consisting of protons and neutrons with the nuclear interaction between them and the electrostatic interaction between the protons.

For a system of particles, on the basis of (10.40), we may write the relation

$$M_0^2c^2 = \frac{E^2}{c^2} - \vec{P}^2, \tag{10.51}$$

where M_0 is the rest mass of the system, E its total energy, and $\vec{P}$ its total momentum. The total energy of the system is a sum of relativistic energies of separate particles, E_i, and the energy of interaction between them, U,

$$E = \sum_i E_i + U; \tag{10.52}$$

and the total momentum, as an additive quantity, equals

$$\vec{P} = \sum_i \vec{P}_i. \tag{10.53}$$

Let us change from the laboratory reference frame, for which relation (10.51) was written, to the reference frame of the center of inertia, where the system, as a whole, is motionless and $\vec{P} = 0$. Then, instead of (10.51), we obtain

$$M_0 = \frac{1}{c^2}\left(\sum_i E_i + U\right). \tag{10.54}$$

The relativistic energy E_i of separate particle consists of the rest energy $m_{0i}c^2$ and the kinetic energy T_i, i.e.

$$E_i = m_{0i}c^2 + T_i.$$

Therefore,

$$M_0 = \sum_i m_{0i} + \frac{1}{c^2}\sum_i T_i + \frac{U}{c^2}. \tag{10.55}$$

One can easily see from (10.55) that the rest mass of a system of particles is not an additive quantity even in the absence of interaction between them.

In the case of a stable (bound) system of particles, $U < 0$ and the value of $|U|$ is called the **binding energy**. Since $U = -|U|$ in this case, from (10.55), it follows that

$$|U| = \left(\sum_i m_{0i} - M_0\right)c^2 + \sum_i T_i. \tag{10.56}$$

The difference

$$\Delta M = \sum_i m_{0i} - M_0 \tag{10.57}$$

is called the **mass defect**. Hence,

$$|U| = \Delta M c^2 + \sum_i T_i. \tag{10.58}$$

If $\sum_i T_i \ll |U|$, we obtain the relation between the binding energy and the mass defect,

$$|U| = \Delta M c^2, \tag{10.59}$$

which is used to calculate, in particular, the binding energy in atomic nuclei.

11

RELATIVISTIC ELECTRODYNAMICS

The special theory of relativity demands that all physical laws must be covariant with respect to the Lorentz transformations. Newton's laws of classical mechanics turned out non-covariant. Therefore, corrections to the determination of such mechanical quantities as the velocity, acceleration, force, mass, momentum, and energy, as well as the equations of motion, were made. The same issue concerning the invariant character of quantities and equations arises in the classical electrodynamics of Maxwell. While solving it, in order to make the consideration not too complicated, let us analyze consider the electromagnetic field in vacuum, because general conclusions are applicable to the field in a material medium as well. For this purpose, let us first write down the required relations for the electromagnetic field in vacuum (see part I.A).

The electromagnetic field of charges and currents in vacuum is described by the vectors of electric field strength $\vec{E}$ and magnetic field inductions $\vec{B}$, which satisfy Maxwell's equations (3.36)–(3.39). Now, we write down them in the form of two equation pairs as follows:

$$\begin{aligned} &\text{a)}\quad \operatorname{rot}\vec{B} = \mu_0\vec{j} + \frac{1}{c^2}\frac{\partial\vec{E}}{\partial t}, &&\text{b)}\quad \operatorname{rot}\vec{E} = -\frac{\partial\vec{B}}{\partial t},\\ &\phantom{\text{a)}}\quad \operatorname{div}\vec{E} = \frac{\rho}{\varepsilon_0}, &&\phantom{\text{b)}}\quad \operatorname{div}\vec{B} = 0. \end{aligned} \tag{11.1}$$

where ρ is the electric charge density, $\vec{j}$ the electric current density, and $\varepsilon_0\mu_0 = 1/c^2$.

The field vectors $\vec{E}$ and $\vec{B}$ can also be defined by means of the scalar, φ, and vector, $\vec{A}$, potentials of electromagnetic field (see (3.80) and (3.82)) as

$$\begin{aligned} \vec{E} &= -\nabla\varphi - \frac{\partial \vec{A}}{\partial t}, && (11.2) \\ \vec{B} &= \operatorname{rot} \vec{A}. && (11.3) \end{aligned}$$

The potentials φ and $\vec{A}$ satisfy d'Alembert equations (3.90) and (3.91),

$$\begin{aligned} \Delta\varphi - \varepsilon_0\mu_0 \frac{\partial^2\varphi}{\partial t^2} &= -\frac{\rho}{\varepsilon_0}, && (11.4) \\ \Delta\vec{A} - \varepsilon_0\mu_0 \frac{\partial^2\vec{A}}{\partial t^2} &= -\mu_0\vec{j}. && (11.5) \end{aligned}$$

and Lorentz gauge condition (3.89),

$$\operatorname{div} \vec{A} + \frac{1}{c^2}\frac{\partial\varphi}{\partial t} = 0. \qquad (11.6)$$

At last, the conservation law of electric charge is written in the form of continuity equation (2.10),

$$\operatorname{div} \vec{j} + \frac{\partial\rho}{\partial t} = 0. \qquad (11.7)$$

For those quantities, as well as the electrodynamics equations, to be covariant with respect to the Lorentz transformations, they have to be written in the four-dimensional form as 4-vectors or 4-tensors. Below, we will demonstrate that the electrodynamics equations allow us to do this without modifications, because they turned out to be correct from the viewpoint of STR, where the speed of light in vacuum is an invariant. At the same time, their expression in the four-dimensional form makes it possible to establish the laws of transformation for electrodynamic quantities, in particular, the field vectors $\vec{E}$ and $\vec{B}$, when changing from one IRF to the other; to substantiate the conclusion about the electromagnetic field as a unique object described by a tensor of electromagnetic field composed of the components of vectors $\vec{E}$ and $\vec{B}$; and to draw a number of other important conclusions.

11.1. 4-potential and 4-current

The scalar potential φ and the vector potential $\vec{A}$ can be united into a single 4-vector, which is called the **4-potential**. For this purpose, let us write the Lorentz gauge condition (11.6) in the expanded form as

$$\frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} + \frac{1}{c^2}\frac{\partial \varphi}{\partial t} = 0. \tag{11.8}$$

Introducing a 4-vector with the components $A_1 = A_x$, $A_2 = A_y$, $A_3 = A_z$, and $A_4 = i\varphi/c$, i.e. the 4-vector

$$\{A_i\} = \left\{\vec{A}, i\frac{\varphi}{c}\right\}, \tag{11.9}$$

equation (11.8) can be rewritten in the form

$$\sum_{i=1}^{4} \frac{\partial A_i}{\partial x_i} = 0. \tag{11.10}$$

The left hand side of this equation is the 4-divergency of 4-vector (11.9), which is called the **4-potential**. In order to prove that quantity (11.9) is a 4-vector, it is sufficient to be convinced that the 4-divergency in (11.10) is a true invariant.

The transformations of 4-potential components in the case of IRF change are described by formulas (9.84). Making allowance for (11.9), these transformations are written in terms of three-dimensional quantities as follows:

$$A'_x = \frac{A_x - \dfrac{\beta}{c}\varphi}{\sqrt{1-\beta^2}}, A'_y = A_y, A'_z = A_z, \varphi' = \frac{\varphi - vA_x}{\sqrt{1-\beta^2}}. \tag{11.11}$$

Continuity equation (11.7) in the expanded form looks like

$$\frac{\partial j_x}{\partial x} + \frac{\partial j_y}{\partial y} + \frac{\partial j_z}{\partial z} + \frac{\partial \rho}{\partial t} = 0. \tag{11.12}$$

If we construct the 4-vector

$$\{j_i\} = \left\{\vec{j}, ic\rho\right\}, \tag{11.13}$$

it is easy to be convinced that equation (11.12) takes the form

$$\sum_{i=1}^{4} \frac{\partial j_i}{\partial x_i} = 0, \tag{11.14}$$

where in the left hand side is the 4-divergence of vector (11.13), which is called the **4-vector of current density** (the **4-current**).

With the help of (9.84) and (11.13), we obtain

$$j'_x = \frac{j_x - v\rho}{\sqrt{1-\beta^2}}, \quad j'_y = j_y, \quad j'_z = j_z, \quad \rho' = \frac{\rho - \dfrac{\beta}{c} j_x}{\sqrt{1-\beta^2}}. \tag{11.15}$$

Let us consider this transformation in more details. Let an electric charge q be distributed over the frame K with the bulk density ρ,

$$q = \int_V \rho dV. \tag{11.16}$$

The charge is at rest, i.e. $j_x = j_y = j_z = 0$. In the frame K', however, this charge moves at the velocity $-v$. Therefore, in accordance with (11.15),

$$j'_x = \frac{-v\rho}{\sqrt{1-\beta^2}}, \quad j'_y = 0, \quad j'_z = 0, \quad \rho' = \frac{\rho}{\sqrt{1-\beta^2}}. \tag{11.17}$$

So, the density of electric charge is not invariant, because it has different values in different IRFs. The total charge calculated in the frame K' equals

$$q' = \int_{V'} \rho' dV', \tag{11.18}$$

where the elementary volume dV' is (see section 9.5)

$$dV' = dx'dy'dz' = dx\sqrt{1-\beta^2}dydz = dV\sqrt{1-\beta^2}.$$

Therefore,

$$q' = \int_{V'} \rho' dV' = \int_V \rho dV, \tag{11.19}$$

i.e. the total electric charge is an invariant quantity. This conclusion agrees with the experimental fact consisting in that a molecule or an atom is an electrically neutral system despite of electron motion in it.

With the help of 4-potential (11.9) and 4-current (11.13), the d'Alembert equations for potentials (11.4) and (11.5) can be written as a single equation for the 4-potential,

$$\Box A_i = -\mu_0 j_i, \quad (i = 1, 2, 3, 4), \tag{11.20}$$

where

$$\Box = \sum_{i=1}^{4} \frac{\partial^2}{\partial x_i^2} = \Delta - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \tag{11.21}$$

is the d'Alembert operator. It is a square of the vector 4-gradient operator, i.e. invariant; therefore, the left and right hand sides of equation (11.20) transform as 4-vectors.

11.2. Tensor of electromagnetic field in vacuum

Formulas (11.2) and (11.3) connect the field vectors $\vec{E}$ and $\vec{B}$ with the electromagnetic potentials φ and $\vec{A}$. Using (11.2) and definition (11.9) of 4-potential, we can show that

$$E_x = ic\left(\frac{\partial A_4}{\partial x_1} - \frac{\partial A_1}{\partial x_4}\right), \tag{11.22}$$

$$E_y = ic\left(\frac{\partial A_4}{\partial x_2} - \frac{\partial A_2}{\partial x_4}\right), \tag{11.23}$$

$$E_z = ic\left(\frac{\partial A_4}{\partial x_3} - \frac{\partial A_3}{\partial x_4}\right). \tag{11.24}$$

Analogously, from (11.3) and (11.9), it follows that

$$B_x = ic\left(\frac{\partial A_3}{\partial x_2} - \frac{\partial A_2}{\partial x_3}\right), \tag{11.25}$$

$$B_y = ic\left(\frac{\partial A_1}{\partial x_3} - \frac{\partial A_3}{\partial x_1}\right), \tag{11.26}$$

$$b_z = ic\left(\frac{\partial A_2}{\partial x_1} - \frac{\partial A_1}{\partial x_2}\right). \tag{11.27}$$

Let us introduce the quantity

$$\mathscr{F}_{ij} = \frac{\partial A_j}{\partial x_i} - \frac{\partial A_i}{\partial x_j}. \tag{11.28}$$

As the derivative of 4-vector with respect to 4-coordinates, this is a 4-tensor of the second rank. As follows from (11.28), this tensor is antisymmetric, because $\mathscr{F}_{ij} = -\mathscr{F}_{ji}$. By definition, the 4-tensor$\mathscr{F}_{ij}$ is covariant with respect to the Lorentz transformations. It unites the electric and magnetic fields into a single object—the electromagnetic field. Therefore, it is called the **tensor of electromagnetic field**. Its components, in accordance with (11.22)–(11.27), are defined in terms of components of vectors $\vec{E}$ and $\vec{B}$. In particular, from (11.22), it follows that

$$E_x = ic\left(\frac{\partial A_4}{\partial x_1} - \frac{\partial A_1}{\partial x_4}\right) = ic\mathscr{F}_{14},$$

i.e.

$$\mathscr{F}_{14} = -\frac{i}{c}E_x.$$

Using the same method to find all components of the tensor $\mathscr{F}$, we will obtain the matrix of its components in the form

$$\{\mathscr{F}_{ij}\} = \begin{pmatrix} 0 & B_z & -B_y & -\frac{i}{c}E_x \\ -B_z & 0 & B_x & -\frac{i}{c}E_y \\ B_y & -B_x & 0 & -\frac{i}{c}E_z \\ \frac{i}{c}E_x & \frac{i}{c}E_y & \frac{i}{c}E_z & 0 \end{pmatrix}. \tag{11.29}$$

The antisymmetric 4-tensor of electromagnetic field (11.29) represents an electromagnetic field, i.e., from the mathematical viewpoint, it is an electromagnetic field. This circumstance agrees with the fact that, as follows directly from (11.29), the tensor concerned has six independent components. At the same time, in the three-dimensional space, the electromagnetic field is described by the vectors $\vec{E}$ and $\vec{B}$, which also have six independent components together. The tensor of electromagnetic field $\mathscr{F}_{ij}$ is the key concept of relativistic electrodynamics and plays an important role in the general theory of relativity.

11.3. Covariance of Maxwell's equations

As was already indicated above, Maxwell's equations (11.1) are covariant with respect to the Lorentz transformations. To be convinced of it, they should be expressed in terms of 4-vectors and 4-tensors. Let us consider first the pair of equations (11.1,a),

$$\begin{aligned} \operatorname{rot}\vec{B} &= \mu_0\vec{j} + \frac{1}{c^2}\frac{\partial\vec{E}}{\partial t}, && (11.30)\\ \operatorname{div}\vec{E} &= \frac{\rho}{\varepsilon_0}. && (11.31) \end{aligned}$$

In the four-dimensional form, these equations can be written as

$$\sum_{k=1}^{4}\frac{\partial\mathscr{F}_{ik}}{\partial x_k} = \mu_0 j_i, \quad (i = 1, 2, 3, 4). \tag{11.32}$$

To be convinced of it, let $i = 1$. Then,

$$\frac{\partial\mathscr{F}_{11}}{\partial x_1} + \frac{\partial\mathscr{F}_{12}}{\partial x_2} + \frac{\partial\mathscr{F}_{13}}{\partial x_3} + \frac{\partial\mathscr{F}_{14}}{\partial x_4} = \mu_0 j_x.$$

Taking into account (11.29) and (11.13), we obtain

$$\frac{\partial B_z}{\partial y} - \frac{\partial B_y}{\partial z} - \frac{i}{c}\frac{\partial E_x}{\partial(ict)} = \mu_0 j_x$$

or

$$\operatorname{rot}_x\vec{B} = \mu_0 j_x + \frac{1}{c^2}\frac{\partial E_x}{\partial t},$$

which agrees with equation (11.30). For $i = 4$, from (11.32), we find

$$\frac{\partial\mathscr{F}_{41}}{\partial x_1} + \frac{\partial\mathscr{F}_{42}}{\partial x_2} + \frac{\partial\mathscr{F}_{43}}{\partial x_3} + \frac{\partial\mathscr{F}_{44}}{\partial x_4} = \mu_0 j_4$$

or

$$\frac{i}{c}\left(\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z}\right) = \mu_0(ic\rho),$$

i.e.

$$\operatorname{div}\vec{E} = \frac{\rho}{\varepsilon_0}.$$

The second pair of Maxwell's equations (11.1,b) look like

$$\operatorname{rot}\vec{E} = -\frac{\partial\vec{B}}{\partial t}, \tag{11.33}$$

$$\operatorname{div}\vec{B} = 0. \tag{11.34}$$

In the four-dimensional form, they are written as

$$\frac{\partial\mathscr{F}_{ik}}{\partial x_k} + \frac{\partial\mathscr{F}_{ki}}{\partial x_j} + \frac{\partial\mathscr{F}_{jk}}{\partial x_i} = 0, \quad (i, j, k = 1, 2, 3, 4). \tag{11.35}$$

To prove this, let us put $i = 1$, $j = 2$, and $k = 3$. In this case, on the basis of (11.35), we obtain the equation

$$\frac{\partial\mathscr{F}_{12}}{\partial x_3} + \frac{\partial\mathscr{F}_{31}}{\partial x_2} + \frac{\partial\mathscr{F}_{23}}{\partial x_1} = 0,$$

which, taking into account (11.29), can be written as

$$\frac{\partial B_z}{\partial z} + \frac{\partial B_y}{\partial y} + \frac{\partial B_x}{\partial x} = 0,$$

or

$$\operatorname{div}\vec{B} = 0. \tag{11.36}$$

For $i = 1$, $j = 4$, and $k = 2$ in (11.35), we obtain

$$\frac{\partial\mathscr{F}_{14}}{\partial x_2} + \frac{\partial\mathscr{F}_{21}}{\partial x_4} + \frac{\partial\mathscr{F}_{42}}{\partial x_1} = 0$$

or, taking into account (11.29),

$$\operatorname{rot}_z\vec{E} = -\frac{\partial B_z}{\partial z}. \tag{11.37}$$

This equation agrees with equation (11.33).

To summarize, notice that Maxwell's equation for the electromagnetic field in vacuum (11.1) can be written in the relativistically covariant form as two 4-equations,

$$\sum_{k=1}^{4}\frac{\partial\mathscr{F}_{ik}}{\partial x_k} = \mu_0 j_i, \quad (i = 1, 2, 3, 4) \tag{11.38}$$

and

$$\frac{\partial\mathscr{F}_{ik}}{\partial x_k} + \frac{\partial\mathscr{F}_{ki}}{\partial x_j} + \frac{\partial\mathscr{F}_{jk}}{\partial x_i} = 0, \quad (i, j, k = 1, 2, 3, 4). \tag{11.39}$$

11.4. Transformations of electromagnetic field components. Electromagnetic field invariants

The representation of electromagnetic field by means of 4-tensor $\mathscr{F}_{ji}$ makes it possible to easily find the transformations for the components of field vectors $\vec{E}$ and $\vec{B}$ when changing the reference frame. It is so because those components define the components of electromagnetic field tensor which transform according to the rule

$$\mathscr{F}'_{ij} = \sum_{k,l=1}^{4} a_{ik} a_{jl} \mathscr{F}_{kl}, \tag{11.40}$$

in which the coefficients a_{ik} are the members of matrix (9.76).

As an example, let us establish the transformation for the component B_z of the magnetic induction vector. According to (11.29), $B_z = \mathscr{F}_{12}$. Therefore, with the help of (11.40) and (9.76), we obtain

$$B'_z = \mathscr{F}'_{12} = \sum_{k,l=1}^{4} a_{1k} a_{2l} \mathscr{F}_{kl} = \sum_{k} a_{1k} \mathscr{F}_{k2} = \frac{\mathscr{F}_{12} + i\beta \mathscr{F}_{42}}{\sqrt{1-\beta^2}}.$$

Using the values of components $\mathscr{F}_{12}$ and $\mathscr{F}_{42}$, we obtain

$$B'_z = \frac{B_z - \dfrac{v}{c^2} E_y}{\sqrt{1-\beta^2}}.$$

In the same way, we can find transformations for every component of vectors $\vec{E}$ and $\vec{B}$. As a result, we obtain

$$\begin{aligned} E'_x &= E_x, & B'_x &= B_x, \\ E'_y &= \frac{E_y - vB_z}{\sqrt{1-\beta^2}}, & B'_y &= \frac{B_y + \dfrac{v}{c^2} E_z}{\sqrt{1-\beta^2}}, \\ E'_z &= \frac{E_z + vB_y}{\sqrt{1-\beta^2}}, & B'_z &= \frac{B_z - \dfrac{v}{c^2} E_y}{\sqrt{1-\beta^2}}. \end{aligned} \tag{11.41}$$

The formulas for inverse field transformations follow from (11.41) if we substitute v by $-v$ there.

In the low-velocity case ($v/c \ll 1$), formulas (11.41) become simpler,

$$\begin{array}{lll} E'_x = E_x, & E'_y = E_y - vB_z, & E'_z = E_z + vB_y, \\ B'_x = B_x, & B'_y = B_y + \dfrac{v}{c^2}E_z, & B'_z = B_z - \dfrac{v}{c^2}E_y; \end{array} \tag{11.42}$$

or, in the vector form,

$$\vec{E}' = \vec{E} + \left[\vec{v}\,\vec{B}\right], \tag{11.43}$$

$$\vec{B}' = \vec{B} - \frac{1}{c^2}\left[\vec{v}\,\vec{E}\right]. \tag{11.44}$$

Formulas (11.41)–(11.44) for the transformation of electric and magnetic fields give rise to a number of important conclusions. The first one concerns the relativity of electromagnetic field separation into electric and magnetic components. For instance, that, in the frame K, there is only an electric field created by fixed charges ($\vec{E} \neq 0$, $\vec{B} = 0$). Then, in the frame K' moving at the velocity v along the axis x, there exist both the electric and magnetic fields. If $v \ll c$, the corresponding field vectors look like

$$\vec{E}' = \vec{E}, \quad \vec{B}' = -\frac{1}{c^2}\left[\vec{v}\,\vec{E}\right]. \tag{11.45}$$

On the other hand, if there is only a magnetic field of moving charges in the frame K ($\vec{E} = 0$, $\vec{B} \neq 0$), both fields exist in the frame $K'\hat{}$

$$\vec{E}' = [\vec{v}\,\vec{B}]; \vec{B}' = \vec{B}.$$

Another conclusion, which follows from transformation formulas (11.41)–(11.44), is trivial: if there is no electromagnetic field ($\vec{E} = 0$, $\vec{B} = 0$) in a certain reference frame (in our case, this is the frame K), it is also absent in all other reference frames. One more conclusion consists in that if there is a frame K, where there exists only one field (electric or magnetic), in all others IRFs, there exist both the electric, and magnetic fields with the mutually perpendicular vectors $\vec{E}'$ and $\vec{B}'$. For instance, if $\vec{B} = 0$ and $\vec{E} \neq 0$ in the frame K, this conclusion follows from (11.45).

When the components $\mathscr{F}_{ji}$ of electromagnetic field tensor transform, some of their combinations remain constant; such combinations are called the **electromagnetic field invariants**. Two of them, in the three-dimensional notation, look like

$$I_1 = c^2B^2 - E^2, \tag{11.46}$$

$$I_2 = \vec{B}\,\vec{E}. \tag{11.47}$$

One can easily be convinced in that, using formulas (11.41).

From the invariance of (11.47), it follows that, if $\vec{B}\vec{E} = 0$, i.e. the field vectors $\vec{E}$ and $\vec{B}$ are mutually orthogonal, in a certain IRF, they are orthogonal in all other IRFs. In turn, if $E = cB$ in a certain IRF, then, according to (11.46), this relation holds true in other IRFs as well. In particular, for the light wave in vacuum, $\vec{E} \perp \vec{B}$ and $E = cB$; therefore, these properties remain valid in all IRFs.

11.5. Wave 4-vector. Doppler effect

For a plane monochromatic electromagnetic wave in vacuum, the vectors of electric field strength and magnetic field induction are written in the form (see section 4.2)

$$\vec{E} = \vec{E}_0 \mathrm{e}^{i(\vec{k}\vec{r}-\omega t)}, \quad \vec{B} = \vec{B}_0 \mathrm{e}^{i(\vec{k}\vec{r}-\omega t)},$$

so that they have the same phase $\vec{k}\vec{r} - \omega t$, where $\vec{k}$ is the wave vector; and ω the frequency. In this case, the components of electromagnetic field tensor (11.29) that are different from zero have to look like

$$\mathscr{F}_{ij} = \mathscr{F}_{ij}^{(0)} \mathrm{e}^{i(\vec{k}\vec{r}-\omega t)}, \tag{11.48}$$

where $\mathscr{F}_{ij}^{(0)}$ is the component magnitude. The corresponding transformation of tensor $\mathscr{F}_{ij}$ components takes the form

$$\mathscr{F}_{ij}^{(0)\prime} \mathrm{e}^{i(\vec{k}'\vec{r}'-\omega' t')} = \sum_{k,l=1}^{4} a_{ik} a_{jl} \mathscr{F}_{kl}^{(0)} \mathrm{e}^{i(\vec{k}\vec{r}-\omega t)}. \tag{11.49}$$

This relation must be satisfied at every point in the space and at any time momentum, which is possible only if the phase is invariant, i.e. if

$$\vec{k}\vec{r} - \omega t = \vec{k}'\vec{r}' - \omega' t'. \tag{11.50}$$

Introducing the notation

$$k_1 = k_x, k_2 = k_y, k_3 = k_z, k_4 = \frac{i}{c}\,\omega, \tag{11.51}$$

condition (11.50) is written as

$$\sum_{i=1}^{4} k_i x_i = \sum_{j=1}^{4} k'_j x'_j. \tag{11.52}$$

The invariance of quantity (11.52) proves that it is a scalar product of two 4-vectors: a coordinate 4-vector $\{x_i\}$ and a 4-vector

$$\{k_i\} = \left\{\vec{k}, \frac{i}{c}\,\omega\right\}, \tag{11.53}$$

which is called the **wave 4-vector**.

The concept of wave 4-vector (11.53) can also be arrived at by considering the 4-vector of energy-momentum for the light quantum (photon) written in the form

$$\{p_i\} = \left\{\vec{p}, \frac{i}{c} E\right\}. \tag{11.54}$$

In quantum physics, the energy E and the 3-momentum $\vec{p}$ of photon look like $E = \hbar\omega$ and $\vec{p} = \hbar\vec{k}$, where $\hbar$ is Planck's constant. Therefore, for the photon,

$$\{p_i\} = \left\{\hbar\vec{k}, \frac{i}{c}\,\hbar\omega\right\},$$

whence, reducing by the common multiplier $\hbar$, we obtain wave 4-vector (11.53).

Let us consider the transformation formulas for the components of wave 4-vector (11.53). Using (9.84), (9.76), and (11.53), it is easy to be convinced that they can be rewritten in the form

$$k'_x = \frac{k_x - \dfrac{\beta}{c}\,\omega}{\sqrt{1-\beta^2}}, k'_y - k_y, k'_z = k_z, \omega' = \frac{\omega - v k_x}{\sqrt{1-\beta^2}}, \tag{11.55}$$

which bring us to the phenomenon of starlight aberration and Doppler effect. The former was dealt with in sections 9.1 and 9.6.

Now, let us consider the Doppler effect. It consists in the frequency change of light emitted by a moving source in comparison with the frequency of radiation emitted by a motionless object. Since $k_x = k\cos\theta$, where θ is the angle between the direction of light beam and the axis x, from the last formula in (11.55), we find

$$\omega' = \omega\frac{1-\beta\cos\theta}{\sqrt{1-\beta^2}}. \tag{11.56}$$

Let the light source be at rest in the frame K', and the corresponding frequency of emitted radiation equals $\omega' = \omega_0$. The observer is in the frame K and registers the frequency ω of light emitted by the moving source. In this case, on the basis of formula (11.56), we obtain

$$\omega = \omega_0\frac{\sqrt{1-\beta^2}}{1-\beta\cos\theta}. \tag{11.57}$$

If the light source moves away from the observer (Fig. 11.1,a), $\theta = \pi$ and

$$\omega = \omega_0\frac{\sqrt{1-\beta^2}}{1+\beta} = \omega_0\sqrt{\frac{1-\beta}{1+\beta}} < \omega_0, \tag{11.58}$$

i.e. the "red shift" of the frequency is observed. If the source approaches the observer (Fig. 11.1,b), $\theta = 0$ and

$$\omega = \omega_0\frac{\sqrt{1-\beta^2}}{1-\beta} = \omega_0\sqrt{\frac{1+\beta}{1-\beta}} > \omega_0, \tag{11.59}$$

so that the observer registers the "blue shift". In the case $\theta = \pi/2$ (Fig. 11.1,c),

$$\omega = \omega_0\sqrt{1-\beta^2}, \tag{11.60}$$

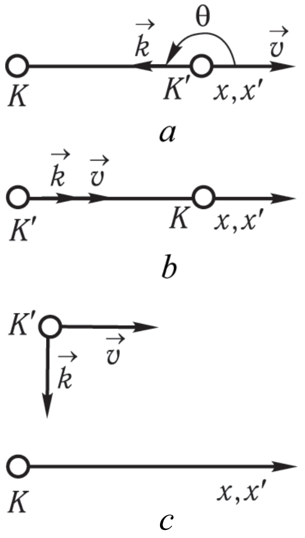

Fig. 11.1.

and we have the **transverse Doppler effect**. Notice that the **longitudinal Doppler effect** (panels *a* and *b* in Fig. 11.1) can be qualitatively explained beyond the scope of STR, but the transverse one is a completely relativistic phenomenon.

EXERCISES TO PART I

1. Calculate the strength and potential of the electric field around a straight infinite thread with a linear charge density η.

   - Solution: $E = \dfrac{\eta}{2\pi\varepsilon_0 R}$ and $\varphi = \dfrac{\eta}{2\pi\varepsilon_0 R^2}$.

2. An electrically neutral bead is suspended by an elastic thread fabricated of a dielectric material with the elastic constant k at the distance d over a conducting surface. What charge should be given to the bead for the distance to diminish by x?

   - Solution: $q = 4\,(d - x)\,\sqrt{\pi\varepsilon_0 k x}$.

3. Calculate the energy of electrostatic field around a charged sphere of radius R, with the charge q being uniformly distributed over the sphere's surface.

   - Solution: $U = \dfrac{q^2}{8\pi\varepsilon_0 R}$.

4. Find the induction of magnetic field created by the ring current I of the radius R at the distance h reckoned from the ring center along the ring axis.

   - Solution: $B = \dfrac{\mu_0 I R^2}{2\left(R^2 - h^2\right)^{3/2}}$.

5. Determine the induction of magnetic field around an infinitely long ($l \to \infty$) linear current I.

   - Solution: $B = \dfrac{\mu_0 I}{2\pi R}$.

6. Prove that a uniform magnetic field with the induction B can be described by the vector potential $\vec{A} = \dfrac{1}{2}\left[\vec{B}\,\vec{r}\right]$.

7. Find the magnetic moment of a sphere with the radius R and the electric charge q uniformly distributed over its volume, provided that the sphere rotates around its diameter with the constant angular velocity ω.

   - Solution: $d_m = \dfrac{1}{5} q\omega R^2$.

8. The current I is uniformly distributed over the cross-section of a long round cylinder of the radius r. Show that the heat released in the conductor equals the energy of electromagnetic field absorbed from the outside.

9. Show that, in the course of the plane air capacitor discharge, the conduction current is looped by the displacement one at the capacitor plates.

10. Write the field components for a plane monochromatic electromagnetic wave linearly polarized along the axis x and propagating along the axis z.

    - Solution: $E_x = E_0 \mathrm{e}^{i(kz-\omega t)}$, $E_y = E_z = 0$, $B_y = \dfrac{E_0}{c}\mathrm{e}^{i(kz-\omega t)}$, $B_x = B_z = 0$.

11. If the electric field strength reaches the value $E = 30$ kV/cm, air begins to ionize. What energy flux density in a plane low-frequency electromagnetic wave would be enough for the ionization in air to begin?

    - Solution: $\langle \Pi \rangle = \dfrac{E_0^2}{2\mu_0 c} = 1.2 \times 10^{10}\ \mathrm{W/m^2}$.

12. Show that the power of electric-dipole radiation emitted by a closed system consisting of particles which identical specific charges equals zero.

13. Calculate the mean-square values of electric field strength and magnetic field induction for the electromagnetic wave emitted by a 100-W lamp at the distance $r = 1$ m from it. Consider the lamp as a point source with the efficiency of electric energy transformation into light $\eta = 0.75$.

- Solution: $\sqrt{\langle E^2\rangle} = \sqrt{\dfrac{\eta N \mu_0 c}{4\pi r^2}} = 33.5$ V/m, $\sqrt{\langle B^2\rangle} = \dfrac{1}{c}\sqrt{\langle E^2\rangle} = 1.12 \times 10^{-7}$ T.

14. For what medium do Maxwell's equations look like as follows?

$$\begin{aligned} &\operatorname{rot} \vec{H} = \sigma \vec{E} + \varepsilon\varepsilon_0 \frac{\partial \vec{E}}{\partial t}; \qquad && \operatorname{div} \vec{H} = 0, \\ &\operatorname{rot} \vec{E} = -\mu\mu_0 \frac{\partial \vec{H}}{\partial t}; && \operatorname{div} \vec{E} = 0. \end{aligned}$$

15. A linearly polarized plane electromagnetic wave falls normally on the surface of non-magnetic medium with the dielectric permittivity ε and the electric conductivity σ. Find the corresponding reflection coefficient R if the conductivity of the medium is high.

- Solution: $R = \dfrac{1 - 2n + 2n^2}{1 + 2n + 2n^2} = 1 - \dfrac{4n}{1 + 2n + 2n^2} \approx 1 - \dfrac{2}{n} = 1 - 2\sqrt{\dfrac{2\varepsilon_0 \omega}{\sigma}}$.

16. A light wave with the frequency $\omega = 3 \times 10^{15}$ s^{-1} falls on the interface between glass ($n_1 = 1.5$) and water ($n_2 = 1.33$) from the glass side. The angle of incidence equals 70°. Calculate the depth of electromagnetic field penetration into water.

- Solution: $|\Delta z| = 2 \times 10^{-7}$ m.

17. Calculate the reflection, R, and transmission, D, coefficients for the light beam that propagates normally from glass ($\varepsilon_1 = 2.25$) into water ($\varepsilon_2 = 1.77$).

- Solution: $R = 0.0037$, $D = 0.9963$.

18. Calculate the dielectric permittivity $\varepsilon(\omega)$ of the medium, provided that $\varepsilon(t) = A \exp(t/a)$, where A and a are constants. Instruction: use formula (8.9).

- Solution: $\varepsilon(\omega) = \dfrac{aA}{1 - ia\omega}$.

19. Find the wavelengths of electromagnetic waves that can propagate in plasma with the free electron concentration $N = 10^{13}$ m^{-3}. Instruction: use formula (8.41).

    - Solution: $\lambda < 2\pi c\sqrt{\dfrac{\varepsilon_0 m_0}{e^2 N}} = 11$ m.

20. Two identical rods of the length l_0 (in the proper reference frames) are arranged along a straight line and approach each other at identical velocities v with respect to the stationary reference frame. Find the length of either rod in the reference frame attached to the other rod.

    - Solution: $l = l_0\sqrt{1 - \dfrac{v_{rel}^2}{c^2}} = l_0 \dfrac{1 - \dfrac{v^2}{c^2}}{1 + \dfrac{v^2}{c^2}}$.

21. Two electron beams move towards each other at the velocity $v = 0/9c$ in the laboratory reference frame. What is the relative velocity of electrons v_{rel} (i) in the laboratory reference frame and (ii) from the viewpoint of an observer moving together with either of electron beams?

    - Solution: (i) $v_{rel} = 1.8\,c$, (ii) $v_{rel} = 0.994\,c$.

22. Prove that the volume element $dx_1dx_2dx_3dx_4$ in the 4-space is invariant with respect to the Lorentz transformations.

23. A particle at rest has the mass m_0. Find its velocity v in terms of (i) the total energy E and (ii) the kinetic energy T.

    - Solution: (i) $v = c\sqrt{1 - \left(\dfrac{m_0c^2}{E}\right)^2}$,

      (ii) $v = c\sqrt{1 - \left(\dfrac{m_0c^2}{m_0c^2 + T}\right)^2}$.

24. Find the kinetic energy T of a particle with the rest mass m_0 in terms of its velocity v to an accuracy of v^4/c^4 in the case $v \ll c$. Instruction: use Newton's binomial formula.

- Solution: $T = E - m_0c^2 \approx \dfrac{m_0v^2}{2} + \dfrac{3m_0v^4}{8c^2}$.

25. Fund the momentum p of relativistic particle in terms of its kinetic energy T.

    - Solution: $p = \dfrac{1}{c}\sqrt{T\,(T + 2m_0c^2)}$.

26. An ultra-relativistic particle ($E \gg E_0$) with the rest mass m_0 has the energy E. Find its velocity v.

    - Solution: $v = c\left\{1 - \dfrac{1}{2}\left(\dfrac{m_0c^2}{E}\right)^2\right\}$.

27. In the reference frame K in vacuum, there are the electric, E, and magnetic, B, fields. The both fields are constant ($E = \text{const}$ and $B = \text{const}$), mutually perpendicular, and $E < cB$. Find the reference frame in which only the magnetic field is available and calculate its induction. Instruction: use formulas (11.41).

    - Solution:.$v = \dfrac{E}{B}$, $B' = \dfrac{1}{c}\sqrt{B^2c^2 - E^2}$.

28. Monochromatic light with the frequency ω falls normally on the surface of a plane mirror that moves at the constant velocity $\vec{v}$ along the direction of light propagation. Find the frequency of reflected light. Instruction: use formula (11.56).

    - Solution: $\omega = \omega_0\dfrac{1-\beta}{1+\beta}$, $\beta = \dfrac{v}{c}$.

29. While analyzing the quasar radiation spectrum, the red shift of spectral lines was found to equal $z = \Delta\lambda/\lambda$. Calculate the velocity of quazar moving away from the terrestrial observer. Instruction: use formula (11.57).

    - Solution: $z = \dfrac{\lambda - \lambda_0}{\lambda_0} = \dfrac{\omega_0}{\omega} - 1 = \dfrac{1+\beta}{\sqrt{1-\beta^2}} - 1$; whence $\beta = 0.9$, so that $v = 0.9\,c$.

PART II

QUANTUM MECHANICS

Quantum mechanics is a physical theory aimed at finding the ways to describe microparticles—molecules, atoms, elementary particles—and their systems, as well as the laws of their motion. The methods applied by quantum mechanics are also used in other physics domains, such as statistical physics, physics of atomic nucleus, solid state physics, and others. The quantum-mechanical laws made it possible to explain the structure of molecules, atoms, and atomic nuclei; the nature of atomic and molecular spectra; the properties of insulators, metals, and semiconductors; ferromagnetism; superfluidity; and so forth.

The formation of quantum mechanics began when the incapability of classical mechanics methods to study the phenomena in the microworld had been recognized. Its key concepts were formulated in 1925–1928, as a result of the analysis why the methods of classical mechanics and electrodynamics are inapplicable to the phenomena of thermal radiation and photoeffect, to the problems of atomic stability and line structure of atomic spectra.

The first step in this direction was made by M. Planck in 1900. He suggested that light emission by atoms occurs discretely, by portions (quanta), the energy of which is proportional to the frequency ω, i.e.

$$E = \hbar\omega, \tag{II.1}$$

where $\hbar = 1.05 \times 10^{-34}$ J × s is Planck's constant.

In 1905, A. Einstein put forward a hypothesis that light is not only emitted by quanta, but it exists as a true set of quanta (photons) with a certain energy and momentum

$$\vec{p} = \hbar\vec{k}, \tag{II.2}$$

where $\vec{k}$ is the wave vector with the absolute value $\left|\vec{k}\right| = 2\pi/\lambda$. In other words, light are characterized by both quantum-mechanical (corpuscular) and wave properties. This phenomenon is called **the corpuscular-wave dualism**. The wave properties of light manifest themselves in the interference and diffraction phenomena, and the quantum-mechanical ones in the photoeffect phenomenon, the Compton effect, and other phenomena.

In 1924, L. de Broglie extended the corpuscular-wave dualism on all microparticles. According to the de Broglie hypothesis, the wave properties are inherent to all "ordinary" microparticles—molecules, atoms, electrons, protons. Lately, this hypothesis was confirmed in experimental studies dealing with diffraction of various microparticles. According to the de Broglie hypothesis, a free microparticle has to be described by a plane monochromatic wave (the de Broglie wave), with the wavelength given by **the de Broglie formula**

$$\lambda = \frac{2\pi\hbar}{p}. \tag{II.3}$$

In the general case, microparticles are described by wave functions (see section 12.1), which are a generalization of de Broglie waves. The equation for wave functions was derived by E. Schrödinger in 1926. The Schrödinger equation is the fundamental equation of quantum mechanics.

An important contribution to the formation of quantum mechanics was made by W. Heisenberg, N. Bohr, P. Dirac, W. Pauli, and other physicists, including the scientists of the Ukrainian school of theoretical physics–M.M. Bogolyubov, L.D. Landau, O.S. Davydov, and others.

12

KEY CONCEPTS OF QUANTUM MECHANICS

12.1. Wave function

In classical mechanics, the state of material point is described using the coordinate $\vec{r}$ and the momentum $\vec{p} = m\vec{v}$. However, experimental data testify that, owing to the wave properties of microparticles, those parameters are inapplicable for the description of microparticle state.

The wave properties are inherent to all microparticles. Under corresponding conditions, in particular, when microparticles pass through a slit or a system of slits (the diffraction grating), the diffraction and interference phenomena are observed. Therefore, to describe the behavior of microparticle, it has to be associated with a wave.

For instance, a free photon as a light particle has to be associated with a plane monochromatic wave with the vector of electric field strength

$$\vec{E}\,(\vec{r},t) = \vec{E}_0 \mathrm{e}^{i(\vec{k}\vec{r} - \omega t)}. \tag{12.1}$$

Other free microparticles, which are not quanta of electromagnetic field, are associated with abstract plane monochromatic waves of the form

$$\Psi_{\vec{P}}\,(\vec{r},t) = A\mathrm{e}^{i(\vec{k}\vec{r} - \omega t)} = A\mathrm{e}^{\dfrac{i}{\hbar}(\vec{P}\vec{r} - Et)}. \tag{12.2}$$

Such plane monochromatic waves are called the **de Broglie waves**, with de Broglie formula (II.3) being applicable to them.

If the microparticle is bound—e.g., this is an electron in the atom—it is described using a function $\Psi\,(\vec{r},t)$ depending on certain conditions and called the **wave function**. Hence, the de Broglie wave is the wave function of free particle.

To understand the meaning of wave function, let us address, e.g., the example known from optics; namely, let us consider diffraction of light particles—photons—at a slit, i.e. the diffraction of light (Fig. 12.1). Beam a of coherent light (a beam of photons) falls normally on opaque screen ξ_1 with slit d oriented perpendicularly to the figure plane. The diffracted light is focused with the help of lenses L in its focal plane on screen ξ_2 (a photographic plate), where the interference pattern of diffracting light is observed. The diffraction pattern looks like alternating light and dark strips oriented perpendicularly to the figure plane. In Fig. 12.1, the distribution of illumination intensity is schematically shown by a curve, the ordinate of which is proportional to the illumination intensity and consisting of a central (main) maximum and secondary maxima and minima.

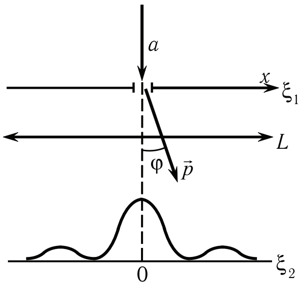

Fig. 12.1.

From optics, it is known that, in the case of light diffraction at one slit, the positions of illumination minima is determined by the relation

$$d\sin\varphi = n\lambda, \quad n = \pm 1, \pm 2, \ldots, \tag{12.3}$$

where φ is the diffraction angle. Therefore,

$$|\sin\varphi| \geq \frac{\lambda}{d}. \tag{12.4}$$

When a photon passes through the slit, the uncertainty of its x-coordinate equals the slit width (see Fig. 12.1), i.e. $\Delta x = d$. Behind the slit, the photon can arrive at various regions on screen ξ_2 corresponding to diffraction maxima of different orders. Therefore, the quantity $p\,|\sin\varphi|$ is the uncertainty in the projection of photon momentum $\vec{p}$ on the axis x, i.e.

$$\Delta p_x = p\,|\sin\varphi| \geq p\frac{\lambda}{d}. \tag{12.5}$$

Taking into account de Broglie formula (II.3), we obtain the relation

$$d\Delta p_x = \Delta x \Delta p_x \geq 2\pi\hbar,$$

which, owing to the smallness of Planck's constant, is written as

$$\Delta x \Delta p_x \geq \hbar. \tag{12.6}$$

Formula (12.6) is called the **uncertainty relation for the coordinate and momentum** of microparticle. It was obtained by W .Heisenberg in 1927. Besides (12.6), we may write another two relations,

$$\Delta y \Delta p_x = \hbar, \quad \Delta z \Delta p_x = \hbar.$$

The Heisenberg uncertainty relations are a consequence of the wave properties of microparticles and testify to the inapplicability of coordinates and momenta for the description of microparticle states. For instance, in order to determine the coordinate x more accurately, i.e. to reduce Δx (see Fig. 12.1), we must make the slip narrower. Then, however, in accordance with (12.6), the momentum uncertainty Δp_x becomes larger and the diffraction pattern on screen ξ_2 becomes wider. And vice versa, in order to measure the momentum more accurately, i.e. to reduce

Δp_x, it is necessary to make the diffraction pattern narrower by widening the slit, so that the coordinate uncertainty Δx increases. Hence, the microparticle cannot be simultaneously characterized by exact coordinate and momentum values in the sense of their simultaneous exact measurement.

If light that falls on the slit is coherent, i.e. monochromatic, its wavelength λ is known exactly (to within the natural width of spectral line). Therefore, in accordance with (II.3), $\Delta p_x \to 0$ and, according to (12.6), $\Delta x \to \infty$, i.e. the location of photon in the incident light beam is unknown.

Photons falling on the slit can be regarded as free. Therefore, they are associated with a wave function in the form of de Broglie wave (12.2). In the case concerned, this is a plane monochromatic electromagnetic wave (12.1). Behind the slit, the wave function of photon so transforms that the electromagnetic field becomes distributed over screen ξ_2 to form a diffraction pattern. It is principally impossible to predict at which point the photon hits the screen. In this case, we must use the probability theory.

While examining the diffraction pattern, it is reasonable to suppose that the probability for a photon to find itself in the region of main maximum is the highest. At the same time, the probabilities that the photon hits the screen in the regions of secondary maxima, where the illumination is lower than in the main maximum, are also lower. Hence, the probability ω for the photon to hit the screen at a certain point $\vec{r}$, calculated per unit area, has to be proportional to the screen illumination at this point, which, in turn, is proportional to the square of the electric field strength magnitude

$$\frac{d\omega\,(\vec{r})}{dS} \sim \left|\vec{E}\,(\vec{r})\right|^2 . \tag{12.7}$$

Taking into account that $\vec{E}(\vec{r}, t)$ has the meaning of photon wave function and extending the scope of applicability of (12.7) onto arbitrary microparticles, it is possible to write the relation

$$\frac{d\omega\,(\vec{r})}{dV} \sim \left|\Psi\,(\vec{r}, t)\right|^2 , \tag{12.8}$$

or

$$d\omega = \left|\Psi\,(\vec{r}, t)\right|^2 dV. \tag{12.9}$$

According to (12.9), the probability $d\omega$ to find a microparticle in the volume dV around a point with the coordinate $\vec{r}$ is equal to the product of the square of the absolute value of wave function at this point and the volume dV. Relationship (1.9) was established by M. Born in 1926.

Hence, the wave function $\Psi(\vec{r}, t)$ is an abstract concept. It has a statistical meaning and determines, according to Born formula (12.9), the probability for a particle to be at the given point in space. In quantum mechanics, the wave function $\Psi(\vec{r}, t)$ is used to describe the state of quantum-mechanical system. According to previous remarks, it has the following properties.

Integrating (12.9) over the coordinates, we obtain the normalization condition for the wave function

$$\int |\Psi(\vec{r}, t)|^2 \, dV = 1. \tag{12.10}$$

However, in the case of free particle, from (12.2) and (12.10), it follows that

$$\int \left|\Psi_{\vec{P}}(\vec{r}, t)\right|^2 dV = A^2 \int_0^\infty dV \to \infty.$$

This result stems from the idealized character of free-particle concept, because there are no completely free particles in the nature. In this case, the wave function of free particle is written in the form

$$\Psi_{\vec{P}}(\vec{r}, t) = (2\pi\hbar)^{-3/2} \exp\left\{\frac{i}{\hbar}(\vec{p}\vec{r} - Et)\right\} \tag{12.11}$$

and normalized to the Dirac delta-function (see Appendix A),

$$\int \Psi^*_{\vec{P}}(\vec{r}, t)\, \Psi_{\vec{P}'}(\vec{r}, t)\, dV = \delta\left(\vec{P}' - \vec{P}\right). \tag{12.12}$$

From (12.10), it follows that the normalized wave function can be determined to within the factor $\mathrm{e}^{i\alpha}$, where α is a real-valued number, because $\left|\mathrm{e}^{i\alpha}\right|^2 = 1$.

The wave function must be single-valued and continuous; otherwise, the probability of particle location would have been ambiguous and discontinuous, which contradicts physical conditions. Besides, in general, the wave function can be either a real- or a complex-valued quantity, but mathematical operations with complex functions are easier. Therefore, in what follows, the wave functions are supposed to be complex. In the exponential form, they read

$$\Psi = |\Psi|\, \mathrm{e}^{i\alpha},$$

where $|\Psi|$ is the absolute value of function Ψ, and $\alpha = \arctan\left(\mathrm{Im}\ \Psi / \mathrm{Re}\ \Psi\right)$ is the phase, which is put equal to zero in most cases.

Provided that normalization condition (12.10) is obeyed, the wave functions Ψ and $C\Psi$, where C is an arbitrary constant multiplier, describe the same state, because

$$\frac{d\omega}{dV} = |\Psi|^2 = \frac{|\Psi|^2}{1} = \frac{|\Psi|^2}{\int |\Psi|^2 \, dV} = \frac{|C\Psi|^2}{\int |C\Psi|^2 \, dV}. \tag{12.13}$$

Of large importance is the property of wave functions, which is called the **principle of superposition states**. It declares that, if a quantum-mechanical system can exist in states described by the wave functions Ψ_1, Ψ_2,..., Ψ_n,..., it can also exist in a state described by the wave function

$$\Psi = \sum_n C_n \Psi_n, \tag{12.14}$$

where C_n are arbitrary constants. The superposition principle gives rise, in particular, to the linearity of the quantum-mechanical equations.

12.2. Operators in quantum mechanics

Uncertainties in the coordinate x and momentum p_x of microparticle (as well as in the (y, p_y) and (z, p_z) pairs) are responsible for the uncertainties in the quantity values, which are functions of coordinates and momenta. In particular, these are the kinetic and potential energies, the angular momentum, and so on. This circumstance, in turn, invokes the application of such concepts as probable (allowable) quantity values, their probabilities and average values. In this connection, there arises an issue concerning the methods of their calculation. Therefore, let us consider first the following example.

Let a particle be in a state described by the wave function $\Psi(\vec{r}, t)$. Then, it is possible to prove (see section 2.1 in [20]) that the average value of its momentum can be calculated by the formula

$$\langle \vec{p} \rangle = \int \Psi^* (\vec{r}, t) \left(-i\hbar \nabla\right) \Psi (\vec{r}, t) \, dV, \tag{12.15}$$

where the asterisk (*) denotes the operation of complex conjugation, and the symbol ∇ means the operator

$$\nabla = \vec{i}\frac{\partial}{\partial x} + \vec{j}\frac{\partial}{\partial y} + \vec{k}\frac{\partial}{\partial z}. \tag{12.16}$$

From (12.15), it follows that, while calculating the average value of momentum, the operator $-i\hbar\nabla$ plays the role of momentum quantity.

Generalizing this example, we may draw a conclusion that, in quantum mechanics, every physical quantity is associated with a certain **operator;** i.e. every physical quantity is represented by the corresponding operator. Let us consider the main properties of quantum-mechanical operators. Below, the operators are denoted by letters with a "hat" over it. For instance, if A is a certain physical quantity, the corresponding quantum-mechanical operator is $\widehat{A}$.

The **operator** symbolizes a definite action (operation). In particular, the equality

$$\widehat{A}u\,(x) = v\,(x) \tag{12.17}$$

means that the function $u\,(x)$, where x is the general notation of coordinates, being undergone some action (the operator $\widehat{A}$), transforms into the function $v(x)$.

Only linear self-adjoint operators are used in quantum mechanics. The operator $\widehat{A}$ is called **linear** if the relation

$$\widehat{A}\,(c_1u_1 + c_2u_2) = c_1\widehat{A}u_1 + c_2\widehat{A}u_2 \tag{12.18}$$

holds true for any functions u_1 and u_2, and any constants c_1 and c_2. This property follows from the principle of superposition states.

The **self-adjoint** (**Hermitian**) operator satisfies the condition

$$\int u_1^*\,(x)\,\widehat{A}u_2\,(x)\,dx = \int u_2\,(x)\,\widehat{A}^*u_1^*\,(x)\,dx, \tag{12.19}$$

where u_1 and u_2 are functions for which the operator $\widehat{A}$ is determined, and integration is carried out over the whole region of variables x. Condition (12.19) is often written in the form

$$A_{12} = A_{21}^*, \tag{12.20}$$

where

$$A_{12} = \int u_1^* \widehat{A} u_2 dx$$

is the matrix element of operator $\widehat{A}$ calculated with the help of functions u_1 and u_2. The self-adjointness of operators in quantum mechanics stems from the requirement that all physical quantities are real-valued (not complex).

The operator $\widehat{C}$ is called the sum of operators $\widehat{A}$ and $\widehat{B}$ (i.e. $\widehat{C} = \widehat{A} + \widehat{B}$) if

$$\widehat{C}u = \left(\widehat{A} + \widehat{B}\right) u = \widehat{A}u + \widehat{B}u. \tag{12.21}$$

Analogously, the product of operators $\widehat{A}$ and $\widehat{B}$ is the operator $\widehat{C} = \widehat{A}\widehat{B}$ if

$$\widehat{C}u = \left(\widehat{A}\widehat{B}\right) u = \widehat{A}\left(\widehat{B}u\right). \tag{12.22}$$

The operator ratio is not determined. The operator raised to some power is defined as the operator product; e.g., $\widehat{A}^3 = \widehat{A}\widehat{A}\widehat{A}$.

In general, operators do not commute, i.e. $\widehat{A}\widehat{B} \neq \widehat{B}\widehat{A}$. The operator

$$\widehat{K} = \widehat{A}\widehat{B} - \widehat{B}\widehat{A}$$

is called the commutator of operators $\widehat{A}$ and $\widehat{B}$.

One should pay attention that operators are symbols denoting actions with respect to quantities rather than the quantities (numbers) themselves. Therefore, the action with operators can be carried out only implying that they act on the corresponding functions.

Now, let us return to relation (12.15). Adopting that

$$\widehat{\vec{p}} = -i\hbar\nabla \tag{12.23}$$

is the operator of momentum, the average value of any physical quantity F must be calculated according to the formula

$$\langle F \rangle = \int \Psi^* \widehat{F} \Psi d\tau, \tag{12.24}$$

where $\widehat{F}$ is the operator of the quantity to be averaged, and $d\tau$ is the product of the differentials of those variables that the wave function Ψ depends on. In particular, according to (12.24),

$$\langle \vec{r} \rangle = \int \Psi^* \widehat{\vec{r}} \Psi dV. \tag{12.25}$$

Therefore, the average value $\langle\vec{r}\rangle$ can be written, in accordance with the probability theory, in the form

$$\langle\vec{r}\rangle = \int \vec{r} d\omega\left(\vec{r}\right), \tag{12.26}$$

where $d\omega\left(\vec{r}\right)$ is the probability that the particle is in the volume dV with the coordinate $\vec{r}$. However, on the basis of (12.9),

$$d\omega\left(\vec{r}\right) = \left|\Psi\left(\vec{r}\right)\right|^2 dV = \Psi^*\left(\vec{r}\right)\Psi\left(\vec{r}\right) dV. \tag{12.27}$$

Therefore, in accordance with (12.25)–(12.27), we may write

$$\int \Psi^* \widehat{\vec{r}} \Psi dV = \int \vec{r} \Psi^* \Psi dV = \int \Psi^* \vec{r} \Psi dV, \tag{12.28}$$

where we took into account that the vector $\vec{r}$ in the product $\vec{r}\Psi^*\Psi$ is an ordinary multiplier. At last, comparing (12.25) and (12.28), we come to a conclusion that the operator of coordinate is the coordinate itself, i.e.

$$\widehat{\vec{r}} = \vec{r}, \quad \widehat{x} = x, \quad \widehat{y} = y, \quad \widehat{z} = z. \tag{12.29}$$

According to (12.23) and (12.16), the operators of momentum and its projections look like

$$\widehat{\vec{p}} = -i\hbar\nabla, \quad \widehat{p}_x = -i\hbar\frac{\partial}{\partial x}, \quad \widehat{p}_y = -i\hbar\frac{\partial}{\partial y}, \quad \widehat{p}_z = -i\hbar\frac{\partial}{\partial z}. \tag{12.30}$$

Let some quantity F be, in classical mechanics, a function $F\left(\vec{r},\vec{p}\right)$ of coordinates and momenta. Then, in quantum mechanics, its operator is defined as

$$\widehat{F}\left(\widehat{\vec{r}},\widehat{\vec{p}}\right) = F\left(\vec{r},\widehat{\vec{p}}\right). \tag{12.31}$$

In particular, in classical mechanics, the projections of angular momentum $\vec{L} = \left[\vec{r}\vec{p}\right]$ look like

$$L_x = yp_z - zp_y, \quad L_y = zP_x - xP_z, \quad L_z = xP_y - yP_x. \tag{12.32}$$

Therefore, in quantum mechanics,

$$\begin{aligned}
\widehat{L}_x &= y\widehat{p}_z - z\widehat{p}_y = -i\hbar\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right), \\
\widehat{L}_y &= z\widehat{p}_x - x\widehat{p}_z = -i\hbar\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right), \\
\widehat{L}_z &= x\widehat{p}_y - y\widehat{p}_x = -i\hbar\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right).
\end{aligned} \tag{12.33}$$

In the same way, we can obtain the operators of potential, $U = U(\vec{r}, t)$, and kinetic, $T = \frac{1}{2m}\vec{p}^{\,2}$, energies for the particle,

$$\begin{aligned} \widehat{U} &= U\left(\widehat{\vec{r}}, t\right) = U\left(\vec{r}, t\right), && (12.34)\\ \widehat{T} &= \frac{1}{2m}\widehat{\vec{p}}^{\,2} = \frac{1}{2m}\left(-i\hbar\nabla\right)^2 = -\frac{\hbar^2}{2m}\Delta, && (12.35)\end{aligned}$$

where Δ is the **Laplace operator** (the **Laplacian**)

The total energy of the particle, $H = T + U$, in classical mechanics is called the **Hamilton function**. Therefore, the operator of total energy, i.e. the **Hamilton operator**, or the **Hamiltonian**, is written as

$$\widehat{H} = \widehat{T} + \widehat{U} = -\frac{\hbar^2}{2m}\Delta + U\left(\vec{r}, t\right). \tag{12.36}$$

If the quantum-mechanical system has the symmetry center, e.g., an atom, the spherical coordinate system is used as a rule. Therefore, below, we present the operators $\widehat{L}_z$, $\widehat{L}^2$, and $\widehat{H}$ in spherical coordinates (see Appendix E) omitting the stage of their derivation:

$$\widehat{L}_z = -i\hbar\frac{\partial}{\partial\varphi}, \tag{12.37}$$

$$\widehat{L}^2 = -\hbar^2\left\{\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right\}, \tag{12.38}$$

$$\widehat{H} = \widehat{T}_r + \frac{1}{2mr^2}\widehat{L}^2 + U\left(\vec{r}\right), \tag{12.39}$$

$$\widehat{T}_r = -\frac{\hbar^2}{2mr^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right). \tag{12.40}$$

12.3. Eigenvalues and eigenfunctions of operators

Let a quantum-mechanical system be in a state Ψ, in which the quantity F has an exact value. Then, $\Delta F = F - \langle F\rangle = 0$, as well as the dispersion

$\left\langle (\Delta F)^2 \right\rangle = \left\langle (F - \langle F \rangle)^2 \right\rangle = 0$. On the basis of (12.24) and using the condition of self-adjointness of operator (12.19), we obtain

$$\begin{aligned} \left\langle (\Delta F)^2 \right\rangle &= \int \Psi^* \left(\Delta \widehat{F}\right)^2 \Psi d\tau = \int \Psi^* \Delta \widehat{F} \left(\Delta \widehat{F} \Psi\right) d\tau \\ &= \int \left(\Delta \widehat{F} \Psi\right) \Delta \widehat{F}^* \Psi^* d\tau = \int \left|\Delta \widehat{F} \Psi\right|^2 d\tau = 0. \end{aligned} \quad (12.41)$$

Whence it follows that $\Delta \widehat{F} \Psi = 0$. The operator $\Delta \widehat{F}$ looks like $\Delta \widehat{F} = \widehat{F} - \langle F \rangle$. Therefore, $\left(\widehat{F} - \langle F \rangle\right) \Psi = 0$. Taking into consideration that $\langle F \rangle = F$ in this case, we find the equation

$$\widehat{F} \Psi = F \Psi, \quad (12.42)$$

in which the action of operator on the function Ψ is reduced to the multiplication of the latter by a certain number. Equations of type (12.42) are called the **eigenvalue-eigenfunction equations** for the operator $\widehat{F}$. In the general form, they are written as

$$\widehat{F} u = F u, \quad (12.43)$$

where u is the eigenfunction and F the eigenvalue of operator $\widehat{F}$. By solving such an equation taking into account the corresponding boundary conditions, we can determine the set (spectrum) of eigenfunctions and the spectrum of eigenvalues, the both can be discrete or continuous depending on the boundary conditions.

From (12.42), it follows that the quantity F has an exact value if the wave function Ψ is the eigenfunction of operator $\widehat{F}$. If the wave function Ψ of the system does not satisfy equation (12.42), the value of quantity F is uncertain, because different measurements carried out using identical objects in the same quantum-mechanical state Ψ produce different F-values, which belong to the spectrum of F-eigenvalues obtained for equation (12.43). This spectrum of values constitutes the physical meaning of eigenvalues for the specific operator.

If the spectra of eigenvalues and eigenfunctions are discrete, equation (12.43) takes the form

$$\widehat{F} u_n = F_n u_n. \quad (12.44)$$

If the eigenvalue F_n is associated with only one eigenfunction u_n, this eigenvalue is called **nondegenerate**, If there are f such eigenfunctions $(u_{n1}, u_{n2}, \ldots, u_{nf})$, the eigenvalue F_n is called f-fold **degenerate**.

If the set of eigenfunctions is discrete, those functions satisfy the **orthonormality condition**

$$\int u_m^* u_n d\tau = \delta_{mn} = \begin{cases} 1, & m = n, \\ 0, & m \neq n. \end{cases} \tag{12.45}$$

In the case, when the spectra of eigenvalues and eigenfunctions are continuous, equation (12.43) is written as

$$\widehat{F} u_F = F u_F, \tag{12.46}$$

and the functions u_F satisfy the condition

$$\int u_F^* u_{F'} d\tau = \delta \left(F - F' \right). \tag{12.47}$$

The system of eigenfunctions for the given operator is complete. Therefore, any function depending on the same variables and obeying the same boundary conditions as the functions u_n's can be expanded in a series of u_n. For the wave function, the corresponding expansion looks like

$$\Psi \left(x \right) = \sum_n C_n u_n \left(x \right), \tag{12.48}$$

where

$$C_n = \int \Psi \left(x \right) u_n^* \left(x \right) dx. \tag{12.49}$$

The quantity

$$w \left(F_n \right) = \left| C_n \right|^2 \tag{12.50}$$

is the probability that, while measuring the quantity F, we obtain the value $F = F_n$. If the spectra of eigenvalues and eigenfunctions of operator are continuous, instead of (1.53)–(1.55), we must write

$$\Psi \left(x \right) = \int C \left(F \right) u_F \left(x \right) dF, \tag{12.51}$$

$$C \left(F \right) = \int \Psi \left(x \right) u_F^* \left(x \right) dx. \tag{12.52}$$

In this case, the quantity $|C(F)|^2\, dF$ is the probability of finding the value of quantity F within the interval from F to $F + dF$.

Besides, another two important circumstances have to be noted. First, according to uncertainty relation (12.6) two canonically conjugate quantities x_i and P_i cannot have exact values simultaneously. They are not measurable simultaneously in the sense that the measurement of either quantity inserts the uncertainty into the other. In this connection, the following question arises: Under what conditions can two quantities, A and B, be measured simultaneously? In other words, when do they have exact values simultaneously? It can be proved (see [20]) that this condition consists in the commutation of the operators of those quantities, i.e. $\widehat{A}\widehat{B} = \widehat{B}\widehat{A}$. It is easy to be convinced that $\widehat{x}\widehat{p}_x - \widehat{p}_x\widehat{x} = i\hbar \neq 0$, so that x and p_x are not simultaneously measurable.

The second circumstance consists in that the commutating operators $\widehat{A}$ and $\widehat{B}$ (i.e. $\widehat{A}\widehat{B} = \widehat{B}\widehat{A}$) have the same set of eigenfunctions (see [20]).

12.4. Schrödinger equation

Every state of quantum-mechanical system is described by a wave function $\Psi(\vec{r}, t)$. This function is used to calculate the average values of physical quantities (see (12.24)), the probabilities of their possible values, and so forth. Therefore, there arises a problem how the wave function can be determined in every specific case. In other words, it is necessary to find an equation, the solution of which gives the wave function.

Such an equation must be the fundamental equation of quantum mechanics. Therefore, in accordance with the principles of theoretical physics, this equation cannot be derived theoretically. Instead, it can only be substantiated on the basis of experimental data. Its validity must be tested by comparing the results following from this equation and experimental data.

There are a number of ways to substantiate the equation concerned. Let us choose the simplest one using the fact that we already know the wave function for a free particle. According to (12.2), this function looks like

$$\Psi_{\vec{p}}(\vec{r}, t) = A\mathrm{e}^{\dfrac{i}{\hbar}(\vec{p}\vec{r} - Et)}. \tag{12.53}$$

Acting on it with the operator of kinetic energy (12.35), we obtain

$$\widehat{T}\Psi_{\vec{p}}\left(\vec{r},t\right)=-\frac{\hbar^2}{2m}\Delta\Psi_{\vec{p}}\left(\vec{r},t\right)=-\frac{\hbar^2}{2m}\left(\frac{\partial}{\partial\vec{r}}\right)^2\Psi_{\vec{p}}\left(\vec{r},t\right)=\frac{\vec{p}^{\,2}}{2m}\Psi_{\vec{p}}\left(\vec{r},t\right). \tag{12.54}$$

Here, we took into account that $\Delta=\nabla^2=\left(\frac{\partial}{\partial\vec{r}}\right)^2$. It is easy to be convinced that

$$i\hbar\frac{\partial}{\partial t}\Psi_{\vec{p}}\left(\vec{r},t\right)=E\Psi_{\vec{p}}\left(\vec{r},t\right). \tag{12.55}$$

In the free-particle case, $E=T=\vec{p}^{\,2}/\left(2m\right)$. Therefore, from (12.54) and (12.55), it follows that

$$i\hbar\frac{\partial\Psi}{\partial t}=\widehat{T}\Psi. \tag{12.56}$$

In the general case, the operator of kinetic energy $\widehat{T}$ has to be supplemented with the operator of potential energy $\widehat{U}$, i.e. to be substituted by the Hamilton operator $\widehat{H}$. As a result, we obtain the sought equation

$$i\hbar\frac{\partial\Psi}{\partial t}=\widehat{H}\Psi. \tag{12.57}$$

It is called the **general** or **time-dependent**or **temporal Schrödinger equation**.

Taking into account (12.36), Schrödinger equations (12.57) for a particle can be written as follows:

$$i\hbar\frac{\partial\Psi\left(\vec{r},t\right)}{\partial t}=-\frac{\hbar^2}{2m}\Delta\Psi\left(\vec{r},t\right)+U\left(\vec{r},t\right)\Psi\left(\vec{r},t\right). \tag{12.58}$$

In quantum mechanics, the Schrödinger equation (see formula (12.57) or (12.58)) plays the role of the equation of motion like the Lagrange (Hamilton) equations in classical mechanics. However, the solutions of Lagrange equations determine unequivocally the coordinates and momenta of the particle, whereas the Schrödinger equation allows one only to obtain the wave function, which can be used to calculate only the average values of physical quantities.

Analyzing Schrödinger equation (12.58), we would like to note that a number of important conclusions can be drawn without solving it. Let us write equation (12.58) and the complex-conjugate one, taking into account that $U^* = U$,

$$\begin{aligned} i\hbar\frac{\partial\Psi}{\partial t} &= -\frac{\hbar^2}{2m}\Delta\Psi + U\Psi, && (12.59)\\ -i\hbar\frac{\partial\Psi^*}{\partial t} &= -\frac{\hbar^2}{2m}\Delta\Psi^* + U\Psi^*. && (12.60)\end{aligned}$$

Multiplying equation (12.59) by Ψ^* from the left and equation (12.60) by Ψ from the right, and subtracting the results, we obtain

$$i\hbar\left(\Psi^*\frac{\partial\Psi}{\partial t} + \Psi\frac{\partial\Psi^*}{\partial t}\right) = -\frac{\hbar^2}{2m}\left(\Psi^*\Delta\Psi - \Psi\Delta\Psi^*\right). \tag{12.61}$$

The expression in the parentheses in the left hand side of (12.61) can be written as

$$\Psi^*\frac{\partial\Psi}{\partial t} + \Psi\frac{\partial\Psi^*}{\partial t} = \frac{\partial}{\partial t}\left|\Psi\right|^2 = \frac{\partial\rho_i}{\partial t}, \tag{12.62}$$

where

$$\rho_i = \left|\Psi\right|^2 = \frac{\partial w}{\partial V}$$

is the probability density for the particle coordinate. At the same time, the expression in the parentheses in the right hand side of (12.61), making use of the following formula of vector analysis (see Appendix B):

$$\operatorname{div}\left(\varphi\vec{a}\right) = \varphi\nabla\vec{a} + \vec{a}\nabla\varphi = \varphi\operatorname{div}\vec{a} + \vec{a}\nabla\varphi, \tag{12.63}$$

in which the substitutions $\varphi = \Psi$ and $\vec{a} = \nabla\Psi^*$ are made, acquires the form

$$\Psi^*\Delta\Psi - \Psi\Delta\Psi^* = \operatorname{div}\left(\Psi\nabla\Psi^* - \Psi^*\nabla\Psi\right). \tag{12.64}$$

On the basis of (12.62) and (12.64), instead of (12.61), we obtain the equation

$$\frac{\partial\rho_i}{\partial t} + \operatorname{div}\vec{j}_i = 0, \tag{12.65}$$

where the notation

$$\vec{j}_i = \frac{i\hbar}{2m}\left(\Psi\nabla\Psi^* - \Psi^*\nabla\Psi\right) \tag{12.66}$$

is used.

Equation (12.65) is called the **continuity equation**. Actually, this is a conservation law, in our case, for the probability density of particle coordinate ρ_i. Therefore, equation (12.65) is the **probability conservation law**, i.e. the **preservation of the particle number**, and vector (12.66) is the **probability flux density vector**.

The quantity

$$\rho = e\rho_i = e\left|\Psi\right|^2, \tag{12.67}$$

where e is the electric charge of the particle, has the meaning of the electric charge density associated with the particle motion, so that the particle charge becomes "smeared" in the space. The vector

$$\vec{j} = e\vec{j}_i = \frac{ie\hbar}{2m}\left(\Psi\nabla\Psi^* - \Psi^*\nabla\varphi\right) \tag{12.68}$$

is the vector of electric current density that emerges owing to the particle motion.

Let us consider an important case, when the potential energy and, hence, the Hamiltonian do not depend on the time, $\widehat{H} \neq \widehat{H}\left(t\right)$. The corresponding system states are called **stationary**, and the wave function $\Psi\left(\vec{r}, t\right)$ can be tried as the product

$$\Psi\left(\vec{r}, t\right) = \Psi\left(\vec{r}\right)\varphi\left(t\right). \tag{12.69}$$

The Hamiltonian $\widehat{H}$ acts only on the multiplier $\Psi\left(\vec{r}\right)$. Therefore, substituting (12.69) into Schrödinger equation (12.57) and dividing the result by product (12.69), we obtain the relation

$$i\hbar\frac{1}{\varphi\left(t\right)}\frac{\partial\varphi\left(t\right)}{\partial t} = \frac{1}{\Psi\left(\vec{r}\right)}\widehat{H}\Psi\left(\vec{r}\right), \tag{12.70}$$

in which the left hand side depends only on the time, and the right hand side only on the coordinates. It is possible only provided that each side equals the same constant, which we designate as E. Then, as a result, we obtain two equations,

$$\frac{\partial\varphi\left(t\right)}{\partial t} = -\frac{i}{\hbar}E\varphi\left(t\right), \tag{12.71}$$

$$\widehat{H}\Psi\left(\vec{r}\right) = E\Psi\left(\vec{r}\right). \tag{12.72}$$

The solution of equation (12.71) is

$$\varphi(t) = \mathrm{e}^{-\frac{i}{\hbar}Et} = \mathrm{e}^{-i\omega t}, \tag{12.73}$$

where $E = \hbar\omega$.

Equation (12.72) cannot be solved in the general case, because the solution depends on the form of potential energy $U(\vec{r})$. By its form, equation (12.72) is an eigenvalue-eigenfunction equation for the Hamilton operator. Therefore, the quantity E has the meaning of energy, and the equation itself is called the **Schrödinger equation for stationary states**.

Hence, the wave function of stationary state (12.69) is written in the form

$$\Psi(\vec{r}, t) = \Psi(\vec{r})\,\mathrm{e}^{-i\omega t}. \tag{12.74}$$

The multiplier $\Psi(\vec{r})$, which is called the **wave function of stationary state**, is determined from equation (12.72).

It is easy to be convinced that, in the stationary state, the average values of quantities (12.24) and the probabilities of their allowed values (12.50) are time-independent. In particular, substituting (12.74) into (12.24), we obtain

$$\langle F\rangle = \int \Psi^*(\vec{r})\,\mathrm{e}^{i\omega t}\widehat{F}\Psi(\vec{r})\,\mathrm{e}^{-i\omega t}dV = \int \Psi^*(\vec{r})\,\widehat{F}\Psi(\vec{r})\,dV. \tag{12.75}$$

12.5. Conservation laws in quantum mechanics

The time-dependent Schrödinger equation (12.57) makes it possible to substantiate the conservation laws of physical quantities in quantum mechanics. Differentiating the average value of a certain quantity F,

$$\langle F\rangle = \int \Psi^*\widehat{F}\Psi d\tau, \tag{12.76}$$

with respect to the time, we obtain

$$\frac{d\langle F\rangle}{dt} = \int \left\{\frac{\partial\Psi^*}{\partial t}\widehat{F}\Psi + \Psi^*\frac{\partial\widehat{F}}{\partial t}\Psi + \widehat{F}\frac{\partial\Psi}{\partial t}\right\} d\tau, \tag{12.77}$$

where we took into account that

$$\frac{d\Psi}{dt} = \frac{\partial\Psi}{\partial t} + (\nabla\Psi)\frac{\partial\vec{r}}{\partial t} = \frac{\partial\Psi}{\partial t}, \tag{12.78}$$

because, as a result of the uncertainty of coordinate $\vec{r}$, the function $\vec{r} = \vec{r}(t)$ does not exist. Opening the brackets in (12.77) , we obtain a sum of integrals

$$\frac{d\langle F\rangle}{dt} = \int \frac{\partial\Psi^*}{\partial t}\widehat{F}\Psi d\tau + \int \Psi^*\frac{\partial\widehat{F}}{\partial t}\Psi d\tau + \int \Psi^*\widehat{F}\frac{\partial\Psi}{\partial t}d\tau. \tag{12.79}$$

From Schrödinger equation (12.57), it follows that

$$\frac{\partial\Psi}{\partial t} = -\frac{i}{\hbar}\widehat{H}\Psi, \quad \frac{\partial\Psi^*}{\partial t} = \frac{i}{\hbar}\widehat{H}^*\Psi^*.$$

Therefore, equation (12.79) can be written in the form

$$\frac{d\langle F\rangle}{dt} = \frac{i}{\hbar}\int\left(\widehat{H}^*\Psi^*\right)\widehat{F}\Psi^* d\tau + \int \Psi^*\frac{\partial\widehat{F}}{\partial t}\Psi d\tau - \frac{i}{\hbar}\int \Psi^*\widehat{F}\widehat{H}\Psi d\tau. \tag{12.80}$$

The first integral is transformed using the self-adjointness of operator $\widehat{H}$ in accordance with (12.19),

$$\int\left(\widehat{H}^*\Psi^*\right)\widehat{F}\Psi d\tau = \int\left(\widehat{F}\Psi\right)\widehat{H}^*\Psi^* d\tau = \int \Psi^*\widehat{H}\widehat{F}\Psi d\tau. \tag{12.81}$$

Then, equation (12.80) can be written as

$$\begin{aligned} \frac{d\langle F\rangle}{dt} &= \int\left\{\frac{i}{\hbar}\Psi^*\widehat{H}\widehat{F}\Psi + \Psi^*\frac{\partial\widehat{F}}{\partial t}\Psi - \frac{i}{\hbar}\Psi^*\widehat{F}\widehat{H}\Psi\right\}d\tau \\ &= \int \Psi^*\left\{\frac{\partial\widehat{F}}{\partial t} + \frac{i}{\hbar}\right\}\left(\widehat{H}\widehat{F} - \widehat{F}H\right)\Psi d\tau. \end{aligned} \tag{12.82}$$

Since the time derivative of the average value equals the average value of the time derivative, we obtain

$$\begin{aligned} \frac{d\langle F\rangle}{dt} &= \left\langle\frac{dF}{dt}\right\rangle = \int \Psi^*\frac{d\widehat{F}}{dt}\Psi d\tau \\ &= \int \Psi^*\left\{\frac{\partial\widehat{F}}{\partial t} + \frac{i}{\hbar}\Psi^*\left(\widehat{H}\widehat{F} - \widehat{F}\widehat{H}\right)\right\}\Psi d\tau. \end{aligned} \tag{12.83}$$

Whence it follows that the time derivative of operator is calculated as

$$\frac{d\widehat{F}}{dt} = \frac{\partial \widehat{F}}{\partial t} + \frac{i}{\hbar}\left(\widehat{H}\widehat{F} - \widehat{F}\widehat{H}\right). \tag{12.84}$$

Another important conclusion following from (12.83) consists in that, provided

$$\frac{\partial \widehat{F}}{\partial t} = 0, \tag{12.85}$$

$$\widehat{H}\widehat{F} = \widehat{F}\widehat{H}, \tag{12.86}$$

we have $\dfrac{d\langle F\rangle}{dt} = 0$. Hence, $\langle F\rangle = \text{const}$, i.e. the average value of F remains constant.

Thus, the average values of quantum-mechanical quantities, which operators do not depend on the time ($\partial\widehat{F}/\partial t = 0$) or commute with the system Hamiltonian ($\widehat{H}\widehat{F} = \widehat{F}\widehat{H}$), are constant.

The conservation laws are closely connected with the system symmetry. As such, we understand transformations (operations), with respect to which the system and its Hamiltonian remain invariant. Symmetry transformations (rotations, reflections, translations in space and time, permutations, and so forth) are presented by corresponding operators. If $\widehat{F}$ is an operator of symmetrical transformation, it must commute with the system Hamiltonian ($\widehat{H}\widehat{F} = \widehat{F}\widehat{H}$), and if $\widehat{F}$ is an operator of continuous (not discrete) transformation and $\partial\widehat{F}/\partial t = 0$, then, in accordance with (12.85) and (12.86), there must exist the corresponding conservation law. This is the essence of **Noether's theorem**. Let us consider a few examples.

Let a quantum-mechanical system be closed. Then, it does not change (i.e. invariant) in time. This property is a consequence of **time homogeneity**, the equivalence of all its moments. The Hamiltonian of such a system does not depend on the time, $\partial\widehat{H}/\partial t = 0$. Then, in accordance with (12.84),

$$\frac{d\widehat{H}}{dt} = \frac{\partial \widehat{H}}{\partial t} + \frac{i}{\hbar}\left(\widehat{H}\widehat{H} - \widehat{H}\widehat{H}\right) = 0,$$

and, according to (12.83), $\langle E\rangle = \text{const}$. Hence, the energy conservation law stems from the time homogeneity.

The momentum conservation law is a consequence of space homogeneity, which means the equivalence of all space points. Therefore, a closed system and its Hamiltonian must be invariant with respect to spatial translations. Consider an infinitesimal shift of the system by the vector $\delta\vec{a}$ (Fig. 12.2). Then, point A of the system, where the wave function is $\Psi\left(\vec{r}-\delta\vec{a}\right)$, shifts to point B, where the wave function was $\Psi\left(\vec{r}\right)$ before the translation; i.e. the wave function transforms as follows:

$$\Psi\left(\vec{r}\right)\to\Psi\left(\vec{r}-\delta\vec{a}\right). \tag{12.87}$$

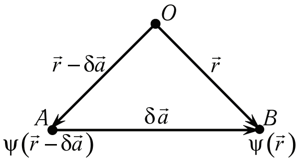

Fig. 12.2.

Hence, the translation operator $\widehat{T}\left(\delta\vec{a}\right)$ has to be defined according to the formula

$$\widehat{T}\left(\delta\vec{a}\right)\Psi\left(\vec{r}\right)=\Psi\left(\vec{r}-\delta\vec{a}\right). \tag{12.88}$$

Expanding the function $\Psi\left(\vec{r}-\delta\vec{a}\right)$ in a power series of $\delta\vec{a}$ and taking into account that $\delta\vec{a}$ is an infinitesimal quantity, we obtain

$$\widehat{T}\left(\delta\vec{a}\right)\Psi\left(\vec{r}\right)=\Psi\left(\vec{r}-\delta\vec{a}\right)=\Psi\left(\vec{r}\right)-\delta\vec{a}\nabla\Psi\left(\vec{r}\right)=\left(1-\delta\vec{a}\nabla\right)\Psi\left(\vec{r}\right).$$

Therefore, the operator $\widehat{T}\left(\delta\vec{a}\right)$ looks like

$$\widehat{T}\left(\delta\vec{a}\right)=1-\delta\vec{a}\nabla=1-\frac{i}{\hbar}\delta\vec{a}\Psi\widehat{\vec{p}}, \tag{12.89}$$

where the relation $\widehat{\vec{p}}=-i\hbar\nabla$ is used. The invariance of closed system with respect to translations means that

$$\widehat{T}\left(\delta\vec{a}\right)\widehat{H}=\widehat{H}\widehat{T}\left(\delta\vec{a}\right). \tag{12.90}$$

Substituting (12.89) into (12.90), after simple transformations, we obtain

$$\widehat{\vec{p}}\widehat{H}=\widehat{H}\widehat{\vec{p}}. \tag{12.91}$$

Thus, we arrive at the conclusion that the operator of momentum $\widehat{p}=-i\hbar\nabla$ does not depend on time and commutes with the Hamiltonian of closed system. Therefore, the momentum $\vec{p}$ is invariant.

In a similar way, it is possible to prove that the conservation law of angular momentum in the closed system is a consequence of space isotropy, i.e. the equivalence of all directions in the space (see section 3.6 in [20]).

13

PARTICLE MOTION IN ONE DIMENSION

13.1. General properties of one-dimensional motion

Proceeding to the solution of problems in quantum mechanics, let us consider first the simplest systems, where the particle moves in one dimension and the corresponding stationary Schrödinger equations have exact solutions. In this section, we analyze the general properties of such a motion.

We suppose that the particle can move only along the axis x. Then, its coordinates y and z do not change, and the Laplace operator Δ can be written as

$$\Delta = \frac{d^2}{dx^2}. \tag{13.1}$$

Accordingly, the Hamiltonian operator for stationary states takes the form

$$\widehat{H} = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + U\left(x\right). \tag{13.2}$$

The, the stationary Schrödinger equation (12.72) looks like

$$-\frac{\hbar^2}{2m}\frac{d^2\Psi\left(x\right)}{dx^2} + U\left(x\right)\Psi\left(x\right) = E\Psi\left(x\right). \tag{13.3}$$

In this case, the energy levels are nondegenerate, i.e. only one wave function corresponds to a definite energy value. To prove this, let us rewrite equation (13.3) in the form

$$\frac{1}{\Psi}\frac{d^2\Psi}{dx^2} = \frac{2m\left(U - E\right)}{\hbar^2}. \tag{13.4}$$

Suppose that two wave functions, Ψ_1 and Ψ_2, correspond to the energy E. This means that

$$\frac{1}{\Psi_1}\frac{d^2\Psi_1}{dx^2} = \frac{2m\left(U - E\right)}{\hbar^2}, \quad \frac{1}{\Psi_2}\frac{d^2\Psi_2}{dx^2} = \frac{2m\left(U - E\right)}{\hbar^2}. \tag{13.5}$$

Since the right hand sides of equations (13.5) are equal, we have

$$\frac{1}{\Psi_1}\frac{d^2\Psi_1}{dx^2} = \frac{1}{\Psi_2}\frac{d^2\Psi_2}{dx^2} \tag{13.6}$$

or

$$\Psi_2\frac{d^2\Psi_1}{dx^2} - \Psi_1\frac{d^2\Psi_2}{dx^2} = 0. \tag{13.7}$$

It is easy to be convinced that equation (13.7) can be rewritten as

$$\frac{d}{dx}\left(\Psi_2\frac{d\Psi_1}{dx} - \Psi_1\frac{d\Psi_2}{dx}\right) = 0, \tag{13.8}$$

i.e.

$$\Psi_2\frac{d\Psi_1}{dx} - \Psi_1\frac{d\Psi_2}{dx} = \text{const}. \tag{13.9}$$

Since $\Psi(\infty) = 0$, it follows that const $= 0$. Therefore,

$$\Psi_2\frac{d\Psi_1}{dx} = \Psi_1\frac{d\Psi_2}{dx} \tag{13.10}$$

or

$$\frac{d\Psi_1}{\Psi_1} = \frac{d\Psi_2}{\Psi_2}. \tag{13.11}$$

Integrating equation (2.11), we obtain

$$\ln\Psi_1 = \ln\Psi_2 + \ln C \text{ or } \Psi_1 = C\Psi_2. \tag{13.12}$$

The multiplication of the wave function by a constant does not change the system state (see (12.14)). Therefore, the functions Ψ_1 and Ψ_2 actually coincide, i.e. the energy level E is nondegenerate.

Now, let us consider the case, when the potential energy of the particle, $U(x)$, changes at the point $x = x_0$ in a jump-like manner from the value $U_1 = U(x < x_0) \neq \infty$ (region 1) to the value $U_2 = U(x > x_0) \neq \infty$ (region 2), being constant to the left and to the right from the barrier (see Fig. 13.1). The probability density for the particle must be continuous across the boundary between the regions. Hence, the

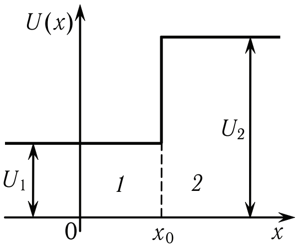

Fig. 13.1.

wave function $\Psi(x)$ must also be continuous across this boundary, i.e. the condition

$$\Psi_1(x_0) = \Psi_2(x_0) \tag{13.13}$$

has to be satisfied, where $\Psi_1(x_0)$ and $\Psi_2(x_0)$ are the wave functions of the particle in region 1 and 2, respectively, to the right and to the left, respectively, from the point x_0. In our case, the derivative $d\Psi/dx$ must also be continuous across the boundary, i.e.

$$\Psi_1'(x_0) = \Psi_2'(x_0)\,. \tag{13.14}$$

13.2. Particle in a potential well

Let us consider a particle with the potential energy

$$U(x) = \begin{cases} U_0 > 0 & (x < 0) \\ 0 & (0 < x < a) \\ U_0 > 0 & (x > a) \end{cases}\,, \tag{13.15}$$

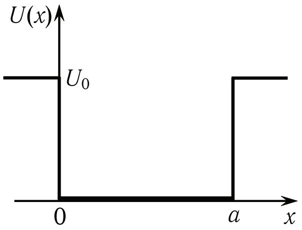

Fig. 13.2.

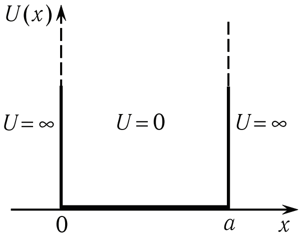

Fig. 13.3.

which is plotted in Fig. 13.2 and is called the **rectangular potential well** of the depth U_0. The problem is to find the spectra of wave functions and energies of the particle in the potential well, i.e. in the interval $0 < x < a$, if the particle energy $E < U_0$.

Let us first confine the consideration to the case $E \ll U_0$. The, we may put $U_0 = \infty$ and consider a potential well with infinitely high walls (Fig. 13.3). Of course, this is an idealized system, but its analysis brings about illustrative results.

Since the particle energy is not infinitely high, it cannot penetrate into the regions $x \leq 0$ and $x \geq a$, being restricted to the interval $0 < x < a$, where $U = 0$. Therefore, such a potential well is called the **potential well with absolutely impenetrable walls**.0.40pt0pt13.213.3

Within the interval $0 < x < a$, $U(x) = 0$ and Hamiltonian (13.2) takes the form

$$\widehat{H} = -\frac{\hbar^2}{2m} = -\frac{d^2}{dx^2} + U\left(x\right) = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2}. \tag{13.16}$$

The corresponding Schrödinger equation (13.3) looks like

$$-\frac{\hbar^2}{2m}\frac{d^2\Psi\left(x\right)}{dx^2} = E\Psi\left(x\right) \tag{13.17}$$

or

$$\frac{d^2\Psi\left(x\right)}{dx^2} = -k^2\Psi, \tag{13.18}$$

where the notation

$$k^2 = \frac{2mE}{\hbar^2} \tag{13.19}$$

was used. Equation (13.18) has to be supplemented with boundary conditions

$$\Psi\left(0\right) = 0, \quad \Psi\left(a\right) = 0, \tag{13.20}$$

which follow from the impossibility for the particle to move outside the interval $0 < x < a$ and the continuity of the wave function Ψ. Equation (13.18) is a linear homogeneous differential equation of the second order. Its two particular solutions look like

$$\Psi_{1,\,2} = \mathrm{e}^{\pm ikx},$$

and the general solution is written as

$$\Psi\left(x\right) = A\mathrm{e}^{ikx} + B\mathrm{e}^{-ikx}, \tag{13.21}$$

where A and B are arbitrary constants. According to the boundary condition $\Psi\left(0\right) = 0$, we obtain the equality $A + B = 0$ or $B = -A$. Then, solution (13.21) takes the form

$$\Psi\left(x\right) = A\left(\mathrm{e}^{ikx} - \mathrm{e}^{-ikx}\right) = C\sin kx,$$

where $C = 2iA$. With the help of boundary condition $\Psi\left(a\right) = 0$, we obtain the equation $C\sin ka = 0$ or

$$\sin ka = 0. \tag{13.22}$$

The corresponding solutions must satisfy the equality $ka = n\pi$, where $n = 0, \pm 1, \pm 2, ...$ Then,

$$k_n = \frac{\pi n}{a}. \tag{13.23}$$

Substituting (13.23) into (13.21), we obtain the wave function in the form

$$\Psi(x) = C \sin\frac{\pi n x}{a}. \tag{13.24}$$

As was already indicated, n is an integer number. However, if $n = 0$, the wave function $\Psi_0(x) \equiv 0$, irrespective of the coordinate x; hence, this state does not exist. Besides, at negative n-values, the wave function simply changes its sign, but the substitution $\Psi \to -\Psi$ does not change the system state. Therefore, we may adopt that $n \in N$, i.e. $n = 1, 2, 3, \ldots$. Hence,

$$\Psi_n(x) = C \sin\frac{\pi n x}{a}, \; n \in N. \tag{13.25}$$

To find the constant C, let us apply the normalization condition for wave function (12.10),

$$\int_0^a |\Psi_n(x)|^2 \, dx = |C|^2 \int_0^a \sin^2\frac{\pi n x}{a} dx = 1.$$

Integrating, we obtain the equation

$$|C|^2 \frac{a}{2} = 1,$$

the solution of which looks like

$$C = \sqrt{\frac{2}{a}} \mathrm{e}^{i\alpha},$$

where a is an arbitrary real value. Hence,

$$\Psi_n(x) = \sqrt{\frac{2}{a}} \mathrm{e}^{i\alpha} \sin\frac{\pi n x}{a}. \tag{13.26}$$

The multiplier $\mathrm{e}^{i\alpha}$ is the phase one (α is the phase), but, when calculating the average values of physical quantities, the probabilities of allowed values, and so forth, the corresponding formulas (see, e.g., (12.24), (12.50), and (12.52)) always include the product $\Psi\Psi^*$, so that $\left|e^{i\alpha}\right|^2 = 1$. Therefore, in what follows, we adopt $\alpha = 0$. Ultimately, the wave function for a particle in the potential well with absolutely impenetrable walls reads

$$\Psi_n(x) = \sqrt{\frac{2}{a}} \sin \frac{\pi n x}{a}, \quad n \in N. \tag{13.27}$$

From (13.19), it follows that

$$E = \frac{\hbar^2 k^2}{2m}. \tag{13.28}$$

Taking into account (2.23), we find the energy spectrum of the particle,

$$E_n = \frac{\pi \hbar^2 n^2}{2ma^2}, \quad n \in N. \tag{13.29}$$

Hence, the energy of a particle in the potential well is quantized, i.e. it can get discrete values. Therefore, the parameter n is called the **quantum number**. The energy difference ΔE_n between the neighbor levels n and $n+1$ grows as n increases,

$$\Delta E_n = E_{n+1} - E_n = \frac{\pi^2 \hbar^2}{2m} (2n+1). \tag{13.30}$$

Notice that, owing to the infinite depth of potential well, the number of energy levels in it is unlimited.

Let us consider now a rectangular potential well with a finite depth depicted in Fig. 13.2. The solution of problem concerning the wave functions and the energies of a particle in it is much more difficult (see, e.g., [20]) than in the case of potential well with infinitely high walls considered above. Therefore, we present only some of the relevant results.

The energy of the particle in this case is also written in the form (see (13.28))

$$E_n = \frac{\hbar^2 k_n^2}{2m}, \tag{13.31}$$

but k_n is now determined from the equation (cf. equation (13.23))

$$ka = \pi n - 2\arcsin\frac{\hbar k}{\sqrt{2mU_0}}, \quad (n = 1, 2, 3, \ldots). \tag{13.32}$$

Besides, owing to the finite depth of potential well, the number n of energy levels in it is confined,

$$n < 1 + \frac{a}{\pi\hbar}\sqrt{2mU_0}. \tag{13.33}$$

The value $n = 1$ always satisfies this inequality, which means that the potential well with a finite depth contains at least one energy level.

If $U_0 \to \infty$, equation (13.32) transforms into $ka = \pi n$, i.e.

$$k_n = \frac{\pi n}{a}.$$

Therefore, from (13.31), we obtain the formula

$$E_n = \frac{\pi^2\hbar^2 n^2}{2ma^2}, \quad (n \in N),$$

which coincides with formula (13.29).

13.3. Tunnel effect

Consider the case, when the potential energy of the particle is high in a confined space region, but rapidly decreases beyond it (Fig. 13.4). Such a potential energy profile is called the **potential barrier**.

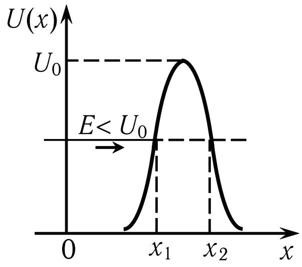

Fig. 13.4.

Let a particle with the energy E lower than the barrier height U_0 move along the axis x towards the barrier from its left hand side. From the viewpoint of classical mechanics, the particle definitely cannot pass through the barrier, because its kinetic energy would have been negative ($T = E - U_0 < 0$) and the momentum $p = \sqrt{2mT}$ imaginary under the barrier. At the same time, provided that $E > U_0$, the classical particle can overcome the barrier for certain.

In quantum mechanics, the situation is different. Even if the microparticle energy is lower than the barrier height ($E < U_0$), there is some probability for the particle to pass beneath the barrier. This phenomenon is called the **tunnel effect**. Moreover, if the particle energy is higher than the barrier height ($E > U_0$), there exists a certain probability that the particle will be reflected from the barrier (the overbarrier reflection).

The tunnel effect is a consequence of the wave properties of microparticles. It can be explained with the help of the uncertainty relation $\Delta x \Delta p_x = \hbar$. If the particle is detected inside the barrier, its coordinate x has the uncertainty Δx equal to the barrier width. However, its momentum p_x is also characterized by the uncertainty $\Delta p_x = \hbar/\Delta x$, so that the kinetic energy $T = p_x^2/(2m)$ becomes indefinite and the relation $T = E - U_0 < 0$ has no meaning. Therefore, the particle can overcome the barrier without violating the energy conservation law.

The main parameter of tunnel effect is the probability for a particle to pass through the barrier, which is called the **transmission coefficient** (the transparency). Numerically, this is the ratio between the flux of particles that passed through the barrier and the flux of particles that fall on it.

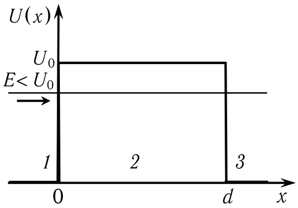

Fig. 13.5.

Let us determine first the transmission coefficient for a rectangular barrier of the height U_0 and the width d (see Fig. 13.5). In region $x < 0$ (region 1), $U(x) = 0$. Therefore, the Schrödinger equation

$$-\frac{\hbar^2}{2m}\frac{d^2\Psi_1}{dx^2} = E\Psi_1$$

can be written as

$$\frac{d^2\Psi_1}{dx^2} = -k^2\Psi_1, \quad k = \frac{\sqrt{2mE}}{\hbar^2}. \tag{13.34}$$

The general solution of this equation looks like

$$\Psi_1(x) = A_1 e^{ikx} + A e^{-ikx}. \tag{13.35}$$

To elucidate the meaning of summands in (13.35), let us write down the total wave function $\Psi_1(x,t)$ with the time-dependent multiplier $e^{-i\omega t}$ (see (12.74)),

$$\Psi_1(x,t) = A_1 e^{i(kx-\omega t)} + A_2 e^{-i(kx+\omega t)}.$$

The first summand in the left hand side is a plane de Broglie wave propagating along the axis x and corresponding to the particle that approaches the barrier. The second summand is a de Broglie wave propagating in the opposite direction. It describes a particle reflected from a potential barrier. Therefore, wave function (13.35) can be written as

$$\Psi_1(x) = \Psi_{1\mathrm{inc}}(x) + \Psi_{1\mathrm{refl}}(x),$$

where

$$\Psi_{1\mathrm{inc}}(x) = A_1 \mathrm{e}^{ikx}, \quad \Psi_{1\mathrm{refl}}(x) = A_2 \mathrm{e}^{-ikx}. \tag{13.36}$$

The Schrödinger equation in region 2 ($U(x) = U_0$) looks like

$$-\frac{\hbar^2}{2m}\frac{d^2\Psi_2}{dx^2} + U_0\Psi_2 = E\Psi_2,$$

or

$$\frac{d^2\Psi_2}{dx^2} = -k_2^2\Psi_2, \quad k_2 = \frac{\sqrt{2m(E-U_0)}}{\hbar}. \tag{13.37}$$

The general solution of this equation is

$$\Psi_2(x) = A_3\mathrm{e}^{ik_2x} + A_4\mathrm{e}^{-ik_2x}. \tag{13.38}$$

The first summand corresponds to a particle that moves along the axis x, i.e. penetrated through the left wall of potential barrier, and the second one to a particle reflected from the right wall of the barrier and moving in the opposite direction to the axis x. However, within region 2, $E < U_0$; therefore,

$$k_2 = \frac{1}{\hbar}\sqrt{2m(E-U_0)} = i\frac{1}{\hbar}\sqrt{2m(U_0-E)} = i\gamma, \quad \gamma = \frac{1}{\hbar}\sqrt{2m(U_0-E)}, \tag{13.39}$$

and, as a result,

$$\Psi_2(x) = A_3\mathrm{e}^{-\gamma x} + A_4\mathrm{e}^{\gamma x}. \tag{13.40}$$

In the region $x > d$ (region 3), we have $U(x) = 0$, and the Schrödinger equation has the same form as (13.34),

$$\frac{d^2\Psi_3}{dx^2} = -k^2\Psi_3.$$

Therefore,

$$\Psi_3(x) = A_5\mathrm{e}^{ikx} + A_6\mathrm{e}^{-ikx}.$$

However, there is no reflected wave in region 3. Hence, $A_6 = 0$ and

$$\Psi_3(x) = A_5 \mathrm{e}^{ikx}. \tag{13.41}$$

According to the definition given above to the transmission coefficient D can be calculated as the ratio between the probability to find the particle behind the barrier (region 3) to the probability to find it in the incident beam (region 1), i.e.

$$D = \frac{|\Psi_3|^2}{|\Psi_{1\mathrm{inc}}|^2}, \tag{13.42}$$

or, taking into account (13.36) and (13.41),

$$D = \frac{|A_5|^2}{|A_1|^2}. \tag{13.43}$$

To simplify the calculations, let us neglect the particle reflection from the right wall of the barrier. Then, $A_4 = 0$ and

$$\Psi_2(x) = A_3 \mathrm{e}^{-\gamma x}. \tag{13.44}$$

In this case, the amplitude A_5 of the wave function Ψ_3 has to be substituted by the value of wave function Ψ_2 at $x = d$,

$$A_5 = A_3 \mathrm{e}^{-\gamma d}. \tag{13.45}$$

Then, the transmission coefficient looks like

$$D = \frac{\left|A_3 \mathrm{e}^{-\gamma d}\right|^2}{|A_1|^2} = \left|\frac{A_3}{A_1}\right|^2 \mathrm{e}^{-2\gamma d}. \tag{13.46}$$

Now, let us determine the ratio A_3/A_1. The wave function and its first derivative must be continuous across the left wall of the barrier (see section 13.1). Therefore, in accordance with (13.13) and (13.14), we may write

$$\Psi_1(0) = \Psi_2(0), \quad \Psi_1'(0) = \Psi_2'(0). \tag{13.47}$$

Substituting wave functions (13.35) and (13.44) into (13.47), we obtain a system of two linear algebraic equations,

$$\begin{cases} A_1 + A_2 = A_3, \\ ikA_1 - ikA_2 = -\gamma A_3, \end{cases} \tag{13.48}$$

whence we can calculate the relationship between A_3 and A_1,

$$A_3 = \frac{2}{1 + i\dfrac{\gamma}{k}} A_1. \tag{13.49}$$

Taking into account (13.49), transmission coefficient (13.46) takes the form

$$D = \frac{4}{1 + \frac{\gamma^2}{k^2}} \mathrm{e}^{-2\gamma d}, \tag{13.50}$$

and, making use of (13.34) and (13.39), we obtain

$$D = D_0 \mathrm{e}^{-\dfrac{2}{\hbar}\sqrt{2m(U_0 - E)}d}, \quad D_0 = 4\frac{E}{U_0}. \tag{13.51}$$

Since $E < U_0$, it is easy to be convinced that $0 < D < 4$. Taking into account that the exponential multiplier in (13.51) is, as a rule, much lesser than unity, let us put $D \approx 1$. Then,

$$D = \mathrm{e}^{-\dfrac{2}{\hbar}\sqrt{2m(U_0 - E)}d}. \tag{13.52}$$

It is easy to prove (see [20]) that, if the profile of potential barrier has an arbitrary form (see Fig. 13.4), the transmission coefficient is calculated by the formula

$$D = \exp\left\{-\frac{2}{\hbar}\int_{x_1}^{x_2} \sqrt{2m\left[U\left(x\right) - E\right]}dx\right\}. \tag{13.53}$$

From (13.52), it follows (comes up) that the transmission coefficient differs from zero only for particles with a very small mass and narrow potential barriers. Therefore, the tunnel effect can be observed only in the microworld. However, since the number of microparticles that have passed through the potential barrier can be considerable, there are a number of physical phenomena, where the tunnel effect reveals itself. In particular, these are the radioactive alpha decay of atomic nuclei, cold emission of electrons from metal (field emission), spontaneous disintegration of atomic nuclei, Josephson effect, nuclear (thermonuclear) fusion, and so forth.

13.4. Linear harmonic oscillator

In quantum mechanics, the oscillator problem is among the most important ones. This is so, because it forms the basis for the analysis of oscillations in various systems: atomic vibrations in molecules, atomics and molecular vibrations in crystals, plasma oscillations, and so on.

In classical mechanics, the **linear harmonic oscillator** is a particle that executes one-dimensional harmonic oscillations around an equilibrium position under the action of elastic force $F = -kx$, where k is the elastic constant. The corresponding law of motion looks like

$$x = a\cos\left(\omega t + \alpha\right), \tag{13.54}$$

where a is the amplitude ($-a \leq x \leq a$), $\omega = k/m$ the frequency, and α the initial phase of oscillations. The kinetic, potential, and total energies of classical oscillator are determined by the formulas

$$T = \frac{p^2}{2m}, \quad U = \frac{kx^2}{2} = \frac{m\omega^2 x^2}{2}, \quad E = \frac{p^2}{2m} + \frac{m\omega^2 x^2}{2} = \frac{m\omega^2 a^2}{2}. \tag{13.55}$$

If the oscillator ceases to oscillate, the particle is located at its equilibrium position and has exact values for the coordinate and the momentum, $x = 0$ and $p_x = 0$.

The properties of quantum oscillator are different. Its oscillations cannot stop completely, because simultaneous exact values of coordinate and momentum are impossible. Moreover, in quantum mechanics, the coordinate cannot depend explicitly on the time, and the motion law (13.54) is inapplicable. For these and other properties of quantum oscillator to be examined in detail, we should determine its wave functions and energy spectrum.

On the basis of (13.55), the stationary Schrödinger equation for a linear harmonic oscillator can be written as

$$-\frac{\hbar^2}{2m}\frac{d^2\Psi}{dx^2} + \frac{1}{2}m\omega^2 x^2\Psi = E\Psi. \tag{13.56}$$

This is a linear nonhomogeneous differential equation of the second order. For convenience, let us introduce the dimensionless variable

$$\xi = \sqrt{\frac{m\omega}{\hbar}}x. \tag{13.57}$$

Then,

$$\frac{d\Psi}{dx} = \frac{d\Psi}{d\xi}\frac{d\xi}{dx} = \frac{d\Psi}{d\xi}\sqrt{\frac{m\omega}{\hbar}}, \quad \frac{d^2\Psi}{dx^2} = \frac{m\omega}{\hbar}\frac{d^2\Psi}{d\xi^2},$$

and equation (2.57) takes the form

$$\frac{d^2\Psi}{d\xi^2} - \left(\xi^2 - \lambda\right)\Psi = 0, \tag{13.58}$$

where

$$\lambda = \frac{2E}{\hbar\omega}. \tag{13.59}$$

First, let us find the asymptotic solution of equation (13.58) for $\xi^2 \gg \lambda$. Then, the parameter λ can be neglected, and we obtain

$$\frac{d^2\Psi_a}{d\xi^2} - \xi^2\Psi_a = 0.$$

Taking into account the boundary conditions $\Psi_a(\pm\infty) = 0$, the solution of this equation is

$$\Psi_a(\xi) = \mathrm{e}^{-\frac{1}{2}\xi^2}. \tag{13.60}$$

The asymptotic solution (2.61) gives ground to try the exact solution of equation (13.58) in the form

$$\Psi(\xi) = f(\xi)\,\mathrm{e}^{-\frac{1}{2}\xi^2}, \tag{13.61}$$

where $f(\xi)$ is an unknown function. Substituting (13.61) into equation (13.58), we obtain an equation for the function $f(\xi)$,

$$\frac{d^2 f}{d\xi^2} - 2\xi\frac{df}{d\xi} + (\lambda - 1) f = 0. \tag{13.62}$$

Its solution is sought as a series

$$f(\xi) = \sum_{i=0}^{\infty} a_i\xi^i = a_0 + a_1\xi + a_2\xi^2 + \ldots. \tag{13.63}$$

To find the coefficients a_i, we substitute (13.63) into equation (13.62) to obtain

$$\sum_i i\,(i-1)\,a_i\xi^{i-2} - 2\sum_i i a_i\xi^i + (\lambda - 1)\sum_i a_i\xi^i = 0.$$

If a power series equals zero at every value of its variable, the coefficients in all members must also equal zero. On this basis, we collect the coefficients of all ξ^k-members and equate the result to zero,

$$(k+2)\,(k+1)\,a_{k+2} - 2ka_k + (\lambda - 1)\,a_k = 0. \tag{13.64}$$

From (13.64), a recurrent formula for calculating the coefficients a_i follows,

$$a_{k+2} = \frac{2k - (\lambda - 1)}{(k+2)\,(k+1)}a_k, \tag{13.65}$$

whence one can easily see that series (13.63) contains either only even or only odd terms, i.e.

$$f\,(\xi) = \begin{cases} a_0 + a_2\xi^2 + a_4\xi^4 + \ldots, \\ a_1\xi + a_3\xi^3 + a_5\xi^5 + \ldots. \end{cases}$$

Wave function (13.61) must satisfy the boundary conditions $\Psi\,(\pm\infty) = 0$, and, hence, the function $f\,(\xi)$ must be polynomial; otherwise, the infinite series is divergent at $\xi \to \pm\infty$. Let the last nonzero summand be $a_n\xi^n$, i.e.

$$f\,(\xi) = f_n\,(\xi) = ... + a_n\xi^n, \quad n \geq 0. \tag{13.66}$$

The polynomials $f_n\,(\xi)$ are designated as $H_n\,(\xi)$ and are called the **Hermite polynomials**. Taking into account that $a_n \neq 0$, but $a_{n+2} = 0$, with the help of recurrent formula (13.65), we obtain the relationship

$$a_{n+2} = \frac{2n - (\lambda - 1)}{(n+2)\,(n+1)}a_n = 0,$$

whence we determine the parameter λ,

$$\lambda = \frac{2E}{\hbar\omega} = 2n + 1, \quad n = 0, 1, 2, \ldots.$$

Hence, the energy spectrum of quantum oscillator looks like

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega, \quad n = 0, 1, 2, \ldots. \tag{13.67}$$

The integer parameter n is the number of the last summand in the Hermite polynomial $H_n(\xi) = -f_n(\xi)$. It can be equal to 0, 1, 2,.... In formula (13.67), it is called the **quantum number**.

The coefficient a_n is equated to 2^n. Then, with the help of formula (13.65), it is possible to calculate previous coefficients and find both the Hermite polynomial $H_n(\xi)$ and wave function (13.61). However, there is a formula, which enables the Hermite polynomials to be determined directly; namely,

$$H_n(\xi) = (-1)^n \mathrm{e}^{-\xi^2} \frac{d^n}{d\xi^n} \mathrm{e}^{-\xi^2}. \tag{13.68}$$

It is easy to be convinced that $H_0 = 1$, $H_1 = 2\xi$, $H_2 = -2 + 4\xi^2$, $H_3 = -12\xi + 8\xi^3$, and so forth.

Now, wave function (13.61) can be written in the form

$$\Psi_n(\xi) = A_n H_n(\xi)\, \mathrm{e}^{-\xi^2/2}, \tag{13.69}$$

where

$$A_n = \frac{1}{\sqrt{2^n n! \sqrt{\pi}}} \tag{13.70}$$

is the normalizing coefficient. The transformation from the wave function representation $\Psi_n(\xi)$ to the $\Psi_n(x)$ one is done as follows:

$$\Psi_n(x) = \sqrt[4]{\frac{m\omega}{\hbar}}\Psi_n(\xi), \quad \xi = \sqrt{\frac{m\omega}{\hbar}}x. \tag{13.71}$$

The following formulas are given without proof:

$$\xi\Psi_n(\xi) = \sqrt{\frac{n}{2}}\Psi_{n-1}(\xi) + \sqrt{\frac{n+1}{2}}\Psi_{n+1}(\xi), \tag{13.72}$$

$$\frac{d\Psi_n(\xi)}{d\xi} = \sqrt{\frac{n}{2}}\Psi_{n-1}(\xi) - \sqrt{\frac{n+1}{2}}\Psi_{n+1}(\xi). \tag{13.73}$$

Important conclusions follow from the analysis of oscillator energy spectrum expressed by formula (13.67),

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega, \quad n = 0, 1, 2, \ldots.$$

This is a system of equidistant levels, with the energy difference between neighbor levels equal to $\hbar\omega$. An important feature of the spectrum is the existence of the **zero-point energy**, E_0, or the **energy of zero-point oscillations**. If the oscillator is in the minimum-energy state ($n = 0$), i.e. in the ground state, its energy differs from zero and equals

$$E_0 = \frac{1}{2}\hbar\omega.$$

From the viewpoint of classical mechanics, this means that the oscillator does not stop to oscillate in the ground state. Therefore, such oscillations are called **zero-point**. Their existence follows from the Heisenberg uncertainty relations, i.e. from the wave properties of microparticles.

The reality of zero-point oscillations is confirmed by experiments on X-ray scattering in crystals, which occurs at non-homogeneities in crystal structure associated with thermally induced vibrations of atoms, i.e. at density fluctuations. When the temperature of the crystal T decreases the intensity of atomic vibrations diminishes. As a result, the degree of crystal non-homogeneity and the intensity of scattered X-rays rays decrease. In the absence of zero-point vibrations, the intensity of scattered rays should have tended to zero at $T \to 0$, but it tends to a value different from zero, i.e. scattering occurs at zero-point vibrations of atoms.

Zero-point oscillations are closely associated with the properties of physical vacuum. In the quantum field theory, all elementary particles are regarded as quanta of corresponding fields: photons are quanta of electromagnetic field, electrons and positrons are quanta of electron-positron field, and so on. Vacuum is considered as the lowest (ground) state for all fields.

For example, let us consider an electromagnetic field. It can be regarded as a superposition (set) of plane monochromatic electromagnetic waves like $\vec{E} = \vec{E}_0 \cos\left(\omega t - \vec{k}\vec{r}\right)$. Formally, such a wave is a linear harmonic oscillator, so that the electromagnetic field is equivalent to a set of oscillators (from the viewpoint of quantum mechanics, these are photons). The lowest state of electromagnetic field is characterized by the absence of observable electromagnetic waves (photons), but there remain non-observable zero-point oscillations of the field. From the viewpoint of quantum mechanics, zero-point oscillations of electromagnetic field are associated with virtual (probable) photons, which permanently emerge and disappear in accordance with the uncertainty relation for energy and time $\Delta E \Delta t = \hbar$ (see section 6.3).

This conclusion is valid for all fields. **Physical vacuum** is a material medium, where virtual pairs of all particles and antiparticles emerge and disappear all the time (in this connection, it is called the **“boiling” medium**). If the vacuum is given the energy $E \geq 2\varepsilon_0$, where ε_0 is the characteristic energy of some elementary particle, there emerges a real pair particle– antiparticle, i.e. the virtual pair becomes real. Virtual particles cannot be registered directly. Therefore, physical vacuum is “empty”, but those particles are responsible for some effects, in particular, the shift of electron energy levels in the hydrogen atom, which was discovered experimentally.

14

PARTICLE IN A CENTRALLY SYMMETRIC FIELD

14.1. General properties of particle motion in a centrally symmetric field

The problem dealing with the atom is one of the most important in quantum mechanics. It will be recalled that the problem of atomic mechanics was one of those which solution stimulated the creation of quantum mechanics.

The simplest atom—the hydrogen one—consists of a nucleus and an electron, and potential energy of their electrostatic interaction depends only on the distance between them. Such an interaction is called **centrally symmetric** or **central**. Notice what not any interaction is central. For instance, the gravitational interaction is central, but the nuclear one is noncentral, because, besides the distance, it also depends on other factors.

First, let us consider general properties of the system of two particles with the central interaction between them. This system is shown in Fig. 14.1, where m_1 and m_2 are the masses of particles, $\vec{r}_1$ and $\vec{r}_2$ their radius vectors, $|\vec{r}_1 - \vec{r}_2|$ is the distance between the particles, O the coordinate origin, and C the center of mass of particles. The Hamiltonian operator of the system looks like

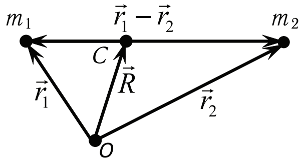

Fig. 14.1.

$$\widehat{H} = -\frac{\hbar^2}{2m_1}\Delta_1 - \frac{\hbar^2}{2m_2}\Delta_2 + U\left(|\vec{r}_1 - \vec{r}_2|\right). \tag{14.1}$$

The Laplace operators Δ_1 and Δ_2 act only on the coordinates of particle 1 and 2, respectively. The corresponding stationary Schrödinger equation reads

$$-\frac{\hbar^2}{2m_1}\Delta_1\Psi\left(\vec{r}_1, \vec{r}_2\right) - \frac{\hbar^2}{2m_2}\Delta_2\Psi\left(\vec{r}_1, \vec{r}_2\right) + U\left(|\vec{r}_1 - \vec{r}_2|\right)\Psi\left(\vec{r}_1, \vec{r}_2\right) =$$

$$= E\Psi\left(\vec{r}_1, \vec{r}_2\right), \tag{14.2}$$

in which the wave function depends on the coordinates of both particles.

Equation (14.2) is rather complicated, but it can be simplified. We are not interested in the motion of the system as a whole; therefore, let us change to the center of mass reference frame. For this purpose, we translate the coordinate origin O to the center of mass C (Fig. 14.2). Then, $\vec{R} = 0$, i.e. we suppose that the center of mass is fixed. Besides, instead of coordinates $\vec{r}_1$ and $\vec{r}_2$, we introduce the relative coordinate $\vec{r} = \vec{r}_1 - \vec{r}_2$. The corresponding transformations in (14.2) look like

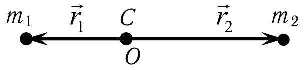

Fig. 14.2.

$$\Delta_1\Psi = \nabla_1\nabla_1\Psi = \nabla_1\nabla\Psi\frac{d\vec{r}}{dr_1} = \nabla_1\nabla\Psi = \nabla^2\Psi\frac{d\vec{r}}{d\vec{r}_1} = \Delta\Psi, \tag{14.3}$$

$$\Delta_2\Psi = \nabla_2\nabla_2\Psi = \nabla_2\nabla\Psi\frac{d\vec{r}}{dr_2} = -\nabla_2\nabla\Psi = -\nabla^2\Psi\frac{d\vec{r}}{d\vec{r}_2} = \Delta\Psi. \tag{14.4}$$

Substituting (14.3) and (14.4) into (14.2), we obtain the equation

$$-\frac{\hbar^2}{2\mu}\Delta\Psi\left(\vec{r}\right)+U\left(r\right)\Psi\left(\vec{r}\right)=E\Psi\left(\vec{r}\right), \tag{14.5}$$

where

$$\mu=\frac{m_1 m_2}{m_1+m_2}. \tag{14.6}$$

The quantity μ has the dimensionality of mass and is called the **reduced mass**. In some cases, the mass of one particle considerably exceeds the mass the other particle, i.e. $m_1 \gg m_2 = m$ (this is true, in particular, for the nucleus and electron in the atom). Then,

$$\mu=\frac{m_1 m_2}{m_1+m_2}=\frac{m_2}{1+\frac{m_2}{m_1}}\approx m,$$

and the reduced mass μ in Schrödinger equation (14.5) can be substituted by the mass of light particle. In this approximation, the heavy particle is considered at rest, the center of mass coincides with it, and the light particle moves in its field.

The Hamiltonian of Schrödinger equation (14.5) is

$$\widehat{H}=-\frac{\hbar^2}{2\mu}\Delta+U\left(r\right). \tag{14.7}$$

Owing to the central symmetry of the field $U(r)$, it can be written in the spherical coordinate system. Then, according to (12.39) and (12.40), we obtain

$$\widehat{H}=\widehat{T}_r+\frac{1}{2\mu r^2}\widehat{L}^2+U\left(r\right), \quad \widehat{T}_r=-\frac{\hbar^2}{2\mu r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right), \tag{14.8}$$

where, in accordance with (12.37) and (12.38),

$$\widehat{L}^2=-\hbar^2\left\{\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right\}, \tag{14.9}$$

$$\widehat{L}_z=-i\hbar\frac{\partial}{\partial\varphi}. \tag{14.10}$$

It is easy to be convinced that the operators $\widehat{H}$, $\widehat{L}^2$, and $\widehat{L}_z$ commute, i.e.

$$\widehat{H}\widehat{L}^2=\widehat{L}^2\widehat{H}, \quad \widehat{H}\widehat{L}_z=\widehat{L}_z\widehat{H}, \quad \widehat{L}_z\widehat{L}^2=\widehat{L}^2\widehat{L}_z.$$

Hence, those three operators have a common system of eigenfunctions (see section 12.3). Besides, the eigenvalues of those operators (E, L^2, and L_z) can have exact values simultaneously (they are simultaneously measurable). Moreover, they are integrals of motion, i.e. the energy E, the square of angular momentum L^2, and the projection of angular momentum L_z remain constant in the system of two particles with the central interaction. Therefore, let us first determine the eigenvalues and eigenfunctions of operators $\widehat{L}_z$ and $\widehat{L}^2$.

The eigenvalue equation for the operator $\widehat{L}_z$ looks like (see section 12.3)

$$-i\hbar\frac{\partial U(\varphi)}{\partial\varphi} = L_z U(\varphi). \tag{14.11}$$

Its solution is

$$U(\varphi) = C\exp\left(\frac{i}{\hbar}L_z\varphi\right). \tag{14.12}$$

Since the azimuthal angle φ is the cyclic coordinate ($0 \le \varphi \le 2\pi$), we obtain the condition

$$U(\varphi) = U(\varphi + 2\pi). \tag{14.13}$$

Substituting (14.12) into (14.13), we obtain the equation

$$\exp\left(\frac{i}{\hbar}L_z 2\pi\right) = 1, \tag{14.14}$$

whence it follows that

$$\frac{i}{\hbar}L_z 2\pi = m2\pi, \quad m = 0, \pm1, \pm2, \ldots. \tag{14.15}$$

From (14.15), we obtain the eigenvalues of operator $\widehat{L}_z$ in the form

$$L_z = m\hbar, \quad m \in Z. \tag{14.16}$$

Then, eigenfunction (14.12) is written as

$$U_m(\varphi) = C\mathrm{e}^{im\varphi}. \tag{14.17}$$

The integration constant C can be determined from the normalization condition for the function $u(\varphi)$,

$$\int_0^{2\pi} |U_m(\varphi)|^2 \, d\varphi = 1. \tag{14.18}$$

Substituting (14.17) into (14.18), we find the equation

$$2\pi \left|C\right|^2 = 1,$$

the solution of which is

$$C = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{i\alpha},$$

where α is an arbitrary real number. Choosing $\alpha = 0$, we obtain $C = 1/\sqrt{2\pi}$, so that

$$U_m\left(\varphi\right) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{im\varphi}. \tag{14.19}$$

Here, the quantum number m is called the **magnetic quantum number** (see section 3.3).

The eigenvalue equation for the operator $\widehat{L}^2$ looks like

$$\widehat{L}^2 Y\left(\theta, \varphi\right) = L^2 Y\left(\theta, \varphi\right). \tag{14.20}$$

The operator $\widehat{L}^2$ in (3.21) is expressed by formula (14.9). The solution of equation (14.20) is rather cumbersome (see Appendix F). It demonstrates that

$$L^2 = l\left(l+1\right)\hbar^2. \tag{14.21}$$

The number l is called the **orbital quantum number**. At a given l, the magnetic quantum number m can acquire the values $m = 0, \pm 1, \pm 2, \ldots, \pm l$; i.e. $2l+1$ values in total.

The eigenfunctions $Y\left(\theta, \varphi\right)$ of equation (14.20) are called the **spherical functions.** They are written as

$$Y_{lm}\left(\theta, \varphi\right) = A_l^{|m|} P_l^{|m|}\left(\cos\theta\right) \mathrm{e}^{im\varphi}; \tag{14.22}$$

where

$$A_l^{|m|} = \sqrt{\frac{\left(l - |m|\right)!\left(2l+1\right)}{\left(l+|m|\right)! 4\pi}}, \tag{14.23}$$

$$P_l^{|m|}\left(\cos\theta\right) = \left(1 - \cos^2\theta\right)^{\frac{1}{2}|m|} \frac{d\,|m|}{d\left(\cos\theta\right)^{|m|}} P_l\left(\cos\theta\right), \tag{14.24}$$

$$P_l\left(\cos\theta\right) = \frac{1}{2^l l!} \frac{d^l}{d\left(\cos\theta\right)^l} \left(\cos^2\theta - 1\right)^l. \tag{14.25}$$

The functions $P_l\left(\cos\theta\right)$ are polynomials, they are called the **Legendre polynomials**; and the polynomials $P_l^{|m|}\left(\cos\theta\right)$ are called the **associated Legendre polynomials**. As an example, the simplest spherical functions look like

$$Y_{0,0}\left(\theta,\varphi\right)=\frac{1}{\sqrt{4\pi}},\tag{14.26}$$

$$Y_{1,0}\left(\theta,\varphi\right)=\sqrt{\frac{3}{4\pi}}\cos\theta,\tag{14.27}$$

$$Y_{1,\pm1}\left(\theta,\varphi\right)=\sqrt{\frac{3}{8\pi}}\sin\theta\mathrm{e}^{\pm i\varphi}.\tag{14.28}$$

According to (14.16) and (14.21), the absolute value of vector $\overrightarrow{L}$ and its projection L_z have exact values and are preserved. However, the vector $\overrightarrow{L}$ itself is not preserved, i.e. its orientation in the space changes (only the angle between the vector $\overrightarrow{L}$ and the axis z remains constant). From the viewpoint of classical mechanics, this circumstance can be explained as a precession of vector $\overrightarrow{L}$ around the axis z.

Now let us return to Schrödinger equation (14.7) and, taking into account (14.8), rewrite it in the form

$$\widehat{T}_z\Psi\left(\vec{r}\right)+\frac{1}{2\mu r^2}\widehat{L}^2\Psi\left(\vec{r}\right)+U\left(r\right)\Psi\left(\vec{r}\right)=E\Psi\left(\vec{r}\right).\tag{14.29}$$

The wave function $\Psi\left(\vec{r}\right)$ is tried as a product of radial, $R(r)$, and spherical, $Y\left(\theta,\varphi\right)$, functions,

$$\Psi\left(\vec{r}\right)=R\left(r\right)Y\left(\theta,\varphi\right).\tag{14.30}$$

While substituting (14.30) into equation (14.29), it is necessary to take into account that the operator $\widehat{T}_z$ acts only on the function $R(r)$, and the operator $\widehat{L}^2$ only on the function $Y\left(\theta,\varphi\right)$. According to (14.20) and (14.21), $\widehat{L}^2Y=l\left(l+1\right)\hbar^2Y$. Then, dividing the resulting equation by RY, we obtain

$$\widehat{T}_rR\left(r\right)+\frac{l\left(l+1\right)\hbar^2}{2\mu r^2}R\left(r\right)+U\left(r\right)R\left(r\right)=ER\left(r\right).\tag{14.31}$$

Equations (14.31) is called the **radial Schrödinger equation**

14.2. Electron in the field of atomic nucleus

Let us consider a system of two particles: an atomic nucleus with the electric charge Ze and an electron with the charge $-e$. This system corresponds to the hydrogen atom ($Z = 1$), as well as to He^{+} ($Z = 2$), Li^{2+} ($Z = 3$), Be^{3+} ($Z = 4$), and so on, ions. The potential energy of interaction between the electron and the nucleus, in this case, looks like

$$U(r) = -\frac{Ze^2}{4\pi\varepsilon_0 r}. \tag{14.32}$$

It is centrally symmetric. Therefore, we may use the methods used in the previous section.

The wave function $\Psi(\vec{r})$ is sought in the form

$$\Psi(\vec{r}) = R(r)\, Y(\theta, \varphi) \tag{14.33}$$

and the radial function $R(r)$ in the form

$$R(r) = -\frac{v(r)}{r}. \tag{14.34}$$

Substituting (14.34) and (14.32) into radial equation (14.31) and changing the reduced mass μ to the electron mass m_0, we obtain the equation for the function $v(r)$,

$$-\frac{\hbar^2}{2m_0}\frac{d^2 v}{dr^2} + \frac{l(l+1)\hbar^2}{2m_0 r^2}v - \frac{Ze^2}{4\pi\varepsilon_0 r}v = Ev. \tag{14.35}$$

Let us introduce the dimensionless variable

$$\rho = \frac{r}{r_1}, \quad r_1 = \frac{4\pi\varepsilon_0\hbar^2}{m_0 e^2}, \tag{14.36}$$

where r_1 is the radius of the first Bohr orbit. The corresponding transformations in (14.35) look like

$$\frac{dv}{dr} = \frac{dv}{d\rho}\frac{d\rho}{dr} = \frac{1}{r_1}\frac{dv}{d\rho}, \quad \frac{d^2 v}{dr^2} = \frac{1}{r_1^2}\frac{d^2 v}{d\rho^2},$$

and we obtain the equation

$$\frac{d^2 v}{d\rho^2} - \frac{l(l+1)}{\rho^2}v + \frac{2Z}{\rho}v = \alpha^2 v, \tag{14.37}$$

where

$$\alpha^2 = -\frac{32\pi^2\varepsilon_0^2\hbar^2}{m_0 Z^2 e^4} E. \tag{14.38}$$

At $\rho \to \infty$, equation (14.37) transforms into

$$\frac{d^2 v}{d\rho^2} = \alpha^2 v. \tag{14.39}$$

This equation has two partial solutions, $v_{1,2} = \mathrm{e}^{\pm\alpha\rho}$, so that radial function (14.34) looks like

$$R(\rho) = \frac{A_1}{r}\mathrm{e}^{-\alpha\rho^2} + \frac{A_2\mathrm{e}^{\alpha\rho}}{r}.$$

For this function to be finite at infinity, we must put $A_2 = 0$. Then, the asymptotic ($\rho \to \infty$) solution of equations (14.37) is

$$v_a(\rho) = A_1\mathrm{e}^{\alpha\rho}.$$

Therefore, the exact solution of equation (14.37) should be sought in the form

$$v(\rho) = \mathrm{e}^{-\alpha\rho} f(\rho). \tag{14.40}$$

Substituting (3.41) into (3.38), we find the equation for the function $f(\rho)$,

$$\frac{d^2 f}{d\rho^2} - 2\alpha\frac{df}{d\rho} - \frac{l(l+1)}{\rho^2} f + \frac{2Z}{\rho} f = 0. \tag{14.41}$$

At $\rho \to 0$, the first and third summands dominates, and

$$\frac{d^2 f}{d\rho^2} - \frac{l(l+1)}{\rho^2} f = 0.$$

The solution of this equation, which is finite at $\rho \to 0$, equals

$$f = \rho^{l+1}.$$

In this connection, the exact solution of equation (14.41) is sought in the form

$$f = \rho^{l+1}\sum_{i=0}^{\infty} a_i\rho^i = \sum_{i=0}^{\infty} a_i\rho^{i+l+1}. \tag{14.42}$$

To find the coefficients a_i, we substitute (14.42) into (14.41) to obtain

$$\sum_i (i+l+1)(i+l)\,a_i\rho^{i+l-1} - 2\alpha\sum_i (i+l+1)\,a_i\rho^{i+l} - l(l+1)\sum_i a_i\rho^{i+l-1} + 2Z\sum_i a_i\rho^{i+l} = 0.$$

A similar equation was considered in the problem of linear harmonic oscillator (see section 13.4). Therefore, we use the method applied there: the combined coefficients at all power terms of ρ must be equal to zero. On this basis, the coefficient at ρ^{k+l} equals

$$(k+l+2)(k+l+1)\,a_{k+1} - 2\alpha(k+l+1)\,a_k - l(l+1)\,a_{k+1} + 2Za_k = 0,$$

whence we obtain the following recurrent formula:

$$a_{k+1} = \frac{2\alpha(k+l+1) - 2Z}{(k+l+2)(k+l+1) - l(l+1)}a_k. \tag{14.43}$$

The finiteness of wave function demands that series (14.42) should terminate at a definite summand with the number n_r. Then, the function $f(\rho)$ is a polynomial of the order n_r+l+1 (the number n_r is called the **radial quantum number**),

$$f(\rho) = \sum_{i=0}^{n_r} a_i\rho^{i+l+1} = a_0\rho^{l+1} + a_1\rho^{l+2} + \ldots + a_{n_r}\rho^{n_r+l+1}. \tag{14.44}$$

Since $a_{n_r} \neq 0$, but $a_{n_r+1} = 0$, we obtain

$$a_{n_r+1} = \frac{2\alpha(n_r+l+1) - 2Z}{(n_r+l+2)(n_r+l+1) - l(l+1)}a_{n_r} = 0, \tag{14.45}$$

whence we find

$$n_r + l + 1 = \frac{Z}{\alpha}. \tag{14.46}$$

Taking into account that the numbers n_r and l can be equal to 0, 1, 2, . . . , we make a conclusion that the sum n_r+l+1 can equal 1, 2, 3, ... Let us introduce the notation

$$n = n_r + l + 1. \tag{14.47}$$

Then, $n = Z/\alpha$, so that $\alpha = Z/n$. On this basis of (14.38), we find the electron energy

$$E_n = -\frac{m_0 Z^2 e^4}{32\pi^2 \varepsilon_0 \hbar^2 n^2}, \quad n = 1, 2, 3, \ldots. \tag{14.48}$$

The number n is called the **principal quantum number**. From (14.47), it follows that at a given n, the maximum value $l_{\max} = n - n_{r\min} - 1 = n - 1$, i.e. the orbital quantum number can be equal to

$$l = 0, 1, 2, \ldots, n - 1.$$

The choice of quantum numbers n and l determines the quantum number $n_r = n - l - 1$ and, hence, polynomial (14.44), which we denote as $f_{nl}(\rho)$. With the help of recurrent formula (14.43), all the coefficients in the polynomial $f_{nl}(\rho)$ can be determined in terms of coefficient a_0. Afterwards, we can find function (14.40) $v_{nl}(\rho) = f_{nl}(\rho)\exp(-\alpha\rho)$ and the radial function $R_{nl}(r) = \dfrac{1}{r} v_{nl}(r)$, which should be normalized by unity,

$$\int_0^\infty R_{nl}^2(r)\, r^2 dr = 1. \tag{14.49}$$

As a result, the uncertainty of coefficient a_0 becomes eliminated.

There is a more convenient way to determine the radial functions $R_{nl}(r)$. It is based on the formulas

$$R_{nl}(\xi) = A_{nl} \mathrm{e}^{-\frac{1}{2}\xi} \xi^l L_{n+l}^{2l+1}(\xi), \xi = \frac{2Z}{nr_1} r; \tag{14.50}$$

$$L_k^s(\xi) = \frac{d^s}{d\xi^s} L_k(\xi); \tag{14.51}$$

$$L_k(\xi) = \mathrm{e}^{\xi} \frac{d^k}{d\xi^k}\left(\mathrm{e}^{-\xi}\xi^k\right). \tag{14.52}$$

The multiplier A_{nl} in (14.50) is determined from normalization condition (14.49). Functions (14.52) are called the **Laguerre polynomials**.

According to (14.30) and (14.22), the wave function of electron has the following structure:

$$\Psi_{nlm}(r, \theta, \varphi) = R_{nl}(r) Y_{lm}(\theta, \varphi) = R_{nl}(r) A_l^{|m|} P_l^{|m|}(\cos\theta)\, \mathrm{e}^{im\varphi}, \tag{14.53}$$

in which the multipliers are expressed by formulas (14.23)–(14.25) and (14.50)–(14.52). As an example and for the following application, we present the explicit form of sought wave functions for a few lowest states,

$$\Psi_{100} = \frac{1}{\sqrt{\pi}}\left(\frac{Z}{r_1}\right)^{\frac{3}{2}} e^{-\frac{Z}{r_1}r}, \tag{14.54}$$

$$\Psi_{200} = \frac{1}{4\sqrt{2\pi}}\left(\frac{Z}{r_1}\right)^{\frac{3}{2}}\left(2-\frac{Z}{r_1}r\right) e^{-\frac{Z}{2r_1}r}, \tag{14.55}$$

$$\Psi_{210} = \frac{1}{4\sqrt{2\pi}}\left(\frac{Z}{r_1}\right)^{\frac{5}{2}} r\, e^{-\frac{Z}{2r_1}r}\cos\theta, \tag{14.56}$$

$$\Psi_{21\pm1} = \frac{1}{8\sqrt{\pi}}\left(\frac{Z}{r_1}\right)^{\frac{5}{2}} r\, e^{-\frac{Z}{2r_1}r}\sin\theta\, e^{\pm i\varphi}. \tag{14.57}$$

14.3. Hydrogen atom

For the hydrogen atom, $Z = 1$. Therefore, on the basis of (14.48), the energy of electron (since the nucleus is assumed to be at rest, this is also the energy of the whole atom) is express in electronvolts,

$$E_n = -\frac{m_0 e^4}{32\pi^2\varepsilon_0^2\hbar^2 n^2} = -\frac{13.55}{n^2}\ \text{eV}. \tag{14.58}$$

The corresponding energy levels are shown in Fig. 14.3, where the hatched region means the continuous spectrum of the electron kinetic energy formed as a result of atomic ionization. The ionization energy of hydrogen atom in the ground state $A_i = -E_1 = 13.55$ eV.

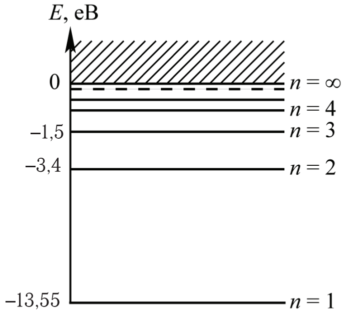

Fig. 14.3.

In the case $Z = 1$, the wave functions of electron (atom) are calculated by formula (14.53). They are completely determined by the quantum numbers n, l, and m. Therefore, these three quantum numbers (respectively, the values E_n, L^2, and L_z) form a complete set of physical parameters that determine the system state in quantum mechanics.

The energy E_n depends only on the principal quantum number n, and the wave functions on the quantum numbers n, l, and m. Therefore, a certain value of energy E_n (the eigenvalue of the Hamilton operator $\widehat{H}$) may be associated with several wave functions (the eigenfunctions of operator $\widehat{H}$). This phenomenon is called the degeneration of operator eigenvalues (see section 12.3) or, in this case, energy levels. At a given n, the quantum number l can obtain values from 0 to $n-1$; and, at a given l, the magnetic quantum number m has $2l+1$ values. Hence, the degeneracy order of energy level E_n equals

$$f = \sum_{l=0}^{n-1} (2l+1) = n^2. \tag{14.59}$$

The wave function $\Psi(\vec{r})$ determines the probability of particle location. Namely, the probability to find the particle in the volume dV around the point with the coordinate $\vec{r}$ amounts to

$$dw(\vec{r}) = |\Psi(\vec{r})|^2 \, dV. \tag{14.60}$$

For the electron in the hydrogen atom, taking into account (14.53), we have

$$\begin{aligned} dw_{nlm}(r,\theta,\varphi) &= |R_{nl}(r) \, Y_{lm}(\theta,\varphi)|^2 \, r^2 dr d\Omega \\ &= R_{nl}^2(r) \, r^2 dr \, |Y_{lm}(\theta,\varphi)|^2 \, d\Omega. \end{aligned}$$

The function

$$w_{lm}(\theta) = |Y_{lm}(\theta,\varphi)|^2 \tag{14.61}$$

is the **angular probability density**. It equals the probability to find the electron inside a unit solid angle in the direction of the angle θ measured from the axis z.

There is no electron trajectory in the atom (see chapter 12). Therefore, the electron motion is represented as an electron cloud with the density equal to the probability density dw_{nlm}/dV. The shape of electron cloud is determined using functions (14.60) and (14.61). In the ground state, $n = 1$ and $l = m = 0$); then, according to (14.26),

$$w_{00}(\theta) = |Y_{00}(\theta)|^2 = \frac{1}{4\pi} = \text{const.}$$

This means that the electron cloud is spherically symmetric in the ground state, i.e. the probability to find the electron is identical in all directions. The radial distribution of probability in this case looks like

$$w_{10}(r) = R_{10}^2(r)\, r^2 = \frac{4}{r_1^3} \mathrm{e}^{-\dfrac{2r}{r_1}} r^2 \tag{14.62}$$

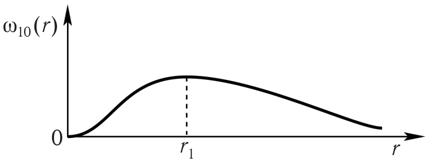

Fig. 14.4.

and is depicted in Fig. 14.4. The analysis shows that the maximum of function $w_{10}(r)$ is located at $r = r_1$. Hence, the radius of the first Bohr orbit in quantum mechanics has the meaning of the most probable distance between the electron and the nucleus, when the system is in the ground state. In a similar way, it can be shown that the most probable distance equals $4r_1$ at $n = 2$, $9r_1$ at $n = 3$, and so on.

From (14.26)–(14.28) or from (14.55)–(14.57), it follows that the state Ψ_{200} is also spherically symmetric, whereas the states Ψ_{210} and $\Psi_{21\pm1}$ are not,

$$w_{10}(\theta) = \frac{3}{4\pi}\cos^2\theta, \quad w_{1\pm1}(\theta) = \frac{3}{8\pi}\sin^2\theta.$$

In the state Ψ_{210}, the electron can be found most probably near the axis z, and, in the state $\Psi_{21\pm1}$, in the plane perpendicular to it.

Notice that there is no selected direction in the atom, e.g., the selected axis z. This direction arises if the atom is located in an external field, e.g., in a magnetic; then the axis z is chosen along the field. Therefore, in the absence of the field, the axis z cannot be distinguished, and, consequently, the angular distribution of probability remains uncertain. Only the radial distribution has sense.

From the reasoning given above, it follows that the concept of "the electron motion in the atom" has to be used carefully, not assigning it the classical content. More correctly in quantum mechanics is to use the term "quantum state".

For quantum states of electron in the atom, there exists special spectroscopic notation in the form nl. For the letter n, the corresponding principal quantum number is used, whereas for l, these are letters s $(l = 0)$, p $(l = 1)$, d $(l = 2)$, f $(l = 3)$, g $(l = 4)$, and h $(l = 5)$. For instance, the state $1s$ means that $n = 1$ and $l = 0$; for the state $2p$, $n = 2$ and $l = 1$; and so on. According to previous remarks, the nl-states are $(2l + 1)$-fold degenerate in the magnetic quantum number m.

14.4. Orbital magnetic moment of electron

Depending on the quantum state of electron, there may be an electric current in the atom, the density of which can be determined by formula (12.68). Taking into account (12.66), this formula looks like

$$\vec{j} = \frac{ie\hbar}{2m}\left(\Psi\nabla\Psi * -\Psi * \nabla\Psi\right).$$

For the electron, $e < 0$, $m = m_0$, and $\Psi = \Psi_{nlm}$; therefore,

$$\vec{j} = -\frac{ie\hbar}{2m_0}\left(\Psi_{nlm}\nabla\Psi^*_{nlm} - \Psi^*_{nlm}\nabla\Psi_{nlm}\right). \tag{14.63}$$

From (14.63), it follows that $\vec{j} = 0$ if the wave function is real. Besides, since the wave functions of electron in the hydrogen atom look like (14.53), i.e.

$$\Psi_{nlm} = R_{nl}(r)\, A_l^{|m|} P_l^{|m|}(\cos\theta)\, \mathrm{e}^{im\varphi}, \tag{14.64}$$

where $R_{nl}(r)$ and $P_l^{|m|}(\cos\theta)$ are real functions, $\vec{j} = 0$ for quantum states with $m = 0$.

At every point of the spherical coordinate system, we can draw three mutually perpendicular unit basis vectors: the vector $\vec{e}_r$ is directed along the radius vector $\vec{r}$, the vector $\vec{e}_\theta$ is tangent to the meridian, and the vector $\vec{e}_\varphi$ is tangent to the parallel (see Appendix C). Let us project the vector of electric current density $\vec{j}$ onto those basis vectors, i.e. let us determine the current densities along the radius vector, meridian, and parallel, taking into account that the projections of operator ∇ onto the vectors $\vec{e}_r$, $\vec{e}_\theta$, and $\vec{e}_\varphi$ look like (see Appendix C)

$$\nabla_r = \frac{\partial}{\partial r}, \quad \nabla_\theta = \frac{1}{r}\frac{\partial}{\partial\theta}, \quad \nabla_\varphi = \frac{1}{r\sin\theta}\frac{\partial}{\partial\varphi}. \tag{14.65}$$

From (14.63) and (14.65), we obtain

$$j_r = -\frac{ie\hbar}{2m_0}\left(\Psi_{nlm}\frac{\partial}{\partial r}\Psi^*_{nlm} - \Psi^*_{nlm}\frac{\partial}{\partial r}\Psi_{nlm}\right), \qquad (14.66)$$

$$j_\theta = -\frac{ie\hbar}{2m_0 r}\left(\Psi_{nlm}\frac{\partial}{\partial \theta}\Psi^*_{nlm} - \Psi^*_{nlm}\frac{\partial}{\partial \theta}\Psi_{nlm}\right), \qquad (14.67)$$

$$j_\varphi = -\frac{ie\hbar}{2m_0 r\sin\theta}\left(\Psi_{nlm}\frac{\partial}{\partial \varphi}\Psi^*_{nlm} - \Psi_{nlm}\frac{\partial}{\partial \varphi}\Psi^*_{nlm}\right). \qquad (14.68)$$

Substituting (14.64) into (14.66) and (14.67), we differentiate the real function $R_{nl}(r)$ in (14.66) with respect to r and the real function $P_l^{|m|}(\cos\theta)$ in (14.67) with respect to θ. Therefore, it is clear that $j_r = j_\theta = 0$.

In (14.68), we differentiate the exponential function $\mathrm{e}^{im\varphi}$ with respect to φ. As a result, we obtain

$$j_\varphi = -\frac{em\hbar}{m_0 r\sin\theta}\left|\Psi_{nlm}\right|^2. \qquad (14.69)$$

Hence, if $m \neq 0$, there are ring currents in the atom that circulate along the parallels. It will be recalled that this conclusion is correct, provided that there is a selected axis z; otherwise, the current direction is uncertain.

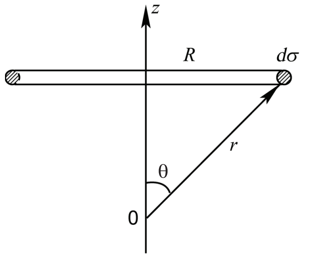

Fig. 14.5.

Now, let us determine the magnetic moment of the atom. For this purpose, let us consider a tube of current locating along the parallel and having the form of a torus with the cross-section $d\sigma$ (Fig. 14.5). The current strength in the torus is $dI = j_\varphi d\sigma$. The magnetic moment of the ring current is defined as the current strength times the contour area. Therefore, the magnetic moment induced by the current dI and directed along the axis z equals

$$dP_{mz} = SdI = \pi R^2 dI = \pi r^2 \sin^2\theta\, j_\varphi d\sigma.$$

Rewriting this expression in the form

$$dP_{mz} = \frac{1}{2} r\sin\theta\, j_\varphi\left(2\pi r\sin\theta\, d\sigma\right) = \frac{1}{2} r\sin\theta\, j_\varphi dV,$$

where $dV = 2\pi R d\sigma = 2\pi r \sin\theta d\sigma$ is the torus volume, and substituting expression (14.69) for j_φ, we obtain

$$dP_{mz} = -\frac{em\hbar}{2m_0} \left|\Psi_{nlm}\right|^2 dV. \tag{14.70}$$

The total projection of atomic magnetic moment onto the axis z equals

$$P_{mz} = -\frac{em\hbar}{2m_0} \int \left|\Psi_{nlm}\right|^2 dV = -\frac{em\hbar}{2m_0} = -m\mu_B. \tag{14.71}$$

The quantity

$$\mu_B = \frac{e\hbar}{2m_0} = 9.27 \times 10^{-24} \text{ J/T}$$

is called the **Bohr magneton**. Equation (14.71) means that the projection of atomic magnetic moment is quantized, being a multiple of the Bohr magneton.

Since $L_z = m\hbar$, equation (14.71) can be rewritten as

$$P_{mz} = -\frac{e}{2m_0} L_z \tag{14.72}$$

or, in the vector form, as

$$\vec{P}_m = -\frac{e}{2m_0} \vec{L}. \tag{14.73}$$

The quantity $-e/\left(2m_0\right)$ is called the **gyromagnetic ratio** between the orbital magnetic and angular (mechanic) momenta of atomic electron.

14.5. Optical-electron model for alkali-metal atoms

Atoms of alkali metals (Li, Na, K, ...) possess one outer electron moving in the electric field created by the nucleus and other atomic electrons. Hence, this is a field of nonpoint charge, which is considered in electrodynamics as a sum of multipole fields: the fields of point charge ($\tilde{}1/r$), electric dipole ($\tilde{}1/r^2$), electric quadrupole, and so on. Therefore, in the dipole approximation, the potential energy of optical electron is written as

$$U\left(r\right) = -\frac{e^2}{4\pi\varepsilon_0 r} - C\frac{e^2}{4\pi\varepsilon_0 r^2}, \tag{14.74}$$

where C is a constant. Energy (14.74) corresponds to the central interaction, which enables us to use the radial Schrödinger equation (14.31). Substituting (14.74) into it, we obtain

$$\widehat{T}_r R(r) + \left\{ \frac{l(l+1)\hbar^2}{2m_0 r^2} - C\frac{e^2}{4\pi\varepsilon_0 r^2} \right\} R(r) - \frac{e^2}{4\pi\varepsilon_0 r} R(r) = ER(r) . \tag{14.75}$$

Introducing the notation

$$l_1(l_1+1) = l(l+1) - C\frac{2m_0 e^2}{4\pi\varepsilon_0 \hbar^2} = l(l+1) - \frac{2C}{r_1}, \tag{14.76}$$

equation (14.75) can be rewritten as

$$\widehat{T}_r R(r) + \frac{l(l+1)\hbar^2}{2m_0 r^2} R(r) - \frac{e^2}{4\pi\varepsilon_0 r} R(r) = ER(r) . \tag{14.77}$$

By its form, equation (3.78) is identical to equation (14.31). Therefore, it has identical solutions at $n \to n_1 = n_r + l_1 + 1$. The number $l_1 > 0$ is determined from quadratic equation (3.77),

$$\begin{aligned} l_1 &= -\frac{1}{2} + \sqrt{\frac{1}{4} + l(l+1) - \frac{2C}{r_1}} = -\frac{1}{2} + \sqrt{\left(l+\frac{1}{2}\right)^2 - \frac{2C}{r_1}} \\ &= -\frac{1}{2} + \left(l+\frac{1}{2}\right)\left\{1 - \frac{2C}{r_1\left(l+\frac{1}{2}\right)^2}\right\} \\ &\approx -\frac{1}{2} + \left(l+\frac{1}{2}\right)\left\{1 - \frac{C}{r_1\left(l+\frac{1}{2}\right)^2}\right\} \approx l - \frac{C}{r_1\left(l+\frac{1}{2}\right)}. \end{aligned} \tag{14.78}$$

Accordingly, the energy of optical electron is expressed by formula (14.58) with the substitution $n \to n_1$,

$$E_{n_1} = E_{nl} = -\frac{m_0 e^4}{32\pi^2\varepsilon_0^2\hbar^2 n_1^2} = -\frac{m_0 e^4}{32\pi^2\varepsilon_0\hbar^2 \left[n - \dfrac{C}{r_1\left(l+\frac{1}{2}\right)}\right]}. \tag{14.79}$$

Hence, in a multielectron atom, the energy of electron depends on both the principal and orbital quantum numbers.

15

ELECTRON SPIN

15.1. Experimental basis

From the viewpoint of classical mechanics, the angular momentum (the mechanical momentum) $\overrightarrow{L}$ and the magnetic moment $\overrightarrow{P}_m$ considered above are a consequence of electron motion around the nucleus. Therefore, they are called the **orbital angular momentum** and the **orbital magnetic moment**, respectively. However, in 1920s, on the basis of experimental data, it was shown that, besides the orbital magnetic moment $\overrightarrow{P}_m$, the electron also possesses the own magnetic moment $\overrightarrow{P}_{ms}$, which is independent of the electron quantum state (motion). In particular, this fact was supported by the splitting of electron energy levels and the availability of lines in the radiation spectra of alkali metal atoms. Namely, the yellow line in the sodium spectrum turned out a doublet, i.e. consisting of two close lines. This splitting is small; therefore, it was called the **fine structure** of energy levels and spectral lines.

An explanation for this phenomenon was proposed by the Dutch physicists S. Goudsmit and G. Uhlenbeck in 1925. They formulated a hypothesis that the electron has a magnetic moment $\overrightarrow{P}_{ms}$ interacting with the magnetic field created by the orbital electron motion. This interaction is called **spin-orbit**. Its energy is an additive to the electron energy E_{nl} and amounts to

$$E_{SO} = -\left(\overrightarrow{P}_{ms}\overrightarrow{B}_{orb}\right) = -P_{msz}B_{orb}, \tag{15.1}$$

where $\overrightarrow{B}_{orb}$ is the magnetic induction vector of orbital magnetic field, and P_{msz} is the projection of the own magnetic moment of electron onto the axis z, which coincides with the direction of vector $\overrightarrow{B}_{orb}$ (if the parameter P_{msz} has f values, the energy E_{SO} also has f values, and the level E_{nl} becomes split into f sublevels).

There is always an interrelation between the magnetic moment and the angular momentum of a charged particle, which is called the **gyromagnetic ratio** and which looks like (14.73) for the electron. Therefore, the own magnetic moment of electron $\overrightarrow{P}_{ms}$ has to be associated with the own angular momentum $\overrightarrow{L}_s$, which is called the **spin**.

An important evidence for the existence of electron spin is the experimental results by O. Stern and W. Gerlach (1922). In those experiments, a beam of monovalent atoms (hydrogens, sodiums, argentums) in the ground state with the zero orbital magnetic moment $\overrightarrow{P}_m$ of valence electron was sent through a transverse inhomogeneous magnetic field. As a result, the beam split into two components. The beam splitting testified that atoms underwent the action of a force that could get two values.

In electrodynamics, it is demonstrated that a magnetic moment in an inhomogeneous magnetic field is subjected to the action of the force

$$\overrightarrow{F} = \left(\overrightarrow{P}_{ms}\nabla\right)\overrightarrow{B}, \tag{15.2}$$

where $\overrightarrow{P}_{ms}$ is the own magnetic moment of electron, because, in our case, $\overrightarrow{P}_m = 0$; and $\overrightarrow{B}$ is the magnetic induction of external inhomogeneous magnetic field. If the magnetic field changes along the axis z,

$$\overrightarrow{F} = P_{msz}\frac{d\overrightarrow{B}}{dz}.$$

Two values of force $\overrightarrow{F}$ correspond to two values of projection P_{msz} and, hence, two spin projections L_{sz}.

The spin obeys the same quantization rules as the orbital angular momentum, but with different quantum numbers,

$$L_s = \sqrt{l_s\left(l_s+1\right)}\hbar; \tag{15.3}$$

$$L_{sz} = m_s\hbar,\ m_s = l_s,\ l_s-1,\ \dots,\ -l_s, \tag{15.4}$$

where l_s is the spin quantum number, and m_s is the quantum number of spin projection, which can take $2l_s+1$ values. Since L_{sz} has two experimental values, i.e. $2l_s+1=2$, this means that $l_s=\frac{1}{2}$ and $m_s=\pm\frac{1}{2}$. Hence, from the Stern–Gerlach experiments, it follows that the absolute value of the spin vector for the electron $L_s=\frac{1}{2}\sqrt{3}\hbar$, and its projection $L_{sz}=\pm\frac{\hbar}{2}$.

The Stern–Gerlach experiments also proved that the projection of own magnetic moment of electron equals the Bohr magneton, i.e.

$$P_{msz} = \pm\mu_B = \pm\frac{e\hbar}{2m_0}. \tag{15.5}$$

This fact means that the gyromagnetic ratio for the spin moments looks like (the **anomalous gyromagnetic ratio**)

$$P_{msz} = -\frac{e}{m_0} L_{sz} \tag{15.6}$$

or, in the vector form,

$$\overrightarrow{P}_{ms} = -\frac{e}{m_0} \overrightarrow{L}_s. \tag{15.7}$$

Notice that, it is the projection P_{msz} or, equivalently, the spin projection $L_{sz} = \pm\frac{\hbar}{2}$ rather than the absolute value $L_s = \frac{1}{2}\sqrt{3}\hbar$ that manifests itself in physical phenomena associated with the electron spin. Therefore, the electron spin is said to equal $\frac{\hbar}{2}$, or, in dimensionless units $\frac{1}{2}$, i.e. l_s.

In the early models, the electron spin was considered as a consequence of its rotation around some axis. However, it was found out that this model has principal difficulties associated with electron sizes and the theory of relativity. On the other hand, P. Dirac showed that the spin automatically follows from the relativistic wave equation for the electron. Therefore, the spin is considered to have a relativistic origin and to be an intrinsic property of electron like the electric charge.

15.2. Spin functions. Spin operators

Making no allowance for the spin, the wave function of electron in the atom is determined by the principal, orbital, and magnetic quantum numbers (see section 14.2),

$$\Psi = \Psi_{nlm}(r, \theta, \varphi), \tag{15.8}$$

and its energy E_{nl} by the principal and orbital quantum numbers. If the spin is taken into account, the electron energy changes owing to the spin-orbit interaction. However, according to experimental data, $|E_{SO}| \ll |E_{nl}|$. Therefore, the spin-orbit interaction can be neglected in some cases. However, the spin orientation (along or opposite to the direction of axis z) can be of great importance concerning other physical properties.

Therefore, if the spin orientation is taken into account, but the spin-orbit interaction is not, the wave function of electron in the atom is written as

$$\Psi = \Psi_{nlmm_s}(r, \theta, \varphi, L_{sz}) = \Psi_{nlm}(r, \theta, \varphi)\, \chi_{m_s}(L_{sz}), \tag{15.9}$$

where the factor $\chi_{m_s}(L_{sz})$, which is called the **spin function**, indicates the spin orientation a back; namely,

$$\chi_{1/2}\left(\frac{\hbar}{2}\right)=\chi_{-1/2}\left(-\frac{\hbar}{2}\right)=1,\quad \chi_{1/2}\left(-\frac{\hbar}{2}\right)=\chi_{-1/2}\left(\frac{\hbar}{2}\right)=0. \tag{15.10}$$

With (15.10) in view, it is easy to be convinced that the spin functions are orthonormalized,

$$\sum_{L_{sz}}\chi^*_{m_s}(L_{sz})\,\chi_{m'_s}(L_{sz})=$$

$$=\chi^*_{m_s}\left(-\frac{\hbar}{2}\right)\chi_{m'_s}\left(-\frac{\hbar}{2}\right)+\chi^*_{m_s}\left(\frac{\hbar}{2}\right)\chi_{m'_s}\left(\frac{\hbar}{2}\right)=\delta_{m_s,\,m'_s}. \tag{15.11}$$

Since the spin variable L_{sz} has two values, wave function (15.9) is written in the matrix form,

$$\Psi=\begin{pmatrix}\Psi_1\\ \Psi_2\end{pmatrix}\equiv\begin{pmatrix}\Psi_1 & 0\\ \Psi_2 & 0\end{pmatrix}, \tag{15.12}$$

where $\Psi_1=\Psi_{nlm}\chi_{1/2}L_{sz}$ and $\Psi_2=\Psi_{nlm}\chi_{-1/2}L_{sz}$. Therefore, according to (15.10), if $L_{sz}=\hbar/2$, we obtain $\Psi_1=\Psi_{nlm}$ and $\Psi_2=0$, i.e.

$$\Psi=\begin{pmatrix}\Psi_1\\ 0\end{pmatrix}=\Psi_1\begin{pmatrix}1\\ 0\end{pmatrix}=\Psi_{nlm}\chi_{1/2},\quad \chi_{1/2}=\chi_{1/2}\left(\frac{\hbar}{2}\right)=\begin{pmatrix}1\\ 0\end{pmatrix},$$

whereas if $L_{sz}=-\hbar/2$, we obtain $\Psi_1=0$ and $\Psi_2=\Psi_{nlm}$, i.e.

$$\Psi=\begin{pmatrix}0\\ \Psi_2\end{pmatrix}=\Psi_1\begin{pmatrix}0\\ 1\end{pmatrix}=\Psi_{nlm}\chi_{-1/2},$$

$$\chi_{-1/2}=\chi_{-1/2}\left(-\frac{\hbar}{2}\right)=\begin{pmatrix}0\\ 1\end{pmatrix},$$

Each of the spin functions

$$\chi_{1/2}=\begin{pmatrix}1\\ 0\end{pmatrix},\quad \chi_{-1/2}\begin{pmatrix}0\\ 1\end{pmatrix}$$

corresponds to the case when the electron has a definite value of spin projection, $L_{sz}=\frac{\hbar}{2}$ or $-\frac{\hbar}{2}$. However, there exist electron states, for which the spin projection is uncertain, so that one can talk only about a probability to obtain either of the value $\frac{\hbar}{2}$ or $-\frac{\hbar}{2}$ at the measurement. In such cases, the spin function is written in the form

$$\chi = \begin{pmatrix} a \\ b \end{pmatrix}$$

that obeys the condition

$$\chi^+\chi = \begin{pmatrix} a \\ b \end{pmatrix}^+ \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a^* & b^* \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = |a|^2 + |b|^2 = 1,$$

whence it follows that $|a|^2$ is a probability to measure the value $\hbar/2$, and $|b|^2$ the value $-\hbar/2$.

The electron spin, as a dynamic variable, is represented by the corresponding operator. There are also operators for the spin projections L_{sx}, L_{sy}, L_{sz}, and the square of absolute value $\left|\overrightarrow{L}_s\right|^2$. It is possible to show (see Appendix G) that the spin operators look like

$$\widehat{L}_{sx} = \frac{\hbar}{2}\widehat{\sigma}_x, \quad \widehat{L}_{sy} = \frac{\hbar}{2}\widehat{\sigma}_y, \quad \widehat{L}_{sz} = \frac{\hbar}{2}\widehat{\sigma}_z, \quad \widehat{L}_s^2 = \frac{1}{4}\hbar^2\widehat{\sigma}^2, \tag{15.13}$$

where $\widehat{\sigma}_x$, $\widehat{\sigma}_y$, and $\widehat{\sigma}_z$ are the **Pauli operators** (matrices), or the **spin matrices**, which look like

$$\widehat{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \widehat{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \widehat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$
$$\widehat{\sigma}^2 = 3\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{15.14}$$

Therefore,

$$\widehat{L}_{sx} = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \widehat{L}_{sy} = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \widehat{L}_{sz} = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$
$$\widehat{L}_s^2 = \frac{3}{4}\hbar^2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{15.15}$$

Let us determine the eigenvalues and eigenfunctions of spin operators (15.15). The corresponding equation for the operator $\widehat{L}_{sx}$ looks like

$$\widehat{L}_{sx}\chi^{(x)} = L_{sx}\chi^{(x)}. \tag{15.16}$$

The eigenfunction $\chi^{(x)}$ is sought as a column matrix

$$\chi^{(x)} = \begin{pmatrix} a \\ b \end{pmatrix}. \tag{15.17}$$

Substituting (15.17) into (15.16) and taken into account (15.15), we obtain the equation

$$\frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = L_{sx}\begin{pmatrix} a \\ b \end{pmatrix}.$$

After calculating the matrix product, we obtain

$$\frac{\hbar}{2}\begin{pmatrix} b \\ a \end{pmatrix} = L_{sx}\begin{pmatrix} a \\ b \end{pmatrix},$$

whence a system of linear equations for the matrix components a and b follows,

$$\frac{\hbar}{2}b = L_{sx}a, \quad \frac{\hbar}{2}a = L_{sx}b. \tag{15.18}$$

Multiplying those equations and dividing the result by ab, we obtain

$$L_{sx}^2 = \frac{\hbar}{4}, \quad L_{sx} = \pm\frac{\hbar}{2}, \tag{15.19}$$

i.e. the eigenvalues of operator $\widehat{L}_{sx}$ are $\hbar/2$ and $-\hbar/2$.

Substituting the eigenvalue $L_{sx} = \hbar/2$ into system of equations (15.18), we obtain $b = a$. Then, eigenfunction (15.17) takes the form

$$\chi_{1/2}^{(x)}\begin{pmatrix} a \\ a \end{pmatrix} = a\begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

The normalization condition for this function,

$$\left(\chi_{1/2}^{(x)}\right)^{+}\chi_{1/2}^{(x)} = a^{*}\begin{pmatrix} 1 \\ 1 \end{pmatrix}^{+} a\begin{pmatrix} 1 \\ 1 \end{pmatrix} = |a|^2\,(1 \quad 1)\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 2\,|a|^2 = 1,$$

makes it possible to find a,

$$|a|^2 = 1/2, \quad a = \frac{1}{\sqrt{2}}\mathrm{e}^{i\alpha}.$$

Neglecting the phase α and putting $\alpha = 0$, we obtain $a = 1/\sqrt{2}$ and

$$\chi_{1/2}^{(x)} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \tag{15.20}$$

It is easy to check that the eigenvalue $L_{sx} = -\hbar/2$ corresponds to the eigenfunction

$$\chi_{-1/2}^{(x)} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \tag{15.21}$$

In a similar way, one can find the eigenvalues and eigenfunctions of operators $\widehat{L}_{sy}$ and $\widehat{L}_{sz}$.

15.3. Total angular momentum and total magnetic moment of electron

The spin-orbit interaction between the orbital, $\overrightarrow{P}_m$, and the intrinsic, $\overrightarrow{P}_{ms}$, magnetic moments of the electron gives rise to their summation into the total magnetic moment $\overrightarrow{P}_{mj}$, and the angular momenta $\overrightarrow{L}$ and $\overrightarrow{L}_s$ into the total angular momentum $\overrightarrow{L}_j$. Experimental and theoretical researches proved that the vectors $\overrightarrow{L}$ and $\overrightarrow{L}_s$ are summed up following the parallelogram rule (this is the essence of the **vector model of atom**),

$$\overrightarrow{L}_j = \overrightarrow{L} + \overrightarrow{L}_s, \tag{15.22}$$

but making allowance for the quantization of the absolute values and projections of the vectors $\overrightarrow{L}$ and $\overrightarrow{L}_s$,

$$L = \sqrt{l\,(l+1)}\hbar, \quad L_z = m\hbar, \quad L_s = \sqrt{l_s\,(l_s+1)}\hbar, \quad L_{sz} = m_s\hbar, \tag{15.23}$$

and the quantization of L_j and L_{jz},

$$L_j = \sqrt{j\,(j+1)}\hbar, \quad j = |l \pm l_s| = \left| l \pm \frac{1}{2} \right|, \tag{15.24}$$

$$L_{jz} = m_j\hbar, \quad m_j = j, j-1, \ldots, -j. \tag{15.25}$$

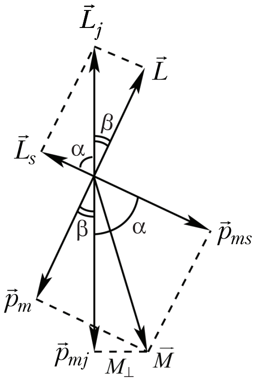

Fig. 15.1.

The total angular momentum of electron, $\overrightarrow{L}_j$, corresponds to the total magnetic moment $\overrightarrow{P}_{mj}$, the formation of which is illustrated in Fig. 15.1. According to gyromagnetic ratios (14.72) and (14.73), the vectors $\overrightarrow{P}_m$ and $\overrightarrow{P}_{ms}$ are directed oppositely to the vectors $\overrightarrow{L}$ and $\overrightarrow{L}_s$, respectively. All exhibited vectors precess around the direction of vector $\overrightarrow{L}_j$. Therefore, only the projection of vector $\overrightarrow{M} = \overrightarrow{P}_m + \overrightarrow{P}_{ms}$ onto the direction of vector $\overrightarrow{L}_j$ differs from zero. This projection is taken as the total magnetic moment of electron (the average value of component $M_\perp$ equals zero).

From Fig. 15.1, it follows that

$$P_{mj} = P_{ms} \cos\alpha + P_m \cos\beta. \tag{15.26}$$

Examining the triangles formed by the vectors $\overrightarrow{L}$, $\overrightarrow{L}_s$, and $\overrightarrow{L}_j$, we see that

$$L^2 = L_j^2 + L_s^2 - 2L_j L_s \cos\alpha, \quad L_s^2 = L_j^2 + L^2 - 2L_j L \cos\beta,$$

i.e.

$$\cos\alpha = \frac{L_j^2 + L_s^2 - L^2}{2L_j L_s}, \quad \cos\beta = \frac{L_j^2 + L^2 - L_s^2}{2L_j L}. \tag{15.27}$$

Substituting (4.29) into (4.28) and taking into account (14.73) and (15.7), we find

$$P_{mj} = -\frac{e}{m_0} \frac{L_j^2 + L_s^2 - L^2}{2L_j} - \frac{e}{2m_0} \frac{L_j^2 + L^2 - L_s^2}{2L_j} = -\frac{e}{2m_0} \frac{L_j^2 + L_s^2 - L^2}{2L_j}.$$

Using formulas (15.23) and (15.24) to express L^2, L_s^2, and L_j^2, we finally obtain

$$P_{mj} = -\frac{e}{2m_0} g L_j, \tag{15.28}$$

where

$$g = 1 + \frac{j\,(j+1) + l_s\,(l_s+1) - l\,(l+1)}{2j\,(j+1)} \tag{15.29}$$

is the **Landé factor**. In the vector form, equation (15.28) looks like

$$\overrightarrow{P}_{mj} = -\frac{e}{2m_0} g \overrightarrow{L}_j. \tag{15.30}$$

The projection of vector $\overrightarrow{P}_{mj}$ onto the selected direction is

$$P_{mjz} = -\frac{e}{2m_0} g L_{jz} = -\frac{e}{2m_0} g m_j \hbar = -\mu_B g m_j. \tag{15.31}$$

15.4. Fine-structure formula

As was already indicated in section 15.1, the electron spin has a relativistic origin. Therefore, the problem of hydrogen atom making allowance for the electron spin is a relativistic quantum-mechanical one. Hence, it falls beyond the scope of this textbook. However, the mentioned problem is of principal importance, we give some of its results.

In 1928, P. Dirac derived a relativistic wave equation, which could be used to find an expression for the energy of hydrogen atom. This expression is called the fine-structure formula, and it looks like

$$E_{nj} = E_n \left\{ 1 + \frac{\alpha^2 Z^2}{n} \left(\frac{1}{j+1/2} - \frac{3}{4n} \right) \right\}, \tag{15.32}$$

where $E_n = -13.55 Z^2/n^2$ eV is the energy of electron in the hydrogen atom determined without taking into account the spin (see section 3.2), and

$$\alpha = -\frac{e^2}{4\pi\varepsilon_0 \hbar c} \approx \frac{1}{137}$$

is a constant, which is respectively called the **fine-structure constant**. At the same time, it is also of great importance in the theory of elementary particles, where it is called the **coupling constant**.

The energy levels of hydrogen atom calculated for $n = 1, 2, 3$ without making allowance for their fine structure (left panel) and according to formula (4.34) (right panel) are depicted in Fig. 15.2. The subscripts (1/2, 3/2, 5/2) in the level notation specify the value of quantum number j. $\int t$is evident that the account of spin (fine structure) results in the level splitting, the magnitude of which decreases as the quantum number n grows. For $n \geq 3$, the level splitting can be neglected. The arrows denote

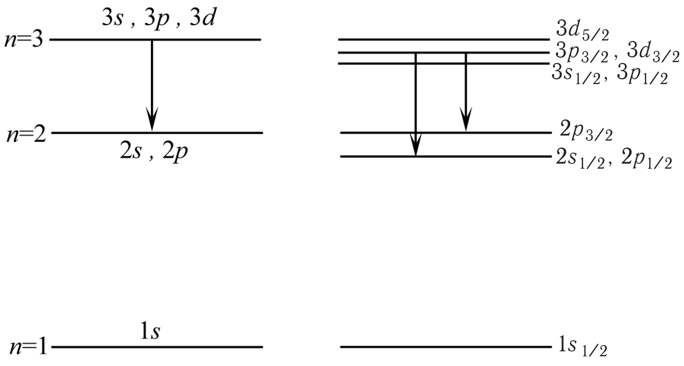

Fig. 15.2.

quantum transitions which are responsible for the red line in the Balmer series (H_α). This line is a doublet with the wavelength difference $\Delta\lambda = 5 \times 10^{-4}$ nm ($\lambda_{H_\alpha} = 656$ nm) originating from the spin-orbit interaction.

16

MULTI-ELECTRON ATOMS

16.1. Principle of identity. Symmetric and antisymmetric states. Bosons and fermions

There are quantum-mechanical systems consisting of identical (of the same kind) microparticles. Such systems include, e.g., electrons in the atom or in the crystal, protons or neutrons in the atomic nucleus, photons with the identical frequency ω and wave vector $\vec{k}$, and so forth. In this connection, there arises a question: To what degree are those particles similar, identical, indistinguishable?

In classical mechanics, which studies the motion of macroscopic objects, there are no absolutely identical (by size, mass, and so on) objects; therefore, they can always be distinguished from one another. Even the idealizing objects of classical mechanics, material points, can be distinguished, because they move along individual paths: at a given time moment, every of them has exact specific values of coordinates, and the motion of every point can be traced.

In quantum mechanics (in the microworld), the situation is different. Owing to the wave properties of microparticles, the absence of paths, and the overlapping of wave functions in the space, it is impossible to trace the motion of any particle in the system. Therefore, having detected a particle at some point in the space, this particle could have been identified (What is its "number"?) only on the basis of its individual attributes. However, if such attributes are absent, i.e. in the case of absolute identity of particles, this question cannot be answered in principle. At the same time, we cannot know a priori whether, e.g., electrons are absolutely identical or not with respect to the mass, electric charge, spin, and so forth. Only experimental data can answer this question, but not the question concerning the direct measurements of indicated quantities, because there are no absolutely exact measurements.

It was found out that whether microparticles of the same kind are distinguishable or not essentially affects the properties of the systems consisting of them. Systems of identical, indistinguishable microparticles possess the properties that do not exist in classical mechanics. Therefore, let us consider them in more details.

Let a system consisting of N identical (indistinguishable) microparticles be in a quantum state described by the wave function

$$\Psi = \Psi\left(1, 2, \ldots, i, j, \ldots, N\right), \tag{16.1}$$

where numbers stand for the coordinates, including the spin variables, of corresponding microparticles. If microparticles are identical (indistinguishable), we may assert that the mutual permutation of two arbitrary particles (the interchange of their coordinates and spin variables, but not the quantum numbers) does not change the quantum state of the system. This statement is called the **principle of identity**. Hence, its Hamiltonian has to be invariant with respect to the operation of particle permutation (see section 1.5). To analyze the consequences that follow from permutation symmetry, let us introduce the permutation operator $\widehat{P}_{ij}$ that swaps the coordinates of the i-th and j-th particles in the wave function,

$$\widehat{P}_{ij}\Psi\left(\ldots, i, \ldots, j, \ldots\right) = \Psi\left(\ldots, j, \ldots, i, \ldots\right). \tag{16.2}$$

The invariance of Hamiltonian with respect to particle permutations results in that the operators $\widehat{P}$ and $\widehat{H}$ commute, i.e.

$$\widehat{P}\widehat{H} = \widehat{H}\widehat{P} \tag{16.3}$$

(hereafter, the subscripts i and j are not indicated). This means that those operators have common eigenfunctions (see section 1.3), i.e. the wave function Ψ (the eigenfunction of operator $\widehat{H}$) must also be the eigenfunction of operator $\widehat{P}$,

$$\widehat{P}\Psi = \lambda\Psi, \tag{16.4}$$

where λ is the eigenvalue of operator $\widehat{P}$. Whence,

$$\widehat{P}^2\Psi = \lambda^2\Psi = \Psi;$$

therefore, $\lambda^2 = 1$ and $\lambda = \pm 1$. The wave function is called **symmetric** (Ψ_s) if it does not change its sign ($\lambda = 1$) at the particle permutation, and **antisymmetric** (Ψ_a) if it does ($\lambda = -1$); i.e.

$$\begin{aligned} \widehat{P}\Psi_s &= \Psi_s, \quad \lambda = 1, \\ \widehat{P}\Psi_a &= -\Psi_a, \quad \lambda = -1. \end{aligned} \tag{16.5}$$

Equality (16.3) brings about the conservation law for the eigenvalue λ of operator $\widehat{P}$ or, equivalently, the symmetry type of the quantum state Ψ. This means that all microparticles in the nature can be divided into two classes. Particles in one of them compose systems described by symmetric wave functions, and particles in the other, systems described by antisymmetric wave functions. Particles that compose systems described by symmetric wave functions are called **bosons**, and particles that compose systems described by antisymmetric wave functions are called **fermions**,. The class of bosons inludes particles with integer spins ($l_s = 0, 1, 2, \ldots$), and that of fermions, particles with half-integer spins ($l_s = 1/2, 3/2, \ldots$). Photons ($l_s = 1$), pions ($l_s = 0$), and kaons ($l_s = 0$) are bosons; whereas electrons ($l_s = 1/2$), protons ($l_s = 1/2$), neutrons ($l_s = 1/2$), and other particles are fermions.

In a system with a very large number of particles, statistical laws are valid. In a system of bosons, the particles occupy the energy levels according to the **Bose–Einstein distribution**

$$\langle n \rangle = \left(\mathrm{e}^{\dfrac{\varepsilon - \mu}{kT}} - 1 \right)^{-1}, \tag{16.6}$$

and in a system of fermions, according to the **Fermi–Dirac distribution**

$$\langle n\rangle = \left(\mathrm{e}^{\dfrac{\varepsilon-\mu}{kT}} + 1 \right)^{-1}, \tag{16.7}$$

where $\langle n\rangle$ is the average number of particles in the quantum state with the kinetic energy ε, μ the chemical potential, k the Boltzmann constant, and T the temperature.

16.2. Wave functions of the systems of bosons and fermions. Pauli principle

In order to make calculations for boson and fermion systems, we must use the corresponding wave functions Ψ_s and Ψ_a. Let us find them in the case of weakly interacting particles, when the interaction can be neglected, and the particles can be regarded to be in individual one-particle states $\Psi_{n_j}(i)$. Here, the argument i means the coordinates (including the spin coordinate L_{sz}) and n_i the set of quantum numbers (including the quantum number m_s) of the particle.

First, let us consider a system of two particles, 1 and 2. Their energies as denoted as E_1 and E_2, the wave functions as $\Psi_{n_1}(1)$ and $\Psi_{n_2}(2)$, the Hamiltonians as $\widehat{H}(1)$ and $\widehat{H}(2)$, respectively, and the wave function of the whole system as $\Psi(1,2)$. Then, it is easy to be convinced that the Schrödinger equation for this system,

$$\left[\widehat{H}\,(1) + \widehat{H}\,(2)\right] \Psi\,(1,2) = (E_1 + E_2)\,\Psi\,(1,2)\,, \tag{16.8}$$

is satisfied by the wave function

$$\Psi\,(1,2) = \Psi_{n_1}\,(1)\,\Psi_{n_2}\,(2)\,. \tag{16.9}$$

According to the principle of identity, the permutation of the particles does not change the system state. Therefore, the wave function

$$\Psi\,(2,1) = \Psi_{n_1}\,(2)\,\Psi_{n_2}\,(1) \tag{16.10}$$

is also the solution of Schrödinger equation (16.8). The general solution of equation (5.8) has to be either a symmetric (for bosons) or an antisymmetric (for fermions) linear combination of solutions (5.9) and (5.10).

Such combinations look like

$$\Psi_s = \frac{1}{\sqrt{2}}\left\{\Psi\left(1,2\right)+\Psi\left(2,1\right)\right\} = \frac{1}{\sqrt{2}}\left\{\Psi_{n_1}\left(1\right)\Psi_{n_2}\left(2\right)+\Psi_{n_1}\left(2\right)\Psi_{n_2}\left(1\right)\right\}, \tag{16.11}$$

$$\Psi_a = \frac{1}{\sqrt{2}}\left\{\Psi\left(1,2\right)-\Psi\left(2,1\right)\right\} = \frac{1}{\sqrt{2}}\left\{\Psi_{n_1}\left(1\right)\Psi_{n_2}\left(2\right)-\Psi_{n_1}\left(2\right)\Psi_{n_2}\left(1\right)\right\}. \tag{16.12}$$

Introducing the permutation operator $\widehat{P}_\nu$, functions (16.11) and (16.12) can be written as

$$\Psi_s = \frac{1}{\sqrt{2!}}\sum_{\nu=0}^{1}\widehat{P}_\nu\Psi_{n_1}\left(1\right)\Psi_{n_2}\left(2\right), \tag{16.13}$$

$$\Psi_a = \frac{1}{\sqrt{2!}}\sum_{\nu=0}^{1}\left(-1\right)^\nu\Psi_{n_1}\left(1\right)\Psi_{n_2}\left(2\right), \tag{16.14}$$

where

$$\widehat{P}_0\Psi_{n_1}\left(1\right)\Psi_{n_2}\left(2\right) = \Psi_{n_1}\left(1\right)\Psi_{n_2}\left(2\right),$$

$$\widehat{P}_1\Psi_{n_1}\left(1\right)\Psi_{n_2}\left(2\right) = \Psi_{n_1}\left(2\right)\Psi_{n_2}\left(1\right).$$

Formulas (16.13) and (16.14) can be easily generalized to a system consisting of N particles,

$$\Psi_s = \frac{1}{\sqrt{N!}}\sum_{\nu}\widehat{P}_\nu\Psi_{n_1}\left(1\right)\Psi_{n_2}\left(2\right)\ldots\Psi_{n_N}\left(N\right), \tag{16.15}$$

$$\Psi_a = \frac{1}{\sqrt{N!}}\sum_{\nu}\left(-1\right)^\nu\widehat{P}_\nu\Psi_{n_1}\left(1\right)\Psi_{n_2}\left(2\right)\ldots\Psi_{n_N}\left(N\right). \tag{16.16}$$

Here, summation is carried out over all possible particle permutations ν's ($\widehat{P}_0 = 1$), i.e. each right hand side contains $N!$ summands. The multiplier $1/N!$ is normalizing. The wave function of fermions (16.16) can also be written in the form of determinant,

$$\Psi_a = \frac{1}{\sqrt{N!}}\begin{vmatrix} \Psi_{n_1}\left(1\right) & \Psi_{n_1}\left(2\right) & \ldots & \Psi_{n_1}\left(N\right) \\ \Psi_{n_2}\left(1\right) & \Psi_{n_2}\left(2\right) & \ldots & \Psi_{n_2}\left(N\right) \\ \ldots & \ldots & \ldots & \ldots \\ \Psi_{n_N}\left(1\right) & \Psi_{n_N}\left(2\right) & \ldots & \Psi_{n_N}\left(N\right) \end{vmatrix}. \tag{16.17}$$

From this formula, an important consequence follows. Let $n_1 = n_2$, i.e. two fermions are in the same quantum state. Then, the first and the second row in the determinant are identical, and the determinant equals zero. However, the equality of the wave function to zero means that the corresponding quantum state of the system does not exist. Hence, in a system of fermions, only one particle is allowed to occupy any quantum state. This is the essence of the **Pauli principle**, which was formulated by W. Pauli in 1925.

The Pauli principle is one of the most fundamental principles of quantum mechanics. He plays an important role for the explanation of structures of electron shells in atoms and molecules, atomic and molecular spectra, in the quantum theories of solids and atomic nucleus, and so on.

16.3. Helium atom. Exchange interaction

An important consequence that follows from the identity of microparticles belonging to the same kind is the **exchange interaction** between them. This interaction makes it possible to explain the peculiarities in atomic spectra, the nature of chemical interaction forces, ferromagnetism, and so forth. Let us consider it using the helium atom as an example.

The helium atom consists of a nucleus ($Z = 2$) and two electrons. Since the nucleus mass is much larger than that of electrons, the nucleus is considered to be at rest, and the system of two electrons is examined in the constant Coulomb field of the nucleus.

Electrons are fermions; therefore, their wave function must be antisymmetric (Ψ_a). Taking into account the spin, but neglecting the spin-orbit interaction, the wave function can be written as a product of the coordinate, $\Phi(\vec{r}_1, \vec{r}_2)$, and spin,

$$\chi_{m_{s1} m_{s2}}(L_{sz1}, L_{sz2}) \equiv \chi_{m_{s1} m_{s2}}(1, 2), \tag{16.18}$$

parts, i.e.

$$\Psi_a = \Phi\chi = \Phi(\vec{r}_1, \vec{r}_2)\, \chi(L_{sz1}, L_{sz2}). \tag{16.19}$$

The antisymmetric character of the product $\Phi\chi$ can be provided in two ways,

$$\Psi_a = \begin{cases} \Phi_s \chi_a; \\ \Phi_a \chi_s. \end{cases} \tag{16.20}$$

If the interaction between the own magnetic moments of electrons (the spin–spin interaction) is neglected, spin function (16.18) is written as a product of one-electron spin functions,

$$\chi_{m_{s1}m_{s2}}(1,2) = \chi_{m_{s1}}(1)\,\chi_{m_{s2}}(2)\,. \tag{16.21}$$

Then, the symmetric spin function χ_s can be written down in three variants,

$$\chi_s^{(1)} = \chi_{1/2}(1)\,\chi_{1/2}(2)\,, \quad \chi_s^{(2)} = \chi_{-1/2}(1)\,\chi_{-1/2}(2)\,, \tag{16.22}$$

$$\chi_s^{(3)} = \frac{1}{\sqrt{2}}\left\{\chi_{1/2}(1)\,\chi_{-1/2}(2) + \chi_{1/2}(2)\,\chi_{-1/2}(1)\right\}. \tag{16.23}$$

In all three cases, the electron spins are parallel, and the total spin equals unity, i.e. $\hbar$. However, in cases (16.25), the total spin is oriented along or in the opposite direction to the axis z, and in case (16.23) perpendicularly to it. The antisymmetric spin function χ_a can only be written as

$$\chi_a = \frac{1}{\sqrt{2}}\left\{\chi_{1/2}(1)\,\chi_{-1/2}(2) - \chi_{1/2}(2)\,\chi_{-1/2}(1)\right\}. \tag{16.24}$$

In this case, the spins of electrons are antiparallel to each other, and the total spin equals zero. Taking into account the normalization of one-electron spin functions (see (15.11)), we can get convinced that spin functions (16.22)–(16.24) are also normalized to unity,

$$\sum_{L_{sz1},\,L_{sz2}} \chi_{s,a}^{+}(L_{sz1}, L_{sz2},) = \chi_{s,a}(L_{sz1}, L_{sz2},) = 1. \tag{16.25}$$

Here, the both subscripts must be equal to either s or a.

Hence, all states of helium atom are grouped into two classes. One of them includes the states with parallel electron spins, **orthohelium states**, and the other includes the states with antiparallel electron spins, **parahelium states**. Experimental data demonstrate that the spectral lines of orthohelium are triplets, and those emitted by parahelium are singlets.

Now let us determine the energy of helium atom, i.e. the energy of its electron subsystem in the field of nucleus. Neglecting the spin-orbit and spin-spin interaction, the Hamilton operator looks like

$$\widehat{H} = \frac{\hbar^2}{2m_0}\Delta_1 - \frac{Ze^2}{4\pi\varepsilon_0 r_1} - \frac{\hbar^2}{2m_0}\Delta_2 - \frac{Ze^2}{4\pi\varepsilon_0 r_2} + \frac{e^2}{4\pi\varepsilon_0 r_{12}}. \tag{16.26}$$

The first and the third summand are the operators of the kinetic energy of electrons, the second and the fourth one are the operators of the potential energy of electrons in the electrostatic field of nucleus ($Z = 2$), and the last summand is the energy operator of Coulomb interaction between the electrons. Formally, the first two summands in (16.26) are the Hamiltonian of electron 1 in a hydrogen-like ion, and the next two summands are the Hamiltonian of electron 2 also in in a hydrogen-like ion. Therefore, we may write

$$\widehat{H} = \widehat{H}_{\mathrm{hydr}}\left(\vec{r}_1\right) + \widehat{H}_{\mathrm{hydr}}\left(\vec{r}_2\right) + \frac{e^2}{4\pi\varepsilon_0 r_{12}}. \tag{16.27}$$

In this case, the energy of electrons is determines as

$$E = E_n + E_k + \langle V\rangle\,, \tag{16.28}$$

where

$$E_n = -\frac{13.55Z^2}{n^2}\ \mathrm{eV}, \quad E_k = -\frac{13.55Z^2}{k^2}\ eV, \quad (n,k) \in N, \tag{16.29}$$

and $\langle V\rangle$ is the average value of the interaction energy of electrons describing by the operator

$$\widehat{V} = \frac{e^2}{4\pi\varepsilon_0 r_{12}}. \tag{16.30}$$

The wave functions of the system of two electrons, according to (16.20), look like

$$\Psi_a = \Phi_{s,a}\chi_{s,a}. \tag{16.31}$$

Here, the spin functions $\chi_{s,a}$ are defined by formulas (16.22)–(16.24) with normalization condition (16.25), and the coordinate functions $\Phi_{s,a}$ by formulas (16.11) and (16.12), i.e.

$$\Phi_{s,a} = \frac{1}{\sqrt{2}}\left\{\Psi_{n_1}\left(\vec{r}_1\right)\Psi_{n_2}\left(\vec{r}_2\right) \pm \Psi_{n_1}\left(\vec{r}_2\right)\Psi_{n_2}\left(\vec{r}_1\right)\right\} = \frac{1}{\sqrt{2}}\left(\Psi_1 \pm \Psi_2\right), \tag{16.32}$$

where $n_1 = (n, l_1, m_1)$ and $n_2 = (k, l_2, m_2)$, and the form of functions Ψ_1 and Ψ_2 follows from (16.32).

Calculating the average energy of interaction between the electrons, $\langle V\rangle$, with the help of functions (5.31), we obtain

$$\langle V\rangle \sum_{L_{sz1},\,L_{sz2}} \Phi^*_{s,a}\left(\vec{r}_1, \vec{r}_2\right)\chi^+_{s,a}\left(L_{sz1}, L_{sz2}\right)\widehat{V}\times$$

$$\times \Phi_{s,a}(\vec{r}_1, \vec{r}_2)\, \chi_{s,a}(L_{sz1}, L_{sz2})\, dv_1 dv_2 =$$

$$= \Phi^{+}_{s,a}(\vec{r}_1, \vec{r}_2)\, \widehat{V} \Phi_{s,a}(\vec{r}_1, \vec{r}_2)\, dv_1 dv_2. \qquad (16.33)$$

Here, we took into account that the operator $\widehat{V}$ does not depend on the spin variables and used the condition of spin function normalization (16.25). Substituting (16.32) into (16.33), we obtain a sum of integrals

$$\langle V \rangle = 1/2 \int \Psi_1^* \widehat{V} \Psi_1 dv_1 dv_2 \pm 1/2 \int \Psi_1^* \widehat{V} \Psi_2 dv_1 dv_2 \pm$$

$$\pm 1/2 \int \Psi_2^* \widehat{V} \Psi_1 dv_1 dv_2 + \frac{1}{2} \int \Psi_2^* \widehat{V} \Psi_2 dv_1 dv_2, \qquad (16.34)$$

where and the first and the last integral have identical values, because the function Ψ_2 is formed as the function Ψ_1 but with the particle (coordinate) swapping, so that the definite integrals different in only the notation of variables are identical. For the same reason, the second and the third integral in (16.34) also have identical values. Therefore, it is possible to write

$$\langle V \rangle = \int \Psi_1^* \widehat{V} \Psi_1 dv_1 dv_2 \pm \int \Psi_1^* \widehat{V} \Psi_2 dv_1 dv_2 = K \pm A. \qquad (16.35)$$

What is the meaning of integrals K and A? On the basis of (16.30) and (16.32), we have

$$K(n_1, n_2) = \int \frac{\left[-e\Psi^*_{n_1}(\vec{r}_1)\, \Psi_{n_1}(\vec{r}_1)\right] \left[-e\Psi^*_{n_2}(\vec{r}_2)\, \Psi_{n_2}(\vec{r}_2)\right]}{4\pi\varepsilon_0 r_{12}} dv_1 dv_2 =$$

$$= \int \frac{\rho_{n_1}(\vec{r}_1)\, \rho_{n_2}(\vec{r}_2)}{4\pi\varepsilon_0 r_{12}} dv_1 dv_2. \qquad (16.36)$$

The quantities $\rho_{n_1}(\vec{r}_1) = -e\left|\Psi_{n_1}(\vec{r}_1)\right|^2$ and $\rho_{n_2}(\vec{r}_2) = -e\left|\Psi_{n_2}(\vec{r}_2)\right|^2$ are the charge densities of electrons in the quantum states $\Psi_{n_1}(\vec{r}_1)$ and $\Psi_{n_2}(\vec{r}_2)$, respectively. Therefore, the integral $K(n_1, n_2)$ is a component of the Coulomb interaction energy between the electrons, which does not take into account their identity and has an analog in classical electrodynamics (see, e.g., [10]). It is called the **Coulomb integral**.

The integral A looks like

$$A(n_1, n_2) = \int \frac{\left[-e\Psi^*_{n_1}(\vec{r}_1)\, \Psi_{n_2}(\vec{r}_1)\right] \left[-e\Psi^*_{n_2}(\vec{r}_2)\, \Psi_{n_1}(\vec{r}_2)\right]}{4\pi\varepsilon_0 r_{12}} dv_1 dv_2 =$$

$$= \int \frac{\rho_{n_1 n_2}(\vec{r}_1)\,\rho_{n_2 n_1}(\vec{r}_2)}{4\pi\varepsilon_0 r_{12}} dv_1 dv_2. \qquad (16.37)$$

It is equal to the energy of Coulomb interaction between the electrons that arises owing to their identity. It is called the **exchange integral**, or the **integral of exchange interaction**. The exchange interaction is a quantum-mechanical phenomenon, which, in contrast to the Coulomb integral K, has no classical analog.

From (16.28) and (16.35), it follows that the energy of helium atom equals

$$E = E_n + E_k + \langle V\rangle = E_n + E_k + K(n_1, n_2) \pm A(n_1, n_2),$$

with the energy for parahelium ($\Psi_a = \Phi_s \chi_a$) being

$$E_s = E_n + E_k + K(n_1, n_2) + A(n_1, n_2)$$

and for orthohelium ($\Psi_a = \Phi_a \chi_s$)

$$E_a = E_n + E_k + K(n_1, n_2) - A(n_1, n_2).$$

The wave functions $\Psi_{n_1}(\vec{r}_1)$ and $\Psi_{n_2}(\vec{r}_2)$ are known; therefore, the integrals K and A can be calculated. The corresponding results show that $K > 0$ and $A > 0$. Therefore, $E_s > E_a$, i.e. the energy levels in parahelium are located higher than the corresponding levels in orthohelium. As a result, we obtain that, when the both electrons are in the state $1s$, it follows from (16.32) that $\Phi_s \neq 0$ and $\Phi_a = 0$, so that the ground state of helium atom is the parahelium state with the energy

$$E_s(1s, 1s) = 2E_1 + K(1_s, 1_s) + A(1_s, 1_s).$$

16.4. One-electron approximation. Periodic table of elements

In the one-electron approximation, electrons in the multielectron atom are regarded to be in individual one-electron states, which are determined by the set of quantum numbers (n, l, m, m_s) (see (15.9)). Despite some shortcomings in this atomic model (see section 16.5), the model remains rather effective for the consideration of a good many issues, in particular, for the substantiation of the periodic table of chemical elements. If we arrange chemical elements in the order when the electric charges of atomic nuclei (the element numbers) grow, we obtain the recurrence of their chemical properties, which is formulated in the form of the **periodic law**:

- The chemical properties of elements periodically depend on their atomic numbers (the nuclear charges).

In this connection, the whole sequence of elements is organized into periods, and the elements are arranged into a table in such a manner that the elements with similar properties should be located in the same table column. As a result, we obtain groups and subgroups (see section 9.2 in [20]). The explanation of the periodic law is a task of quantum mechanics, in particular, quantum chemistry.

Let an atom with Z electrons be in the ground state, i.e. in the state with the minimum energy. Let us consider the distribution of electrons over one-electron quantum states. The energy of electron depends on the quantum numbers n and l (see section 14.5), and its wave function on the quantum numbers n, l, m, and m_s. Therefore, the energy level E_{nl} is $2(2l+1)$-fold degenerate, i.e. $2(2l+1)$ different quantum states correspond to it. Hence, the quantum numbers n and l determine the energy of electron and the set of its quantum states, and will further be referred to as the **state**.

Table 1

n	1	2		3			4				5					6						7						
l	0	0	1	0	1	2	0	1	2	3	0	1	2	3	4	0	1	2	3	4	5	0	1	2	3	4	5	...
nl	1s	2s	2p	3s	3p	3d	4s	4p	4d	4f	5s	5p	5d	5f	5g	6s	6p	6d	6f	6g	6h	*7s*	*7p*	*7d*	*7f*	*7g*	*7h*	...

The sequence of states (n, l) arranged according to the growth of the most probable distance of electron from the nucleus (see section 14.3) is shown in Table 1. At the same time, the energy of states increases in a somewhat different order in connection with the effects of screening and penetration. The **screening effect** consists in that the inner electrons screen, i.e. weaken, the electric field of the nucleus by their electric field. Therefore, the outer electrons undergo the field action of not the total nucleus charge Ze, but a reduced one Z^*e $(Z^* < Z)$, which is called **effective**. The screening effect gives rise to the growth of the energy of outer electrons. The **penetration effect** is associated with the uncertainty in the electron coordinates, the "smear" of electron position in space. As a result, an outer electron can sometimes be located near the nucleus, where its energy is lower. The inner electrons also can reach the atom periphery, but they are coupled with the nucleus more strongly, and the penetration effect for them is weaker.

Table 2

$n+l$	1	2	3	3	4	4	5	5	5	6	6	6	7	7	7	7	8	8
nl	$1s$	$2s$	$2p$	$3s$	$3p$	$4s$	$\sim 3d$	$4p$	$5s$	$\sim 4d$	$5p$	$6s$	$\sim 4f$	$\sim 5d$	$6p$	$7s$	$\sim 5f$	$\sim 6d$
l	0	0	1	0	1	0	2	1	0	2	1	0	3	2	1	0	3	2
N_{max}	2	2	6	2	6	2	10	6	2	10	6	2	14	10	6	2	14	10

Hence, the energy of states (n, l) grows as the sum $n+l$ increases. The corresponding sequence of states is quoted in Table 2, Here, $N_{\max}$ means the maximum number of electrons in the corresponding state ($N_{\max} = 2\,(2l+1)$), and the tilde indicates the close values of state energies.

If an atom is in the ground state, we can easily, using Table 2, determine the distribution of electrons over the states (the **electron formula** or the **electron configuration**). For instance, the electron configuration of for the sodium atom $(Z = 11)$ looks like [Na]:$1s^2 2s^2 2p^6 3s^1$, where the subscripts indicate the number of electrons in the corresponding states.

The set of electrons in the states with the same principal quantum number n is called the **electron shell**. Electron shells are designated using the letters K $(n = 1)$, L $(n = 2)$, M $(n = 3)$, N $(n = 4)$, O $(n = 5)$, P $(n = 6)$, and Q $(n = 7)$. The maximum numbers of electrons in them equal $2n^2$, i.e. 2 (K-shell), 8 (L-shell), 18 (M-shell), and so on. Electrons in states $1s$, $2s$, $2p$, etc. form **subshells**.

The electron formulas for the atoms of almost all chemical elements are listed in Table 3. Here, the underlined element symbols mark the last

Table 3

No.	Atom	Electron formula	No.	Atom	Electron formula
1	H	$1s^1$	35	Br	$4s^2 3d^{10} 4p^5$
2	He	$1s^2$	36	Kr	$4s^2 3d^{10} 4p^6$
3	Li	$1s^2 2s^1$	37	Rb	$[\mathrm{Kr}]5s^1$
4	Be	$1s^2 2s^2$	38	Sr	$5s^2$
5	B	$1s^2 2s^2 2p^1$	39	Y	$5s^2 4d^1$
6	C	$1s^2 2s^2 2p^2$	40	Zr	$5s^2 4d^2$
7	N	$1s^2 2s^2 2p^3$	41	Nb	$5s^1 4d^4$
8	O	$1s^2 2s^2 2p^4$	42	Mo	$5s^1 4d^5$
9	F	$1s^2 2s^2 2p^5$	43	Tc	$5s^2 4d^5$
10	Ne	$1s^2 2s^2 2p^6$	44	Ru	$5s^1 4d^7$
11	Na	$[\mathrm{Ne}]3s^1$	45	Rh	$5s^1 4d^8$
12	Mg	$3s^2$	46	Pd	$4d^{10}$
13	Al	$3s^2 3p^1$	47	Ag	$5s^1 4d^{10}$
14	Si	$3s^2 3p^2$	48	Cd	$5s^2 4d^{10}$
15	P	$3s^2 3p^3$	49	In	$5s^2 4d^{10} 5p^1$
16	S	$3s^2 3p^4$	50	Sn	$5s^2 4d^{10} 5p^2$
17	Cl	$3s^2 3p^5$	51	Sb	$5s^2 4d^{10} 5p^3$
18	Ar	$3s^2 3p^6$	52	Te	$5s^2 4d^{10} 5p^4$
19	K	$[\mathrm{Ar}]4s^1$	53	I	$5s^2 4d^{10} 5p^5$
20	Ca	$4s^2$	54	Xe	$5s^2 4d^{10} 5p^6$
21	Sc	$4s^2 3d^1$	55	Cs	$[\mathrm{Xe}]6s^1$
22	Ti	$4s^2 3d^2$	56	Ba	$6s^2$
23	V	$4s^2 3d^3$	57	La	$6s^2 5d^1$
24	Cr	$4s^1 3d^5$	58	Ce	$6s^2 4f^2$
25	Mn	$4s^2 3d^5$	59	Pr	$6s^2 4f^3$
26	Fe	$4s^2 3d^6$	60	Nd	$6s^2 4f^4$
27	Co	$4s^2 3d^7$	61	Pm	$6s^2 4f^5$
28	Ni	$4s^2 3d^8$	62	Sm	$6s^2 4f^6$
29	Cu	$4s^1 3d^{10}$	63	Eu	$6s^2 4f^7$
30	Zn	$4s^2 3d^{10}$	64	Gd	$6s^2 4f^8$
31	Ga	$4s^2 3d^{10} 4p^1$	65	Tb	$6s^2 4f^9$
32	Ge	$4s^2 3d^{10} 4p^2$	66	Dy	$6s^2 4f^{10}$
33	As	$4s^2 3d^{10} 4p^3$	67	Ho	$6s^2 4f^{11}$
34	Se	$4s^2 3d^{10} 4p^4$	68	Er	$6s^2 4f^{12}$

Table 3 (cont.)

No.	Atom	Electron formula	No.	Atom	Electron formula
69	Tm	$6s^24f^{13}$	86	Rn	$6s^25d^{10}4f^{14}6p^6$
70	Yb	$6s^24f^{14}$	87	Fr	$[\mathrm{Rn}]7s^1$
71	Lu	$6s^25d^14f^{14}$	88	Ra	$7s^2$
72	Hf	$6s^25d^24f^{14}$	89	Ac	$7s^26d^1$
73	Ta	$6s^25d^34f^{14}$	90	Th	$7s^26d^2$
74	W	$6s^25d^44f^{14}$	91	Pa	$7s^26d^15f^2$
75	Re	$6s^25d^54f^{14}$	92	U	$7s^26d^15f^3$
76	Os	$6s^25d^64f^{14}$	93	Np	$7s^25f^5$
77	Ir	$6s^25d^74f^{14}$	94	Pu	$7s^25f^6$
78	Pt	$6s^25d^84f^{14}$	95	Am	$7s^25f^7$
79	Au	$6s^25d^{10}4f^{14}$	96	Cm	$7s^26d^15f^7$
80	Hg	$6s^25d^{10}4f^{14}$	97	Bk	$7s^25f^9$
81	Tl	$6s^25d^{10}4f^{14}6p^1$	98	Cf	$7s^25f^{10}$
82	Pb	$6s^25d^{10}4f^{14}6p^2$	99	Es	$7s^25f^{11}$
83	Bi	$6s^25d^{10}4f^{14}6p^3$	100	Fm	$7s^25f^{12}$
84	Po	$6s^25d^{10}4f^{14}6p^4$	101	Md	$7s^25f^{13}$
85	At	$6s^25d^{10}4f^{14}6p^5$	102	No	$7s^25f^{14}$

elements in the periods, and the brackets designate the electron configurations of the inert gas atoms. One can see that the outer electrons form identical configurations in the atoms of chemical elements with similar chemical properties. For example, atoms of alkali metals (Li, Na, K, ...) have one outer electron in the s-state, each of halogen atoms (F, Cl, Br, ...) has five outer electrons in the p-states, and atoms of inert gases (Ne, Ar, Kr, ..., but He) six electrons in the p-states.

Thus, the periodic law is based on the fact that the order in which outer electrons occupy their atomic states (s, p, ...) is repeated. As the **period**, we call a set of elements arranged according to the growth of their nuclear charges, which have the same value of principal quantum number n for outer electrons; this number is equal to the period number (except for palladium, $Z = 46$). The number of elements in the n-th period is expressed by the formula $2n^2$, i.e. it is equal to the number of electron quantum states.

The periodic table of elements has seven periods, with the last, seventh, period being incomplete. Outer electrons fill ns-states in the atoms of first two elements in every period, and np-states in the atoms of last six elements in complete periods. Those eight elements of each group form eight **main groups** (A). Hence, the main subgroups consist of the elements that are arranged vertically in the periodic table; their atoms have the same number of electrons at the outer level, which is equal to the group number (except for helium). The other elements form the **groups of transition elements** (B).

The first main group (group IA) contains elements, which atoms have one outer electron in the state ns: $1s^1$ (H), $2s^!$ (Li), ..., $7s^1$ (Fr). Atoms of elements in the second main group (group IIA) have two outer electrons in the state ns: $2s^2$ (Be), $3s^2$ (Mg), ..., $7s^2$ (Ra). The third main group (group IIIA) consists of atoms that have three outer electrons in the state ns^2np^1: from boron ($2s^22p^1$) to thallium ($6s^26p^1$). Atoms of the group VIIA have seven electrons in the state ns^2np^5: from fluorine ($2s^22p^5$) to astatine ($6s^26p^5$).

16.5. Multielectron atomic states. Terms

While analyzing the helium atom (see section 16.3), it was found that its state and energy depend on the total electron spin (para- and orthohelium). Therefore, when characterizing the state of multielectron atom as a whole, it is necessary, besides the electron formula, to take into account spins and orbital momenta of electrons that form the total mechanical momentum of the atom (its electron subsystem).

In atoms of most chemical elements atoms, the spins of electrons, owing to the spin-spin interaction, are summed up into the total spin of atom,

$$\vec{L}_s^a = \sum_{i=1}^{z} \vec{L}_{si}, \tag{16.38}$$

and, owing to the interaction between orbital magnetic moments, the total orbital angular momentum of the atom is formed,

$$\vec{L}^a = \sum_{i=1}^{z} \vec{L}_i. \tag{16.39}$$

The momenta $\overrightarrow{L}_s^a$ and $\overrightarrow{L}^a$ compose the total mechanical momentum of the atom (the vector model),

$$\overrightarrow{L}_J = \overrightarrow{L}_s^a + \overrightarrow{L}^a. \tag{16.40}$$

This scheme of formation of the total angular momentum of the atom is called the **normal coupling** or the **LS-coupling**.

At the LS-coupling, the absolute values of vectors (16.38)–(16.40) and their projections are quantized according to the general rules,

$$L_s^a = \sqrt{L_s\,(L_s+1)}\hbar, \quad L_s = \frac{z}{2}, \frac{z}{2} - 1 \quad (L_s \geq 0)\,; \tag{16.41}$$

$$L_{sz}^a = m_{L_s}\hbar, \quad m_{L_s} = L_s, L_s - 1, \ldots, -L_s; \tag{16.42}$$

$$L^a = \sqrt{L\,(L+1)}\hbar, \quad L = \sum_{i=1}^{z} l_i, \sum_{i=1}^{z} l_i - 1 \quad (L \geq 0)\,; \tag{16.43}$$

$$L_z^a = m_L\hbar, \quad m_L = L, L-1, \ldots, -L; \tag{16.44}$$

$$L_J = \sqrt{J\,(J+1)}\hbar, \quad J = L + L_s, L + L_s - 1, \ldots, |L - L_s| \quad (J \geq 0)\,; \tag{16.45}$$

$$L_{Jz} = m_J\hbar, \quad m_J = J, J-1, \ldots, -J. \tag{16.46}$$

Hence, the quantum numbers m_{L_s}, m_L, and m_J can acquire one of $2L_s+1$, $2L+1$, and $2J+1$ values, respectively, and the quantum number J either $2L_s+1$ values if $L > L_s$ or $2L+1$ values if $L_s > L$.

The total angular momentum L_J corresponds to the total magnetic moment

$$\overrightarrow{P}_{mJ} = -\frac{e}{2m_0} g \overrightarrow{L}_J, \tag{16.47}$$

$$P_{mJz} = -\frac{e}{2m_0} g L_{Jz} = -\mu_B g m_J, \tag{16.48}$$

where g is the Landé factor,

$$g = 1 + \frac{J(J+1) - L(L+1) + L_s(L_s+1)}{2J(J+1)}. \tag{16.49}$$

Formulas (16.47)–(16.49) can be derived similarly to the formulas for the total magnetic moment of electron $\overrightarrow{P}_{mj}$ (see section 15.3).

The quantum state of atom is defined by the electron configuration and the quantum numbers L, L_s, and m_{L_s} (or L, L_s, J, and m_J), and its energy by the electron configuration and the quantum numbers L and L_s (neglecting the spin-orbit and spin-spin interaction, as well as relativistic corrections). Hence, the energy level is degenerate with respect to the quantum numbers m_L and m_{L_s} (or J and m_J). Since the quantum numbers m_L and m_{L_s} can acquire $2L+1$ and $2L_s+1$ values, respectively, the state degeneracy equals $(2L+1)(2L_s+1)$. Therefore, the state degeneracy with respect to the quantum numbers J and m_J also equals $(2L+1)(2L_s+1)$.

The set of $(2L+1)(2L_s+1)$ states belonging to a certain electron configuration with the given quantum numbers L and L_s are called the **term**. Terms are marked as

$$^{2L_s+1}L_J,$$

where L_J is the quantum number of the total orbital angular momentum of the atom (see (16.43)), which is denoted with the help of letters S ($L=0$), P ($L=1$), D ($L=2$), F ($L=3$), G ($L=4$), and so on (see Tables 1 and 2). The subscript J equals the quantum number of the total angular momentum (see (16.45)).

The quantity $2L_s+1$ is called the **multiplicity of** the **term** or the **spin multiplicity**. If $2L_s+1=1$, the term is called **singlet**; if $2L_s+1=2$, **doublet**; if $2L_s+1=3$, **triplet**, and so on. If $L > L_s$, the multiplicity is equal to the number of possible values of the quantum number J, i.e. it indicates how many sublevels the energy level becomes split into if the spin-orbit interaction and relativistic corrections are taken into account. If $L < L_s$, the level is split into $2L+1$ sublevels ($2L+1 < 2L_s+1$). In particular, the electron configurations, quantum numbers, and terms for the helium atom are quoted in Table 4.

The ground term of the atom at a given configuration is determined with the help of **Hund's rules**:

1. The term with the maximum L_s value that does not contradict the Pauli principle has the lowest.
2. For a given L_s, the term with the largest L-value has the lowest energy.
3. The quantum number J equals $J=|L-L_s|$ if the electron subshell is half-filled or less, or $J=L+L_s$ otherwise.

Table 4

EC[*]	l_1	l_2	m_{s1}	m_{s2}	L	L_s	J	PH[†]	OH[‡]
$1s^2$	0	0	1/2	–1/2	0	0	0	$^1S_0$	
$1s^1 2s^1$	0	0	1/2	–1/2	0	0	0	$^1S_0$	
	0	0	1/2	1/2	0	1	1		$^3S_1$
$1s^1 2p^1$	0	1	1/2	–1/2	1	0	1	$^1P_1$	
	0	1	1/2	1/2	1	1	2		$^3P_2$
							1		$^3P_1$
							0		$^3P_0$
$1s^1 3s^1$	0	0	1/2	–1/2	0	0	0	$^1S_0$	
	0	0	1/2	1/2	0	1	1		$^3S_1$

[*] EC: electron configuration
[†] PH: parahelium
[‡] OH: orthohelium

As an example, let us determine the ground term of nitrogen atom ($Z = 7$). Its electron formula is [N] :$1s^2 2s^2 2p^3$. The states $1s$ and 2s are filled completely, and the spins of electrons in them are compensated. There can be six electrons in the state $2p$, but the nitrogen atom has only three p-electrons. For their total spin to be maximum, the three p-electrons should be so arranged in accordance with the Pauli principle in the states with $m = 1, 0, -1$ that their spins would be parallel. Then, $L_s = 3 \times 1/2 = 3/2$ and $L = 0$, because the sum of magnetic quantum numbers equals $1 + 0 - 1 = 0$. In this case, the $2p$-subshell is half-filled. Therefore, $J = |L - L_s| = L + L_s = 3/2$, and the ground term of N atom is $^4S_{3/2}$.

17

ELEMENTS OF RADIATION THEORY

17.1. Probability of quantum transition

Light emission is a quantum-mechanical phenomenon. Therefore, only quantum mechanics made it possible to create a consecutive and consistent theory of radiation.

An elementary event of emission consists in that an atom emits a photon. In quantum mechanics, it is considered as a result of the **quantum transition** by the atom from the quantum state $\Psi_k^{(0)}(\vec{r})$ with the energy $E_k^{(0)}$ into the quantum state $\Psi_m^{(0)}(\vec{r})$ with the energy $E_m^{(0)} < E_k^{(0)}$. As a result of this transition, the atom emits a photon with the energy $\hbar\omega = E_k^{(0)} - E_m^{(0)}$.

Any quantum transition, as any other physical phenomenon, is causal: it occurs owing to the interaction of the atom with another system, in particular, with other atoms, an external electromagnetic field, and so forth. Such an interaction is called the **perturbation**.

Let N_k atoms be in the quantum state $\Psi_k^{(0)}(\vec{r})$ before perturbation, and the perturbation act for the time interval τ. Then, each of the atoms has a certain probability $W_{mk}(\tau)$ to make a quantum transition $k \to m$ within this interval and emit a photon with the energy $\hbar\omega = E_k^{(0)} - E_m^{(0)} = \hbar\omega_k - \hbar\omega_m = \hbar\omega_{km}$. The intensity of radiation emission by this system can be expressed by the formula

$$I(\omega_{km}) = N_k \frac{W_{mk}(\tau)}{\tau} \hbar\omega_{km}. \tag{17.1}$$

Hence, the key quantity is the probability of quantum transition, which should be determined first.

In the course of quantum transition, the wave function changes. Therefore, here we must apply the general (time-dependent) Schrödinger equation

$$i\hbar \frac{\partial \Psi(\vec{r}, t)}{\partial t} = \widehat{H} \Psi(\vec{r}, t). \tag{17.2}$$

As a rule, this equation cannot be solved exactly. However, under certain conditions, its approximate analytical solution can be obtained.

Suppose that the Hamiltonian $\widehat{H}$ of the quantum-mechanical system can be written in the form

$$\widehat{H} = \widehat{H}_0(\vec{r}) + \widehat{V}(\vec{r}, t), \tag{17.3}$$

where

$$\widehat{V}(\vec{r}, t) \ll \widehat{H}_0(\vec{r}). \tag{17.4}$$

Then, the operator $\widehat{H}_0(\vec{r})$ is called the **Hamiltonian of unperturbed system** (e.g., the Hamiltonian of an atom), and the operator $\widehat{V}(\vec{r}, t)$ the time-dependent **perturbation operator** (e.g., the operator of interaction between the atom and the electromagnetic field).

Let the perturbation $\widehat{V}(\vec{r}, t)$ begin at the time moment $t = 0$ and terminate at the moment $t = \tau$,

$$\widehat{V}(\vec{r},\ t) = \begin{cases} 0, & t < 0 \text{ and } t > \tau, \\ \widehat{V}(\vec{r}, t), & 0 \leq t \leq \tau. \end{cases} \tag{17.5}$$

Let the wave functions and the energy spectrum of unperturbed system, i.e. the solutions of the equation

$$\widehat{H}_0(\vec{r})\,\Psi_n^{(0)}(\vec{r}) = E_n^{(0)}\Psi_n^{(0)}(\vec{r}), \tag{17.6}$$

be known. Suppose that, before perturbation ($t < 0$), the system was in the state k described by the wave function $\Psi_k^{(0)}(\vec{r})$ and the energy $E_k^{(0)}$. Taking into account (6.3), Schrödinger equation (17.2) takes the form

$$i\hbar\frac{\partial\Psi(\vec{r}, t)}{\partial t} = \left[\widehat{H}_0(\vec{r}) + \widehat{V}(\vec{r}, t)\right]\Psi(\vec{r}, t). \tag{17.7}$$

The wave function $\Psi(\vec{r}, t)$ should be expanded in a series of eigenfunctions of operator $\widehat{H}_0$ taking into account their dependence on the time, i.e. in terms of functions

$$\Psi_n^{(0)}(\vec{r}, t) = \Psi_n^{(0)}(\vec{r})\,\mathrm{e}^{-i\omega_n t}, \quad \omega_n = \frac{1}{\hbar} = E_n^{(0)}. \tag{17.8}$$

The corresponding expansion looks like

$$\Psi(\vec{r}, t) = \sum_n C_n(t)\,\Psi_k^{(0)}(\vec{r})\,\mathrm{e}^{-i\omega_n t}. \tag{17.9}$$

Substituting (17.9) into equation (17.7) and differentiating the result with respect to the time, we obtain

$$\begin{aligned} i\hbar\sum_n\frac{\partial C_n(t)}{\partial t} &= \Psi_k^{(0)}(\vec{r})\,\mathrm{e}^{-i\omega_n t} + i\hbar\sum_n C_n(t)\,\Psi_n^{(0)}(\vec{r})\,(-i\omega_n)\,\mathrm{e}^{-i\omega_n t} \\ &= \widehat{H}_0(\vec{r})\sum_n C_n(t)\,\Psi_n^{(0)}(\vec{r})\,\mathrm{e}^{-i\omega_n t} + \widehat{V}(\vec{r}, t)\sum_n C_n(t)\,\Psi_n^{(0)}(\vec{r})\,\mathrm{e}^{-i\omega_n t}. \end{aligned} \tag{17.10}$$

This equation should be multiplied from the left by the function

$$\Psi_m^{(0)*}(\vec{r},t) = \Psi_m^{(0)*}(\vec{r})\,\mathrm{e}^{-i\omega_m t} \tag{17.11}$$

and integrated over the coordinates, using the following relations and notations:

$$\int \Psi_m^{(0)*}(\vec{r}) = \Psi_n^{(0)}(\vec{r})\,d^3\tau = \delta_{mn}, \quad \hbar\omega_m = E_m^{(0)},$$

$$\widehat{H}_0(\vec{r})\,\Psi_n^{(0)*}(\vec{r}) = E_n^{(0)}\Psi_n^{(0)}(\vec{r}),$$

$$\int \Psi_m^{(0)*}(\vec{r})\,\widehat{V}(\vec{r},t)\,\Psi_n^{(0)}(\vec{r})\,d^3\tau = V_{mn}(t). \tag{17.12}$$

As a result, instead of Schrödinger equation (17.7) in the coordinate representation, we obtain an equivalent equation in the energy representation (see section 7.1 in [20]),

$$i\hbar\frac{\partial C(t)}{\partial t} = \sum_n C_n(t)\,V_{mn}(t)\,\mathrm{e}^{-i\omega_{mn}t}, \quad \omega_{mn} = \omega_m - \omega_n, \tag{17.13}$$

where the wave function is represented by the column matrix of coefficients $C_m(t)$.

The solution of equation (17.13) is sought in the form

$$C_m(t) = \delta_{mk} + C_m^{(1)}(t) + C_m^{(2)}(t) + \ldots . \tag{17.14}$$

Series (17.14) must quickly converge, i.e. the following conditions must be satisfied:

$$C_m^{(1)}(t) \ll 1, \quad C_m^{(2)}(t) \ll C_m^{(1)}(t), \quad \ldots . \tag{17.15}$$

Substituting series (17.14) into equation (17.13) and taking into account that $\frac{\partial}{\partial t}\delta_{mk} = 0$, we obtain the equations

$$i\hbar\frac{\partial C_m^{(1)}(t)}{\partial t} + i\hbar\frac{\partial C_m^{(2)}(t)}{\partial t} + \ldots = \sum_n \delta_{nk}V_{mn}(t)\,\mathrm{e}^{i\omega_{mn}t} +$$

$$+ \sum_n C_n^{(1)}(t)\,V_{mn}(t)\,\mathrm{e}^{i\omega_{mn}t} + \ldots . \tag{17.16}$$

Provided that conditions (17.15) are satisfied, we may confine the series in (17.16) to the summands of the first order of smallness and below. Then, instead of (17.16), we obtain the equation

$$i\hbar\frac{\partial C_m^{(1)}(t)}{\partial t}=\sum_n \delta_{nk}V_{mn}(t)\,\mathrm{e}^{i\omega_{mn}t}=V_{mk}(t)\,\mathrm{e}^{i\omega_{mk}t}. \tag{17.17}$$

Integrating it within the time interval $0 \le t \le \tau$, we find

$$C_m^{(1)}(\tau)=\frac{1}{i\hbar}\int_0^\tau V_{mk}(t)\,\mathrm{e}^{i\omega_{mk}t}dt. \tag{17.18}$$

Hence, on the basis of(17.14) and (17.18), we write

$$C_m(\tau)=\delta_{mk}+\frac{1}{i\hbar}\int_0^\tau V_{mk}(t)\,\mathrm{e}^{i\omega_{mk}t}dt. \tag{17.19}$$

In accordance with (12.50), the squared absolute value of this coefficient equals the probability to find the system in the m-th state. Before perturbation, the system was in the k-th state. Therefore, the probability to reveal it in the m-th state is the probability of quantum transition $k \to m$. Designating this probability as $W_{mk}(\tau)$ and taking into account that $m \neq k$, we obtain

$$W_{mk}(\tau)=|C_m(\tau)|^2=\left|C_m^{(1)}(\tau)\right|^2=\frac{1}{\hbar^2}\left|\int_0^\tau V_{mk}(t)\,\mathrm{e}^{i\omega_{mk}t}dt\right|^2. \tag{17.20}$$

This formula expresses the probability of quantum transition during the time interval τ of perturbation action.

17.2. Electric dipole radiation

Let us consider an important case, when the quantum-mechanical system is an atom interacting with a plane monochromatic electromagnetic wave. The latter creates an electric field with the strength

$$\vec{\mathscr{E}}=\vec{\mathscr{E}}_0\left(\vec{k}\right)\cos\left(\omega t-\vec{k}\vec{r}\right)=\frac{1}{2}\vec{\mathscr{E}}_0\left(\vec{k}\right)\left[\mathrm{e}^{i(\omega t-\vec{k}\vec{r})}+\mathrm{e}^{-i(\omega t-\vec{k}\vec{r})}\right]. \tag{17.21}$$

The atom interacts with the electromagnetic wave due to its electric multipole moments: the electric dipole moment $\vec{d}$, the electric quadrupole moment D_{ij}, and so forth. Confining the consideration to the dipole approximation, the operator of interaction between the atom and the plane electromagnetic wave is written in the form

$$\begin{aligned} \widehat{V}\left(\vec{r},t\right) &= -\vec{d}\vec{\mathscr{E}} = -\vec{d}\vec{\mathscr{E}}_0 \cos\left(\omega t - \vec{k}\vec{r}\right) \\ &= \widehat{v}\left(\vec{r}\right) \mathrm{e}^{i\omega t} + \widehat{v}\left(\vec{r}\right) \mathrm{e}^{-i\omega t}, \end{aligned} \tag{17.22}$$

where

$$\widehat{v}\left(\vec{r}\right) = -\frac{1}{2}\vec{d}\vec{\mathscr{E}}_0\left(\vec{k}\right)\mathrm{e}^{-i\vec{k}\vec{r}}. \tag{17.23}$$

According to (17.12), the matrix element $V_{mk}\left(t\right)$ (t) is calculated by integrating over the coordinates. Therefore,

$$V_{mk}\left(t\right) = v_{mk}\mathrm{e}^{i\omega t} + v_{mk}^{*}\mathrm{e}^{-i\omega t}. \tag{17.24}$$

Substituting (17.24) into (17.20) and calculating the integral, we obtain

$$\begin{aligned} W_{mk}\left(\tau\right) = \frac{1}{\hbar^2}\left| v_{mk}\frac{\mathrm{e}^{i(\omega_{mk}+\omega)\tau}-1}{i\left(\omega_{mk}+\omega\right)} + v_{mk}^{*}\frac{\mathrm{e}^{i(\omega_{mk}-\omega)\tau}-1}{i\left(\omega_{mk}-\omega\right)}\right|^2 = \\ = \frac{2\pi\tau}{\hbar^2}\left|v_{mk}\right|^2 f\left(\omega,\tau\right) + \frac{2\pi\tau}{\hbar^2}\left|v_{mk}\right|^2 f\left(-\omega,\tau\right) + \\ + \frac{4\pi\tau}{\hbar^2}\left[\left(\mathrm{Re}\ v_{mk}^2\right)\cos\omega\tau - \left(\mathrm{Im}\ v_{mk}^2\right)\sin\omega\tau\right]F\left(\omega,\tau\right), \end{aligned} \tag{17.25}$$

where

$$f\left(\pm\omega,\tau\right) = \frac{\tau}{2\pi}\frac{\sin^2\left[\frac{1}{2}\left(\omega_{mk}\pm\omega\right)\tau\right]}{\left[\frac{1}{2}\left(\omega_{mk}\pm\omega\right)\tau\right]^2}. \tag{17.26}$$

The function $f\left(\omega,\tau\right)$ is plotted in Fig. 17.1. It has the maximum value $\frac{\tau}{2\pi}$ if the sine argument equals zero, i.e. at $\omega = -\omega_{mk} = \omega_{km}$. The other maxima are considerably lower and quickly decrease as the frequency ω goes away from ω_{km}. The minima take place at frequencies $\omega_{km} \pm \frac{2\pi}{\tau}$, $\omega_{km} \pm \frac{4\pi}{\tau}$, and so on. The function $f\left(-\omega,\tau\right)$ has a similar plot, but with the main maximum at the frequency $\omega = \omega_{mk} = -\omega_{km}$. Therefore, it is easy to be convinced that the product $f(\omega, \tau) f(-\omega, \tau)$ equals zero. Hence, formula (17.25) takes the form

$$W_{mk}(\tau) = W_{mk}^{(+)}(\tau) + W_{mk}^{(-)}(\tau), \tag{17.27}$$

where

$$W_{mk}^{(\pm)}(\tau) = \frac{2\pi\tau}{\hbar^2} |v_{mk}|^2 f(\pm\omega, \tau). \tag{17.28}$$

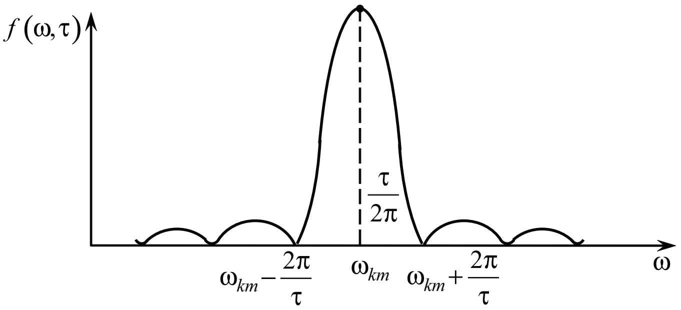

Fig. 17.1.

Now, let us elucidate the meaning of summands in (17.27). The both summands determine the probability of quantum transition $k \to m$ during the time interval τ. However, the summand $W_{mk}^{(+)}(\tau)$ is the probability of quantum transition provided that $\omega = \omega_{km} = \frac{1}{\hbar}\left(E_k^{(0)} - E_m^{(0)}\right)$, i.e. under the condition

$$E_k^{(0)} = E_m^{(0)} + \hbar\omega, \tag{17.29}$$

which corresponds to the emission of a photon with the energy $\hbar\omega$ by the atom. At the same time, the summand $W_{mk}^{(-)}(\tau)$ is the probability of quantum transition provided that $\omega = \omega_{mk} = \frac{1}{\hbar}\left(E_m^{(0)} - E_k^{(0)}\right)$, i.e. under the condition

$$E_m^{(0)} = E_k^{(0)} + \hbar\omega, \tag{17.30}$$

which corresponds to the absorption of a photon.

In what follows, we confine the consideration radiation emissions, i.e. to the summand

$$W_{mk}^{(+)}(\tau) = \frac{2\pi\tau}{\hbar^2} |v_{mk}|^2 f(\omega, \tau). \tag{17.31}$$

According to it, the probability of photon emission per unit time is expressed by the formula

$$w_{mk}^{(+)} = \frac{1}{\tau} W_{mk}^{(\pm)}(\tau) = \frac{2\pi}{\hbar^2} |v_{mk}|^2 f(\omega, \tau). \tag{17.32}$$

Let us consider the matrix element v_{mk} in more details.

According to (17.23), it can be written as

$$v_{mk} = -\frac{1}{2} \vec{\mathscr{E}}_0 \left(\vec{k}\right) \int \Psi_m^{(0)*}(\vec{r}) \left(\vec{d} \mathrm{e}^{-i\vec{k}\vec{r}}\right) \Psi_k^{(0)}(\vec{r})\, d^3\tau. \tag{17.33}$$

Let us estimate the product $\left(\vec{k}\vec{r}\right)$ in the exponential function. In the optical region, $\lambda \approx 5 \times 10^{-7}$ m; therefore, $k = \frac{2\pi}{\lambda} \approx 10^{-7}$ m^{-1}. The coordinate $\vec{r}$ is an argument of the wave functions $\Psi_m^{(0)}(\vec{r})$ and $\Psi_k^{(0)}(\vec{r})$, which differ from zero only in the region, the size of which has an order of atomic dimensions, i.e. 10^{-10} m. Therefore, $r \approx 10^{-10}$ m, and $\left(\vec{k}\vec{r}\right) \approx 10^7\ \mathrm{m}^{-1} \times 10^{-10}\ \mathrm{m} = 10^{-3}$. The smallness of the power in the exponential function means that, while expanding this function in a Taylor series of $\left(-i\vec{k}\vec{r}\right)$,

$$\mathrm{e}^{-i\vec{k}\vec{r}} = 1 - i\vec{k}\vec{r} + \frac{1}{2}\left(-i\vec{k}\vec{r}\right)^2 + \ldots, \tag{17.34}$$

we may confine the expansion to the first term, i.e. $\mathrm{e}^{-i\vec{k}\vec{r}} \approx 1$. Then, matrix element (17.33) looks like

$$v_{mk} = -\frac{1}{2} \vec{\mathscr{E}}_0 \left(\vec{k}\right) \int \Psi_m^{(0)*}(\vec{r})\, \vec{d}\, \Psi_k^{(0)}(\vec{r})\, d^3\tau = -\frac{1}{2} \vec{\mathscr{E}}_0 \left(\vec{k}\right) \vec{d}_{mk}, \tag{17.35}$$

i.e. it is determined by the matrix element of electric dipole moment $\vec{d}_{mk}$, which corresponds to the electric dipole approximation.

Hence, in the electric dipole approximation, the probability per unit time of photon emission by an atom interacting with a plane monochromatic electromagnetic wave is determined, in accordance with (17.32) and (17.35), as

$$w_{mk}^{(+)} = \frac{\pi}{2\hbar^2} \left| \vec{\mathscr{E}}_0 \left(\vec{k}\right) \vec{d}_{mk} \right|^2 f(\omega, \tau). \tag{17.36}$$

Notice that, if the second and third summands should be retained in expansion (17.34), we would obtain the probabilities of magnetic dipole and electric quadrupole radiation emission (see chapter 5).

17.3. Uncertainty relation for energy and time. Natural width of energy levels

From Fig. 17.1, it follows that the function $f(\omega, \tau)$ and, consequently, the probability of photon emission by the atom differs from zero not only at the frequency $\omega = \omega_{mk}$, but also at other frequencies within an interval, the width of which is not less than

$$\Delta\omega = \left(\omega_{km} + \frac{2\pi}{\tau}\right) - \left(\omega_{km} - \frac{2\pi}{\tau}\right) = \frac{4\pi}{\tau}. \tag{17.37}$$

Multiplying (17.37) by $\hbar$ and omitting 4π as an insignificant multiplier, we obtain

$$\Delta(\hbar\omega)\,\tau = \hbar. \tag{17.38}$$

In this relationship, τ is the time interval, within which the atom makes the quantum transition $k \to m$. Therefore, the parameter τ is regarded as the lifetime of the atom in the quantum state $\Psi_k^{(0)}$. In addition, τ has a sense of time uncertainty Δt.

The quantity $\Delta(\hbar\omega)$ is the uncertainty of photon energy, which emerges as a result of energy uncertainty ΔE of the atom in the state $\Psi_k^{(0)}$Then, formula (17.38) can be generalized in the form of **energy–time uncertainty relation**,

$$\Delta E \Delta t = \hbar. \tag{17.39}$$

Hence, in uncertainty relation (17.39), Δt has the meaning of quantum system lifetime in a certain quantum state, and ΔE should be considered as the uncertainty of system energy in this state, i.e. as the width of corresponding energy level.

The lifetime of atom in every quantum state (status) is always limited owing to the radiation emission and absorption processes, so that $\Delta t \neq 0$ and $\Delta E \neq 0$. This means that every quantum-mechanical energy level has an obligatory width ΔE associated with spontaneous quantum transitions, which is called **natural** or **radiation width**,

$$\Delta E = \frac{\hbar}{\Delta t} = \frac{\hbar}{\tau}. \tag{17.40}$$

The natural frequency width of the level is written in the form

$$\Gamma = \frac{1}{\hbar}\Delta E = \frac{1}{\tau}. \tag{17.41}$$

The level width Γ and the lifetime τ have the meaning of average quantum-mechanical values. The natural width of levels is responsible for the natural width of spectral lines, which is also of the order of Γ.

Experimental data testify that a typical lifetime of the atom in the excited state equals $\tau \approx 10^{-8}$ s. Therefore, a typical value of natural width for levels and spectral lines amounts to $\Gamma \approx 10^{8}\ \mathrm{s}^{-1}$. It will be recalled that that the frequencies of spectral lines emitted by atoms equal $\omega \approx 10^{15}\ \mathrm{s}^{-1}$.

Different atomic energy levels have different lifetimes and natural widths. The higher is the level, the larger is the number of different variants of atomic transition onto lower levels, which means the shorter lifetime of the atom on this level and the larger width of the latter. Therefore, the ground energy level of the atom has the smallest width. However, there exist metastable levels for which $\tau \approx 10^{-8}$ s, and $\Gamma \approx 10^{8}\ \mathrm{s}^{-1}$.

The broadening of energy levels and spectral lines is also induced by a number of other factors: the Doppler effect, interatomic interaction, and others. They are responsible for the broadening of atomic levels and spectral lines that considerably exceeds their natural width.

17.4. Selection rules

According to (17.36), the probability of photon emission by the atom differs from zero in the electric dipole approximation if the matrix element of the electric dipole moment of the atom

$$\vec{d}_{mk} = \int \Psi_k^{(0)*}\left(\vec{r}\right) \vec{d}\,\Psi_k^{(0)}\left(\vec{r}\right) d^3\tau \tag{17.42}$$

also differs from zero. To calculate of this integral, it is necessary to know the explicit form of atomic wave functions. The wave functions for an atom with one optical (outer) electron were described in section 14.4,

$$\Psi_{nlm}\left(r,\theta,\varphi\right) = R_{nl}\left(r\right) A_l^{|m|} P_l^{|m|}\left(\cos\theta\right) \mathrm{e}^{im\varphi}. \tag{17.43}$$

The expression for the polynomials $P_l^{|m|}\left(\cos\theta\right)$ and $R_{nl}\left(r\right)$ were given in sections 16.1 and 16.3. Hence, the quantum subscripts m and k in (17.42) correspond to three quantum numbers n, l, and m.

The electric dipole moment of optical electron in the spherical coordinate system is written as

$$\begin{aligned} \vec{d} &= -e\vec{r} = -e\left(\vec{i}x + \vec{j}y + \vec{k}z\right) \\ &= -er\left(\vec{i}\sin\theta\cos\varphi + \vec{j}\sin\theta\sin\varphi + \vec{k}\cos\theta\right), \end{aligned} \tag{17.44}$$

where $\vec{i}$, $\vec{j}$, and $\vec{k}$ are the unit vectors of Cartesian coordinate system. Substituting (17.44) and (17.43) into (17.42), changing m by (n_1, l_1, m_1) and k by (n_1, l_1, m_1), and taking into account that $d^3\tau = r^2 dr \sin\theta d\theta d\varphi$, we obtain

$$\vec{d}_{n_1 l_1 m_1;\, n_2 l_2 m_2} = -eA_{l_1}^{|m_1|} A_{l_2}^{|m_2|} \int dr R_{n_1 l_1}(r)\, r^3 R_{n_2 l_2}(r) \times$$

$$\times \int d\theta d\varphi P_{l_1}^{|m_1|}(\cos\theta)\, \mathrm{e}^{-im_1\varphi} \left(\vec{i}\sin\theta\cos\varphi + \vec{j}\sin\theta\sin\varphi + \vec{k}\cos\theta\right) \times .$$

$$\times P_{l_2}^{|m_2|}(\cos\theta)\, \mathrm{e}^{im_2\varphi} \sin\theta d\theta d\varphi. \tag{17.45}$$

The calculation of integrals in (17.45) is rather a complicated task. Therefore, we present here only the final result, which consists in that matrix element (17.45) differs from zero under the conditions

$$\begin{aligned} \Delta n = n_2 - n_1 &= \Delta n \in Z, \\ l_2 - l_1 &= \Delta l = \pm 1, \\ m_2 - m_1 &= \Delta m = 0, \pm 1, \end{aligned} \tag{17.46}$$

which are called the **selection rules** for the optical electron in the atom.

Selection rules (17.46) make it possible to explain the form of emission spectra produced by atoms with one optical electron. For an example, let us consider the emission spectra of hydrogen and sodium atoms. The energy of hydrogen atom depends only on the principal quantum number n (see (14.48)) associated with several values of quantum numbers l and m. Therefore, for any two energy levels, there always exist such values of numbers l and m that satisfy the selection rules $\Delta l = \pm 1$ and $\Delta m = 0, \pm 1$. Therefore, quantum transitions in the hydrogen atom are allowed between arbitrary levels depicted in Fig. 14.3, and the wave numbers of corresponding spectral lines can be calculated by the **Rydberg formula**

$$\nu^*_{mn} = \frac{1}{\lambda_{mn}} = R\left(\frac{1}{n^2} - \frac{1}{m^2}\right), \tag{17.47}$$

where

$$R = \frac{m_0 e^4}{64\pi^3\hbar^3\varepsilon_0^2 c} = 1.09737 \times 10^7 \text{ m}^{-1}$$

is called the **Rydberg constant**, and m and n are the principal quantum number of the initial and final, respectively, transition states. If $n = 1$ and $m = 2, 3, \ldots$, we obtain the **Lyman spectral series** with its lines in the ultra-violet spectral interval. For $n = 2$ and $m = 3, 4, \ldots$, we obtain the **Balmer series** with four main lines (H_α, H_β, H_γ, and H_δ) in the visible and the others in the ultra-violet spectral range. The values $n = 3$ and $m = 4, 5, \ldots$ correspond to the infra-red **Paschen series**.

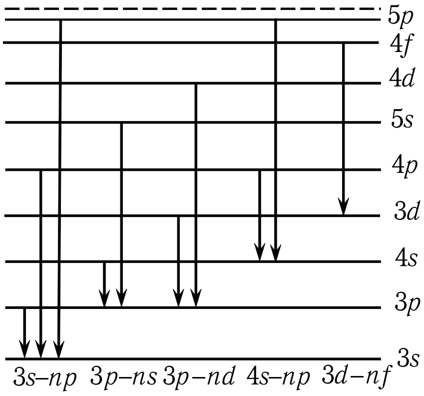

Fig. 17.2.

In the sodium atom, the energy of optical electron is determined by the quantum numbers n and l. The corresponding energy levels are schematically depicted in Fig. 17.2 making no allowance for their fine structure. According to the rule $\Delta l = \pm 1$, the s-levels are connected with the p-levels, the p-levels with the s- and d-levels, the d-levels with the p- and f-levels, and so forth. As a result, the following spectral series are formed:

principal ($3s - np$), sharp ($3p - ns$), diffuse ($3p - nd$), and fundamental ($3d - nf$). The names of those series gave the corresponding designations for the states: s, p, d, f, ...

Now, let us find the selection rule for a linear harmonic oscillator. Taking into account relationship (2.73), the matrix element of the electric dipole moment of oscillator with the mass m and the electric charge e looks like

$$d_{nk} = (ex)_{nk} = e \int \Psi_n^* (x) \, x \Psi_k (x) \, dx = e \sqrt{\frac{\hbar}{m\omega}} \int \Psi_n^* (\xi) \, \xi \Psi_k (\xi) \, d\xi =$$

$$= e \sqrt{\frac{\hbar}{m\omega}} \int \Psi_n^* (\xi) \left[\sqrt{\frac{k}{2}} \Psi_{k-1} (\xi) + \sqrt{\frac{k+1}{2}} \Psi_{k+1} (\xi) \right] d\xi =$$

$$= e \sqrt{\frac{\hbar}{m\omega}} \left[\sqrt{\frac{k}{2}} \delta_{n,\,k-1} + \sqrt{\frac{k+1}{2}} \delta_{n,\,k+1} \right], \tag{17.48}$$

where n and k are the quantum numbers of oscillator. Therefore, electric dipole quantum transitions of oscillator are possible at $n = k \pm 1$, i.e. between the neighbor levels, so that the selection rule is written as $\Delta n = \pm 1$.

17.5. Spontaneous and stimulated emission of radiation

Formula (17.36) gives the probability of photon emission by the atom per unit time as a result of atomic quantum transition from the state $\Psi_k^{(0)}$ into the state $\Psi_m^{(0)}$. The emitted photon has the frequency $\omega = \frac{1}{\hbar} \left(E_k^{(0)} - E_m^{(0)} \right)$, and its wave vector has a certain orientation. However, plenty of different quantum states of photon correspond to the same frequency ω. This means the availability of a good many emission channels (possibilities) and, hence, the increase of emission probability. Therefore, in order to find the total probability of photon emission by the atom, probability (17.36) has to be multiplied by the density of photon quantum states, and the result to be integrated over all possible frequencies and wave vector directions.

The number of photon quantum states in the frequency interval $d\omega$ and with the wave vector direction in the solid angle $d\Omega$ is determined by the formula (see section 11.5 in [20])

$$dg = \frac{\omega^2 V d\omega d\Omega}{(2\pi c)^3}. \tag{17.49}$$

Multiplying (17.36) by (17.49), we obtain the probability per unit time that the atom emits a photon with a definite polarization, with a frequency in the interval from ω to $\omega + d\omega$, and a wave vector directed into the solid angle $d\Omega$,

$$d\omega_{mk} = \omega_{mk}^{(+)} dg = \frac{\omega^2 V}{16\pi^2 \hbar^2 c^3} \left| \mathscr{E}_0 \left(\vec{k} \right) \vec{d}_{mk} \right|^2 f(\omega, \tau)\, d\omega d\Omega. \tag{17.50}$$

The vector $\vec{\mathscr{E}}_0 \left(\vec{k} \right)$ can be written in the form $\vec{\mathscr{E}}_0 \left(\vec{k} \right) = \mathscr{E}_0 (k)\, \vec{u} \left(\vec{k} \right)$, where $\vec{u} \left(\vec{k} \right)$ is the polarization unit vector. Therefore,

$$\left| \mathscr{E}_0 \left(\vec{k} \right) \vec{d}_{mk} \right|^2 = \mathscr{E}_0^2 \left(\vec{k} \right) \left| \vec{u} \left(\vec{k} \right) d_{mk} \right|^2 = \mathscr{E}_0^2 \left(\vec{k} \right) \left| \vec{d}_{mk} \right|^2 \cos^2 \alpha, \tag{17.51}$$

where α is the angle between the vectors $\vec{u} \left(\vec{k} \right)$ and $\vec{d}_{mk}$. Substituting (17.51) into (17.50), we obtain

$$d\omega_{mk} = \frac{\omega^2 V \mathscr{E}_0^2 \left(\vec{k} \right)}{16\pi^2 \hbar^2 c^3} \left| \vec{d}_{mk} \right|^2 \cos^2 \alpha\, f(\omega, \tau)\, d\omega d\Omega. \tag{17.52}$$

Let the function $N \left(\omega, \vec{k} \right)$ denotes the number of photons in the state $\left(\omega, \vec{k} \right)$ that interact with the atom. We should also take into account that the atom interacts with the emitted photon (the radiation reaction). Then, the density of electromagnetic field energy is

$$\rho(\omega, k) = \frac{1}{V} \left[N \left(\omega, \vec{k} \right) + 1 \right] \hbar\omega. \tag{17.53}$$

In electrodynamics, this quantity can be expressed in the form

$$\rho \left(\omega, \vec{k} \right) = \frac{1}{2} \varepsilon_0 \mathscr{E}_0^2 \left(\vec{k} \right), \tag{17.54}$$

where ε_0 is the dielectric constant of vacuum. From (17.53) and (17.54), we obtain the formula

$$\mathscr{E}_0^2\left(\vec{k}\right) = \frac{2\left[N\left(\omega, \vec{k}\right) + 1\right]}{\varepsilon_0 V}\hbar\omega, \tag{17.55}$$

and substituting it into (17.52), we have

$$d\omega_{mk} = \frac{\left[N\left(\omega, \vec{k}\right) + 1\right]}{8\pi^2\varepsilon_0\hbar c^3}\left|\vec{d}_{mk}\right|^2 \cos^2\alpha\, f\left(\omega, \vec{k}\right) d\omega d\Omega. \tag{17.56}$$

If the external electromagnetic field is equilibrium and, hence, isotropic, then, $N\left(\omega, \vec{k}\right) = \overrightarrow{N}(\omega)$ and expression (17.56) reads

$$d\omega_{mk} = \frac{\left(\overline{N}(\omega) + 1\right)\omega^3}{8\pi^2\varepsilon_0\hbar c^3}\left|\vec{d}_{mk}\right|^2 \cos^2\alpha\, f(\omega, \tau)\, d\omega d\Omega. \tag{17.57}$$

Since $d\Omega = \sin\theta d\theta d\varphi$, this formula can be integrated over the angles θ and φ. Additionally, since the radiation polarization is arbitrary, the relative orientation of vectors $\vec{k}$, $\vec{u}\left(\vec{k}\right)$, and $\vec{d}_{mk}$ can be illustrated by Fig. 17.3, whence it follows that $\cos\alpha = \sin\theta$. Then,

$$d\omega_{mk} = \frac{\left[\overline{N}(\omega) + 1\right]}{8\pi^2\varepsilon_0\hbar c^3}\left|\vec{d}_{mk}\right|^2 \sin^3\theta f(\omega, \tau)\, d\omega d\theta d\varphi. \tag{17.58}$$

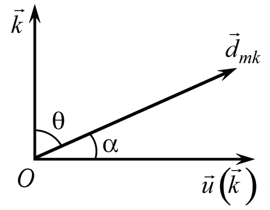

Fig. 17.3.

Integrating expression (17.58) over the angles θ and φ, we find

$$dW_{mk} = \frac{\left(\overline{N}(\omega) + 1\right)\omega^3}{3\pi\varepsilon_0\hbar c^3}\left|\vec{d}_{mk}\right|^2 f(\omega, \tau)\, d\omega. \tag{17.59}$$

This formula describes the probability per unit time (1 s) that the atom emits a photon with the frequency ω within the frequency interval $d\omega$ irrespective of the photon wave vector direction.

Now, let us consider the plot of function $f(\omega, \tau)$ exhibited in Fig. 17.1. In the optical region, $\omega_{km} \approx 10^{15}$ s^{-1}. Therefore, if $\frac{2\pi}{\tau} \ll \omega_{km}$, i.e. if

$\tau \gg \frac{2\pi}{\omega_{km}} \approx 10^{-15}$ s, we may formally put $\tau \to \infty$, and the function $f(\omega, \tau)$ becomes one of the representations of the Dirac delta-function,

$$\lim_{\tau\to\infty} f(\omega,\ \tau) = \lim_{\tau\to\infty} \left\{ \frac{\tau}{2\pi} \frac{\sin^2\left[\frac{1}{2}(\omega_{mk}+\omega)\tau\right]}{\left[\frac{1}{2}\cdot\omega_{mk}+\tau\right]^2} \right\} = \delta(\omega_{mk}+\omega)\,. \tag{17.60}$$

Substituting the function $f(\omega, \tau)$ in (17.59) by $\delta(\omega_{mk}+\omega)$, integrating the result over ω, and taking into account that

$$\int_{-\infty}^{\infty} F(x)\,\delta(x-x_0)\,dx = F(x_0)\,, \tag{17.61}$$

we finally obtain

$$W_{mk} = \frac{\left[\overline{N}(\omega_{km})+1\right]\omega_{km}^3}{3\pi\varepsilon_0\hbar c^3}\left|\vec{d}_{mk}\right|^2. \tag{17.62}$$

Formula (17.62) describes the total probability per unit time of photon emission by the atom in the electric dipole approximation. It can be rewritten as a sum of two terms,

$$W_{mk} = W_{mk}^{\text{stim}} + W_{mk}^{\text{spont}}, \tag{17.63}$$

where

$$W_{mk}^{\text{stim}} = \frac{\overline{N}(\omega_{km})\,\omega_{mk}^3}{3\pi\varepsilon_0\hbar c^3}\left|\overleftarrow{d}_{mk}\right|^2,$$

$$W_{mk}^{\text{spont}} = \frac{\omega_{km}^3}{3\pi\varepsilon_0\hbar c^3}\left|\vec{d}_{mk}\right|^2.$$

The summand W_{mk}^{stim} is the probability of stimulated (forced) photon emission during 1 s under the action of external photons interacting with the atom. It is easy to prove (see section 11.5 in [20]) that

$$W_{mk}^{\text{stim}} = B_{mk}\rho(\omega_{mk})\,, \quad B_{mk} = \frac{\pi\left|\vec{d}_{mk}\right|^2}{3\varepsilon_0\hbar^2}, \tag{17.64}$$

where

$$\rho(\omega) = \frac{\overline{N}(\omega)\,\hbar\omega^3}{\pi^2 c^3}$$

is the energy density (per unit volume) of the external photon field, and the coefficient B_{mk} is called the **Einstein coefficient for stimulated (induced) emission**.

The summand W_{mk}^{spont} differs from zero even in the absence of external electromagnetic field. It determines the probability of spontaneous (not forced) emission. It is called the **Einstein coefficient for spontaneous emission**), which is a result of interaction between the atom and vacuum, i.e. zero-oscillations of the electromagnetic field.

EXERCISES TO PART II

1. Calculate the de Broglie wavelength for electrons with a kinetic energy of (a) 100 eV and (b) 1 MeV.

   - Solution: (a) $\lambda = 0.123$ nm, (b) $\lambda = 87.3$ pm.

2. Evaluate the velocity of electron in the hydrogen atom taking into account that the radius of the first Bohr orbit $r_1 = 0.053$ nm.

   - Solution: $v \approx 10^6$ m/s.

3. Find the result of the action of the operator $\left(\frac{d}{dx}\mathrm{e}^{2x}\right)$ on the function sin x.

   - Solution: $\left(\frac{d}{dx}\mathrm{e}^{2x}\right)\,\mathrm{sin}x = \mathrm{e}^{4x}\,(7\mathrm{sin}x + 6\cos x)$.

4. Find the square of the operator $\frac{d}{dx} + x$.

   - Solution: $\left(\frac{d}{dx} + x\right)^2 = \frac{d^2}{dx^2} + 2x\frac{d}{dx} + x^2 + 1$.

5. Prove the self-adjointness of the operator $\hat{p}_x = -ih\frac{\partial}{\partial x}$.

6. Find the commutator of the operator $\hat{x}$ and the Laplace operator Δ.

   - Solution: $\widehat{x}\Delta - \Delta\widehat{x} = -2\frac{\partial}{\partial x}$.

7. Find the eigenvalue of the operator $-\frac{d^2}{dx^2} + x^2$, which corresponds to the eigenfunction $xe^{-\frac{1}{2}x^2}$.

   - Solution: $\lambda = 3$.

8. Show that, in stationary states, the average values of physical quantities and the probabilities of their allowed values do not depend on time.

9. With the help of Schrödinger equations, prove that the condition of wave function normalization $\dfrac{d}{dt}\int \psi^*\psi d\tau = 0$ is preserved.

10. Find, which of the following quantities—the energy E; the momentum projections p_x, p_y, and p_z; the angular momentum projections L_x, L_y, and L_z; and the square of angular momentum L_2—remain constant for a free moving particle.

11. Give a quantum-mechanical substantiation for the first Bohr postulate.

12. A particle moves in a one-dimensional potential well with absolutely opaque walls. What is the probability to find it in the interval $\frac{1}{4}l \leq x \leq \frac{1}{2}l$, if the particle is in the ground state?

    - Solution: $\omega = 0.41$.

13. Find the wave function of a linear harmonic oscillator with the energy $E = \frac{3}{2}\hbar\omega$.

    - Solution: $\psi_1(\xi) = \sqrt{\dfrac{2}{\sqrt{\pi}}}\xi \mathrm{e}^{-\dfrac{1}{2}\xi^2}$, $\xi = \sqrt{\dfrac{m\omega}{\hbar}}x$.

14. Calculate the transmission factors for a rectangular potential barrier with the height $U = 15$ eV and the width $a = 100$ pm for an electron and a proton, both with a kinetic energy of 10 eV.

    - Solution: $D_e = 0.1$, $D_p = 1.7 \times 10^{-43}$.

15. Find the most probable distance between the electron and the nucleus in the hydrogen atom and other similar ions if the electron is in the ground state.

    - Solution:$r_H = \dfrac{4\pi\varepsilon_0\hbar^2}{m_0 z e^2}$.

16. Find the average distance between the electron and the nucleus in the hydrogen atom if the electron is in the ground state.

    - Solution: $\langle r \rangle = \dfrac{3}{2} r_1$, $r_1 = \dfrac{4\pi\varepsilon_0 \hbar^2}{m_0 e^2}$.

17. Find the average value of the electrostatic potential created by the $1s$-electron at the center of hydrogen atom.

    - Solution: $\varphi = -\dfrac{e}{4\pi\varepsilon_0 r_1}$.

18. Find the ionization energy of helium ion He^+ in the $3s$-state.

    - Solution: $E_i = 6.04$ eV.

19. Prove that the interval between the neighbor energy levels in the hydrogen atom decreases as the principal quantum number n grows.

20. Find the splitting magnitude for a beam of silver atoms after passing through a magnetic field in the Stern–Gerlach experiment if the gradient of magnetic field is $dB/dz = 100$ T/m, the velocity of atoms $v = 200$ m/s, and their path length in the magnetic field $l = 15$ cm.

    - Solution: $2\Delta z = 2{,}9 \times 10^{-3}$ m.

21. According to the Uhlenbeck–Goudsmit hypothesis (1925), the spin of electron should be considered as a result of electron rotation around it axis, and the electron itself as a ball of the radius r_0. Taking the electron spin equal to $\hbar/2$ and its classical radius to $r_0 = 2.8 \times 10^{-15}$ m, find the linear velocity of points at the electron's "equator" and prove that this model of electron spin is inconsistent.

    - Solution: $v = \dfrac{5\hbar}{4 m_0 r_0} = 5 \times 10^{10}$ m/s $\gg 3 \times 10^8$ m/s.

22. Considering the Pauli matrices $\widehat{\sigma}_x$, $\widehat{\sigma}_y$, and $\widehat{\sigma}_z$ as the components of vector operator $\widehat{\vec{\sigma}}$, find the scalar, $\widehat{\vec{\sigma}} \cdot \widehat{\vec{\sigma}}$, and vector, $\left[\widehat{\vec{\sigma}}, \widehat{\vec{\sigma}}\right]$, products.

- Solution: $\widehat{\vec{\sigma}} \cdot \widehat{\vec{\sigma}} = 3\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $\left[\widehat{\vec{\sigma}}, \widehat{\vec{\sigma}}\right] = 2i\widehat{\vec{\sigma}}$.

23. Find the eigenvalues and eigenfunctions of the Pauli operators (matrices) $\widehat{\sigma}_x$, $\widehat{\sigma}_y$, and $\widehat{\sigma}_z$.

    - Solution: $\sigma_x = \pm 1$, $\chi_1^{(1)} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\chi_{-1}^{(1)} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$; $\sigma_y = \pm 1$, $\chi_1^{(2)} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \end{pmatrix}$, $\chi_{-1}^{(2)} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -i \end{pmatrix}$; $\sigma_z = \pm 1$, $\chi_1^{(3)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\chi_{-1}^{(3)} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

24. Find the possible values of magnetic moment and its projections for an atomic electron in the nd-state in terms of Bohr magnetons.

    - Solution: $P_{mj_1} = 2\sqrt{\frac{3}{5}}\mu_B$, $P_{mj_1z} = \pm\frac{6}{5}\mu_B$, $\pm\frac{2}{5}\mu_B$; $P_{mj_2} = 3\sqrt{\frac{7}{5}}\mu_B$, $P_{mj_2z} = \pm 3\mu_B$, $\pm\frac{9}{5}\mu_B$, $\pm\frac{3}{5}\mu_B$.

25. Find the wavelength difference between the fine structure components of the first line in the spectral Lyman series ($\lambda = 121.6$ nm).

    - Solution: $\Delta\lambda = -\frac{E_2\alpha^2\lambda^2}{8\pi\hbar c} = 5.6 \times 10^{-13}$ m.

26. For a system of two identical particles with the spins $L_S = 1$, write down the spin functions that are symmetric with respect to the swapping of spin variables.

    - Solution: $\chi_S^{(1)} = \chi_1(1)\chi_1(2)$, $\chi_S^{(3)} = \chi_{-1}(1)\chi_{-1}(2)$,
      $\chi_S^{(4)} = \frac{1}{\sqrt{2}}\{\chi_0(1)\chi_1(2) + \chi_0(2)\chi_1(1)\}$,
      $\chi_S^{(5)} = \frac{1}{\sqrt{2}}\left\{\chi_0(1)\chi_{-1}(2) + \chi_0(2)\chi_{-1}(1)\right\}$, and
      $\chi_S^{(6)} = \frac{1}{\sqrt{2}}\left\{\chi_1(1)\chi_{-1}(2) + \chi_1(2)\chi_{-1}(1)\right\}$, where the subscripts 1 and 2 mean L_{S_z1} and L_{S_z2}, respectively.

27. For a system of two identical particles with the spins $L_S = 1$, write down the spin functions that are anti-symmetric with respect to the swapping of spin variables.

    - Solution: $\chi_0^{(1)} = \frac{1}{\sqrt{2}} \{\chi_0(1)\chi_1(2) - \chi_0(2)\chi_1(1)\}$,
      $\chi_0^{(2)} = \frac{1}{\sqrt{2}} \left\{\chi_0(1)\chi_{-1}(2) - \chi_0(2)\chi_{-1}(1)\right\}$, and
      $\chi_0^{(3)} = \frac{1}{\sqrt{2}} \left\{\chi_1(1)\chi_{-1}(2) - \chi_1(2)\chi_{-1}(1)\right\}$, where
      the subscripts 1 and 2 mean $L_{S_z 1}$ and $L_{S_z 2}$, respectively.

28. For a system of two identical bosons with the spins $l_S = 0$ and described by the normalized wave function $\psi(\vec{r}_1, \vec{r}_2)$, find the probabilities that (a) either of the particles is in the volume dV_1, and the other in the volume dV_2, and (b) the both particles are in the same volume V.

    - Solution: (a) $d\omega = 2\left|\psi(\vec{r}_1,\ \vec{r}_2)\right|^2 dV_1 dV_2$,
      (b) $d\omega = \int_V \int_V \psi(\vec{r}_1, \vec{r}_2)\, dV_1 dV_2$.

29. What are the electron configurations of iron ($z = 26$), cobalt ($z = 27$), and nickel ($z = 28$) atoms? Which electrons of those atoms are outer?

    - Solution: [Fe]: [Ar]$4s^2 3d^6$, [Co]: [Ar]$4s^2 3d^7$, [Ni]: [Ar]$4s^2 3d^8$. The $4s$-electrons are outer ones.

30. How many components is the term ${}^4D_{1/2}$ split into if the spin-orbit interaction and relativistic corrections are taken into account?

    - Solution: $f\left({}^4D_{1/2}\right) = 4$.

31. Find the magnetic moment of an atom in the state ${}^4D_{1/2}$.

    - Solution: $P_{mJ} = \mu_B g \sqrt{J(J+1)} = 0$.

32. With the help of Hund's rule, find the ground term for an atom with the electron configuration nd^7 above its filled subshells.

    - Solution: ${}^4F_{9/2}$.

33. Write down the term, for which $L_s = 1/2$, $L = 1$, and $g = 4/3$.

    - Solution: ${}^2P_{3/2}$.

34. Find the magnetic moment and its projections for an atom in the state ${}^3D_2$.

    - Solution: $P_{mJ} = \mu_B g \sqrt{J(J+1)} = \dfrac{7}{\sqrt{6}}\mu_B$,
      $P_{mJz} = 0,\ \pm\dfrac{7}{\sqrt{6}}\mu_B,\ \pm\dfrac{14}{\sqrt{6}}\mu_B$.

35. The natural width of a certain spectral line equals $\Delta\omega = 10^8$ s^{-1}. Find the uncertainty of photon coordinate in the direction of photon motion.

    - Solution: $\Delta x = \dfrac{c}{\Delta\omega} = 3$ m.

PART III

STATISTICAL PHYSICS AND THERMODYNAMICS

Statistical physics and thermodynamics study macroscopic systems with a huge number of particles—molecules, atoms, ions, electrons, photons, and so forth. Such systems include ordinary gases, liquids, solids, and free electrons in metal.

A huge number of particles give rise to the existence of macroscopic parameters, which characterize the system as a whole, but are inapplicable to separate particles. Such parameters are, for example, pressure, volume, and temperature. The calculation of macroparameters and the determination of relations between them are the main task of both statistical physics and thermodynamics, although the statistical and thermodynamic approaches differ from each other.

The thermodynamic approach is not based on any model representation concerning the atomic-molecular structure of the matter. Instead, it is based on experimental (phenomenological) relations between macroscopic (measured) quantities, as well as fundamental laws established by generalizing experimental facts. Therefore, the methods of phenomenological thermodynamics are universal and applicable to a very considerable number of macroscopic systems: from ordinary gases, liquids, and solids to the objects in the Universe. This is a substantial advantage of thermodynamics.

At the same time, phenomenological thermodynamics has a substantial fault, because it cannot reveal internal mechanisms of the phenomena and processes. Therefore, some of them, in particular, fluctuation phenomena are beyond its scope. Such phenomena are explained with the help of statistical physics.

The statistical approach makes allowance for the discrete structure of the matter. It is based on corresponding model ideas and formulates the behavior laws for a macrosystem on the basis of behavior laws for its particles. This approach makes it possible to understand the mechanisms of physical phenomena and processes in macrosystems, to substantiate the laws of thermodynamics, and to solve the problems which cannot be solved in the framework of phenomenological thermodynamics.

At the same time, it should be noted that conclusions drawn in statistical physics are valid only provided that the assumptions made about the behavior of system particles are correct. Hence, each of the approach has the corresponding advantages and shortcomings. Therefore, the most fruitful is a combination of both approaches, i.e. the formulation of thermodynamics on the statistical basis. This combined approach is called **statistical thermodynamics**.

The phenomenological thermodynamics was founded in the work of French engineer S. Carnot (1824). This work contained the main principles of the second law of thermodynamics. The modern formulation of this law belongs to the German physicist R. Clausius (1850) and the English physicist W. Thomson (lord Kelvin) (1851). During the forties and fifties of the 19th century, the German scientists R. Mayer and H. von Helmholtz together with the English physicist J.P. Joule, in a series of their works, established the law of energy conservation and transformation, which is called the **first law of thermodynamics**. In 1865, R. Clausius introduced one of the most important concepts of thermodynamics—entropy. In the middle of the 19th century, the concept of “absolute temperature” was formulated for the first time (Kelvin, 1848). In 1906, the German physical chemist and physicist W. Nernst formulated the heat theorem (the Nernst theorem), which is also called the **third law of thermodynamics**.

The origins of statistical physics were associated with the establishment of the atomic-molecular structure of the matter. In this connection, notice that the definitions of the molecule as the smallest particle of the substance and the atom as the smallest particle of the chemical element was adopted at the congress of chemists in 1860. The ultimate evidence for the existence of molecules was given by experiments of the French physicist J. Perrin on the Brownian motion. The molecular-kinetic theory of gases (the Maxwell distribution, 1859; the Boltzmann principle, 1872) became the first stage in the formation of statistical physics. The creation of classical statistical physics was finished by the works of the American physicist J.W. Gibbs, who developed the most general and consistent of its formulations (the microcanonical and canonical Gibbs distributions, 1902).

The discovery of quantum properties in microparticles required the creation of statistics for the systems of particles, to which the concepts of coordinate and momentum are inapplicable, i.e. the creation of quantum statistics. This task was fulfilled in the works by A. Einstein and the Indian physicist S.N. Bose, who developed the quantum statistics for systems of particles with integer spins (the Bose–Einstein distribution, 1924), and in works by the Italian physicist E. Fermi and the English physicist P. Dirac, who formulated the quantum statistics for particles with half-integer spins (the Fermi–Dirac distribution, 1925). The quantum statistics made it possible to explain the properties of the free electrons in metals, the photon gas of equilibrium radiation, the gas of photons and excitons in crystals, and so forth.

A sound contribution to the development of statistical physics and thermodynamics was made by Ukrainian scientists; in particular, these were M.M. Bogolyubov, who obtained important results in the quantum statistics, L.D. Landau, who developed the theory for the phase transitions of the second kind, and others.

18

BASIC CONCEPTS OF THERMODYNAMICS

18.1. Thermodynamic system and thermodynamic parameters

Macroscopic systems in thermodynamics are called **thermodynamic systems**, and those quantities that characterize them in whole, **macroscopic parameters** (volume, pressure, density, and so forth). The latter, in turn, are divided into external and internal parameters.

External macroparameters are determined by the position and the action of objects that are external with respect to the system. A typical example is the gas volume, which is determined by the vessel walls. Below, the external parameters will be denoted as a_i $(i = 1, 2, \ldots)$. **Internal parameters** are determined by the properties of the system, i.e. by the motion and distribution of particles (pressure, density, magnetization, and so forth).They will be denoted as b_j $(j = 1, 2, \dots)$.

The set of independent macroparameters is called the **state of thermodynamic system**. Those quantities that are determined by the system state at a definite time moment and do not depend on the system states at earlier time moments are called the **state functions**.

If the macroscopic parameters do not change in time, the corresponding state is called **stationary**. If, additionally, there are no fluxes (of energy, mass, and so forth) in the system, the state is called **equilibrium** or, in full, the **state of thermodynamic equilibrium**. If the system is in the state of thermodynamic equilibrium, its macroscopic parameters are called the **thermodynamic parameters**.

If the thermodynamic system is in the equilibrium state, its separate macroscopic parts are also in the equilibrium state.

Thermodynamics mainly studies equilibrium thermodynamic systems. Its main task consists in calculating the thermodynamic parameters and relations between them.

Systems that do not exchange energy and particles with external bodies are called **closed**. On the basis of experimental facts in the following **first postulate of thermodynamics** was formulated:

- any closed system always tends in time towards an equilibrium state and, if being in this state, never leaves it spontaneously.

The time required for a closed system to transit from the nonequilibrium state to the equilibrium one is called the **relaxation time**. Since this transition has an asymptotic character, the time interval, after which the deviation of the corresponding parameter from its equilibrium value decreases by a factor of e, is taken as the relaxation time τ.

18.2. Internal energy. Heat and work in thermodynamics

One of the main physical characteristics of any system is its energy. It has a fundamental value in thermodynamics as well.

The total energy of the system is divided into the external and internal energies. The **external energy** includes the energy of system motion as a whole and its potential energy in the field of external forces. All other components comprise the **internal energy**. However, thermodynamics does not consider the system motion as a whole and the variation of system potential energy in external fields. Therefore, in thermodynamics, the energy of thermodynamic system is reduced to its internal energy.

The phenomenological thermodynamics does not explain the nature of internal energy. It only asserts that that this energy does exist. However, it should be noticed that the internal energy includes the energies of all forms of particle motions and interactions: the energies of translational, rotational, and vibrational motions of molecules and atoms, the energy of intermolecular interaction, the intratomic and intranuclear energies, and so on.

At the same time, the internal energy itself has no meaning in thermodynamics, but its variation owing to the changes in the system state does. Therefore, only the varying components of internal energy are taken into account The most often, these are the energies of translational, rotational, and vibrational motions of molecules, the energy of atomic vibrations in molecules, and the energy of intermolecular interaction. In what follows, the internal energy will be denoted as U.

The internal energy in thermodynamics is an additive quantity and can change due to the interaction of the system with external objects. There are two ways to transfer the energy from the system to external objects in the opposite direction. The first way is associated with the variation of external parameters, e.g., the volume. In this case, the energy transferred by the system to external objects is called the **work** A executed by the system. In the second case, the energy obtained by the system from external objects without the variation of external parameters is called the **heat** Q obtained by the system, and the process of energy transfer, the **heat exchange**.

The work A done by the system and the heat Q obtained by it are considered to be positive. If external objects do the work A on the system, then $A < 0$; and if the system gives up a certain amount of heat, then, $Q < 0$ again. One should distinguish between the term “heat” understood as the form of chaotic (thermal) motion of particles (atoms, molecules, electrons, and so forth) in the system and the same term understood as the quantity of energy obtained or given up by the system in the course of heat exchange.

In mechanics, the work done by the force $\overrightarrow{F}$ to shift an object by an infinitesimal vector $d\vec{s}$ is calculated using the formula

$$\delta A = \overrightarrow{F} d\vec{s} = \sum_{i=1}^{3} F_i ds_i.$$

In thermodynamics, the work δA done by the system at infinitesimal variations of external parameters a_i is analogously defined in the form

$$\delta A = \sum_i \Lambda_i da_i, \tag{18.1}$$

where $\Lambda_i = \dfrac{\partial A}{\partial a_i}$ is the generalized force conjugate with respect to the parameter a_i.

In the simplest case, when the thermodynamic system is a gas, the external parameter is the volume V of the gas, and the generalized force is the gas pressure p, the elementary work done by the gas equals

$$\delta A = pdV. \tag{18.2}$$

The work done by the forces of surface tension in liquids when the area of liquid surface increases by $d\Sigma$ is found by the formula

$$\delta A = -\sigma d\Sigma, \tag{18.3}$$

where σ is the surface tension coefficient; i.e. here, $a = \Sigma$ and $\Lambda = -\sigma$. In this case, the work done by the liquid is negative, and the work done by external forces positive.

The work is not a state function, but a process function. Therefore, in the general case, $\delta A \neq dA$. The method to calculation the heat Q will be considered in section 19.2.

18.3. Temperature. Equation of state

The temperature is the main characteristic of equilibrium state in thermodynamics, The statement about its existence is called the **zeroth law**. The temperature is a special parameter that belongs neither to external parameters nor internal ones.

The temperature concept is introduced as follows. Let us consider two closed thermodynamic systems with the internal energies U_1 and U_2. Each of the systems is the equilibrium state. Let us put those systems in the state of heat exchange, i.e. we allow them to exchange their internal energies provided that the external parameters are fixed. Then, we may consider them as subsystems of a single closed system. Irrespective of the system external parameters, two variants are possible. In the first case, the united system remains in the equilibrium state, i.e. there is no heat exchange between the subsystems. In the second case, the united system turns out nonequilibrium, and there emerges a heat exchange between the subsystems.

Hence, the state of thermodynamic equilibrium is determined not only by external parameters, but it also depends on a parameter that characterizes the internal state of the system and is called the **temperature**, T. In the absence of heat exchange between the subsystems, their temperatures are considered to be identical. If the heat exchange takes place, the temperature of the system that loses its internal energy, i.e. it gives up a certain amount of the heat, is considered to be higher.

Temperature measurements are based on the temperature dependence of a certain system parameter—volume, length, electric conductivity, thermoelectromotive force—and the **transitive property of thermal equilibrium**. The latter consists in that, if each of the systems A and B is in the state of thermodynamic equilibrium with the system C, they are also in equilibrium with each other. An opportunity to use thermometers is based just on this property.

In household mercury and red spirit thermometers, the dependence liquid volume on the temperature is used; as a result, if the temperature varies, the length of the mercury or spirit column changes. The modern centigrade temperature scale was proposed by the Swedish astronomer and physicist A. Celsius in 1742. However, the physicists use the **absolute** (thermodynamic) **temperature** introduced on the basis of the second law of thermodynamics (see section 19.2) by the English physicist W. Thomson (lord Kelvin). According to the International system of units (SI), the reference point on the thermodynamic temperature scale is the temperature of the triple point of water (see section 20.5) exactly equal to 273.16 K, and the unit of thermodynamic temperature is kelvin (K), the magnitude of which is equal to the degree of the Celsius (centigrade) scale. The absolute (thermodynamic) temperature T is coupled with the temperature t on the Celsius scale by the relationship

$$T\,(K) = t\,(^\circ C) + 273.15, \tag{18.4}$$

since, at the normal atmospheric pressure, the temperature $t = 0$ °C corresponds to the absolute temperature $T = 273.15$ K.

Taking into account the internal and external parameters, as well as the temperature, the **second postulate of thermodynamics** can be formulated as follows:

- all equilibrium internal parameters of the thermodynamic system are functions of external parameters and the temperature.

In other words,

$$b_j = b_j\,(a_1, a_2, \ldots, T)\,. \tag{18.5}$$

The internal energy U is one of internal parameters. Therefore, in accordance with (18.5),

$$U = U\,(a_1, a_2, \ldots, T)\,. \tag{18.6}$$

Equation (18.6) is called the **caloric equation of state** for the system. An example of the caloric equation of state is the formula for the internal energy of ideal one-atomic gas,

$$U = \frac{3}{2} N k_0 T, \tag{18.7}$$

where k_0 is the Boltzmann constant.

The generalized forces Λ_i are also internal parameters. Therefore, it is possible to write

$$\Lambda_i = \Lambda_i\,(a_1, a_2, \ldots\,, T)\,. \tag{18.8}$$

Equation (18.8) is called the **thermal equation of state**. An important generalized force is the pressure p. Hence,

$$p = p\,(a_1, a_2, \ldots\,, T)\,. \tag{18.9}$$

This is the most important thermal equation of state. Usually, it is called the **equation of state for the system**. Some examples of the equation of state are the Mendeleev-Clapeyron equation for an ideal gas,

$$p = \frac{m}{M} \frac{RT}{V},$$

where m and M are the mass and the molar mass, respectively, of the gas and R is the universal gas constant; and the Van der Waals equation for a real gas,

$$\left(p + \frac{a}{V^2}\right)(V - b) = RT,$$

where a and b are Van der Waals constants.

18.4. Thermodynamic processes

The thermodynamic system is called simple if it is characterized by a single external parameter, so that there is a single generalize force Λ. The equation of state of this system looks like

$$\Lambda = \Lambda(a, T), \tag{18.10}$$

i.e. the state is determined by the parameters a and T, and it can be represent graphically as a point in the $\Lambda - a$ diagram. For example, point 1 in Fig. 18.1 corresponds in the state of ideal gas with the pressure p_1 and the volume V_1 in the $p - V$ diagram.

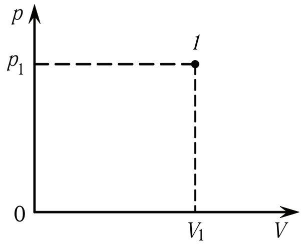

Fig. 18.1.

The thermodynamic parameters may change in time. This change is called the **process**. A considerable value at that has the rate of parameter change.

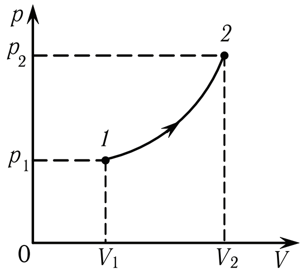

Fig. 18.2.

First, let us consider the limiting case when parameters change very slowly in time. Then, the rate of parameter variation da/dt tends to zero, so that the parameter can be considered as constant and system to be in the corresponding equilibrium state at every time moment t. Such processes are called **equilibrium** or **quasistatic**. The existence of an equilibrium state at every time moment makes it possible to represent the equilibrium process graphically (Fig. 18.2).

Real processes can be considered approximately equilibrium if the rate of parameter a variation, da/dt, is much smaller than the rate of parameter a relaxation $\Delta a/\tau$, i.e. under the condition

$$\left|\frac{da}{dt}\right| \ll \frac{|\Delta a|}{\tau}. \tag{18.11}$$

Condition (18.11) is usually satisfied. For instance, in gases, $\tau < 10^{-5}$ s. The rates of real processes satisfy condition (18.11); therefore, they can be approximately considered as equilibrium.

If condition (18.11) is not satisfied, the process is called **nonequilibrium**. At every time moment t, the system is not at equilibrium, and the process cannot be represented graphically, because there are no exact values of parameters over the whole system at every time moment. In particular, different parts of the system are characterized by different pressures, temperatures, and so forth (Fig. 18.3). In this connection, the nonequilibrium process is conditionally represented with the help of a dashed curve, which means that there are no exact transient states.

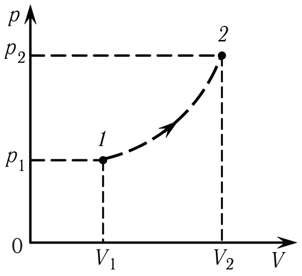

Fig. 18.3.

Since equilibrium processes are infinitely slow, all real processes nonequilibrium, which gives rise to some important consequences (see section 19.2).

19

THERMODYNAMIC LAWS

19.1. First law of thermodynamics

Thermodynamics is based on fundamental principles formulated by generalizing experimental facts. These principles are applicable to all macroscopic systems and are called the **thermodynamics laws**. There are three thermodynamics laws.

The **first law of thermodynamics** is the most easy to understand, because it is the law of energy conservation and transformation for processes with heat exchange. The energy conservation law consists in that the energy in the closed system can change its form (mechanical, electric, magnetic energy, and so on), but its amount remains constant.

As was already mentioned in section 18.2, the internal energy of thermodynamic system can change owing to either some work is done, which is accompanied by the variation of external parameters, or a heat exchange occurs and the external parameters remain constant. The first process (without heat exchange) is called **adiabatic**. In this case, according to the energy conservation law, the system does some work of which a reduction of its internal energy, i.e.

$$(\Delta U)_{\text{ad}} = -A. \tag{19.1}$$

In the case of heat exchange, the work is not done, but the system receives (or loses) some amount of heat Q, and, as a results, its internal energy increases (or decreases). Hence,

$$(\Delta U)_{\text{heat}} = Q. \tag{19.2}$$

In the general case, we must write

$$\Delta U = Q - A. \tag{19.3}$$

This relation is the expression for the first law of thermodynamic in the integral form.

Turning to the first law of thermodynamics in the differential form, we must take into account that the internal energy is a function of state; otherwise, a perpetuum mobile of the first kind, which would have produced work without the input of energy, could have been created. Therefore, its infinitesimal variation is a total differential ($\delta U = dU$). However,, the heat Q and the work A are not functions of state; hence, their infinitesimal changes δQ and δA are not total differentials. Therefore, the first law of thermodynamics in the differential form has to be written as follows:

$$dU = \delta Q - \delta A. \tag{19.4}$$

In certain cases, δQ and δA can be total differentials. In particular, for the isochoric process, $\delta A = 0$, and, accordingly, $dU = -dA$.

19.2. Second law of thermodynamics. Entropy

The second law of thermodynamics is much more difficult for understanding then the first one. It is a result of experimental facts obtained in the course of researches that were mainly stimulated by the invention and improvement of steam engine.

One of the important conclusions that summarized those researches was the establishment of the fact that the heat taken away from a heater cannot be completely transformed into the mechanical work in principle. This conclusion is a special case of the general property of real thermodynamic processes, which consists in their irreversibility, and is a basis of the **second law of thermodynamics**.

The process is called **reversible** if it can be run in a forward direction through some sequence of states and, afterwards, in a backward direction through the same states, so that the system returns into the initial state, with no changes taking place in surrounding objects (in the environment). Therefore, this state is called the **compensation**.

Reversible processes are equilibrium, i.e. infinitely slow, ones, when equilibrium states exist at every time moment, and there is a possibility to run the inverse infinitely slow process through the same states. All real processes are irreversible, i.e. associated with the compensation.

As an example, let us consider the expansion of ideal gas into vacuum. In this case, the gas does no work, since the resistance to the expansion is absent. In order to bring the gas into the initial state, it has to be compressed, so that a certain work A_{ext} has to be done (at the expense of the mechanical energy ΔW of external objects). However, the gas heats up at compression ($\Delta U = A_{\text{ext}}$). Therefore, the heat $Q = A_{\text{ext}}$ has to be taken away from it to the environment. This can be done, and the gas returns into the initial state. However, the state of the environment changes, because, instead of the mechanical energy ΔW, it has the equivalent amount of heat $Q = \Delta W$, which cannot be transformed completely into the mechanical energy. Hence, the gas expansion into vacuum occurs with compensation; therefore, this is an irreversible process.

For the gas expansion to be reversible, it has to occur infinitely slow. Therefore, there must be a piston that moves infinitely slowly; hence, external forces must act on it. However, in this case, the gas does the work A against the external forces at the expense of its internal energy. As a result, the gas cools down, and the mechanical energy of surrounding objects increase by $\Delta W = A$. Later, this energy can be completely used to compress the gas and heat it up to its initial temperature.

The impossibility of the complete conversion of the heat into the work, i.e. the impossibility of getting rid of compensation, dictates a necessity to have a heater, a working medium, and a refrigerator in the heat engine. The heat Q_1 taken away from the heater characterized by the temperature T_1 is partially converted by the expanding working medium into the work A, and the rest heat $Q_2 = Q_1 - A$ is transferred to the refrigerator with the temperature $T_2 < T_1$. Then, the working medium returns into the initial state. Therefore, the heat engine must work **cyclically** (the cyclic process).

The cycle efficiency is determined by the formula

$$\eta = \frac{A}{Q_1} = \frac{Q_1 - Q_2}{Q_1}. \tag{19.5}$$

Hence, to increase this parameter at a given Q_1, a cycle corresponding to the maximum work has to be used. The maximum work is obtained in the reversible cycle composed of equilibrium processes (see section 20.4). Therefore, let us confine ourselves to the qualitative consideration.

If the gas expands infinitely slowly (the reversible process) by moving the piston, the pressure is identical throughout the gas volume. However, in the real expansion process, when the piston moves at a finite velocity, the pressure under the piston is lower than at other points, so that the work done by the gas is also lower. Therefore, for the cycle efficiency to be maximal, the cycle should be equilibrium (reversible), i.e. the process of heat, Q_1, transfer from the heater to the working medium must take place at their identical temperature T_1; otherwise, the system heater–working medium is nonequilibrium. Hence, this process must be isothermal. Similarly, the process of heat, Q_2, transfer from the working medium to the refrigerator must also be isothermal and occur at the temperature T_2.

After the isothermal expansion at the temperature T_1, the working medium has to be cooled down to the refrigerator temperature T_2, and this process must be equilibrium, i.e. without a thermal contact with the heater or the refrigerator. This can be done if the expansion is adiabatic.

At last, after the working medium has transferred the heat Q_2 to the refrigerator, it has to be returned into the initial state with the temperature T_1. This can be done using adiabatic compression.

As a result, we obtain a cycle consisting of two isotherms and two adiabatic curves and called the **Carnot cycle** (Fig. 19.1). The efficiency of this was determined in 1824 by S. Carnot, who formulated the **Carnot theorem** including two statements:

- The efficiency of reversible Carnot cycle does not depend on the working medium, being defined only by the heater and refrigerator temperatures,
$$\eta_{\mathrm{rev}} = \frac{T_1 - T_2}{T_1}. \tag{19.6}$$
- The efficiency of any irreversible cycle is lower than that of reversible cycle,
$$\eta_{\mathrm{irrev}} = \frac{Q_1 - Q_2}{Q_1} < \frac{T_1 - T_2}{T_1}. \tag{19.7}$$

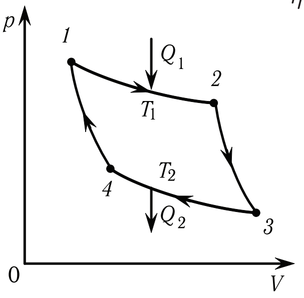

Fig. 19.1.

The proof of the Carnot theorem is given in Appendix H.

On the basis of expressions (19.6) and (19.7), we may write

$$\frac{Q_1 - Q_2}{Q_1} \le \frac{T_1 - T_2}{T_1}, \tag{19.8}$$

where the signs "=" and "<" correspond to the reversible and irreversible Carnot cycles, respectively.

The Carnot theorem brings about the mathematical formation of the second law of thermodynamics. However, let us first formulate it at the qualitative level. The first formulation belongs to R. Clausius (1850): a process when heat spontaneously flows from warmer bodies to colder ones is impossible (or, equivalently, heat always spontaneous flows from more heated to less heated objects). The second, independent, formulation belongs to W. Thomson (lord Kelvin, 1851): it is impossible to construct a cyclically operating engine, the only result of which would be an execution of mechanical work owing to the cooling of a thermal reservoir. The heat engine of this type is called the **perpetuum mobile of the second kind**.

Now, let us find the mathematical formulation of the second law of thermodynamics. From (19.8), it follows that

$$\frac{Q_1}{T_1} - \frac{Q_2}{T_2} \le 0. \tag{19.9}$$

If Q_1 and Q_2 are the heats obtained by the system (in this case, $Q_1 > 0$ and $Q_2 < 0$), relation (19.9) takes the form

$$\frac{Q_1}{T_1} + \frac{Q_2}{T_2} \leq 0. \tag{19.10}$$

The quantity Q/T, where Q is the heat obtained by the system at the temperatures T, is called the **reduced heat**. Hence, the sum of reduced heats in the Carnot cycle does not exceed zero. Generalizing this conclusion for an arbitrary cycle, we may write

$$\sum_i \frac{\Delta Q_i}{T_i} \leq 0, \tag{19.11}$$

where ΔQ_i is the amount of heat obtained by the system at the temperature T_i. In the integral form, relationship (19.11) reads

$$\lim_{\Delta Q_i \to 0} \sum_i \frac{\Delta Q_i}{T_i} = \oint \frac{\delta Q}{T} \leq 0. \tag{19.12}$$

In the case of reversible process (cycle), we have the **Clausius equality**

$$\oint \frac{\delta Q}{T} = 0. \tag{19.13}$$

Then, the integrand has to be a total differential of a certain function S, i.e.

$$dS = \frac{\delta Q}{T}. \tag{19.14}$$

The function S was introduced into thermodynamics by R. Clausius. It is called the **entropy** (from the Greek work “transformation”).

In the case of irreversible process, from (19.12), we obtain the **Clausius inequality**

$$\oint \frac{\delta Q}{T} < 0, \tag{19.15}$$

Therefore,

$$dS > \frac{\delta Q}{T}. \tag{19.16}$$

In the general case, combining (19.15) and (19.16), we obtain

$$dS \geq \frac{\delta Q}{T}. \tag{19.17}$$

The mathematical formulation of the second law of thermodynamics in the integral form is the Clausius “equality–inequality” (19.12),

$$\oint \frac{\delta Q}{T} \leq 0, \tag{19.18}$$

and, in the differential form, relation (19.17).

$$dS \geq \frac{\delta Q}{T}. \tag{19.19}$$

If the system is closed, $\delta Q = 0$. Then,

$$dS \geq 0. \tag{19.20}$$

this means that any real process in the closed system is always accompanied by the entropy growth (the **law of entropy growth**), until the equilibrium state is established, which corresponds to the maximum entropy value $S = S_{\text{max}}$ and in which $dS = 0$.

Hence, heat and work are not equivalent in that sense that work can be completely transformed into heat, i.e. into the energy of chaotic motion of particles, whereas heat cannot be completely transformed into work. Therefore, any processes in the closed system have a certain direction, which consists in that the forms of energy associated with an ordered motion or an ordered state of the system (mechanical, electric, magnetic, and gravitational energies) are ultimately converted into the internal energy, i.e. they are depreciated. Such processes are called **dissipative** and associated with irreversibility. Therefore, the entropy growth in the course of some process is a measure of its irreversibility, and the entropy itself is a measure of chaos and disorder in the system.

According to (19.14), the entropy and its change can be calculated by the formulas

$$S = \int \frac{\delta Q}{T} + S_0, \tag{19.21}$$

$$\Delta S = \int_1^2 \frac{\delta Q}{T}. \tag{19.22}$$

In order to use them, we must calculate δQ. The simplest way to calculate the amount of heat is based on the application of the **heat capacity**. The heat capacity if defined as the quantity

$$C = \frac{\delta Q}{dT}. \tag{19.23}$$

In other words, the heat capacity numerically equals to the heat required to heat up the system (an object) by one kelvin.

The heat capacity substantially depends on the body's mass and the heating conditions. The specific and molar heat capacities at constant volume or a constant pressure are used the most often. In the case of molar heat capacities,

$$C_V = \frac{\delta Q_V}{dT}, \quad C_p = \frac{\delta Q_p}{dT}, \tag{19.24}$$

where δQ_V and δQ_p are the heats required to increase the temperature of one mole of the substance by dT at constant volume or constant pressure, respectively.

On the basis of (19.24), we may write

$$\delta Q_V = C_V dT, \quad \delta Q_p = C_p dT. \tag{19.25}$$

Then,

$$Q_V = \int_{T_1}^{T_2} C_V dT, \quad Q_p = \int_{T_1}^{T_2} C_p dT. \tag{19.26}$$

In general, the heat capacities C_V and C_p depend on the temperature, i.e.

$$C_V = C_V(T), \quad C_p = C_p(T). \tag{19.27}$$

Then, instead of (2.21) and (2.25), we obtain

$$S = \int \frac{C(T)\,dT}{T} + S_0, \tag{19.28}$$

$$\Delta S = \int_1^2 \frac{C(T)\,dT}{T}. \tag{19.29}$$

19.3. Third law of thermodynamics

The first and second laws play a dominant role in thermodynamics. However, they do not solve all its problems, with the uncertainty of integration constant S_0 in formula (19.12) for entropy calculation being one of them. At the same time, a number of quantities depend on its value. Similar problems are resolved with the help of the **third law of thermodynamics**, which follows from neither the first nor the second one.

In 1906, the German physical chemist W. Nernst, on the basis of experimental material obtained while studying the behavior of various substances at low temperatures, found that, when $T \to 0$, the processes run without the entropy change, i.e. the entropy ceases to be the function of the state and tends to a certain constant. Hence,

$$\lim_{T\to 0} \Delta S = \lim_{T\to 0} (S_2 - S_1) = 0. \qquad (19.30)$$

This statement is called the **Nernst theorem**. Then, it follows that

$$\lim_{T\to 0} S_1 = \lim_{T\to 0} S_2 = S_0.$$

In 1912, M. Planck specified the Nernst theorem by formulating the postulate that, at $T \to 0$, not only the entropy ΔS but the entropy S itself tends to zero, i.e.

$$\lim_{T\to 0} S = 0. \qquad (19.31)$$

The Planck postulate is substantiated with the help of quantum-mechanical theory (see section 22.4). It is called the **third law of thermodynamics**.

From formulas (19.21) and (19.31), it becomes clear that $S_0 = 0$. Hence,

$$S = \int \frac{\delta Q}{T} = \int \frac{C\,(T)\,dT}{T}. \qquad (19.32)$$

The third law of thermodynamics brings about a number of important consequences. One of them is the conclusion about the inaccessibility of the absolute zero temperature, which is another formulation of the third law of thermodynamics.

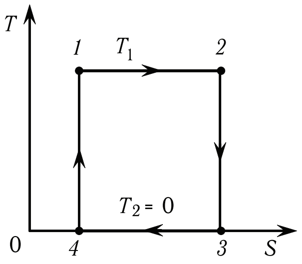

Fig. 19.2.

To substantiate this conclusion, let us consider a reversible Carnot cycle with the heater temperature $T_1 \neq 0$ and the refrigerator temperature $T_2 = 0$, depicted in the $T-S$ diagram (Fig. 19.2). According to the second law of thermodynamics (see (19.13) and (19.14)),

$$\oint \frac{\delta Q}{T} = \oint dS = \Delta S_{12} + \Delta S_{23} + \Delta S_{34} + \Delta S_{41} = 0. \qquad (19.33)$$

Process 1–2 is isotherrmal; therefore,

$$\Delta S_{12} = \int_1^2 \frac{\delta Q}{T} = \frac{1}{T_1} \int_1^2 \delta Q = \frac{Q_1}{T}, \tag{19.34}$$

where Q_1 is the heat obtained by the system in the course of this process. Processes 2–3 and 4–1 are adiabatic ($\delta Q = 0$), the entropy in them does not change, and hence $\Delta S_{23} = \Delta S_{41} = 0$. At last, process 3–4 occurs at $T = 0$; therefore, according to the third law of thermodynamics, $\Delta S_{34} = 0$. Substituting those values into (19.33), we obtain an inconsistent equation

$$\frac{Q_1}{T} = 0,$$

since $Q_1 \neq 0$. This fact testifies that the refrigerator temperature cannot be equal to zero, i.e. the absolute zero temperature is inaccessible.

Another conclusion following from the third law of thermodynamics consists in that the heat capacity of the system $C(T)$ also tends to zero at $T \to 0$. Really, in accordance with (19.32),

$$\lim_{T\to 0} S = \lim_{T\to 0} \int \frac{C\,(T)\,dT}{T} = 0.$$

For the integral not to diverge at $T \to 0$, we must put

$$\lim_{T\to 0} C\,(T) = 0. \tag{19.35}$$

20

THERMODYNAMIC METHODS

20.1. Method of cycles

While solving thermodynamic problems, two methods are used: the method of cyclic processes (the method of cycles) and the method of thermodynamic functions.

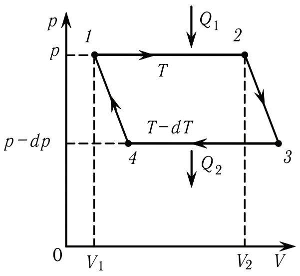

Fig. 20.1.

The **method of cycles** consists in that, when studying the regularities of a certain phenomenon, a specially selected cycle is analyzed by applying thermodynamics laws to it. The application of the Carnot cycle to determine the efficiency of the heat engine (see Appendix H) is an example. Here, we consider another simple example: the temperature dependence of saturated vapor pressure.

Let us consider a Carnot cycle for a working medium consisting of 1 mol of a substance in the form of the liquid and saturated vapor phases (Fig. 20.1). In initial state 1, we have 1 mol of liquid with the volume V_1 and the temperature T. First, let the system isothermally expands to state 2, in which we have 1 mol of saturated vapor with the volume V_2. To do so, the system has to obtain the heat Q_1, which equals to the molar heat of evaporation. Then, we allow the saturated vapor to infinitesimally expand adiabatically to state 3, in which the pressure equals $p-dp$ and the temperature $T-dT$. Afterwards, the system is isothermally compressed to state 4; as a result, it gives the heat Q_2 back, and the saturated vapor is condensed into the liquid. From state 4, the system is adiabatically compressed to initial state 1.

Hence, in sections 1–2–3, the liquid transforms into the saturated vapor, whereas in sections 3–4–1, the saturated vapor transforms into tthe liquid. The substance mass that transforms into the vapor (section 2–3) or liquid (section 4–1) is infinitesimally small.

The cycle efficiency equals

$$\eta = \frac{T-(T-dT)}{T} = \frac{dA}{Q_1} = \frac{(V_2-V_1)\,dp}{Q_1}.$$

Therefore,

$$\frac{dp}{dT} = \frac{Q_1}{T\,(V_2-V_1)}. \tag{20.1}$$

Equations (20.1) is called the **Clapeyron–Clausius equation** (see section 20.4). Its solution gives the dependence $p = p(T)$.

The method of cycles was widely used during the early development period of thermodynamics. However, its application often becomes complicated owing to difficulties associated with the requirement to select the required cyclic process. In modern thermodynamics, a more effective analytical method—the method of thermodynamic functions—is mostly used.

20.2. Method of thermodynamic functions

The thermodynamic functions, which are also called the thermodynamic potentials, can be regarded as analogs of electromagnetic potentials in electrodynamics (whence their name follows), because the macroscopic quantities that characterize the system state are calculated with the help of corresponding potentials in both cases. The thermodynamic potentials were introduced into thermodynamics by J.W. Gibbs. They are characterized by the following properties.

- Thermodynamic functions are functions of state. Therefore, their infinitesimal variations are total differentials.
- Thermodynamic functions are additive quantities.
- If any of thermodynamic functions is known, the thermal and caloric equations of state can be found, and the thermodynamic parameters of the system can be determined.
- If the system is in the equilibrium state, each of thermodynamic functions has an extreme value.

There are various thermodynamic functions. However, J.W. Gibbs distinguished four, the most important, of them: the internal energy U, the free energy F, the enthalpy H (the thermal function), and the thermodynamic Gibbs potential Φ. Let us consider them one by one, assuming the system, for the sake of simplicity, to be characterized by the parameters volume V, pressure p, and temperature T, and the processes in it to be reversible.

In this case, the first law of thermodynamics looks like

$$dU = \delta Q - \delta A = \delta Q - pdV, \tag{20.2}$$

and the second one

$$dS = \frac{\delta Q}{T}, \tag{20.3}$$

i.e. $\delta Q = TdS$. Substituting δQ into (20.2), we obtain the **basic thermodynamic equation**

$$dU = TdS - pdV. \tag{20.4}$$

As was already mentioned, dU is the total differential, and U the state function. From (20.4), the thermal equation of state follows,

$$p = -\left(\frac{\partial U}{\partial V}\right)_S, \tag{20.5}$$

and the temperature is

$$T = \left(\frac{\partial U}{\partial S}\right)_V. \tag{20.6}$$

Therefore, the internal energy is a thermodynamic function.

The right hand side of equation (20.4) contains the total differentials dS and dV, and the corresponding coefficient T and p are thermodynamic parameters. Therefore, S and V are natural (convenient) variables for the internal energy, i.e. we adopt that $U = U(S, V)$. Sometimes, the dependence $U = U(T, V)$ is adopted, but then

$$dU = \left(\frac{\partial U}{\partial V}\right)_V dT + \left(\frac{\partial U}{\partial V}\right)_T dV,$$

and the coefficient before dT has a more complicated sense.

If we subtract the total differential $d\,(TS) = TdS + SdT$ from (20.4), we obtain

$$d\,(U - TS) = -SdT - pdV \tag{20.7}$$

or

$$dF = -SdT - pdV. \tag{20.8}$$

The function

$$F = U - TS \tag{20.9}$$

is called the **Helmholtz free energy or simply free energy**. From (20.8), it follows that

$$S = -\left(\frac{\partial F}{\partial T}\right)_V, \tag{20.10}$$

$$p = -\left(\frac{\partial F}{\partial V}\right)_T. \tag{20.11}$$

From (20.8), it also follows that the natural variables for the free energy are the temperature T and the volume V:

$$F = F(T, V). \tag{20.12}$$

For an isothermal process, $dT = 0$. Then, on the basis of (20.8), we obtain

$$\delta A_T = pdV = -dF, \tag{20.13}$$

i.e. the system does a work owing to a reduction of its free energy, whence the name "free energy" follows.

The third thermodynamic function can be found by adding the differential $d(pV) = pdV + Vdp$ to (3.4),

$$d(U + pV) = TdS + Vdp. \tag{20.14}$$

The function

$$H = U + pV \tag{20.15}$$

is called the **enthalpy** or the **thermal function**. Hence,

$$dH = TdS + Vdp, \tag{20.16}$$

$$T = \left(\frac{\partial H}{\partial S}\right)_p, \tag{20.17}$$

$$V = \left(\frac{\partial H}{\partial p}\right)_S. \tag{20.18}$$

The natural variables for the enthalpy are the entropy and the pressure, i.e.

$$H = H(S, p) \tag{20.19}$$

For an isobaric process, when $dp = 0$, it follows from (20.16) that

$$dH = TdS = \delta Q_p. \tag{20.20}$$

Hence, the enthalpy grows owing to the heat obtained by the system, whence its name follows

By adding the differential $d(pV) = pdV + Vdp$ (20.7), we obtain the fourth thermodynamic function

$$d(U - TS + pdV) = -SdT + Vdp \tag{20.21}$$

or

$$d\Phi = -SdT + Vdp. \tag{20.22}$$

The function

$$\Phi = U - TS + pV = F + pV \tag{20.23}$$

is called the **Gibbs free energy**. On the basis of (20.22), we may write

$$S = -\left(\frac{\partial \Phi}{\partial T}\right)_p, \tag{20.24}$$

$$V = \left(\frac{\partial \Phi}{\partial p}\right)_T. \tag{20.25}$$

Besides, from (3.22), it follows that

$$\Phi = \Phi(T, p). \tag{20.26}$$

If the system is not simple, and, besides the volume, is characterized by other external parameters, the elementary work done by the system looks like

$$\delta A = pdV + \delta A_H, \tag{20.27}$$

where δA_H is the work associated with the change of other external parameters and is called the **work of nonmechanical forces**, e.g., chemical interaction forces. Then, the master thermodynamic equation (20.4) takes the form

$$duA = TdS - pdV - \delta A_H, \tag{20.28}$$

and instead of equations (20.7), (20.16), and (20.22), we obtain

$$dF = -SdT - pdV - \delta A_H, \tag{20.29}$$

$$dH = TdS + Vdp - \delta A_H, \tag{20.30}$$

$$d\Phi = -SdT + Vdp - \delta A_H. \tag{20.31}$$

In this case, for the isothermal and isobaric process ($dT = dp = 0$), we have

$$\delta A_H = -d\Phi, \tag{20.32}$$

i.e. the work of nonmechanical forces is done by the system owing to the reduction of its Gibbs free energy.

The thermodynamic functions are interrelated. For instance, on the basis of (20.9) and (20.10), we may write

$$F = U + T\left(\frac{\partial F}{\partial T}\right)_V , \tag{20.33}$$

i.e.

$$\frac{U}{T^2} = \frac{F}{T^2} - \frac{1}{T}\left(\frac{\partial F}{\partial T}\right)_V = -\frac{\partial}{\partial T}\left(\frac{F}{T}\right)_V , \tag{20.34}$$

whereas, on the basis (20.15), (20.23), and (20.24), we obtain

$$\Phi = H + T\left(\frac{\partial \Phi}{\partial T}\right)_p . \tag{20.35}$$

Equations (20.34) and (20.35) are called the **Gibbs–Helmholtz equations**.

20.3. Systems with a variable number of particles. Chemical potential

Besides the thermodynamic functions, an important quantity in thermodynamics and other domains of physics is the chemical potential. In order to analyze it, it is convenient to include the number of particles in the system, N, into consideration. So, let a simple system consist of N particles of the same kind (the generalization on the case of particles of several kinds is trivial). Let this number be able to vary, i.e. the system can exchange particles with the environment. Then, the internal energy of the system U is a function of the volume V, the entropy S, and particle number N, so that

$$dU = dU(S, V, N) = \left(\frac{\partial U}{\partial S}\right)_{V,N} dS + \left(\frac{\partial U}{\partial V}\right)_{S,N} dV + \left(\frac{\partial U}{\partial N}\right)_{S,V} dN. \tag{20.36}$$

Taking (20.5) and (20.6) into account and introducing the notation

$$\mu = \left(\frac{\partial U}{\partial N}\right)_{S,V} , \tag{20.37}$$

we obtain

$$dU = TdS - pdV + \mu dN. \tag{20.38}$$

The function μ is called the **chemical potential**. From (20.38), it follows that the chemical potential equals the change in the internal energy of the system if the number of particles in it varies, but the entropy and the volume remain invariable.

Making use of (20.38), we obtain the relations

$$dF = -SdT - pdV + \mu dN, \quad \mu = \left(\frac{\partial F}{\partial N}\right)_{T,V}; \tag{20.39}$$

$$dH = TdS + Vdp + \mu dN, \quad \mu = \left(\frac{\partial H}{\partial N}\right)_{S,p}; \tag{20.40}$$

$$d\Phi = -SdT + Vdp + \mu dN, \quad \mu = \left(\frac{\partial \Phi}{\partial N}\right)_{T,p}. \tag{20.41}$$

Hence, the chemical potential can be obtained by differentiating any thermodynamic function with respect to the particle number, provided that other (different) parameters remain constant.

If the system consists of particles of different kinds, then

$$dU = TdS - pdV + \sum_i \mu_i dN_i, \tag{20.42}$$

where

$$\mu_i = \left(\frac{\partial U}{\partial N_i}\right)_{S,\,V,N_J}, \quad (j \neq i) \tag{20.43}$$

is the chemical potential of particles of i-th kind.

The thermodynamic functions are additive. Therefore, we may write

$$\Phi = \Phi\left(T, p, N\right) = N\Phi_0\left(T, p\right), \tag{20.44}$$

where $\Phi_0\left(T, p\right)$ is the Gibbs free energy related to one particle. Here, we took into account that the temperature T and the pressure p are not additive quantities and, therefore, do not depend on N. From (20.41) and (20.44), it follows that

$$\mu = \mu\left(T, p\right) = \left(\frac{\partial \Phi}{\partial N}\right)_{T,p} = \Phi_0\left(T, p\right) = \frac{\Phi}{N}, \tag{20.45}$$

i.e. the chemical potential is identical to the Gibbs free energy related to one particle.

In the case of systems with the variable particle number, the thermodynamic function

$$G = F - \mu N = U - TS - \mu N \tag{20.46}$$

is often used, which is called the **grand potential**. Taking into account that $\mu N = \Phi = U - TS + pV$, we may write

$$G = -pV. \tag{20.47}$$

On the basis of (20.46) and (20.39), we have

$$dG = dF - d\left(\mu N\right) = -SdT - pdV - Nd\mu, \tag{20.48}$$

i.e.

$$G = G\left(T, V, \mu\right).$$

From (20.48), it follows that

$$S = -\left(\frac{\partial G}{\partial T}\right)_{V,\mu}, \quad p = -\left(\frac{\partial G}{\partial V}\right)_{T,\mu}, \quad N = -\left(\frac{\partial G}{\partial \mu}\right)_{T,V}. \tag{20.49}$$

20.4. Thermodynamic equilibrium conditions

According to the first postulate of thermodynamics, a closed system always transforms to an equilibrium state as time elapses. However, experimental data testifies that an open system also can be in the equilibrium state under certain external conditions. In this connection, let us consider the issue concerning the conditions of thermodynamic equilibrium in open systems at various external factors.

In the general case, the second law of thermodynamics looks like

$$dS \geq \frac{\delta Q}{T}, \tag{20.50}$$

where the equality corresponds to reversible processes and the inequality to irreversible ones. Therefore, $\delta Q \leq TdS$. Substituting this δQ-value into the first law of thermodynamics for simple systems,

$$dU = \delta Q - pdV,$$

we obtain the **main thermodynamic inequality**

$$dU \leq TdS - pdV. \tag{20.51}$$

Then, instead of (20.8), (20.16), and (20.22), we have

$$dF \leq -SdT - pdV, \tag{20.52}$$

$$dH \leq -TdS + Vdp, \tag{20.53}$$

$$d\Phi \leq -SdT + Vdp. \tag{20.54}$$

Inequalities (20.51)–(20.54) bring about a number of important conclusions concerning the general conditions of equilibrium in open systems and some other issues. Let us write (20.51) in the form

$$pdV = \delta A \leq TdS - dU. \tag{20.55}$$

From this inequality, it follows that the *maximum work* is done by the system in case of reversible process. This conclusion was used in section 2.2.

1. *The system is closed.* Then, according to the first postulate and the second law of thermodynamics, only processes with the entropy growth are possible in the system, in which give rise to state of thermodynamic equilibrium, i.e. $S = S_{\max}$. Therefore, the necessary and sufficient conditions of equilibrium in the closed system are written as

$$\delta S = 0, \quad \delta^2 S < 0, \tag{20.56}$$

where δS and $\delta^2 S$ are the first and the second variation of entropy.

2. *The system is open, the process is isothermal and isochoric.* In this case, $dT = 0$ and $dV = 0$. Relation (20.52) takes the form

$$dF \leq 0, \tag{20.57}$$

whence it follows that, in this system, only spontaneous processes with a reduction of free energy are possible. In the equilibrium state, $F = F_{\min}$. Therefore, the equilibrium conditions look like

$$\delta F = 0, \quad \delta^2 F < 0. \tag{20.58}$$

3. *The system is opem, the process is isothermal and isobaric* ($dT = 0$ and $dp = 0$). Then, on the basis of (20.54), we obtain

$$d\Phi \leq 0. \tag{20.59}$$

Hence, in the equilibrium state, $\Phi = \Phi_{\min}$, and the equilibrium conditions are written as

$$\delta\Phi = 0, \quad \delta^2\Phi > 0. \tag{20.60}$$

4. *The system is open, the process is isoentropic and isobaric* ($dS = 0$ and $dp = 0$). In this case, in accordance with (20.53),

$$dH \leq 0. \tag{20.61}$$

Therefore, in the equilibrium state, $H = H_{\min}$, and the equilibrium conditions look like

$$\delta H = 0, \quad \delta^2 H > 0. \tag{20.62}$$

5. *The system is open, the process is isoentropic and isochoric.* Then, on the basis of (20.51),

$$dU \leq 0. \tag{20.63}$$

Hence, in the equilibrium state, $U = U_{\min}$, and the equilibrium conditions are written as follows:

$$\delta U = 0, \quad \delta^2 U > 0. \tag{20.64}$$

20.5. Phase equilibrium and phase transitions. Clapeyron–Clausius equation

A thermodynamic system can be homogeneous or heterogeneous. In the latter case, it consists of the several homogeneous regions separated from one another by visible dividing surfaces (interfaces). From this viewpoint, all thermodynamic systems are divided into two classes: homogeneous and heterogeneous.

A system is **homogeneous** if its physical properties are identical at all points. Examples of such systems are a gas or a mixture of gases, liquids, solutions, and so forth.

A system consisting of several homogeneous parts, so that the physical properties change in a jump-like manner at thire interfaces, is **heterogeneous**. Heterogeneous systems can be two-phase, three-phase, and so on

The **phase** is a homogeneous part of the system separated from other parts by a pronounced interface. As various phases of the system, there can be its various aggregate states (gas, liquid, solid), various crystal modifications of the solid, various magnetic and electric states (para- and ferromagnets, conductors and superconductors), and so forth.

Of large importance is also the concept of component. The **component** is a constituent (a substance) of the system, the content of which does not depend on the contents of other constituents (substances). For instance; air is a single-phase system, but it has some components: oxygen, nitrogen, carbon dioxide, and so on. Should oxygen be removed from air, it would not affect the nitrogen content, and vice versa. The system "water–ice" is two-phase, but it has only one component (H_2O): oxygen can be removed only with hydrogen.

In order to determine the equilibrium conditions for the heterogeneous system, let us first consider a single-component two-phase closed system. The general conditions of its equilibrium look like (see section 20.4)

$$\delta S = 0, \quad \delta^2 S < 0.$$

Designating the entropies of two phases as S_1 and S_2, on the basis of their additivity, we may write

$$\delta S = \delta\left(S_1 + S_2\right) = \delta S_1 + \delta S_2 = 0. \tag{20.65}$$

The phases can exchange particles. Therefore, let us apply the master thermodynamic equation (20.38) to them tol obtain

$$\delta S_1 = \frac{1}{T_1}\delta U_1 + \delta U_1 + \frac{p_1}{T_1}\delta V_1 + \frac{\mu_1}{T_1}\delta N_1,$$

$$\delta S_2 = \frac{1}{T_2}\delta U_2 + \delta U_2 + \frac{p_2}{T_2}\delta V_2 + \frac{\mu_2}{T_2}\delta N_2.$$

Hence, equation (20.65) reads

$$\frac{1}{T_1}\delta U_1 + \frac{1}{T_2}\delta U_2 + \frac{p_1}{T_1}\delta V_1 + \frac{p_2}{T_2}\delta V_2 + \frac{\mu_1}{T_1}\delta N_1 + \frac{\mu_2}{T_2}\delta N_2 = 0. \tag{20.66}$$

The internal energy $U = U_1 + U_2$, the volume $V = V_1 + V_2$, and the number of particles $N = N_1 + N_2$ in the closed system are constant. Therefore,

$$\delta U = \delta U_1 + \delta U_2 = 0, \quad \delta\nabla = \delta V_1 + \delta V_2 = 0, \quad \delta N = \delta N_1 + \delta N_2 = 0,$$

or

$$\delta U_2 = -\delta U_1, \quad \delta V_2 = -\delta V_1, \quad \delta N_2 = -\delta N_1. \tag{20.67}$$

Taking into account (20.67), equation (20.66) can be written as

$$\left(\frac{1}{T_1}-\frac{1}{T_2}\right)\delta U_1+\left(\frac{p_1}{T_1}-\frac{p_2}{T_2}\right)\delta V_1+\left(\frac{\mu_1}{T_1}-\frac{\mu_2}{T_2}\right)\delta N_1=0. \tag{20.68}$$

The variations δU_1, δV_1, and δN_1 differ from zero. Therefore, all three expressions in the parentheses vanish. Hence,

$$T_1=T_2, \quad p_1=p_2, \quad \mu_1=\mu_2. \tag{20.69}$$

One can see that the equilibrium conditions for the one-component two-phase systems consist in the identity of the temperatures and the pressures in the phases, as well as in the identity of the chemical potentials of the component in both phases.

Equilibrium conditions (20.69) have a simple physical meaning. The equality of temperatures means the thermal equilibrium, the equality of pressures the mechanical equilibrium, and the equality of chemical potentials the diffusion (dynamic) equilibrium.

From the equality of chemical potentials

$$\mu_1(T,p)=\mu_2(T,p)\,, \tag{20.70}$$

it follows the dependence

$$p=p(T) \tag{20.71}$$

between the pressures and the temperatures, at which two phases are in the equilibrium with each other. The curve that corresponds to dependence (20.71) in the $p-T$ diagram is called the **phase equilibrium curve**. It is shown schematically in Fig. 20.2.

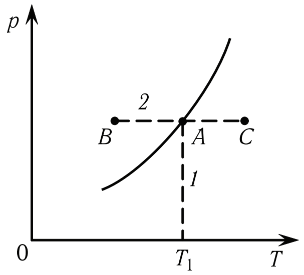

Fig. 20.2.

Two phases are in the equilibrium with each other only at those values of the temperature and pressure that correspond to the points in the curve. All other T- and p-values correspond to the equilibrium states of the whole system either as phase 1 or phase 2. An example is the system "liquid–saturated vapor". In this case, phase 1 is the saturated vapor, and phase 2 the liquid; and the function $p(T)$ determines the temperature dependence of the saturated vapor pressure.

In state B, the system is one-phase (the liquid, phase 2). If the temperature grows under isobaric conditions (process BC) and reaches the value

T_1, phase 1 (the saturated vapor) emerges at point A, and phase 2 transforms into phase 1 (the liquid transforms into the saturated vapor) at the constant temperature T_1. This process is called the **phase transition.** Then, the system becomes single-phase again (the vapor).

An equation, which can be used to determine the function $p(T)$ can be find from condition (20.70). On its basis, we may write

$$\left(\frac{\partial \mu_1}{\partial T}\right)_p dT + \left(\frac{\partial \mu_1}{\partial p}\right)_T dp = \left(\frac{\partial \mu_2}{\partial T}\right)_p dT + \left(\frac{\partial \mu_2}{\partial p}\right)_T dp, \tag{20.72}$$

whence we obtain the equation

$$\frac{dp}{dT} = \frac{\left(\frac{\partial \mu_2}{\partial T}\right)_p - \left(\frac{\partial \mu_1}{\partial T}\right)_p}{\left(\frac{\partial \mu_1}{\partial p}\right)_T - \left(\frac{\partial \mu_2}{\partial p}\right)_T}. \tag{20.73}$$

According to (20.45), (20.24), and (20.25),

$$\left(\frac{\partial \mu}{\partial T}\right)_p = \frac{\partial}{\partial T}\left(\frac{\Phi}{N}\right)_p = \frac{1}{N}\left(\frac{\partial \Phi}{\partial T}\right)_p = -\frac{S}{N},$$

$$\left(\frac{\partial \mu}{\partial T}\right)_T = \frac{\partial}{\partial p}\left(\frac{\Phi}{N}\right)_T = \frac{1}{N}\left(\frac{\partial \Phi}{\partial p}\right)_T = \frac{V}{N}.$$

Therefore, instead of (20.73), we may write

$$\frac{dp}{dT} = \frac{S_1 - S_2}{V_1 - V_2} = \frac{-\Delta S}{V_1 - V_2}. \tag{20.74}$$

Experimental data testify that the phase transition occurs at a constant temperature. Then,

$$\Delta S = \frac{Q}{T}$$

and

$$\frac{dp}{dT} = \frac{-Q}{T\left(V_1 - V_2\right)}, \tag{20.75}$$

where Q is the heat obtained by the system in the course of phase transition, when some mass m of the substance transformed from phase 1 into phase 2; and V_1 and V_2 are the volumes of mass m in phase states 1 and 2, respectively.

However, the heat $\lambda = -Q$ released in the course of phase transition and called the transition heat is used more often. It is usually related to a mass unit (the specific transition heat) or to one mole (the molar heat transition) of the substance. Then, equation (20.75) takes the form

$$\frac{dp}{dT} = \frac{\lambda}{T\left(V_1 - V_2\right)}. \tag{20.76}$$

It is called the **Clapeyron–Clausius equation** (see section 20.1). It can be used to find the function $p(T)$ that describes the phase equilibrium curve or the dependence of phase transition temperature on the pressure.

Equation (20.76) cannot be solved in the general form, because the quantities λ, V_1, and V_2 are functions of the temperature possessing different forms in different cases. However, in some cases, certain simplifications are possible, so that the equation can be solved analytically

As an example, let us consider the equilibrium system "liquid–saturated vapor" and apply the Clapeyron–Clausius equation to the phase transition saturated vapor → liquid, at which, as experimental data show, the transition heat λ is released ($\lambda > 0$). In this case, V_1 and V_2 are the molar volumes of saturated vapor and liquid, respectively. If the system state is not close to critical, then $V_2 \ll V_1 = V$, and we may put $V_2 = 0$, i.e.

$$\frac{dp}{dT} = \frac{\lambda}{TV}.$$

Applying the equation of state for the ideal gas, $pV = RT$, to the saturated vapor, we obtain the equation

$$\frac{dp}{dT} = \frac{\lambda p}{RT^2}. \tag{20.77}$$

In a narrow temperature interval, the transition heat λ can be considered as constant. Then, equation (20.77) can be solved using the method of separation of variables to obtain

$$p = Ce^{-\frac{\lambda}{RT}}, \tag{20.78}$$

where C is the integration constant. From this dependence, it follows that the pressure of saturated vapor grows together with the temperature.

20.6. Gibbs phase rule. Ternary point

How many phases can be in the state of thermodynamic equilibrium with one another simultaneously? To answer this question, we should first generalize phase equilibrium conditions (20.69) on the case when the system consists of n phases and r components.

In the thermodynamic equilibrium state, the temperatures and pressures mist be identical in all phases,

$$T_1 = T_2 = \ldots = T_n, \quad p_1 = p_2 = \ldots = p_n. \tag{20.79}$$

Besides, the chemical potentials of every component in all phases must also be identical, i.e.

$$\begin{aligned}
\mu_1^{(1)} &= \mu_2^{(1)} = ... = \mu_n^{(1)}, \\
\mu_1^{(2)} &= \mu_2^{(2)} = ... = \mu_n^{(2)}, \\
&\ldots \\
\mu_1^{(r)} &= \mu_2^{(r)} = ... = \mu_n^{(r)}.
\end{aligned} \tag{20.80}$$

Here, the superscripts enumerate the components, and the subscripts the phases. The number of independent equations in every row in (20.80) equals $n-1$, and the number of rows equals r. Therefore, the number of independent equations in (20.80) equals $r(n-1)$.

As the variables that the chemical potentials depend on, we select the temperature T and the pressure p, what are identical in all phases, as well as the relative concentrations of components in every phase. The number of components equals r; therefore, the number of concentrations in every phase also equals r. However, the sum of relative concentrations of all components in every phase equals unity; hence, the number of independent concentrations in every phase equals $r-1$, and in all phases $n(r-1)$. Taking into account the temperature and the pressure, the total number of independent variables equals $n(r-1)+2$.

Those variables must satisfy the system of equations (20.80), i.e. they must be its solutions. However, the system of equations has solutions provided that the number of independent equations does not exceed the number of variables. Therefore, the inequality

$$r\,(n-1) \leq n\,(r-1) + 2$$

must be satisfied, whence it follows that

$$n \leq r + 2. \tag{20.81}$$

Relation (20.84) determines the number of phases that can be in the equilibrium state simultaneously and is called the **Gibbs phase rule**.

According to this rule, the maximum number of phases in a one-component system at equilibrium equals three. An example of such a system is a system water–ice–saturated vapor. In this case, the conditions of phase equilibrium (20.79) and (20.80) look like

$$T_1 = T_2 = T_3 = T,$$

$$p_1 = p_2 = p_3 = p,$$

$$\mu_1 = \mu_2 = \mu_3.$$

Since there is only one component, the variables that the chemical potentials depend on are the temperature and the pressure. Therefore,

$$\mu_1 (T, p) = \mu_2 (T, p) = \mu_3 (T, p) ,$$

i.e. we have a system of two independent equations with two unknown variables. Its solution is a set of the temperature, T_0, and pressure, p_0, values, because three phases can be in the thermodynamic equilibrium state only at $T = T_0$ and $p = p_0$. In the $p-T$ diagram, three curves of phase equilibrium start from the point (p_0, T_0); therefore, it is called the **ternary point**.

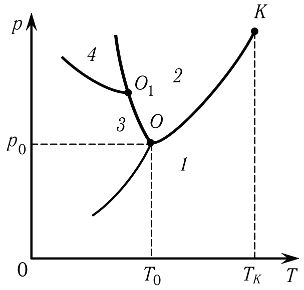

Fig. 20.3.

The phase equilibrium curves and the ternary point for the system water–ice–saturated vapor is schematically shown in Fig. 20.3, where phase 1 is saturated vapor, phase 2 water, phase 3 ice (ice 1), and phase 4 ice (ice 2). Phases ice 1 and ice 2 are different crystal modifications of ice. In the case concerned, $p_0 = 609$ Pa and $T = 273.16$ K (exactly) is the basic reference point of the absolute temperature scale. The curve of phase equilibrium water–saturated vapor ends at point *K,* which corresponds to the critical state, and the curve of phase equilibrium water–ice 1 ends with another ternary point O_1 of the system water–ice 1–ice 2. The curve of phase equilibrium ice 1–saturated vapor requires an additional analysis in the low-temperature interval associated with the quantum-mechanical properties of the substance.

20.7. Phase transitions of the first and second kinds

Phase transitions considered in section 20.6 concerned the change in the aggregate state of the substance associated with the release (or absorption) of the transition heat λ and the variation of the specific (molar) volume $\Delta V = V_2 - V_1$. However, experimental data testify that there are phase transitions for which $\lambda = 0$ and $\Delta V = 0$. In this connection, an issue arises concerning the classification of phase transitions. Such a classification based on general attributes was proposed by P. Ehrenfest in 1933. According to him, phase transitions are divided into transitions of the first and second kinds on the basis of the behavior of Gibbs thermodynamic potential.

The **phase transitions of the first kind** are those transitions for which the first derivatives of the Gibbs thermodynamic potential with respect to the temperature and pressure change in a jump-like manner, i.e. they have a break

$$\Delta\left(\frac{\partial\Phi}{\partial T}\right)_p \neq 0, \quad \Delta\left(\frac{\partial\Phi}{\partial p}\right)_T \neq 0. \tag{20.82}$$

According to (20.24) and (20.25) ,

$$\left(\frac{\partial\Phi}{\partial T}\right)_p = -S, \quad \left(\frac{\partial\Phi}{\partial p}\right)_T = V;$$

therefore,

$$\Delta\left(\frac{\partial\Phi}{\partial T}\right)_p = -\Delta S = -\frac{Q}{T} = \frac{\lambda}{T} \neq 0,$$

$$\Delta\left(\frac{\partial\Phi}{\partial p}\right)_T = \Delta V \neq 0;$$

i.e., according to section 20.6, the transition heat ($\lambda \neq 0$) is released (or absorbed) and the specific volume changes stepwise ($\Delta V \neq 0$) at phase transitions of the first kind. The phase transitions of the first kind include the changes of the aggregate state and transitions between some crystal modifications in solids.

The **phase transitions of the second kind** are those transitions for which $\lambda = 0$ and $\Delta V = 0$, but the second derivatives of the Gibbs free energy change stepwise

$$\Delta \left(\frac{\partial^2 \Phi}{\partial T^2} \right)_p \neq 0, \tag{20.83}$$

$$\Delta \left(\frac{\partial^2 \Phi}{\partial p^2} \right)_T \neq 0, \tag{20.84}$$

$$\Delta \left(\frac{\partial^2 \Phi}{\partial T \partial p} \right) \neq 0. \tag{20.85}$$

Let us consider the consequences that follow from inequalities (20.83)–(20.85). In particular, from condition (20.83), it follows (see (20.24)) that

$$-\Delta \left(\frac{\partial^2 \Phi}{\partial T^2} \right)_p = \Delta \left(\frac{\partial S}{\partial T} \right)_p = \Delta \left(\frac{\partial S_p}{dT} \right) =$$

$$= \Delta \left(\frac{1}{T} \frac{\delta Q_p}{dT} \right) = \Delta \left(\frac{1}{T} \frac{C_p dT}{dT} \right) = \frac{\Delta C_p}{T} \neq 0, \tag{20.86}$$

i.e. the heat capacity has a break.

From condition (20.84), we obtain (see (20.25))

$$\Delta \left(\frac{\partial^2 \Phi}{\partial p^2} \right)_T = \Delta \left(\frac{\partial V}{\partial p} \right)_T \neq 0,$$

However, the derivative $(\partial V / \partial p)_T$ determines the isothermal compressibility,

$$\beta = -\frac{1}{V} \left(\frac{\partial V}{\partial p} \right)_T .$$

Therefore, $\Delta \beta \neq 0$, i.e. the isothermal compressibility also changes stepwise.

At last, from condition (20.85), we find (see (20.24) and (20.25)) that

$$\Delta \left(\frac{\partial^2 \Phi}{\partial T \partial p} \right) = \Delta \left(\frac{\partial V}{\partial T} \right)_p = V \Delta \alpha \neq 0.$$

The derivative $(\partial V/\partial T)_p$ determines the coefficient of thermal expansion,

$$\alpha = \frac{1}{V}\Delta\left(\frac{\partial V}{\partial T}\right)_p .$$

Hence, at the point of the phase transition of the second kind, the heat capacity, the isothermal compressibility, and the coefficient of thermal expansion of the substance change in a jump-like manner. The phase transitions of the second kind include the transitions ferromagnet ⇄ paramagnet, ferrimagnet ⇄ paramagnet, antiferromagnet ⇄ paramagnet, conductor ⇄ superconductor, helium I ⇄ helium II.

Notice that the Clapeyron–Clausius equation is inapplicable to phase transitions of the second kind, because its right hand side (see (20.76)) has a 0/0-singularity at $\lambda = 0$ and $V_1 - V_2 = 0$. However, it can be proved (see section 2.6 in [22]) that

$$\frac{dp}{dT} = \frac{\Delta C_p}{TV\Delta\alpha} = \frac{\Delta\alpha}{\Delta\beta}, \tag{20.87}$$

$$\Delta C_p = TV\frac{(\Delta\alpha)^2}{\Delta\beta}. \tag{20.88}$$

Equations (20.87) and (20.88) are called the **Ehrenfest equations**. While studying the Ehrenfest phase transitions of the second kind, they are of the same importance as the Clapeyron–Clausius equation for the study of phase transitions of the first kind.

20.8. Temperature dependence of the surface tension in liquids

Let us apply the method of thermodynamic potentials to study the temperature dependence of the surface tension in liquids. Experimental data testify that, in order to increase in the area of liquid surface by stretching the surface and keeping the mass of liquids constant, it is necessary to do a certain work against the surface tension forces. Besides, as the surface area increases, the liquid is cooled down. Hence, under isothermal conditions, it exchanges heat with the thermostat.

The surface tension in liquids is characterized by the surface tension coefficient σ, which is defined as the work of external forces required to increase the area Σ of liquid surface by unity at the temperature T and the constant volume V. This work is equal to the corresponding work done by the liquid, but taken with the opposite sign, i.e.

$$\sigma = -\frac{\delta A_{T,V}}{d\Sigma}. \tag{20.89}$$

The heat q_T that the liquid has to obtain while creating a unit area of the surface at $T = \text{const}$ is called the **specific latent heat of surface formation**,

$$q_T = -\frac{\delta Q_T}{d\Sigma}. \tag{20.90}$$

In section 20.2, it was shown that, in the course of isothermal process, the system does a work owing to the reduction of its free energy. Therefore,

$$dF_T = -\delta A_{T,V} = \sigma d\Sigma,$$

i.e.

$$\sigma = \frac{dF_T}{d\Sigma}. \tag{20.91}$$

Hence, the surface tension coefficient can be determined as the variation of free energy in the case when the surface area changes by unity under the condition $T = \text{const}$.

If the liquid volume is constant, it follows from (20.29) that

$$dF = -SdT - \delta A_{T,\,V} = -SdT + \sigma d\Sigma.$$

Therefore, since the differential dF is total, we may write

$$\left(\frac{\partial \sigma}{\partial T}\right)_{V,\,\Sigma} = -\left(\frac{\partial S}{\partial \Sigma}\right)_T. \tag{20.92}$$

Taking into account that $\delta Q = TdS$, equation (20.92) takes the form

$$\left(\frac{\partial \sigma}{\partial T}\right)_{V,\,\Sigma} = -\frac{1}{T}\frac{\delta Q_T}{d\Sigma} = -\frac{q_T}{T}. \tag{20.93}$$

As was already mentioned, if the surface area increases, the liquid must obtain some heat from the thermostat in order to keep the temperature constant, i.e. $q_T > 0$. Then, it follows from (20.93) that

$$\frac{d\sigma}{dT} < 0,$$

so that the surface tension decreases as the temperature grows. This conclusion is confirmed by experimental data.

21

BASIC PRINCIPLES OF STATISTICAL PHYSICS

21.1. Macro- and microscopic states of macroscopic system

The main task of statistical physics is the calculation of **macroscopic parameters** (**macroparameters**)—such as the pressure p, the volume V, the temperature T, and so forth—which characterize the macroscopic system as a whole and determine its **macroscopic state** (the **macrostate**). This task is tackled on the basis of the existence of microparticles (atoms, molecules, ions, and so on), which the macrosystem is composed of, and on the basis of their properties and behavior features.

Microparticles have quantum-mechanical properties. Therefore, the macrosystem is quantum-mechanical in principle, and its state has to be determined by the wave function. If the system contains N particles, each of which has f degrees of freedom, the wave function must look like

$$\psi = \psi_{n_1, n_2, \ldots, n_{fN}} \left(q_1, q_2, \ldots, q_{fN}, t \right), \tag{21.1}$$

where n_1, n_2,..., and n_{fN} are the set of quantum numbers (**microscopic parameters**), with their number being equal to the number of the system degrees of freedom fN; and q_1, q_2,..., and q_{fN} are the coordinates of particles. Wave function (21.1) or the set of quantum numbers n_1, n_2,..., and n_{fN} determine the quantum state of macrosystem, which is called the **microscopic state** (the **microstate**) in the statistical physics.

It is clear that macroparameters depend on microparameters, with the same set of macroparameters being usually associated with a huge number of different combinations of quantum numbers n_i, i.e. different microstates. The microstates, in which the macrosystem can be under certain conditions (for the given set of macroparameters p, V, T, and so on) are called the **accessible (allowable) microstates**. The number of accessible microstates in the macrosystem will be designated as Ω. This is a multiplicative quantity, being one of the key concepts of statistical physics. In this connection, it is necessary to find a way of its calculation.

21.2. Classical system. Phase space

Let us first consider a macroscopic system, the particles of which can be described by coordinates and momenta. The system of this kind is called **classical**. If it consists of N particles and each particle has f degrees of freedom, we have fN coordinates $(q_1, q_2, \ldots, q_{fN})$ and fN momenta $(p_1, p_2, \ldots, p_{fN})$, i.e. $2fN$ microparameters in total, which determine a microstate of classical system.

Let us introduce an abstract $2fN$-dimensional space, which axes correspond to fN coordinates and fN momenta. Such an abstract space is called the **phase space** (Γ-space) of the system. In this space, the microstate of classical system, i.e. the set of all coordinates and momenta, is represented by a single point, which is called the **phase point**. When the system changes its microstates, this can be imagined as the motion of the phase point in the phase space, describing a **phase trajectory**.

At given macroparameters, the phase point can move only in certain phase space regions, the volume $\Delta\Gamma$ of which is called the **accessible phase volume**. It is a classical analog of the number of accessible microstates Ω in the quantum-mechanical system. Therefore, $\Delta\Gamma$ and Ω are interrelated (see section 21.3).

An infinitesimally small (elementary) volume of the phase space $d\Gamma$ is a product of differentials of all coordinates and momenta, i.e.

$$d\Gamma = \prod_{i=1}^{fN} dq_i dp_i. \tag{21.2}$$

Therefore,

$$\Delta\Gamma = \int \prod_{i=1}^{fN} dq_i dp_i, \tag{21.3}$$

there the integration is carried out within the limits corresponding to certain conditions of the system existence.

In case of one particle, $N = 1$, the phase space will be denoted as the γ-space. If the particle has only three spatial degrees of freedom, $f = 3$, the elementary volume in the γ-space equals

$$d\gamma_0 = dxdydzdp_xdp_ydp_z = dVdV_p, \tag{21.4}$$

where $dV = dxdydz$ and $dV_p = dp_xdp_ydp_z$ are infinitesimally small volumes in the coordinate and momentum subspaces, respectively. In the spherical coordinate system,

$$dV_p = p^2dpd\Omega_S, \tag{21.5}$$

where $d\Omega_S$ is the elementary solid angle. Hence,

$$d\gamma_0 = p^2dpdVd\Omega_S. \tag{21.6}$$

Formula (21.6) determines the accessible phase volume for a particle located in the volume dV and having a momentum vector with the absolute value ranging from p to $p + dp$ and directed within the limits of the solid angle $d\Omega_S$

If all directions of particle motion and all coordinates of particle location in the volume V are equiprobable, the accessible phase volume of the particle can be calculated,

$$d\gamma = p^2dp\int d\Omega_S\int dV = 4\pi p^2Vdp. \tag{21.7}$$

Besides, if the particle is not relativistic and has a rest mass m different from zero, its kinetic energy is

$$\varepsilon = \frac{p^2}{2m}.$$

Determining p and dp in terms of ε and $d\varepsilon$, and substituting them into (21.7), we obtain

$$d\gamma = 4\pi m\sqrt{2m\varepsilon}Vd\varepsilon, \tag{21.8}$$

which evaluate the accessible phase volume $d\gamma$ for a particle with a nonzero rest mass that freely moves in the volume V and has the kinetic energy within the interval from ε to $\varepsilon + d\varepsilon$.

Expression (21.8) is inapplicable for zero-mass particles (photons, phonons, and so forth). Consequently, formula (21.7) has to be used in such cases.

Now, let us find the accessible phase volume for the classical ideal gas consisting of N atoms and occupying the volume V. On the basis of (21.2) and (21.4), we may write

$$d\Gamma = \prod_{i=1}^{3N} dq_i dp_i = \prod_{i=1}^{N} dV_i dV_{pi}. \tag{21.9}$$

The phase volume the corresponds to all possible coordinates and momenta equals

$$\Gamma = \int \prod_{i=1}^{N} dV_i dV_{pi} = \prod_{i=1}^{N} \int dV_i \int \prod_{j=1}^{N} dV_{pj}. \tag{21.10}$$

Since all atoms move in the same volume V,

$$\prod_{i=1}^{N} \int_0^V dV_i = \prod_{i=1}^{N} V = V^N,$$

therefore,

$$\Gamma = V^N \int \prod_{j=1}^{N} dV_{pj}. \tag{21.11}$$

The internal energy of ideal single-atomic gas is a sum of the kinetic energies of all atoms, i.e.

$$E = \sum_{i=1}^{N} \frac{p_i^2}{2m} = \frac{1}{2m} \sum_{i=1}^{N} \left(p_{xi}^2 + p_{yi}^2 + p_{zi}^2\right). \tag{21.12}$$

From formula (21.12), we obtain the relation

$$\sum_{i=1}^{N} \left(p_{xi}^2 + p_{yi}^2 + p_{zi}^2\right) = 2mE, \tag{21.13}$$

which is an equation of a $3N$-dimensional spherical surface of the radius $R = \sqrt{2mE}$ in the momentum space. For comparison, the equation for the spherical surface with the radius R in the three-dimensional space looks like $x^2 + y^2 + z^2 = R^2$.

Hence, the integral in (21.11) is equal to the volume of $3N$-dimensional sphere with the radius $R = \sqrt{2mE}$. In the case of three-dimensional sphere, its volume equals $\frac{4}{3}\pi R^3$, i.e. it is proportional to R^3. Then, by analogy, we may assert that the volume of $3N$-dimensional sphere has to be proportional to $R^{3N} = (2mE)^{3N/2}$ or to $E^{3N/2}$. Let the coefficient of proportionality be designated as A_N. Its calculation requires more complicated calculations, which we do not consider (see problem 4.14 in [22]). Therefore,

$$\int \prod_{i=1}^{N} dV_{pi} = A_N E^{\frac{3N}{2}} \tag{21.14}$$

and

$$\Gamma = A_N E^{\frac{3N}{2}} . \tag{21.15}$$

Formula (21.15) evaluates the phase volume of the gas with the energy changing from 0 to E. However, the most important case is when the gas energy changes in a narrow interval ΔE around the value E ($\Delta E \ll E$). The, the accessible phase volume equals

$$\Delta\Gamma = \frac{d\Gamma}{dE}\Delta E = \frac{3N}{2} A_N V^N E^{\frac{3N}{2} - 1} \Delta E. \tag{21.16}$$

21.3. Quasiclassical approximation

Any macrosystem consisting of microparticles can be regarded as classical only approximately, provided that the quantum numbers describing it are large. However, even in this case, the quantum (wave) properties of microparticles, which are responsible for the uncertainty of particle coordinates and momenta, have to be take into account. Hence, the Heisenberg uncertainty relation (see [20]) must be to taken into consideration. In statistical physics, this relation is written as follows:

$$\Delta q_i \Delta p_i = 2\pi\hbar. \tag{21.17}$$

As a result of the coordinate and momentum uncertainty, the phase point, which represents a microstate in the phase space, “spreads” in a so-called

quantum cell with the volume

$$\Delta\Gamma_0 = \prod_{i=1}^{fN} \Delta q_i \Delta p_i = \prod_{i=1}^{fN} (2\pi\hbar) = (2\pi\hbar)^{fN} . \tag{21.18}$$

Therefore, the microstate is represented by a quantum cell rather than a phase point. Then, the number of accessible microstates Ω for the macrosystem can be determined by the formula

$$\Omega = \frac{\Delta\Gamma}{\Delta\Gamma_0} = \frac{\Delta\Gamma}{(2\pi\hbar)^{fN}}, \tag{21.19}$$

where $\Delta\Gamma$ is the accessible phase volume. Hence, the account of the uncertainty relation makes it possible to combine the application of coordinates, momenta; and the phase space, i.e. relate the attributes of classical system with the quantum-mechanical properties of microparticles. This method used in statistical physics is called the **quasiclassical approximation**.

If the macroscopic system consists of N identical particles, the identity principle, according to which the mutual particle permutations do not change the quantum state (the microstate), has also to be taken into account. The number of possible permutations for N particles equals $N!$. Therefore, instead of (21.19), we must write

$$\Omega = \frac{\Delta\Gamma}{N!\,(2\pi\hbar)^{fN}}. \tag{21.20}$$

Substituting expression (21.16) for $\Delta\Gamma$ into (21.20), we obtain the number of accessible microstates for the classical single-atomic ideal gas with the energy in a narrow interval ΔE,

$$\Omega = \frac{3NA_N V^N E^{\frac{3N}{2}-1} \Delta E}{2N!\,(2\pi\hbar)^{3N}}. \tag{21.21}$$

Formula (21.21) allows one to estimate the number of accessible microstates. Leaving the details of calculation aside (see section 4.3 in [22]), we note that this number is extremely large ($\Omega \sim 10^{10^{23}}$), i.e. the macrostate is realized by a huge number of different microstates.

In the case of one particle with three degrees of freedom, its phase space is six-dimensional, and the volume of quantum cell, in accordance with (21.18), equals

$$\Delta\gamma_0 = (2\pi\hbar)^3 . \tag{21.22}$$

Then, according to (21.19), the number of accessible microstates for the particle is

$$dg_0 = \frac{d\gamma}{\Delta\gamma_0}, \tag{21.23}$$

where $d\gamma$ is the accessible phase volume (21.7) or (21.8).

If the particle has a non-zero rest mass and a kinetic energy within the interval from ε to $\varepsilon + d\varepsilon$, then, in accordance with (21.8),

$$dg_0 = \frac{4\pi m\sqrt{2m\varepsilon}V\,d\varepsilon}{(2\pi\hbar)^3}. \tag{21.24}$$

Taking into account the particle spin l_s, expression (21.24) has to be multiplied by the number of spin projections, which equals $2l_s + 1$. As a result, instead of (21.24), we obtain the number of accessible microstates for a particle with the spin l_s and the kinetic energy ε in the interval $d\varepsilon$ in the form

$$dg = aV\sqrt{\varepsilon}d\varepsilon, \tag{21.25}$$

where

$$a = \frac{\sqrt{m^3}\,(2l_s + 1)}{\sqrt{2}\pi^2\hbar^3}. \tag{21.26}$$

In the case of a zero-mass particle (photon, phonon, and so forth), after substituting (21.7) into (21.23), the result has to be multiplied by the number of independent polarizations of the particle χ, which equals 2 for the photon, and 3 for the phonon. As a result, we obtain

$$dg = \frac{4\pi p^2 V\,dp}{(2\pi\hbar)^3}\chi = \frac{V\chi}{2\pi^2\hbar^3}p^2 dp. \tag{21.27}$$

21.4. Statistical entropy

In the closed system, all accessible microstates correspond to the same energy. Therefore, it is quite reasonable to suppose that all accessible microstates are equiprobable in an equilibrium state. On the contrary, if the accessible microstates of the closed system are not equiprobable, this system is in a nonequilibrium state.

As an example, let us consider a vessel of the volume V separated by a partition into two equal parts. Let one of them is occupied by a single-atomic ideal gas in the equilibrium state and consisting of N molecules, and the other contains vacuum. According to (21.21), the number of accessible microstates for the gas is proportional to the volume raised to the power N. Therefore, in this case, we may write

$$\Omega_1 = B\left(\frac{V}{2}\right)^N,$$

where the letter B stands for all other multipliers. All Ω_1 accessible microstates are equiprobable.

Now, let the partition be instantly removed. Then, the number of accessible microstates changes to

$$\Omega_2 = BV^N.$$

However, not all of them are equiprobable at the first moment: in particular, the probability of microstates that correspond to the occupation of the whole volume V by the gas still equals zero. Therefore, right after the partition has been removed, the gas turns out in the nonequilibrium state and starts to expand into vacuum, keeping its energy constant. After the relaxation time τ, the gas fills the vacuum part, an all its accessible microstates become equiprobable, so that the gas is in the equilibrium state.

It is important that, after the equilibrium state has been established, the number of accessible and equiprobable states grows enormously; namely, by a factor of $\Omega_2/\Omega_1 = 2^N$ ($N \sim 10^{23}$). The number of new microstates that arose after the gas expansion,

$$\Delta\Omega = \Omega_2 - \Omega_1 = \left(2^N - 1\right)\Omega_1,$$

is much larger than their initial number Ω_1. Therefore, after the expansion, the gas is almost always in new microstates that correspond to the filling of the whole volume. In other words, the irreversibility of gas expansion is explained by the fact that the macrostate, in which the gas fills the whole volume, is realized by an incomparably larger number of microstates than the macrostate, in which the gas fills half the volume.

However, both the number of accessible microstates and its growth are determined by too huge numbers: of an order of $10^{10^{23}}$. Besides, the number of accessible microstates is multiplicative, which is also associated with certain disadvantages. In this connection, a new dimensionless quantity is introduced,

$$\sigma = \ln \Omega, \tag{21.28}$$

which is additive, because the logarithm of the product equals the sum of the logarithms of multipliers. This quantity is of an order of 10^{23}. When the equilibrium state is established, Ω becomes larger and, hence, σ grows ($d\sigma > 0$), whereas in the equilibrium state σ has the maximum value ($d\sigma = 0$). In general,

$$d\sigma \geq 0. \tag{21.29}$$

The quantity σ is called the **statistical entropy**, and relation (21.29) the **law of entropy growth**.

In the case of classical system, it is necessary to use the quasiclassical approximation (see section 21.3). Therefore, in accordance with (21.19), the statistical entropy looks like

$$\sigma = \ln \frac{\Delta\Gamma}{(2\pi\hbar)^{fN}}. \tag{21.30}$$

On the basis of (21.28) and (21.21), the statistical entropy for the classical single-atomic ideal gas reads

$$\sigma = \ln \frac{3NA_N V^N E^{\frac{3N}{2} - 1} \Delta E}{2N!\,(2\pi\hbar)^{3N}}. \tag{21.31}$$

21.5. Statistical distribution function

As was already mentioned, the main task of statistical physics and thermodynamics is the calculation of equilibrium macroscopic parameters of the system. In thermodynamics, this task is solved with the help of equations that are based on thermodynamics laws. It should be emphasized separately that, in the framework of the phenomenological thermodynamics and according to its first postulate, spontaneous deviations of macroparameters from their equilibrium values are impossible.

Let us consider the calculation procedure for equilibrium macroparameters in statistical physics. The equilibrium state of the system is associated with a certain number of accessible microstates Ω. However, any equilibrium system always interacts with its environment, by exchanging the energy, particles, and so on with it. This statement concerns the closed systems as well, because the absolutely closed system is an idealized concept. They do not exist, as well as harmonic oscillations or the ideal gas, being one of the models of macroscopic systems.

Hence, in the case of a system in the equilibrium state, the number of its accessible microstates and their probabilities permanently and very quickly change within certain, relatively small limits. As a result, the instant values of macroparameters, which are functions of microscopic variables, also change very quickly ("quiver") around some average values. Such short-time deviations of macroparameters from their average values are called **fluctuations**. At the same time, devices intended to measure parameters are macroscopic units. Owing to the inertia at measurements, they cannot response quickly to the change of instant macroparameter values. Hence, they measure their average values. Therefore, we arrive at an important principle of statistical physics consisting un that the equilibrium (measured) values of macroparameters have the meaning of averaged ones.

If the measured value of some macroparameter is designated as F, we may write

$$F = \overline{F(t)} = \frac{1}{\Delta t} \int_0^{\Delta t} F(t)\, dt, \tag{21.32}$$

where $F(t)$ is the instant parameter value, and Δt is the time interval within which the averaging is carried out. However, the instant values $F(t)$ change chaotically as a rule. Therefore, the function $F(t)$ is unknown and formula (21.32) cannot be used to calculate average values. In this connection, J.W. Gibbs proposed a method, in which the averaging in time is substituted by averaging over a statistical ensemble.

By a **statistical ensemble** is meant a collection of a very large number N_0 ($N_0 \to \infty$) of identical systems (the "copies" of the given system), which are in the same macroscopic state. At the same time, their microstates change independently, an at every time moment they are in different microstates, i.e. they are characterized by different instant values of microparameters.

Suppose first that the systems in the ensemble are quantum-mechanical, i.e. their microstates are determined by a set of quantum numbers (n_1, $n_2, \ldots, n_{fN}$). Below, this set will be designated as n, and the value of some macroparameter corresponding to this state as F_n. Let us imagine that we simultaneously fixed the microstates of all N_0 systems in the ensemble and calculated the number N_n of systems in the n-th microstate. Then, the ratio

$$w_n = \frac{N_n}{N_0}$$

determines the probability to find any system of the ensemble in the n-th microstate with the macroparameter value F_n. The average value of this parameter equals

$$\langle F\rangle = \sum_n F_n w_n. \tag{21.33}$$

where summation is carried out overall accessible microstates. The probability of microstate w_n is called the **statistical distribution function**. It is a key concept of statistical physics.

In the case of classical system, the microstate is determined by a set of fN coordinates q_i and fN momenta p_i, which will be designated as (q,p). Then, the macroparameter $F(q,p)$ is a function of coordinates and momenta, and, instead of (21.33), we should write

$$\langle F\rangle = \int F\,(q,p)\,dW\,(q,p)\,, \tag{21.34}$$

where $dW\,(q,p)$ is the probability for the phase point corresponding to the microstate (q,p) to be in the phase volume $d\Gamma$. Taking into account that

$$dW\,(q,p) = \frac{dW\,(q,p)}{d\Gamma}d\Gamma = \rho\,(q,p)\,d\Gamma,$$

we obtain

$$\langle F\rangle = \int F\,(q,p)\,\rho\,(q,p)\,d\Gamma, \tag{21.35}$$

where integration is carried out over thw whole phase space. The function $\rho\,(q,p)$ has the sense of probability density for microstates in the phase space and is a statistical distribution function for the classical system. The statistical distribution function (w_n or $\rho\,(q,p)$) is normalized to unity,

$$\sum_B w_n = 1, \qquad \int \rho\,(q,p)\,d\Gamma = 1. \tag{21.36}$$

The quantity $\langle F\rangle$ is the parameter value averaged over the statistical ensemble, and the quantity $\bar{F}$ the value averaged in time. Therefore, there arises a question concerning their relation. In statistical physics, it is adopted that $\langle F\rangle = \bar{F}$. This statement is called the **ergodic hypothesis**. In what follows, will consider that $\langle F\rangle$ and $\bar{F}$ concern the same average value.

Hence, the statistical distribution function is determined by a set of quantum numbers n for the quantum-mechanical system and by a set a set of coordinates and momenta (q, p) for the classical one. Besides, those sets also determine the system energy, $E = E_n$ or $E(q, p)$. Therefore, it can be proved (see, e.g., [10]) that the statistical distribution function has to depend on the system energy, which, in turn, depends on either the quantum numbers or the coordinates and momenta, i.e.

$$w_n = w\left(E_n\right), \tag{21.37}$$

$$\rho\left(q, p\right) = \rho\left(E\left(q, p\right)\right). \tag{21.38}$$

As was already mentioned above, the definition of macroparameters (the measured values) as the statistically averaged quantities brings us to a conclusion about the deviation of their instant values from the average ones. Those deviations are called **fluctuations**. Let us designate the deviation of the parameter F from its average value $\bar{F}$ as

$$\Delta F = F - \bar{F}.$$

At different time moments, those deviations have different numerical values and different signs; hence, $\overline{\Delta F} = 0$. Therefore, a measure for the average value of deviation is the root-mean-square deviation, which is called the **root-mean-square fluctuation,**

$$\delta_F = \sqrt{\overline{\left(\Delta F\right)^2}} = \sqrt{\overline{\left(F - \bar{F}\right)^2}} = \sqrt{\overline{F^2} - \bar{F}^2}. \tag{21.39}$$

While estimating fluctuations, the concept of **relative fluctuation,**

$$\eta_F = \frac{\delta_F}{\bar{F}}, \tag{21.40}$$

is also of great importance. According to the definition, if the parameter F is additive,

$$F = \sum_{i=1}^{N} f_i,$$

where f_i is the contribution to F from the i-th particle. Then,

$$\Delta F = \sum_{i=1}^{N} \Delta f_i .$$

Therefore,

$$\overline{(\Delta F)^2} = \overline{\left(\sum_i \Delta f_i\right)^2} = \overline{\sum_{i,\,j} \Delta f_i \Delta f_j} = \sum_{i,\,j} \overline{\Delta f_i \Delta f_j}.$$

Grouping the summands with $j = i$ and the summands with $j \neq i$, and taking into account that, if $j \neq i$,

$$\overline{\Delta f_i \Delta f_j} = 0,$$

we obtain

$$\overline{(\Delta F)^2} = \sum_{i=1}^{N} (\Delta f_i)^2 \sim N. \tag{21.41}$$

Hence, $\delta_F \sim N$. On the other hand,

$$\overline{F} = \overline{\sum_i f_i} = \sum_i \bar{f}_i \sim N,$$

so that

$$\eta_F \sim \frac{1}{\sqrt{N}}. \tag{21.42}$$

From (21.41), it follows that fluctuations grows by absolute vaue as $\sqrt{N}$ in the macroscopic system, but the relative fluctuations diminish as $1/\sqrt{N}$ (see (21.42)), becoming very small at large N.

21.6. Microcanonical Gibbs distribution

In order to calculate macroparameters with the help of formulas (21.33) and (21.35), we must know the statistical distribution function: $w_n = w\,(E_n)$ for quantum-mechanical systems or $\rho = \rho\,(E)$ for classical ones.

First, let us determine the statistical distribution function for the closed quantum-mechanical system. As was already mentioned in section 21.5, ideally closed systems do not exist, because it is impossible to completely avoid interaction between the system and its environment. Moreover, the system cannot have only one, absolutely exact value of the energy owing to the uncertainty relation for energy and time. Therefore, as the closed system, let us consider a system, the energy of which energy can changein a very narrow interval from E to $E + \delta E$. The magnitude of δE has to be very small in comparison with the energy E. Then, from the macroscopic viewpoint, the energy E of this system can be regardeded as constant, and all its accessible microstates equiprobable, in accordance with the postulate about the equal probability of microstates with the same energy. At the same time, from the microscopic viewpoint, the magnitude of δE has to be large: much larger than the distance between the neighbor energy levels. Then, the interval δE includes a very large number of energy levels and, hence, microstates.

Let us consider an energy level E belonging to the interval δE ($E \in \delta E$) and suppose that it includes $\Omega\,(E)$ quantum states (microstates). Then, the total numberof accessible microstates within the energy interval δE can be written as

$$\Omega_0 = \sum_{E \in \delta E} \Omega\,(E)\,. \tag{21.43}$$

All those microstates are equiprobable. Therefore, the probability for every of them equals $1/\Omega_0$, and the probability for any microstate from those corresponding to the energy E

$$w\,(E) = \frac{1}{\Omega_0}\Omega\,(E) = C\Omega\,(E)\,, \quad C = \frac{1}{\Omega_0} = \text{const}, \tag{21.44}$$

i.e. it is proportional to the number of microstates $\Omega\,(E)$. Formula (21.44) is called the **microcanonical Gibbs distribution** for the quantum-mechanical system. The energy profile of microcanonical distribution is schematically shown in Fig. 21.1.

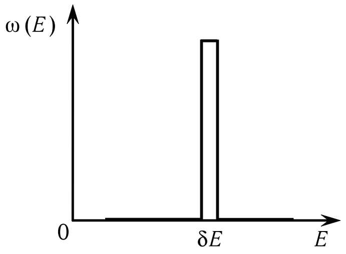

Fig. 21.1.

In the case of classical system, the energy interval δE corresponds to the accessible phase volume $\Delta\Gamma$ in the phase space.

Therefore, the statistical distribution function looks like

$$\rho = \frac{1}{\Delta\Gamma},$$

and the probability of microstate, the phase point of which is in the volume $d\Gamma$, equals

$$dW = \frac{d\Gamma}{\Delta\Gamma} = \frac{1}{\Delta\Gamma}\frac{d\Gamma}{dE}dE. \tag{21.45}$$

One can see that the microcanonical distribution is normalized to unity. For the quantum-mechanical system, it follows from (21.43) and (21.44) that

$$\sum_E w\left(E\right) = \frac{1}{\Omega_0}\sum_E \Omega\left(E\right) = \frac{1}{\Omega_0}\Omega_0 = 1.$$

On the other hand, for the classical system, on the basis of (21.45), we obtain

$$\int dW = \frac{1}{\Delta\Gamma}\int d\Gamma = \frac{1}{\Delta\Gamma} = 1.$$

21.7. Canonical Gibbs distribution

The microcanonical distribution considered in section 21.6 is applicable to closed systems. However, much more important is the case of equilibrium system that freely exchanges the energy with its environment; the latter is conditionally called the **thermostat**. The thermostat and the examined system form a larger system, which is assumed to be closed and at equilibrium. The analyzed system becomes a small subsystem of the large closed system.

Let the energy of the large closed system be designated as E, and the energy of the analyzed subsystem as ε. Then, neglecting the energy of interaction between the subsystem and the thermostat, the energy of the latter can be written in the form $E - \varepsilon$ According to (21.44), the microcanonical distribution for the large closed system (the subsystem and the thermostat together) looks like

$$w = C\Omega\left(E\right).$$

Taking into account that the number of accessible microstates is a multiplicative quantity, i.e.

$$\Omega\left(E\right)=\Omega\left(\varepsilon\right)\Omega_0\left(E-\varepsilon\right),$$

where $\Omega\left(\varepsilon\right)$ is the number of accessible microstates in the subsystem, and $\Omega_0\left(E-\varepsilon\right)$ the number of accessible microstates in the thermostat, we obtain

$$w\left(\varepsilon\right)=C\Omega\left(\varepsilon\right)\Omega_0\left(E-\varepsilon\right). \tag{21.46}$$

Here, we took into account that the energy of closed system E can be considered constant (see section 21.6), and the energy of subsystem ε variable. Normalizing probability (21.46), we have

$$\sum_{\varepsilon}w\left(\varepsilon\right)=C\sum_{\varepsilon}\Omega\left(\varepsilon\right)\Omega_0\left(E-\varepsilon\right)=1,$$

hence,

$$w\left(\varepsilon\right)=\frac{\Omega\left(\varepsilon\right)\Omega_0\left(E-\varepsilon\right)}{\sum_{\varepsilon}\Omega\left(\varepsilon\right)\Omega_0\left(E-\varepsilon\right)}. \tag{21.47}$$

The entropy of thermostat, in accordance with (21.28), equals

$$\sigma\left(E-\varepsilon\right)=\ln\Omega_0\left(E-\varepsilon\right).$$

Therefore,

$$\Omega_0\left(E-\varepsilon\right)=\mathrm{e}^{\sigma(E-\varepsilon)}.$$

Under the condition $\varepsilon\ll E$, the entropy $\sigma\left(E-\varepsilon\right)$ can be expanded in a power series of ε up to the linear term,

$$\sigma\left(E-\varepsilon\right)\approx\sigma\left(E\right)-\varepsilon\left(\frac{\partial\sigma}{\partial E}\right)_{\varepsilon=0}.$$

Then, we obtain

$$\Omega_0\left(E-\varepsilon\right)=\mathrm{e}^{\sigma(E)}\mathrm{e}^{-\dfrac{\varepsilon}{\theta}}, \tag{21.48}$$

where the notation

$$\frac{1}{\theta}=\left(\frac{\partial\sigma}{\partial E}\right)_{\varepsilon=0} \tag{21.49}$$

was used. Substituting (4.50) into (4.49), we ultimately obtain

$$w\left(\varepsilon\right)=\frac{\Omega\left(\varepsilon\right)\mathrm{e}^{-\dfrac{\varepsilon}{\theta}}}{\displaystyle\sum_{\varepsilon}\Omega\left(\varepsilon\right)\mathrm{e}^{-\dfrac{\varepsilon}{\theta}}}. \tag{21.50}$$

Distribution (21.50) describing the probability for a system in the thermostat to be in a state with the energy ε is called the **canonical Gibbs distribution**, and the parameter θ the **modulus of canonical distribution**.

The function $w\left(\varepsilon\right)$ gives the probability of any microstate among those with the energy ε, i.e. from $\Omega\left(\varepsilon\right)$ microstates. The number of microstates $\Omega\left(\varepsilon\right)$ equals the number of different quantum states with the same energy ε. Therefore, the probability of a definite quantum state equals

$$w\left(\varepsilon_{\alpha}\right)=\frac{\mathrm{e}^{-\dfrac{\varepsilon_{\alpha}}{\theta}}}{\displaystyle\sum_{\alpha}\mathrm{e}^{-\dfrac{\varepsilon_{\alpha}}{\theta}}}, \tag{21.51}$$

where α enumerates the quantum states.

Formula (21.50) determines the distribution over the energies, and formula (21.51) over the quantum states. The denominator in canonical distribution (21.50),

$$Z=\sum_{\varepsilon}\Omega\left(\varepsilon\right)\mathrm{e}^{-\dfrac{\varepsilon}{\theta}}, \tag{21.52}$$

is called the **partition function**. It plays an important role, because the thermodynamic parameters of the system can be directly calculated from it.

Let us determine the energy profile of canonical distribution (21.50). The denominator in (21.50) does not depend on the energy ε; therefore, the profile is determined by the numerator, which is a product of two functions. On the basis of formula (21.21), the function $\Omega\left(\varepsilon\right)$ has the general form $\Omega\left(\varepsilon\right)\sim\varepsilon^{\alpha N}$, where α is a coefficient close to unity, and $N\sim10^{23}$.

Therefore, the number of accessible microstates $\Omega(\varepsilon)$ increases extremely repidly as the energy ε grows. On the contrary, the function $\mathrm{e}^{-\varepsilon/\theta}$ rapidly decreases as ε increases. Therefore, the product $\Omega(\varepsilon)\,\mathrm{e}^{-\varepsilon/\theta}$ turns out considerably different from zero in a very narrow interval of energy, and its plot looks like a very sharp and narrow peak, which cannot be exhibited in any scale. Schematically, the plot of canonical distribution is depicted in Fig. 21.2.

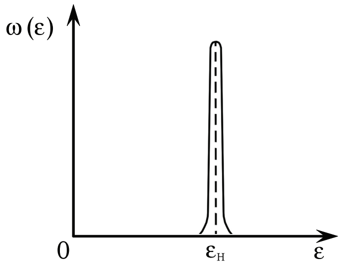

Fig. 21.2.

By comparing Figs. 21.1 and 21.2, we come to a conclusion that the divergences between the microcanonical and canonical Gibbs distributions are insignificant. The curve maximum in Fig. 21.2 is the most probable energy ε_m, and the peak is so narrow that $w(\varepsilon) \approx 0$ if $\varepsilon \neq \varepsilon_m$. In other words,

$$w(\varepsilon) \approx \delta_{\varepsilon,\,\varepsilon_m}, \tag{21.53}$$

where $\delta_{\varepsilon,\,\varepsilon_m}$ is the Kronecker delta-symbol,

$$\delta_{\varepsilon,\,\varepsilon_m} = \begin{cases} 1, & \varepsilon = \varepsilon_m, \\ 0, & \varepsilon \neq \varepsilon_m. \end{cases}$$

As a result of relation (21.53), the average value of system energy in the thermostat, $\bar{\varepsilon}$, coincides with its most probable value, ε_m:

$$\bar{\varepsilon} = \sum_{\varepsilon} \varepsilon w(\varepsilon) \approx \sum_{\varepsilon} \varepsilon \delta_{\varepsilon,\,\varepsilon_m} = \varepsilon_m. \tag{21.54}$$

Expression (21.54) gives rise to an important conclusion that the average value of any quantity depending on the energy ε, the following relation is obeyed:

$$\bar{F} = \sum_{\varepsilon} F(\varepsilon)\, w(\varepsilon) \approx \sum_{\varepsilon} F(\varepsilon)\, \delta_{\varepsilon,\,\varepsilon_m} = F(\varepsilon_m) \approx F(\bar{\varepsilon}). \tag{21.55}$$

On this basis, partition function (21.52) can be rewritten in the form

$$Z = \sum_{\varepsilon} \Omega(\varepsilon)\, \mathrm{e}^{-\dfrac{\varepsilon}{\theta}} \approx \Omega(\bar{\varepsilon})\, \mathrm{e}^{-\dfrac{\bar{\varepsilon}}{\theta}}. \tag{21.56}$$

21.8. Transition to classical statistics

The density of quantum energy levels in macroscopic systems is usually so high that the energy of the system in the thermostat, which will by denoted as E, can be considered as a continuous quantity, and the system itself as classical (in the quasiclassical approximation). This makes it possible to substitute the summation over the energy by the integration. Therefore, let us change from the quantum-mechanical canonical Gibbs distribution (21.50) to the classical canonical distribution, which forms the basis of **classical statistics**.

Supposing the energy E to be a continuous quantity, we may introduce an infinitesimal interval of energy, dE, which includes $d\Omega\,(E)$ accessible microstates. Then, the probability of a microstate with the energy within the interval from E to $E + dE$ can be determined making use of (21.50), in which $\Omega\,(E)$ is substituted by $d\Omega\,(E)$ and the summation in the denominator by the integration,

$$dW\,(E) = \frac{\mathrm{e}^{-\dfrac{E}{\theta}}\,d\Omega\,(E)}{\displaystyle\int \mathrm{e}^{-\dfrac{E}{\theta}}\,d\Omega\,(E)}. \tag{21.57}$$

If the system consists of N identical particles, on the basis of (21.20), we write

$$d\Omega\,(E) = \frac{d\Gamma\,(E)}{N!\,(2\pi\hbar)^{fN}}. \tag{21.58}$$

Substituting (21.58) into (21.57), we obtain the **canonical Gibbs distribution for the classical system**,

$$dW\,(E) = \frac{\mathrm{e}^{-\dfrac{E}{\theta}}\,d\Gamma}{\displaystyle\int \mathrm{e}^{-\dfrac{E}{\theta}}\,d\Gamma}. \tag{21.59}$$

It evaluates the probability that the energy E of the system in the thermostat is in the interval dE, and the phase point that corresponds to the system microstate is in the element $d\Gamma\left(E\right)$ of phase space. The denominator of the classical canonical distribution, i.e. the integral

$$J=\int \mathrm{e}^{-\dfrac{E}{\theta}}\,d\Gamma \tag{21.60}$$

is called the **statistical integral**.

Changing to classical statistics, instead of (21.52), we may write (see (21.57))

$$Z=\int \mathrm{e}^{-\dfrac{E}{\theta}}\,d\Omega\left(E\right).$$

Then, taking into account (21.58), we obtain a formula for the calculation of partition function in the case of classical system (in the quasiclassical approximation),

$$Z=\frac{1}{N!\left(2\pi\hbar\right)^{fN}}\int \mathrm{e}^{-\dfrac{E}{\theta}}\,d\Gamma. \tag{21.61}$$

From distribution (21.59), it follows an expression for the classical statistical distribution function,

$$\rho\left(E\right)=\frac{dW}{dE}=\frac{\mathrm{e}^{-\dfrac{E}{\theta}}\dfrac{d\Gamma}{dE}}{\displaystyle\int \mathrm{e}^{-\dfrac{E}{\theta}}\dfrac{d\Gamma}{dE}dE}. \tag{21.62}$$

By applying the reasonings used in section 21.7, it is easy to prove that the plot of distribution function (21.62) has the same appearance as the plot in Fig. 21.2.

22

STATISTICAL THERMODYNAMICS

22.1. Statistical temperature. Boltzmann principle

Proceeding to the statistical analysis and the substantiation of thermodynamic concepts and laws, we must, first of all, determine the meaning and the properties of the modulus of canonical distribution θ. According to (21.49), this parameter is introduced by the relation

$$\frac{1}{\theta} = \left(\frac{\partial\sigma}{\partial E}\right)_{\varepsilon=0}, \tag{22.1}$$

where $\sigma = \sigma\,(E-\varepsilon)$ is the statistical entropy of thermostat; E the energy of closed system; and ε the energy of the subsystem in the closed system. Hence, the parameter θ characterizes the properties of the thermostat, with which the subsystem is at equilibrium, i.e. it is a characteristic of the equilibrium macroscopic system.

From the canonical Gibbs distribution (21.50),

$$w\,(\varepsilon) = \frac{1}{Z}\,\mathrm{e}^{-\dfrac{\varepsilon}{\theta}}\,\Omega\,(\varepsilon)\,,$$

where

$$Z = Z\,(\theta) = \sum_{\varepsilon} \Omega\,(\varepsilon)\,\mathrm{e}^{-\dfrac{\varepsilon}{\theta}}$$

is the partition function (21.52), it follows that θ is an essentially positive quantity; otherwise, the probability of microstate would infinitely grow with the increase of the microstate energy ε.

Let us demonstrate that the parameter θ characterizes the equilibrium state. For this purpose, let us consider two equilibrium subsystems with the energies ε_1 and ε_2 in thermostats with the moduli of canonical Gibbs distributions θ_1 and θ_2, respectively, for which

$$w_1\,(\varepsilon_1) = \frac{1}{Z_1\,(\theta_1)}\,\mathrm{e}^{-\dfrac{\varepsilon_1}{\theta_1}}\,\Omega_1\,(\varepsilon_1)\,;$$

$$w_2(\varepsilon_2) = \frac{1}{Z_2(\theta_2)} \mathrm{e}^{-\frac{\varepsilon_2}{\theta_2}} \Omega_2(\varepsilon_2).$$

Let those systems exchange the energy and let us consider them as a single combined subsystem with the energy $\varepsilon = \varepsilon_1 + \varepsilon_2$ (the energy of interaction is neglected). Then, the probability of a microstate of combined subsystem, according to the theorem on the product of probabilities, can be written as

$$w(\varepsilon_1 + \varepsilon_2) = w_1(\varepsilon_1)\, w_2(\varepsilon_2) = \frac{\Omega_1(\varepsilon_1)\,\Omega_2(\varepsilon_2)}{Z_1(\theta_1)\, Z_2(\theta_2)} \mathrm{e}^{-\frac{\varepsilon_1}{\theta_1}} \mathrm{e}^{-\frac{\varepsilon_2}{\theta_2}}. \qquad (22.2)$$

As a result of the multiplicativity of the number of accessible microstates,

$$\Omega_1(\varepsilon_1)\,\Omega_2(\varepsilon_2) = \Omega(\varepsilon_1 + \varepsilon_2).$$

Moreover,

$$Z_1(\theta_1)\, Z_2(\theta_2) = \sum_{\varepsilon_1,\, \varepsilon_2} \Omega_1(\varepsilon_1)\,\Omega_2(\varepsilon_2)\, \mathrm{e}^{-\frac{\varepsilon_1}{\theta_1}} \mathrm{e}^{-\frac{\varepsilon_2}{\theta_2}} =$$

$$= \sum_{\varepsilon_1,\, \varepsilon_2} \Omega(\varepsilon_1 + \varepsilon_2) = \mathrm{e}^{-\frac{\varepsilon_1}{\theta_1}} \mathrm{e}^{-\frac{\varepsilon_2}{\theta_2}}.$$

Hence, we obtain

$$w(\varepsilon_1 + \varepsilon_2) = \frac{\Omega_1(\varepsilon_1 + \varepsilon_2)}{\sum\limits_{\varepsilon_1, \varepsilon_2} \Omega(\varepsilon_1 + \varepsilon_2)\, \mathrm{e}^{-\frac{\varepsilon_1}{\theta_1}} \mathrm{e}^{-\frac{\varepsilon_2}{\theta_2}}} \mathrm{e}^{-\frac{\varepsilon_1}{\theta_1}} \mathrm{e}^{-\frac{\varepsilon_2}{\theta_2}}. \qquad (22.3)$$

Let $\theta_1 = \theta_2 = \theta$. Then, taking into account that $\varepsilon_1 + \varepsilon_2 = \varepsilon$, takes the form of canonical Gibbs distribution,

$$w(\varepsilon) = \frac{\Omega(\varepsilon)\, \mathrm{e}^{-\frac{\varepsilon}{\theta}}}{\sum\limits_{\varepsilon} \Omega(\varepsilon)\, \mathrm{e}^{-\frac{\varepsilon}{\theta}}}.$$

Hence, under the condition $\theta_1 = \theta_2 = \theta$, the combined subsystem is in the equilibrium state. If $\theta_1 \neq \theta_2$, distribution function (22.3) has not the form of canonical distribution; therefore, the combined subsystem is nonequilibrium.

From the reasoning given above, it follows that the equilibrium condition consists in the equality of the parameters θ_1 and θ_2. Hence, the parameter θ has the meaning of the temperature (see section 18.3), and it is called the **statistical temperature**.

According to (22.1), the statistical temperature θ is a function of energy E, i.e. $\theta = \theta(E)$, although the inverse dependence $E = E(\theta))$ is mostly used. Let us determine this relation for the classical single-atomic ideal gas. In this case, the statistical entropy, in accordance with (21.31), looks like

$$\sigma(E) = \ln \frac{3NA_N V^N E^{\frac{3N}{2} - 1} \Delta E}{2N! \left(2\pi\hbar\right)^{3N}}. \tag{22.4}$$

Therefore,

$$\frac{1}{\theta} = \frac{\partial\sigma}{\partial E} = \left(\frac{3N}{2} - 1\right)\frac{1}{E}. \tag{22.5}$$

Neglecting unity in comparison with $3N/2$, we obtain

$$\theta = \frac{2}{3}\frac{E}{N} \tag{22.6}$$

or

$$E = \frac{3}{2}N\theta. \tag{22.7}$$

Identifying the energy E with the internal energy of classical single-atomic ideal gas known from the course of general physics,

$$U = \frac{3}{2}Nk_0T, \tag{22.8}$$

where k_0 is the Boltzmann constant, we obtain the relation between the statistical and thermodynamic temperatures,

$$\theta = k_0T. \tag{22.9}$$

From the master thermodynamic equation (22.4), it follows that

$$\frac{1}{T} = \frac{\partial S}{\partial U}, \tag{22.10}$$

where U is the internal energy; and S the thermodynamic entropy of the system. Comparing (22.10) with the definition of statistical temperature,

$$\frac{1}{\theta} = \frac{\partial \sigma}{\partial E},$$

and taking into account that $E = U$ and $\theta = k_0 T$, we find the relation between the statistical and thermodynamic entropies,

$$S = k_0 \sigma. \tag{22.11}$$

Since $\sigma = \ln \Omega$, formula (22.11) looks like

$$S = k_0 \ln \Omega. \tag{22.12}$$

The number of accessible microstates Ω (the other names of this parameter are the number of quantum states, the degeneracy order of the energy level E, and the statistical weight of microstate) can be considered as the number of ways to realize the macroscopic state with the energy E, the entropy S, and so on. Therefore, formula (22.12) is a link between statistical physics and thermodynamics and is called the **Boltzmann principle**. L. Boltzmann established it in 1872 using a method that reminded the method of quantum-mechanical cells and obtained the formula

$$S = k_0 \ln W_B, \tag{22.13}$$

where W_B is the number of ways to realize a macrostate. M. Planck called this parameter the **thermodynamic probability**, and it is an analog of the number of accessible microstates Ω.

As the temperature decreases, the internal energy of the system also decreases. However, it has a certain minimum value $E = E_{\text{min}}$. Since the function derivative equals zero in the minimum, we obtain

$$\theta_{\text{min}} = \left(\frac{\partial E}{\partial \sigma} \right)_{E = E_{\text{min}}} = 0.$$

Hence, the minimum value of statistical temperature equals zero. According to (22.9), the minimum value of thermodynamic temperature also equals zero. This value is called the **absolute zero temperature**.

22.2. Statistical meaning of the work and heat in the first law of thermodynamics

In section 21.1, it was shown that thermodynamic parameters have the sense of averaged statistical quantities. Therefore, the thermodynamic internal energy equals

$$U = \sum_E E w(E), \tag{22.14}$$

where E is the energy of macroscopic system, and $w(E)$ the probability of its microstate. For a system in the thermostat, the probability $w(E)$ is given by the canonical Gibbs distribution (21.50).

The quantum-mechanical energy levels depend on both the properties of the system and the external parameters. Therefore, their variation results in the change of quantum-mechanical energy levels as well.

Let the system have one external parameter a. The infinitesimally small variation δU of the internal energy owing to the variations of the parameter a itself and the probability $w(E)$ equals

$$\delta U = \sum_E w(E)(\delta E)_w + \sum_E E(\delta w)_a, \tag{22.15}$$

where $(\delta E)_w$ is the change of energy level owing to the variation of the parameter a at the constant probability of microstates $w(E)$, and $(\delta w)_a$ is the change of probability $w(E)$ at the constant external parameter a. First, let us elucidate the statistical meaning of the first summand in (22.15). The variation of energy E at an infinitesimally small variation δa of the parameter a equals

$$(\delta E)_w = \frac{\partial E}{\partial a}\delta a = -f(E)\,\delta a, \tag{22.16}$$

where

$$f(E) = -\frac{\partial E}{\partial a} \tag{22.17}$$

is a generalized force associated with the change of the energy level E. Then, the first summand in (22.15) takes the form

$$\sum_E w(E)(\delta E)_w = (\delta U)_w = -\sum_E f(E)\,w(E)\,\delta a = -\Lambda\delta a. \tag{22.18}$$

Here, the quantity

$$\Lambda = \sum_{E} f\left(E\right) w\left(E\right) \tag{22.19}$$

is the average generalized force that acts on the system when the parameter a varies (see section 18.2). The product

$$-\Lambda \delta a = \delta A_{\text{ext}} \tag{22.20}$$

equals the work done by the generalized force on the system. At the same time, the work done by the system equals

$$\delta A = -\delta A_{\text{ext}} = \Lambda \delta a. \tag{22.21}$$

If the volume V is the external parameter, the pressure p is the average generalized force. Then,

$$\delta A = pdV. \tag{22.22}$$

On the bases (22.18) and (22.21), we may write

$$\sum_{E} w\left(E\right) \left(\delta E\right)_{w} = -\delta A. \tag{22.23}$$

The second summand in (22.15) corresponds to the variation of system energy that is not associated with the change of external parameters, but is a consequence the direct energy transfer from environment particles to system ones, which changes the probability of microstates. The corresponding variation of the internal energy of the system is called the **heat**,

$$\delta Q = \sum_{E} E\left(\delta w\right)_{a}. \tag{22.24}$$

Hence, equation (22.15) takes the form

$$\delta U = \delta Q - \delta A. \tag{22.25}$$

This is an expression for the first law of thermodynamics in the differential form (see (19.4)).

It can be proved (see, e.g., section 5.2 in [22]) that the heat δQ can be obtained in terms of the statistical parameters of the system in the form

$$\delta Q = \theta \delta \sigma\left(U\right) = \theta d\sigma\left(U\right). \tag{22.26}$$

Taking into account that $\theta = k_0 T$ and $k_0 \sigma = S$, we obtain

$$\delta Q = TdS. \tag{22.27}$$

22.3. Statistical meaning of the second law of thermodynamics

The second law of thermodynamics (see section 19.2) declares the irreversibility of macroscopic processes that run at a finite rate and determines their direction. It will be recalled that the process is called **reversible** (**equilibrium**, **quasistatic**), if, on the one hand, it can be first executed in some direction through a sequence of states and, afterwards, in the opposite direction through the same states and, on the other hand, when the system returns to the initial state, no changes should take place in the environment (in surrounding objects).

The concept of state is applicable only to the equilibrium system. Therefore, the quoted definition of reversible process means that the system should be at equilibrium at every time momentum. However, this is possible only provided that the process runs infinitely slowly, because, only under this condition, the equilibrium state "has enough time" to be established at every moment.

However, all real processes run at a finite rate. Therefore, the equilibrium has no time to be established at every moment. Hence, real processes are nonequilibrium. As a result, when the system returns into the initial state, some changes take place in the environment, which are called **compensation**. From this viewpoint, gas expansion in vacuum was analyzed in section 19.2. Let us consider some more examples.

According to the first postulate of thermodynamics, only those processes can run in the closed system, which bring it to the equilibrium state. If the system consists of objects with different temperatures, heat always flows spontaneously from a more heated object to a less heated one. This is the essence of the formulation of the second law of thermodynamics by R. Clausius.

Of course, heat can be forced to from a cold object to a hot one: this process takes place, e.g., in the refrigerator. However, in this case, some external factors are engaged, which do a certain work and, as a result, the environment becomes heated up, i.e. process is accompanied by compensation.

Another important example of the irreversibility of processes consists in that the work (the mechanical work, the work done by the electric current, and so forth) can be totally transformed into heat. However, no heat taken from the heater can be totally transformed into work. This fact is reflected in the formulation of the second law of thermodynamics by W. Thomson (see section 19.2). In general, all forms of energy associated with the ordered motion or state (mechanical, electric, magnetic, gravitational energies) tend to be transformed into the thermal form. Such processes are called **dissipative**.

Quantitatively, the second law of thermodynamics is formulated as the law of entropy growth: at any real process, the entropy in the closed macroscopic system either grows or remains constant ($\delta S \geq 0$). In an equilibrium state of the system, the entropy of the closed system has the maximum value, so that any macroprocesses in it are impossible.

The phenomenological thermodynamics does not explain the origin of process irreversibility, i.e. the basis of the second law. The explanation is possible with the help of statistical physics, with the number of accessible microstates $\Omega(E)$ that realize that or another macrostate being the key issue.

The quantity $\Omega(E)$ is called the **statistical weight** of macrostate or its **thermodynamic probability** $W_B = \Omega$ coupled with the system entropy S by the Boltzmann principle

$$S = k_0 \ln \Omega = k_0 \ln W_B. \tag{22.28}$$

Therefore, a macrostate with the maximum statistical weight, i.e. with the maximum thermodynamic probability and, hence, with the maximum entropy $S_{\max}$ is realized in the closed system at equilibrium.

In section 19.2, it was shown that, when the gas expands into vacuum, the statistical weight of the macrostate, in which the gas fills the whole accessible volume, is much larger than the statistical weight of the macrostate, when the gas fills half of the volume. Therefore, the system almost always stays in the microstates, for which the gas occupies the whole volume. The self-compression of the gas is almost impossible, although the probability of every microstate with the volume $V/2$ is the same as with the volume V (the postulate about the equiprobability of microstates with the same energy). Hence, any closed system always transits in due course to the equilibrium state, because the latter is realized by an incomparably larger number of accessible microstates than any nonequilibrium state.

Accordingly, the dissipative character of the processes is a consequence of the fact that heat is associated with the chaotic, disordered motion of particles in the macrosystem, and the chaotic motion is related with a much larger number of accessible microstates, i.e. the statistical weight, than the ordered motion.

The number of accessible microstates essentially depends on the number of particles N (see (21.21)). Therefore, irreversible processes occur only in macroscopic systems with a large number of particles. In this connection, let us consider one more example.

Let a vessel of the volume V contains only one gas molecule. All positions of the molecule in the vessel are equiprobable. Therefore, the probability to find it in either of two vessel halves equals 0.5; hence, the motion of the molecule is reversible. The probability of molecule transition from one half of the volume to the other ("gas" expansion into vacuum) is identical to the probability of the reverse transition ("gas" compression). If the vessel contains two molecules, the probability to find them in the same half of the vessel equals $(0.5)^2 = 0.25$, i.e. the probability of "gas" self-compression equals 0.25, because both molecules are in this half of the vessel for 0.25 of the whole time. For $N = 10$, this time amounts to $(0.5)^{10} \approx 10^{-3}$, and for $N = 100$ it equals $(0.5)^{100} \approx 10^{-31}$ of the observation time. If the gas is a macroscopic system ($N \sim 10^{23}$), the duration of stay of all molecules in one half of the volume equals $0.5^{10^{23}} \sim 10^{-10^{23}}$ of the observation time, i.e., in effect, equals zero. Hence, the gas self-compression does not occur.

22.4. Statistical substantiation of the third law of thermodynamics

As was already mentioned, the macroscopic system, generally speaking, is quantum-mechanical. But, at high enough temperatures, its quantum character becomes obscured, which makes it possible to use the quasi-classical approximation. However, at low temperatures (at $T \to 0$), the situation changes drastically.

Let us designate the energy of the lowest (ground) energy level as E_0, and the energy of the first excited level as E_1. The energy difference $E_1 - E_0$ is very small, but finite. Therefore, at a low enough temperature T, the average thermal energy k_0T will be lower than $E_1 - E_0$, and the system will not be able to transit from the ground energy level E_0 onto the first excited level E_1. Therefore, at $T \to 0$, the system has to be in the ground state with the energy $E = E_0$.

The quantum-mechanical theory proves that the ground energy level of the system is nondegenerate: only one quantum state corresponds to it, i.e. $\Omega(E_0) = \Omega_0 = 1$. Then, according to the Boltzmann principle (22.28),

$$\lim_{T \to 0} S = k_0 \lim_{T \to 0} \ln \Omega = k_0 \ln \Omega_0 = k_0 \ln 1 = 0. \tag{22.29}$$

Hence, the entropy of the system at absolute zero equals zero. This conclusion is called the **Nernst theorem** or the **third law of thermodynamics**. From the viewpoint of phenomenological thermodynamics, it has already been considered in section 19.3.

In some cases, the ground level of the system is degenerate, and a few quantum states correspond to it. However, since the value of the Boltzmann constant is small ($k_0 = 1.38 \times 10^{-23}$ J/K), this circumstance does not change conclusion (22.29).

While substantiating condition (22.29), we considered the system to be in the equilibrium state and $T \to 0$. However, some systems do not satisfy this requirement, in particular, amorphous objects. It will be recalled that that such objects are in a metastable state; their true equilibrium state is a crystal, but crystallization occurs very slowly for hundreds of years, e.g., ordinary glass. Therefore, their entropy does not satisfy condition (22.29) and differs from zero at $T \to 0$.

22.5. Calculation of thermodynamic parameters using the canonical distribution

The analytical method of thermodynamics is based on thermodynamic functions used to calculate thermodynamic parameters. In turn, thermodynamic functions can be determined in the framework of statistical physics methods.

The first thermodynamic function introduced in section 20.2 was the internal energy U. Let us consider the procedure of its calculation using the canonical Gibbs distribution (21.50). The latter can be written in the form

$$w(E) = \frac{1}{Z} \Omega(E)\, \mathrm{e}^{-\dfrac{E}{k_0 T}}, \tag{22.30}$$

where

$$Z = \sum_E \Omega(E)\, \mathrm{e}^{-\dfrac{E}{k_0 T}} \tag{22.31}$$

is the partition function. According to the general rule, thermodynamic parameters have the sense of average statistical values. Therefore,

$$U = \bar{E} = \frac{\sum\limits_{E} E\Omega\left(E\right) \mathrm{e}^{-\dfrac{E}{k_0 T}}}{\sum\limits_{E} \Omega\left(E\right) \mathrm{e}^{-\dfrac{E}{k_0 T}}}. \tag{22.32}$$

It is easy to see that the numerator in (22.32) can be obtained by differentiating the denominator with respect to the temperature,

$$\sum_{E} E\Omega\left(E\right) \mathrm{e}^{-\dfrac{E}{k_0 T}} = k_0 T^2 \frac{\partial}{\partial T} \sum_{E} \Omega\left(E\right) \mathrm{e}^{-\dfrac{E}{k_0 T}} = k_0 T^2 \frac{\partial Z}{\partial T}. \tag{22.33}$$

From formulas (22.32) and (22.33), it follows that

$$U = k_0 T^2 \frac{1}{Z} \frac{\partial Z}{\partial T} = k_0 T^2 \frac{\partial \ln Z}{\partial T}. \tag{22.34}$$

Hence, knowing the energy spectrum of the system, i.e. the dependence of the energy E on quantum numbers, it is possible to calculate the partition function (22.31) and, using it, the internal energy (22.34).

Now, let us consider relation (21.56). Taking into account that $\bar{\varepsilon} = U$, one can see that it describes the relation between the partition function Z and the internal energy U and the temperature T,

$$Z = \Omega\left(U\right) \mathrm{e}^{-\dfrac{U}{k_0 T}}. \tag{22.35}$$

Since Z is a positive quantity, we may write

$$Z = \mathrm{e}^{-\dfrac{F}{k_0 T}} \tag{22.36}$$

Let us show that F in (22.36) is the free energy of the system (see section 20.2). Really, taking the logarithm of the relation

$$\mathrm{e}^{-\dfrac{F}{k_0 T}} = \Omega\left(U\right) \mathrm{e}^{-\dfrac{U}{k_0 T}},$$

we obtain

$$-\frac{F}{k_0 T} = \ln \Omega\left(U\right) - \frac{U}{k_0 T} = \sigma\left(U\right) - \frac{U}{k_0 T},$$

i.e.

$$F = U - k_0 T \sigma = U - TS. \tag{22.37}$$

This equality is identical to the determination of free energy (20.9). Then, on the basis of (22.36), we obtain a formula for the calculation of free energy with the help of partition function,

$$F = -k_0 T \ln Z. \tag{22.38}$$

Now, using expression (22.38) for the free energy, we can, using formulas (20.10) and (20.11), write the equation of state,

$$p = -\left(\frac{\partial F}{\partial V}\right)_T = k_0 T \left(\frac{\partial \ln Z}{\partial V}\right)_T \tag{22.39}$$

and the entropy,

$$S = -\left(\frac{\partial F}{\partial T}\right)_V = k_0 \ln Z + k_0 T \left(\frac{\partial \ln Z}{\partial T}\right)_V . \tag{22.40}$$

At last, on the basis of (20.15) and (20.23), we can calculate an enthalpy,

$$H = U + pV = k_0 T^2 \left(\frac{\partial \ln Z}{\partial T}\right)_V + k_0 T V \left(\frac{\partial \ln Z}{\partial V}\right)_T , \tag{22.41}$$

the thermodynamic Gibbs potential,

$$\Phi = F + pV = -k_0 T \ln Z - k_0 T V \left(\frac{\partial \ln Z}{\partial V}\right)_T , \tag{22.42}$$

and the chemical potential,

$$\mu = \frac{\Phi}{N}. \tag{22.43}$$

22.6. Grand canonical Gibbs distribution

The procedure of calculation of thermodynamic parameter considered in section 22.5 is based on the canonical Gibbs distribution, which concerns systems with a constant number of particles. However, in the general case, the system can exchange both the energy and particles with the thermostat, i.e. to be in the thermal and diffusion equilibrium with it. The state of this system is characterized by the average energy and the average number of particles. In this connection, it is necessary to make corrections to the procedure of calculation of thermodynamic parameters. First of all, we must find the statistical distribution function for systems with the variable number of particles.

Let a system with the energy ε and the number of particles n be a subsystem of a closed system; we consider the energy E and the number of particles N in the latter to be constant. The other part of the closed system is a thermostat with respect to the selected subsystem, which is characterized by the energy $E-\varepsilon$ (the energy of interaction between the subsystem and the thermostat is neglected) and the number of particles $N-n$. In this case, microcanonical distribution (21.44) can be written as

$$w = C\Omega\left(E, N\right). \tag{22.44}$$

On the basis of the multiplicativity of the number of accessible microstates, we have

$$\Omega\left(E, N\right) = \Omega\left(\varepsilon, n\right)\Omega_0\left(E-\varepsilon, N-n\right), \tag{22.45}$$

where $\Omega\left(\varepsilon, n\right)$ is the number of accessible microstates in the subsystem, and $\Omega_0\left(E-\varepsilon, N-n\right)$ the number of accessible microstates in the thermostat. In the case concerned, the variable quantities are ε and n. Hence, we may write

$$w\left(\varepsilon, n\right) = C\Omega\left(\varepsilon, n\right)\Omega_0\left(E-\varepsilon, N-n\right).$$

The constant C is determined from the normalization condition

$$\sum_{\varepsilon,n} w\left(\varepsilon, n\right) = C\sum_{\varepsilon,n}\Omega\left(\varepsilon, n\right)\Omega_0\left(E-\varepsilon, N-n\right) = 1.$$

Therefore,

$$w\left(\varepsilon, n\right) = \frac{\Omega\left(\varepsilon, n\right)\Omega_0\left(E-\varepsilon, N-n\right)}{\sum\limits_{\varepsilon,\, n}\Omega\left(\varepsilon, n\right)\Omega_0\left(E-\varepsilon, N-n\right)}. \tag{22.46}$$

By expanding the entropy of thermostat

$$\sigma\left(E-\varepsilon,N-n\right)=\ln\Omega_0\left(E-\varepsilon,N-n\right)\tag{22.47}$$

in a power series of small quantities ε and n ($\varepsilon \ll E$, $n \ll N$) up to the linear terms including, we obtain

$$\sigma\left(E-\varepsilon,N-n\right)\approx\sigma\left(E,N\right)-\frac{\partial\sigma}{\partial E}\varepsilon-\frac{\partial\sigma}{\partial N}n.\tag{22.48}$$

The derivative $\partial\sigma/\partial E$ determines the statistical temperature of the thermostat and the subsystem, which are in the equilibrium state with each other,

$$\left(\frac{\partial\sigma}{\partial E}\right)_{\varepsilon=0}=\frac{1}{\theta}=\frac{1}{k_0T}.\tag{22.49}$$

To elucidate the meaning of the derivative $\partial\sigma/\partial N$, let us use the thermodynamic master equation (20.38), on the basis of which we may write

$$dS=\frac{1}{T}dU+\frac{p}{T}dV-\frac{\mu}{T}dN,\tag{22.50}$$

where μ is the chemical potential (see section 3.3) of particles in the thermostat and the subsystem . From equation (22.50), it follow that

$$\left(\frac{\partial\sigma}{\partial N}\right)_{n=0}=\frac{1}{k_0}\left(\frac{\partial S}{\partial N}\right)_{n=0}=-\frac{\mu}{k_0T}.\tag{22.51}$$

Hence,

$$\sigma\left(E-\varepsilon,N-n\right)=\sigma\left(E,N\right)+\frac{\mu n-\varepsilon}{k_0T}.\tag{22.52}$$

On the basis of (22.47) and (22.52), we have

$$\Omega_0\left(E-\varepsilon,N-n\right)=\mathrm{e}^{\sigma(E,\ N)}\mathrm{e}^{-\frac{\mu n-\varepsilon}{k_0T}}.\tag{22.53}$$

Substituting (22.53) into (22.46) and reducing the result by $\mathrm{e}^{\sigma(E,\ N)}$, we obtain the final expression,

$$w\left(\varepsilon,n\right)=\frac{\Omega\left(\varepsilon,n\right)\mathrm{e}^{-\frac{\mu n-\varepsilon}{k_0T}}}{\sum\limits_{\varepsilon,n}\Omega\left(\varepsilon,n\right)\mathrm{e}^{-\frac{\mu n-\varepsilon}{k_0T}}},\tag{22.54}$$

which is called the **grand canonical Gibbs distribution**. It determines the probability for the system in the thermostat to be in a microstate with the energy ε and the number of particles n.

The sum in the denominator of (22.54),

$$\tilde{Z} = \sum_{\varepsilon,n} \Omega(\varepsilon, n)\, \mathrm{e}^{-\dfrac{\mu n - \varepsilon}{k_0 T}}, \tag{22.55}$$

is called the **grand partition function**. Similarly to the partition function (see section 22.5), the grand partition function is used to calculate the thermodynamic parameters, but in the systems with the variable number of particles. As an example (see the corresponding proof in section 5.6 of [22]), we give the formulas for the calculation of the average number of particles, $\bar{n}$, and energy, $\bar{\varepsilon}$:

$$\bar{n} = k_0 T \frac{1}{\tilde{Z}} \frac{\partial \tilde{Z}}{\partial \mu} = k_0 T \frac{\partial \ln \tilde{Z}}{\partial \mu}, \tag{22.56}$$

$$\bar{\varepsilon} = U = k_0 T^2 \frac{\partial \ln \tilde{Z}}{\partial T} + \mu \bar{n}. \tag{22.57}$$

23

CLASSICAL IDEAL AND REAL GASES

In this chapter, we intend to apply the main principles and methods of statistical physics, which were considered above, to some macroscopic systems. The simplest for studying are gases, i.e. the systems in which the average potential energy of interaction between particles is much lower than their average kinetic energy. As a result, the particles are weakly coupled with one another by interaction forces and move freely, filling the whole accessible volume, so that the gas acquires the shape of the vessel, in which it is contained.

We confine the consideration of such gases as oxygen, hydrogen, helium, neon, and so forth. Their molecules can consist of a single atom; then, the gas is called single-atomic. Such gasses include, e.g., inert gases and vapor of various metals. If the gas molecule consists of two atoms, such gases are called two-atomic. These are, e.g., oxygen and hydrogen. Besides, we will consider some properties of the gases that do not depend on the number of atoms in the gas molecule; in this case the gas may be considered as single-atomic.

If the temperature is not too low and the pressure is not too high, the intermolecular interaction can be sometimes neglected, i.e. interaction can be considered to occur only at a direct collision of molecules. In this case, the size of molecules is also neglected. This statistical system is called the **ideal gas**. Of course, the ideal gas is an approximate model. At more exact calculations, it is necessary to take into account both the intermolecular interaction and the own volume of the molecules. Then, the gas is called **real**.

Since ordinary gases are rather rarified under ambient conditions and their molecules are rather massive, they can be analyzed using classical statistics in the quasiclassical approximation, which is based on the classical canonical Gibbs distribution (21.57) and the partition function in form (21.61).

23.1. Maxwell–Boltzmann distribution

Let us consider a classical single-atomic ideal gas consisting of N molecules (atoms) is an external potential field. The energy of this gas can be written in the form

$$E = \sum_{i=1}^{N} \left(\frac{1}{2m} p_i^2 + u_i \right), \tag{23.1}$$

where the first summand in the parentheses is the kinetic energy of the i-th molecules and $u_i\left(\vec{r}_i\right)$ is its potential energy in the external field. Substituting expression (23.1) into the classical canonical Gibbs distribution (21.57),

$$dW = \frac{\mathrm{e}^{-\dfrac{E}{k_0 T}}\, d\Gamma}{\displaystyle\int \mathrm{e}^{-\dfrac{E}{k_0 T}}\, d\Gamma}, \tag{23.2}$$

and taking into account that

$$\mathrm{e}^{-\dfrac{E}{k_0T}\sum_i\left(\dfrac{1}{2m}p_i^2+u_i\right)}=\prod_{i=1}^{N}\mathrm{e}^{-\dfrac{p_i^2}{2mk_0T}-\dfrac{u_i}{k_0T}} \tag{23.3}$$

and

$$d\Gamma=\prod_{i=1}^{N}d\gamma_{0i}=\prod_{i=1}^{N}dV_i dV_{pi},$$

where

$$d\gamma_{0i}=dV dV_p=dxdydzdp_xdp_ydp_z \tag{23.4}$$

is the elementary phase volume in the phase space of single molecule, we obtain

$$dW=\frac{\prod\limits_i \mathrm{e}^{-\dfrac{p_i^2}{2mk_0T}-\dfrac{u_i}{k_0T}}d\gamma_{0i}}{\int\prod\limits_i \mathrm{e}^{-\dfrac{p_i^2}{2mk_0T}-\dfrac{u_i}{k_0T}}d\gamma_{0i}}=$$

$$=\prod_{i=1}^{N}\frac{\mathrm{e}^{-\dfrac{p_i^2}{2mk_0T}-\dfrac{u_i}{k_0T}}d\gamma_{0i}}{\int\mathrm{e}^{-\dfrac{p_i^2}{2mk_0T}-\dfrac{u_i}{k_0T}}d\gamma_{0i}}=\prod_{i=1}^{N}dW_i. \tag{23.5}$$

According to the theorem about the product of probabilities, the quantity

$$dw=\frac{\mathrm{e}^{-\dfrac{p^2}{2mk_0T}-\dfrac{u}{k_0T}}d\gamma_0}{\int\mathrm{e}^{-\dfrac{p^2}{2mk_0T}-\dfrac{u}{k_0T}}d\gamma_0} \tag{23.6}$$

can be regareded as the probability related to a single (arbitrary) molecule.

Formula (23.6) looks like a classical canonical Gibbs distribution. Therefore, any molecule of the ideal gas can be considered as a subsystem in the thermostat, the role of which is played by other gas molecules.

Taking into account (23.4), expression (23.6) can be written as a product of two multipliers,

$$dw = \frac{\mathrm{e}^{-\dfrac{p_i^2}{2mk_0T} - \dfrac{u_i}{k_0T}} dV_p}{\displaystyle\int \mathrm{e}^{-\dfrac{p_i^2}{2mk_0T} - \dfrac{u_i}{k_0T}} dV_p} \cdot \frac{\mathrm{e}^{-\dfrac{u}{k_0T}} dV}{\displaystyle\int \mathrm{e}^{-\dfrac{u}{k_0T}} dV}. \tag{23.7}$$

The integral in the denominator of the first multiplier can be easily calculated. In the spherical coordinate system,

$$dV_p = p^2 dp d\Omega_s,$$

where Ω_s is the solid angle. Therefore,

$$\begin{aligned} \int \mathrm{e}^{-\dfrac{p^2}{2mk_0T} - \dfrac{u_i}{k_0T}} dV_p &= \int_0^\infty \mathrm{e}^{-\dfrac{p^2}{2mk_0T}} p^2 dp \int_0^{4\pi} d\Omega_x \\ &= 4\pi \int_0^\infty \mathrm{e}^{-\dfrac{p^2}{2mk_0T}} p^2 dp. \end{aligned}$$

The last integral belongs to the so-called Poisson integrals (see Appendix K) and equals $\frac{1}{4}\sqrt{\pi\,(2mk_0T)^3}$. Therefore,

$$\int \mathrm{e}^{-\dfrac{p^2}{2mk_0T}} dV_p = (2\pi mk_0T)^{3/2}. \tag{23.8}$$

Hence, expression (6.7) reads

$$dw_{\mathrm{MB}} = \frac{\mathrm{e}^{-\dfrac{p^2}{2mk_0T}} p^2 dp d\Omega_s}{(2\pi mk_0T)^{3/2}} \times \frac{\mathrm{e}^{-\dfrac{u}{k_0T}} dV}{\displaystyle\int \mathrm{e}^{-\dfrac{u}{k_0T}} dV} = dw_{\mathrm{M}} dw_{\mathrm{B}}. \tag{23.9}$$

Formula (23.9) is called the **Maxwell–Boltzmann distribution.** It indicates the probability for a gas molecule to be contained in the volume dV and have the momentum vector directed into the solid angle $d\Omega_s$ and the absolute value within the interval from p to $p + dp$. The multiplier

$$dw_{\mathrm{M}} = \frac{1}{(2\pi m k_0 T)^{3/2}} \mathrm{e}^{-\dfrac{p^2}{2mk_0T}} p^2 dp d\Omega_s \tag{23.10}$$

is called the **Maxwell distribution**, and the multiplier

$$dw_B = \frac{1}{\displaystyle\int \mathrm{e}^{-\dfrac{u}{k_0T}} dV} \mathrm{e}^{-\dfrac{u}{k_0T}} dV \tag{23.11}$$

the **Boltzmann distribution**. Maxwell distribution (23.10) determines the distribution of gas molecules of over the momenta p, whereas Boltzmann distribution (23.11) over the coordinates $\vec{r}$. Their product (23.9) means that the distributions of molecules over their momenta and coordinates are mutually independent, i.e. the probability of a certain momentum value and its direction do not depend on the molecule position and vice versa, the probability for a molecule to be located somewhere does not depend on its momentum.

23.2. Maxwell distribution

Let us consider the Maxwell distribution in more details. From formula (23.10), it follows that the probability of molecule's momentum does not depend on its direction, because the angular dependence is absent. In other words, all orientations of the momentum vector $\vec{p}$ within the solid angle $d\Omega_s$ are equiprobable. Therefore, integrating (23.10) over the whole solid angle $\Omega_s = 4\pi$, we obtain the probability for the momentum to be within the interval from p to $p + dp$ irrespective of its direction and the molecule position,

$$dw_{\mathrm{M}}(p) = \frac{4\pi}{(2\pi m k_0 T)^{3/2}} \mathrm{e}^{-\dfrac{p^2}{2mk_0T}} p^2 dp. \tag{23.12}$$

Using the relation $p = mv$, we obtain the distribution over molecular velocities,

$$dw_{\mathrm{M}}(v) = 4\pi \left(\frac{m}{2\pi m k_0 T}\right)^{3/2} \mathrm{e}^{-\dfrac{mv^2}{2k_0T}} v^2 dv. \tag{23.13}$$

The Maxwell distribution in form (23.13) determines the probability that the molecule has a velocity vector with the absolute value in the interval from v to $v + dv$ irrespective of its direction and the molecule position. Just this formula is usually meant under the term “Maxwell distribution”.

If the gas consists of N molecules, their average number with the velocities in the interval dv can be determine by the formula

$$dN = N dw_{\mathrm{M}}(v) = 4\pi N \left(\frac{m}{2\pi k_0 T}\right)^{3/2} \mathrm{e}^{-\dfrac{mv^2}{2k_0T}} v^2 dv. \tag{23.14}$$

The function

$$f_{\mathrm{M}}(v) = \frac{dw_{\mathrm{M}}(v)}{dv} = \frac{dN}{N dv} = 4\pi \left(\frac{m}{2\pi k_0 T}\right)^{3/2} \mathrm{e}^{-\dfrac{mv^2}{2k_0T}} v^2 \tag{23.15}$$

is called the **Maxwell distribution function.** It is plotted in Fig. 23.1.

The velocity of molecule v that corresponds to the maximum of distribution function (23.15) is called the **most probable velocity**. It can be calculated from the equation

$$\frac{df_{\mathrm{M}}}{dv} = 0,$$

and it equals

$$v_m = \sqrt{\frac{2k_0T}{m}} \approx 1.41\sqrt{\frac{k_0T}{m}}. \tag{23.16}$$

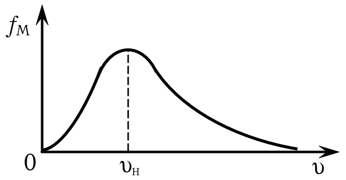

Fig. 23.1.

In terms of the most probable velocity v_m, Maxwell distribution (23.10) looks like

$$dw_{\mathrm{M}}(v) = \frac{4}{\sqrt{\pi}} \frac{v^2}{v_m^3} \mathrm{e}^{-\dfrac{v^2}{v_m^2}} dv. \tag{23.17}$$

The **average** (**mean**) **velocity** of the molecule $\bar{v}$ is determined by the formula

$$\bar{v} = \int v dw_{\mathrm{M}}(v) = 4\pi \left(\frac{m}{2\pi k_0 T}\right)^{3/2} \int_0^\infty v^3 \mathrm{e}^{-\frac{mv^2}{2k_0 T}} dv,$$

where the integral is a Poisson integrals (see Appedix K) and equals $2k_0^2 T^2/m^2$. Therefore,

$$\bar{v} = \sqrt{\frac{8k_0 T}{\pi m}} \approx 1.60\sqrt{\frac{k_0 T}{m}}. \tag{23.18}$$

Besides the most probable and mean velocities, the **root-mean-square velocity** of the molecule $v_{sq} = \sqrt{\overline{v^2}}$ is also used. Using one of Poisson integrals in this case as well, we find

$$\overline{v^2} = \int v^2 dw_{\mathrm{M}}(v) = 4\pi N \left(\frac{m}{2\pi k_0 T}\right)^{3/2} \int_0^\infty v^4 \mathrm{e}^{-\frac{mv^2}{2k_0 T}} dv = \frac{3k_0 T}{m},$$

so that

$$v_{sq} = \sqrt{\frac{3k_0 T}{m}} \approx 1.73\sqrt{\frac{k_0 T}{m}}. \tag{23.19}$$

One can see that

$$v_m < \bar{v} < v_{sq}.$$

However, the difference between those velocities is small. Therefore, the root-mean-square velocity v_{sq} is used the most often as the average one.

23.3. Boltzmann distribution for the gas in an external field

The probabilities of various coordinate values for a gas molecule in an external field are determined by the Boltzmann distribution (23.11),

$$dw_{\mathrm{B}}(\vec{r}) = \frac{1}{\int \mathrm{e}^{-\frac{u(\vec{r})}{k_0 T}} dV} \mathrm{e}^{-\frac{u(\vec{r})}{k_0 T}} dV, \tag{23.20}$$

where $u(\vec{r})$ is the potential energy of the molecule in the external field. Let us introduce the notation

$$\left[\int \mathrm{e}^{-\dfrac{u(\vec{r})}{k_0 T}} dV\right]^{-1} = A(T).$$

Then, expression (23.20) looks like

$$dw_{\mathrm{M}}(\vec{r}) = A(T)\,\mathrm{e}^{-\dfrac{u(\vec{r})}{k_0 T}} dV. \tag{23.21}$$

Taking into account that dw_{M}/dV is the probability density, the average number of molecules in a unit volume can be determined as

$$n = N\frac{dw_{\mathrm{M}}}{dV} = C(T)\,\mathrm{e}^{-\dfrac{u(\vec{r})}{k_0 T}}, \tag{23.22}$$

where $C(T) = NA(T)$, and N is the number of molecules in the selected volume.

Let us apply distribution (23.22) to a gas in the field of Earth gravitation. The energy of a molecule the gravitational field is written in the following general form:

$$u(R) = m\varphi(R), \tag{23.23}$$

where R is the distance from the Earth center, m the molecule mass,

$$\varphi(R) = -G\frac{M}{R} \tag{23.24}$$

is the gravitational field potential; M the Earth mass, and G the gravitational constant. The free fall acceleration on the Earth surface equals

$$g = -G\frac{M}{R_0^2},$$

where R_0 is the Earth radius. Therefore, formula (23.24) can be rewritten as

$$\varphi(R) = -\frac{gR_0^2}{R}, \tag{23.25}$$

and the molecule energy (23.23) as

$$u\,(R) = -\frac{mgR_0^2}{R}. \tag{23.26}$$

Substituting (23.26) into (23.22), we obtain

$$n = C\,(T)\,\mathrm{e}^{-\dfrac{mgR_0^2}{Rk_0T}}. \tag{23.27}$$

Denoting the concentration of molecules at $R = R_0$ as n_0, we have

$$C\,(T) = n_0\mathrm{e}^{-\dfrac{mgR_0^2}{Rk_0T}}.$$

Then, expression (23.27) takes the form

$$n = n_0\mathrm{e}^{-\dfrac{mgR_0}{k_0T}\left(1-\dfrac{R_0}{R}\right)}. \tag{23.28}$$

For the altitude h above the Earth surface, $R = R_0 + h$, and, instead of (23.28), we write

$$n = n_0\mathrm{e}^{-\dfrac{mgR_0}{k_0T}\dfrac{h}{R_0+h}}. \tag{23.29}$$

Under the condition $h \ll R_0$, from (23.29), the **barometric formula** follows:

$$n = n_0\mathrm{e}^{-\dfrac{mgh}{k_0T}}. \tag{23.30}$$

Taking into account that $nk_0T = p$, it can be rewritten as

$$p = p_0\mathrm{e}^{-\dfrac{mgh}{k_0T}}. \tag{23.31}$$

Formula (23.31) correctly describes, in general, the dependence of atmospheric pressure on the altitude h. However, the real dependence can differ from law (23.31), because the atmosphere is not in the equilibrium state, it is nonuniform; in particular, its temperature at various altitudes is different.

Moreover, at $h \gg R_0$, we obtain from (23.29) that

$$n\left(h \gg R_0\right) = n_0 \mathrm{e}^{-\dfrac{mgR_0}{k_0 T}} \neq 0. \tag{23.32}$$

This means the scattering of the Earth atmosphere into the outer space, because the upper atmosphere always contains molecules with the velocities exceeding the escape velocity from the Earth $v_2 = \sqrt{2gR_0} \approx 11.2$ m/s. From (23.32), it follows that the scattering rate is higher for gases with smaller molecular masses and for planets with smaller free fall acceleration.

23.4. Partition function for the classical single-atomic ideal gas

In the case of classical system, the partition function is calculated by formula (21.61),

$$Z = \frac{1}{N!\left(2\pi\hbar\right)^{fN}} \int \mathrm{e}^{-\dfrac{E}{k_0 T}} d\Gamma. \tag{23.33}$$

Let us calculate it for a classical single-atomic ideal gas with N identical molecules in the volume V. In the absence of external field,

$$E = \sum_{i=1}^{N} \frac{1}{2m} p_i^2$$

and

$$d\Gamma = \prod_{i=1}^{N} dV_i dV_{pi}.$$

Hence, expression (23.33) takes the form

$$Z = \frac{1}{N!\left(2\pi\hbar\right)^{3N}} \int \prod_{i=1}^{N} \mathrm{e}^{-\frac{p_i^2}{2mk_0 T}} dV_i dV_{pi}. \tag{23.34}$$

Since all molecules are identical, located in the same volume V, and have the absolute values of momentum vectors within the same limits $0 \leq p \leq$

∞, we obtain

$$\int \prod_{i=1}^{N} \mathrm{e}^{-\dfrac{p_i^2}{2mk_0T}} dV_i dV_{pi} = \left[\int \mathrm{e}^{-\dfrac{p^2}{2mk_0T}} dV dV_p \right]^N . \tag{23.35}$$

Then, for partition function (23.34), we obtain the expression (see (23.8))

$$Z = \frac{1}{N!\,(2\pi\hbar)^{3N}} \left\{ \int \mathrm{e}^{-\dfrac{p_i^2}{2mk_0T}} dV dV_p \right\}^N =$$

$$= \frac{1}{N!\,(2\pi\hbar)^{3N}} \left\{ \int_0^{\infty} \mathrm{e}^{-\dfrac{p^2}{2mk_0T}} dV_p \int_0^{\nabla} dV \right\}^N =$$

$$= \frac{V^N (2\pi m k_0 T)^{3N/2}}{N!\,(2\pi\hbar)^{3N}} . \tag{23.36}$$

The number of gas molecules N is large enough. Therefore, according to the Stirling formula, $N! \approx (N/\mathrm{e})^N$ (see Appendix L). Hence,

$$Z = \left[\frac{V \mathrm{e}\,(2\pi m k_0 T)^{3/2}}{N\,(2\pi\hbar)^3} \right]^N . \tag{23.37}$$

23.5. Thermodynamic functions and the equation of state for the classical single-atomic ideal gas

First, on the basis of (23.37), we obtain the expression for the partition function,

$$\ln Z = N \ln \frac{V e\,(2\pi m k_0 T)^{3/2}}{N\,(2\pi\hbar)^3} , \tag{23.38}$$

required for the calculation of thermodynamic parameters. Then, making use of formulas (21.34) , (21.38), and (21.40), we calculate the internal

U, the free energy F, and the entropy S, respectively, the classical single-atomic ideal gas:

$$U = k_0 T^2 \left(\frac{\partial \ln Z}{\partial T} \right)_V = \frac{3}{2} N k_0 T, \tag{23.39}$$

$$F = -k_0 T \ln Z = -k_0 T N \ln \frac{V e \left(2\pi m k_0 T\right)^{3/2}}{N \left(2\pi\hbar\right)^3}, \tag{23.40}$$

$$S = -\left(\frac{\partial F}{\partial T} \right)_V = k_0 N \ln \frac{V e^{5/2} \left(2\pi m k_0 T\right)^{3/2}}{N \left(2\pi\hbar\right)^3}. \tag{23.41}$$

Using expression (23.40) for the free energy, we find the equation of state for the classical ideal gas,

$$p = -\left(\frac{\partial F}{\partial V} \right)_T = \frac{N k_0 T}{V}, \tag{23.42}$$

and, with its help, the enthalpy,

$$H = U + pV = \frac{5}{2} N k_0 T, \tag{23.43}$$

and the thermodynamic Gibbs potential,

$$\Phi = F + pV = -N k_0 T \ln \frac{V e \left(2\pi m k_0 T\right)^{3/2}}{N \left(2\pi\hbar\right)^3} + N k_0 T. \tag{23.44}$$

Formula (23.44) can be rewritten in the form

$$\Phi = N k_0 T \left\{ \ln \frac{N \left(2\pi\hbar\right)^3}{V e \left(2\pi m k_0 T\right)^{3/2}} + 1 \right\}.$$

Then, taking into account that ln$e$= 1, we obtain

$$\Phi = N k_0 T \ln \frac{N \left(2\pi\hbar\right)^3}{V e \left(2\pi m k_0 T\right)^{3/2}}. \tag{23.45}$$

On the basis of (23.45), we determine the chemical potential of the classical single-atomic ideal gas,

$$\mu = \frac{\Phi}{N} = k_0 T \ln \frac{n_0 \left(2\pi\hbar\right)^3}{\left(2\pi m k_0 T\right)^{3/2}}, \tag{23.46}$$

where $n_0 = N/V$ is the number of molecules in a unit volume (the concentration).

Let us write one more relation useful for following calculations. It follows from (23.46),

$$\frac{n_0 \left(2\pi\hbar\right)^3}{\left(2\pi m k_0 T\right)^{3/2}} = \mathrm{e}^{-\dfrac{\mu}{k_0 T}} . \tag{23.47}$$

By analyzing formula (23.41), one can see that, according to it, $S \to -\infty$ if $T \to 0$, which contradicts to the third law of thermodynamics. This circumstance brings us to a conclusion about the inapplicability of classical statistics at low temperatures, when the quantum-mechanical properties of the system reveal themselves. In the case, formula (23.33) used to calculate the partition function becomes invalid.

23.6. Distribution of molecules over microstates in the classical single-atomic ideal gas

In the quasiclassical approximation, the microstate of a separate molecule in the classical single-atomic ideal gas is represented in the phase space of the molecule by a quantum cell with the volume $\Delta\gamma_0 = \left(2\pi\hbar\right)^3$ (see section 21.3) and characterized by some momentum $\vec{p}$. The probability of this microstate can be written as the product

$$\rho_0\left(\vec{p}\right)\Delta\gamma_0 = \rho_0\left(\vec{p}\right)\left(2\pi\hbar\right)^3 ,$$

where $\rho_0\left(\vec{p}\right)$ is the probability density of microstate distribution in the phase space. Then, the average number of molecules in the microstate with the momentum $\vec{p}$ equals

$$\bar{n} = N\rho_0\left(\vec{p}\right)\left(2\pi\hbar\right)^3 , \tag{23.48}$$

where N is the number of molecules.

The probability density ρ_0 follows from distribution (23.6), which in the absence of external field looks like

$$dw = \frac{e^{-\frac{\vec{p}^{\,2}}{2mk_0T}}\,d\gamma_0}{\int e^{-\frac{\vec{p}^{\,2}}{2mk_0T}}\,d\gamma_0} = \rho_0\left(\vec{p}\right)d\gamma_0. \tag{23.49}$$

Since $d\gamma_0 = dV\,dV_p$ (see (21.4)), the integral in the denominator is known to equal $V\left(2\pi mk_0T\right)^{3/2}$ (see (23.8)). Therefore,

$$\rho_0\left(\vec{p}\right) = \frac{dw}{d\gamma_0} = \frac{1}{V\left(2\pi mk_0T\right)^{3/2}}\,e^{-\frac{\vec{p}^{\,2}}{2mk_0T}}. \tag{23.50}$$

Substituting (23.50) into (23.48), we obtain

$$\bar{n} = \frac{N\left(2\pi\hbar\right)^3}{V\left(2\pi mk_0T\right)^{3/2}}\,e^{-\frac{\vec{p}^{\,2}}{2mk_0T}}. \tag{23.51}$$

Taking into account that $N/V = n_0$ and $\vec{p}^{\,2}/\left(2m\right) = \varepsilon\left(\vec{p}\right)$, and using formula (23.47), we ultimately find

$$\bar{n} = e^{-\frac{\mu - \varepsilon\left(\vec{p}\right)}{k_0T}}. \tag{23.52}$$

Formula (23.52) gives the average number of molecules in the classical ideal gas, which are in the microstate determined by the momentum $\vec{p}$ and the kinetic energy $\varepsilon\left(\vec{p}\right)$. This formula is called the **Boltzmann distribution** (do not confuse with the Boltzmann distribution of molecules with respect to their coordinates in an external field (23.20)!). It forms the basis of the Boltzmann statistics for the classical ideal gas, which we will not consider.

Boltzmann distribution (23.52) agrees with the Maxwell distribution. Really, multiplying (23.51) by the number of molecule microstates (see section 21.3),

$$dg = \frac{4\pi p^2 V\,dp}{\left(2\pi\hbar\right)^3},$$

we obtain the relationship

$$dN = \bar{n}dg = \frac{4\pi N}{\left(2\pi m k_0 T\right)^{3/2}}\,\mathrm{e}^{-\dfrac{p^2}{2mk_0T}}p^2dp,$$

whence Maxwell distribution (23.12) follows,

$$\frac{dN}{N} = dw_{\mathrm{M}} = \frac{4\pi}{\left(2\pi m k_0 T\right)^{3/2}}\,\mathrm{e}^{-\dfrac{p^2}{2mk_0T}}p^2dp.$$

23.7. Classical real gas

The thermodynamic functions and the equation of state that were found in section 23.5 concerned the classical single-atomic ideal gas, in which the interaction between atoms and their size is not taken into account. This approximation is valid if the gas is so rarefied that the average distances between the atoms considerably exceed their dimensions. However, as the density of gas increases, its property start to depend on the interaction between atoms, and, as a result, the equation of state $pV = RT$ becomes inaccurate.

In this connection, there arises a task of calculating the partition function for the classical single-atomic real gas making allowance for the interaction between atoms and their size in order to find the corresponding equation of state and the thermodynamic functions. For this purpose, we must know the character of intermolecular (interatomic) interaction, which considerably depends on the distance r between the centers of molecules (atoms)

At $r > d$, where d is the atomic diameter, the interaction is characterized by mutual attraction. The quantum-mechanical calculation (see, section 7.5 in [20]) shows that, in this case, the potential energy of interaction between two atoms, which are in the ground state and have no electric dipole moment, takes the form

$$u\left(r\right) = -\frac{A}{r^6},$$

where A is a constant. If $r \leq d$, there arises a strong mutual repulsion of atoms, which is mainly associated with the overlapping of their electron shells. The exact law of repulsion is unknown in the general case. The most often, it is written as

$$u(r) \sim \mathrm{e}^{-\alpha r}$$

or as

$$u(r) \sim r^{-n}, \quad 8 \leq n \leq 12.$$

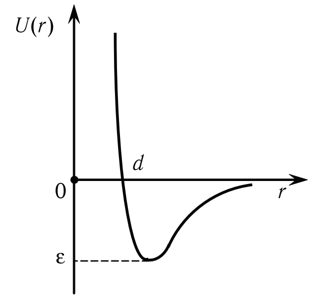

Fig. 23.2.

Experimental and theoretical researches showed that dependence the $u(r)$ within the whole interval $0 < r < \infty$ looks like the curve depicted in Fig. 23.2, where the value $r = d$ is considered to be the effective diameter of the atom, and ε is the energy of interaction between the atoms located at the equilibrium distance (the depth of the potential well). Notice that the qualitative dependence exhibited in Fig. 23.2 is applicable to both the interaction between the atoms of the single-atomic gas and the intermolecular interaction in multiatomic gases.

As a rule, the interaction graphically illustrated in Fig. 23.2 is described analytically with the help of Lennard–Jones potential

$$u(r) = 4\varepsilon \left\{ \left(\frac{d}{r}\right)^{12} - \left(\frac{d}{r}\right)^{6} \right\}. \tag{23.53}$$

The first summand describes the repulsion between the atoms at $r < d$ and dominates in the left branch of the plot. The second summand is the energy of mutual attraction of atoms at $r > d$ and dominates in the right branch of the plot.

Now, let us calculate the partition function for the classical real gas. The energy of the classical single-atomic real gas is a sum of the kinetic energies of atomic translational motions and the potential energies of their interaction with one another, i.e.

$$E = \sum_{i=1}^{N} \frac{p_i^2}{2m} + \sum_{i,j>i}^{N} u(r_{ij}). \tag{23.54}$$

In the double sum over i and j, the summands with $j = i$, which correspond to the self-interaction of molecules, and the duplicating of identical summands $u\left(r_{ij}\right) = u\left(r_{ji}\right)$ must be excluded; in this connection, this summation includes the summands with $jj > i$.

Substituting (23.54) into (23.33) and taking into account that

$$d\Gamma = \prod_{k=1}^{N} dV_k dV_{pk},$$

we obtain

$$Z = \frac{1}{N!\left(2\pi\hbar\right)^{3N}} \prod_{i=1}^{N} \mathrm{e}^{-\dfrac{p_i^2}{2mk_0T}} dV_{pi} \int \prod_{i,j>i}^{N} \mathrm{e}^{-\dfrac{u(r_{ij})}{k_0T}} \prod_{k=1}^{N} dV_k. \tag{23.55}$$

The integral over the momenta was calculated in section 23.4. Therefore, on the basis of the results obtained, we write

$$Z = \frac{\left(2\pi mk_0T\right)^{3/2}}{N!\left(2\pi\hbar\right)^{3N}} J_V, \tag{23.56}$$

where

$$J_V = \int \prod_{i,j>i}^{N} \mathrm{e}^{-\dfrac{u(r_{ij})}{k_0T}} \prod_{k=1}^{N} dV_k. \tag{23.57}$$

Let us introduce the auxiliary function

$$f_{if} = \mathrm{e}^{-\dfrac{u(r_{ij})}{k_0T}} - 1. \tag{23.58}$$

With its help, integral (23.57) is written as

$$J_V = \int \prod_{i,\ j>i}^{N} \left(1 + f_{ij}\right) \prod_{k=1}^{N} dV_k. \tag{23.59}$$

Let us consider the product in the integrand,

$$\prod_{i,\ j>i}^{N} \left(1 + f_{ij}\right) = \left(1 + f_{12}\right)\left(1 + f_{13}\right) ... \left(1 + f_{23}\right)\left(1 + f_{24}\right) ... \left(1 + f_{34}\right) ... .$$

Removing the parentheses, it can be written in the form

$$\prod_{i,\, j>i}^{N}(1+f_{ij}) = 1 + f_{12} + f_{13} + \ldots + f_{23} + \ldots +$$
$$+ f_{12}f_{13} + \ldots + f_{12}f_{13}f_{14} + \ldots . \tag{23.60}$$

From (23.58), it follows that function the f_{ij} differs from zero only at small values of r_{ij} ($r_{ij} \leq 3d$, see Fig. 23.2). Therefore, provided that $d < r_{ij} < 3d$, the exponential function in (23.58) is considerably larger than unity. Small values of r_{ij} correspond to the approach of two atoms. However, in this case, the products like $f_{12}f_{13}$, $f_{12}f_{13}f_{14}$, and so on are different from zero only at the simultaneous approach of three, four, and more atoms, which is almost improbable in the ordinary gas. Hence,

$$\prod_{i,j>i}^{N}(1+f_{ij}) \approx 1 + \sum_{i,j>i}^{N} f_{ij}. \tag{23.61}$$

Substituting (23.61) into (23.59), we find

$$J_V = \int\left(1 + \sum_{i,j>i} f_{ij}\right)\prod_{k=1}^{N} dV_k = V^N + \sum_{i,j>i}\int f_{ij}\prod_{k=1}^{N} dV_k. \tag{23.62}$$

The functions $u(r_{ij})$ and $u(r_{ji})$, as well as f_{ij} and f_{ji}, look identically for any pair of atoms; therefore,

$$J_V = V^N + \frac{N^2}{2}\int f_{ij}dV_i dV_j \int \prod_{k\neq i,j} dV_k =$$
$$= V^N + \frac{N^2}{2}V^{N-2}\int f_{ij}dV_i dV_j, \tag{23.63}$$

where $N\,(N-1)\,/2 \approx N^2/2$ is the number of molecular pairs (the number of combinations of 2 elements from a set of N elements).

Now let us use a spherical coordinate system with the origin at the i-th atom. Then, the function $f_{ij} = f\,(r_{ij})$ depends only on the coordinates of j-th atom. Hence,

$$J_V = V^N + \frac{N^2}{2}V^{N-2}\int f\,(r_j)\,dV_i\int dV_i = V^N\left\{1 + \frac{N^2}{2V}\int f\,(r)\,dV\right\}. \tag{23.64}$$

Introducing the notation

$$B\left(T\right)=\int f\left(r\right)dV, \tag{23.65}$$

we rewrite (23.64) as

$$J_{V}=V^{N}\int\left\{1+\frac{N^{2}}{2V}B\left(T\right)\right\}. \tag{23.66}$$

Substituting (23.66) into (23.56) we obtain the partition function in the form

$$Z=Z_{\mathrm{id}}\left\{1+\frac{N^{2}}{2V}B\left(T\right)\right\}, \tag{23.67}$$

where Z_{id} is the partition function for the ideal gas (23.37),

$$Z_{\mathrm{id}}=\left\{\frac{Ve\left(2\pi mk_{0}T\right)^{3/2}}{N\left(2\pi\hbar\right)^{3}}\right\}^{N}. \tag{23.68}$$

Making use of (23.67), we obntain the free energy of real gas,

$$F=-k_{0}T\ln Z=-k_{0}T\ln Z_{\mathrm{id}}-k_{0}T\ln\left\{1+\frac{N^{2}}{2V}B\left(T\right)\right\}. \tag{23.69}$$

According to previous remarks, $f(r)\ll 1$ as a rule; therefore, $B\left(T\right)\ll 1$ as well. In this connection, we assume that

$$\frac{N^{2}}{2V}B\left(T\right)\ll 1. \tag{23.70}$$

Taking into account that, for $|x|\ll 1$, the function $\ln\left(1+x\right)\approx x$, instead of (23.69), we write

$$F=F_{\mathrm{id}}-\frac{k_{0}TN^{2}}{2V}B\left(T\right).$$

Then, the equation of state looks like

$$p=-\left(\frac{\partial F}{\partial V}\right)_{T}=\frac{Nk_{0}T}{V}-\frac{k_{0}TN^{2}}{V^{2}}B\left(T\right). \tag{23.71}$$

Now let us consider the function $B\,(T)$. Taking into account (23.58), we have

$$\begin{aligned} B\,(T) &= \int_0^\infty \left\{ \mathrm{e}^{-\dfrac{u\,(r)}{k_0T}} - 1 \right\} r^2 dr \int_0^{4\pi} d\Omega_s \\ &= 4\pi \int_0^\infty \left\{ \mathrm{e}^{-\dfrac{u(r)}{k_0T}} - 1 \right\} r^2 dr = \end{aligned}$$

$$= 4\pi \int_0^d \left\{ \mathrm{e}^{-\dfrac{u\,(r)}{k_0T}} - 1 \right\} r^2 dr + 4\pi \int_d^\infty \left\{ \mathrm{e}^{-\dfrac{u(r)}{k_0T}} - 1 \right\} r^2 dr.$$

Within the interval $0 \leq r \leq d$, the function $u\,(r) \to \infty$. Therefore, $\exp\left(-\dfrac{u\,(\vec{r})}{k_0T}\right) \approx 0$ here; and, at $r > d$, the ratio $\dfrac{u\,(\vec{r})}{k_0T} \ll 1$. Hence,

$$\mathrm{e}^{-\dfrac{u\,(r)}{k_0T}} \approx 1 - \frac{u(r)}{k_0T}$$

within the whole interval $0 \leq r < \infty$. Bearing those remarks in mind, we find

$$\begin{aligned} B\,(T) = 4\pi \int_0^d (-1)\, r^2 dr - 4\pi \frac{1}{k_0T} \int_d^\infty u(r) r^2 dr = \\ = -\frac{4}{3}\pi d^3 - \frac{4\pi}{k_0T} \int_d^\infty u\,(r)\, r^2 dr. \end{aligned} \tag{23.72}$$

Substituting (23.72) into (23.71), we obtain the equation of state for the real gas in the form

$$p = \frac{Nk_0T}{V} + \frac{Nk_0T}{V^2} b - \frac{a}{V^2}, \tag{23.73}$$

where the notation

$$\begin{aligned} a &= -2\pi N^2 \int_d^\infty u\,(r)\, r^2 dr > 0, \\ b &= \frac{2\pi N}{3} d^3 = 4Nv_0 \end{aligned} \tag{23.74}$$

is used, and v_0 is the volume of a molecule (an atom).

At last, let us rewrite equation (23.73) in the form

$$p + \frac{a}{V^2} = \frac{Nk_0T}{V^2}\,(V + b)$$

and multiply the result by $V - b$,

$$\left(p + \frac{a}{V^2}\right)(V - b) = \frac{Nk_0T}{V^2}\left(V^2 - b^2\right). \tag{23.75}$$

Neglecting the quantity b^2 in the right hand side, we obtain the **Van der Waals equation**

$$\left(p + \frac{a}{V^2}\right)(V - b) = Nk_0T. \tag{23.76}$$

For one mole of gas, it looks like

$$\left(p + \frac{a}{V^2}\right)(V - b) = RT. \tag{23.77}$$

The Van der Waals constants a and b in equation (23.77) are determined by formulas (23.74), in which $N = N_A$.

Let us examine the meaning of the Van der Waals constants a and b. From the plot depicted in Fig. 23.1, it follows that, for r-values within the interval $d < r < \infty$, the function $u(r) < 0$. Therefore, in accordance with (23.74), the constant a stems from the mutual attraction of atoms, so that the gas pressure on the vessel walls decreases as is shown by formula (23.73). On the other hand, the Van der Waals constant $b = 4Nv_0$ takes into account the own volume of atoms. As a result, the volume accessible for the atomic motion becomes smaller, and the pressure, accordingly, higher.

24

IDEAL QUANTUM GAS

In the previous chapter, the methods of statistical physics were applied to calculate the thermodynamic functions of the macroscopic system, the ideal classical and real gases. However, in section 21.1, it was noticed that any macrosystem is quantum-mechanical in principle; it may be considered as classical only in the limiting case, when the state of microparticles that it consists of can be characterized with the help of coordinates and momenta. However, even in this case, in order to obtain correct results while calculating the thermodynamic functions, the uncertainty relation for coordinates and momenta, i.e. the quantum-mechanical properties of microparticles (see section 21.3), has to be taken into account.

Besides, there exist macroscopic systems, which can be consider as ideal gases, i.e. when the interaction between particles can be neglected, but their state cannon be described with the help of coordinates and momenta. Therefore, attempts to apply the methods of classical statistics to them give rise to results that do not agree with experimental facts. Such systems include, e.g., free electrons in metals, electromagnetic radiation considered as the photon gas, and so forth. The ideal gas, the particles of which cannot be described by coordinate and momenta, and their state is determined by the wave function (quantum numbers), is called the **ideal quantum gas**.

The methods of statistical physics in both the cases of classical and quantum-mechanical systems are based on the canonical Gibbs distribution, which looks like (21.50) for quantum systems and like (21.57) for classical ones. If the system exchanges the energy and particles with the thermostat, the grand canonical Gibbs distribution (22.54) is used.

When applying the methods of statistical physics to the quantum ideal gas, not only the discreteness of states, but also the identity principle has to be taken into account. The latter gives rise to the separation of microparticles between bosons and fermions, and to the Pauli exclusion principle, according to which only one fermion can occupy a certain quantum state. In classical statistics, in particular, when deriving the Boltzmann distribution (23.52), this separation of particles is not taken into consideration. This circumstance can be explained by the fact that the number of accessible microstates for a particle in the classical ideal gas considerably exceeds the number of particles themselves. Therefore, the probability for two or more particles to stay in the same microstate almost equals zero. Hence, the difference between bosons and fermions is insignificant in this case (the criterion of applicability of classical statistics is given in section 24.2).

If the ideal gas is quantum, the number of microstates accessible for particles insignificantly exceeds the number of particles. Therefore, the probability for two or more particles to be in the same quantum state is now different from zero. As a result, the difference between bosons and fermions becomes quite important.

24.1. Fermi–Dirac and Bose–Einstein distributions

Let us consider a quantum ideal gas with N particles in the state of thermodynamic equilibrium. The wave functions of particles are denoted as ψ_α, where α stands for the set of quantum numbers that determine the wave function, i.e. the particle state. Those N gas particles are somehow distributed over the quantum states, with the number of particles n_α in the α-th quantum state changing permanently owing to collisions between the particles. However, if the gas is in the equilibrium state, the number n_α changes around a certain average value $\bar{n}_\alpha$, which is to be calculated.

With that end in view, let us consider the set of n_α particles as a subsystem with the variable number of particles in the thermostat. As the latter, we consider particles in other quantum states. Then, we may apply the grand canonical Gibbs distribution (22.54), to analyze the state of subsystem,

$$w\left(\varepsilon, n\right) = \frac{\Omega\left(\varepsilon, n\right) \mathrm{e}^{-\dfrac{\mu n - \varepsilon}{k_0 T}}}{\sum_{\varepsilon, n} \Omega\left(\varepsilon, n\right) \mathrm{e}^{-\dfrac{\mu n - \varepsilon}{k_0 T}}}. \tag{24.1}$$

Since all n_α particles are in the same α-th state and have the same energy ε_α, then $n = n_\alpha$, $\varepsilon = \varepsilon_\alpha n_\alpha$, and $\Omega = 1$, so that instead of (24.1) we obtain

$$w\left(n_\alpha\right) = \frac{\mathrm{e}^{-\dfrac{\mu - \varepsilon_\alpha}{k_0 T} n_\alpha}}{\sum_{n_\alpha} \mathrm{e}^{-\dfrac{\mu - \varepsilon_\alpha}{k_0 T} n_\alpha}}. \tag{24.2}$$

Formula (24.2) evaluates the probability for n_α particles to stay in the quantum state ψ_α. Therefore, the average value $\bar{n}_\alpha$ is calculated as

$$\bar{n}_\alpha = \sum_{n_\alpha} n_\alpha w\left(n_\alpha\right) = \frac{\displaystyle\sum_{n_\alpha} n_\alpha \mathrm{e}^{-\frac{\mu-\varepsilon_\alpha}{k_0 T} n_\alpha}}{\displaystyle\sum_{n_\alpha} \mathrm{e}^{-\frac{\mu-\varepsilon_\alpha}{k_0 T} n_\alpha}}. \tag{24.3}$$

For the sake of convenience in subsequent calculations, let us introduce the notation

$$\frac{\mu-\varepsilon_\alpha}{k_0 T} = \beta. \tag{24.4}$$

Then, instead of (24.3), we have

$$\bar{n}_\alpha = \frac{\sum_{n_\alpha} n_\alpha \mathrm{e}^{\beta n_\alpha}}{\sum_{n_\alpha} \mathrm{e}^{\beta n_\alpha}} = \frac{\dfrac{\partial}{\partial \beta} \displaystyle\sum_{n_\alpha} \mathrm{e}^{\beta n_\alpha}}{\sum_{n_\alpha} \mathrm{e}^{\beta n_\alpha}}. \tag{24.5}$$

Further calculations will be carried out separately for the gases of fermions and bosons.

First, let us consider the quantum ideal gas of fermions. According to the Pauli principle, only one fermion is allowed in the α-th quantum state (i.e. $n_\alpha = 0$ and 1). Hence,

$$\sum_{n_\alpha} \mathrm{e}^{\beta n_\alpha} = \sum_{n_\alpha = 0} \mathrm{e}^{\beta n_\alpha} = 1 + \mathrm{e}^\beta, \tag{24.6}$$

and, on the basis of (24.5) and (24.6), we obtain

$$\bar{n}_\alpha = \frac{\frac{\partial}{\partial \beta}\left(1+\mathrm{e}^\beta\right)}{1+\mathrm{e}^\beta} = \frac{\mathrm{e}^\beta}{1+\mathrm{e}^\beta} = \frac{1}{\mathrm{e}^{-\beta}+1}. \tag{24.7}$$

Substituting (24.4) and omitting the subscript α, we obtain the formula for the average number of fermions in the quantum state with the energy ε in the form of the **Fermi–Dirac distribution,**

$$\bar{n}_\alpha = \frac{1}{\mathrm{e}^{-\frac{\varepsilon-\mu}{k_0 T}} + 1}. \tag{24.8}$$

On the other hand, if the quantum ideal gas consists of bosons, which do not obey the Pauli principle, the number of particles in the α-th quantum state can by arbitrary within the limits $0 \leq n_\alpha \leq N$. Hence,

$$\sum_{n_\alpha} \mathrm{e}^{\beta n_\alpha} = \sum_{n_\alpha=0}^{N} \mathrm{e}^{\beta n_\alpha}. \tag{24.9}$$

The number of particles is very large. Therefore, we may assume that $N \to \infty$. However, sum (24.9) converges in this case, if $\mathrm{e}^{\beta} \leq 1$, i.e. under the condition

$$\beta = \frac{\mu - \varepsilon_\alpha}{k_0 T} \leq 0. \tag{24.10}$$

Since $0 \leq \varepsilon_\alpha < \infty$, it follows from (24.10) that the chemical potential of bosons has the satisfy the condition

$$\mu \leq 0. \tag{24.11}$$

Then, sum (24.9) is a sum of an infinite descending geometric series,

$$\sum_{n_\alpha} \mathrm{e}^{\beta n_\alpha} = \sum_{n_a=0} \mathrm{e}^{\beta n_\alpha} = \frac{1}{1 - \mathrm{e}^{\beta n_\alpha}}. \tag{24.12}$$

Substituting (24.12) into (24.5), we obtain

$$\bar{n}_\alpha = \frac{1}{1 - \mathrm{e}^{\beta}} \frac{\partial}{\partial \beta} \frac{1}{1 - \mathrm{e}^{\beta}} = \frac{1}{\mathrm{e}^{-\beta} - 1}.$$

Taking into account (24.4), the average number of bosons in the quantum state (status) with the energy ε equals

$$\bar{n}_\alpha = \frac{1}{\mathrm{e}^{\frac{\varepsilon - \mu}{k_0 T}} - 1}. \tag{24.13}$$

Formula (24.13) is called the **Bose–Einstein distribution**.

24.2. Transition to classical statistics. Gas degeneration

The classical statistics of ideal gas is based on the Boltzmann distribution (23.52),

$$\bar{n} = \mathrm{e}^{\frac{\mu - \varepsilon}{k_0 T}}, \tag{24.14}$$

or the equivalent Maxwell distribution (23.12). At the same time, the statistics of quantum ideal gas is based on the Fermi–Dirac distribution (24.8) for fermions and the Bose–Einstein distribution (24.13) for bosons,

$$\bar{n}_\alpha = \frac{1}{\mathrm{e}^{\frac{\varepsilon - \mu}{k_0 T}} \pm 1}. \tag{24.15}$$

Moreover, the classical system is a limiting case of the quantum system. Therefore, a question arises concerning conditions, under which quantum distributions (24.15) transform into classical distribution (24.14).

Comparing those two formulas, one can see that the transition to the Boltzmann distribution can be made if

$$\mathrm{e}^{\frac{\varepsilon - \mu}{k_0 T}} \gg 1 \tag{24.16}$$

or, taking into account the value $\varepsilon = 0$, if

$$\mathrm{e}^{-\frac{\mu}{k_0 T}} \gg 1. \tag{24.17}$$

If condition (24.16) is satisfied, it follows from (24.15) that $\bar{n} \ll 1$. Hence, the number of accessible microstates should considerably exceed the number of particles, which was discussed at the beginning of this chapter. Let us analyze condition (24.17) in more details.

The chemical potential μ is determined using the relation

$$N = \int \bar{n} dg.$$

Taking into account (24.25) and (24.14), we obtain

$$N = aV e^{\frac{\mu}{k_0 T}} \int_0^\infty \varepsilon^{\frac{1}{2}} \mathrm{e}^{-\frac{\varepsilon}{k_0 T}} d\varepsilon.$$

Making the substitution $\varepsilon = x^2$, the integral is reduced to one of Poisson integrals. Therefore,

$$N = 2aV\mathrm{e}^{\dfrac{\mu}{k_0T}} \int_0^\infty x^2 \mathrm{e}^{-\dfrac{x^2}{k_0T}}\,dx = \frac{1}{2}aV\mathrm{e}^{\dfrac{\mu}{k_0T}}\sqrt{\pi\left(k_0T\right)^3}. \tag{24.18}$$

The multiplier a in (24.18) looks like (see (21.26))

$$a = \frac{m^{3/2}\left(2l_s+1\right)}{\sqrt{2}\pi^2\hbar^3}.$$

From (24.18), it follows that

$$\mathrm{e}^{-\dfrac{\mu}{k_0T}} = \left(\frac{mk_0T}{2\pi\hbar^2}\right)^{3/2}\frac{2l_s+1}{n_0}, \tag{24.19}$$

where $n_0 = N/V$.

With the help of (24.17) and (24.19), we obtain a condition for the applicability of classical statistics,

$$T \gg \frac{2\pi\hbar^2}{mk_0}\left(\frac{n_0}{2l_s+1}\right) = T_0. \tag{24.20}$$

The temperature

$$T_0 = \frac{2\pi\hbar^2}{mk_0}\left(\frac{n_0}{2l_s+1}\right)^{2/3} \tag{24.21}$$

is called the **temperature of gas degeneration**.

If the gas parameters satisfy the condition $T \gg T_0$, classical statistics (the Maxwell and Boltzmann distributions) are applicable to it, and the gas is called **nondegenerate**. Otherwise, if $T \leq T_0$, the gas is called **degenerate**. In the latter case, classical statistics is inapplicable, and one has to use the quantum Fermi–Dirac or Bose–Einstein distribution. In this connection, let us consider some examples.

First, let us consider a usual gas, e.g., oxygen. For the oxygen molecule, $m = M/N_A = 5.3 \times 10^{-26}$ kg, $l_s = 0$, and, under normal conditions, $n_0 = 2.7 \times 10^{25}$ m^{-3}. Then, with the help of formula (24.21), we obtain $T_0\left(\mathrm{O}_2\right) = 7.7 \times 10^{-3}$ K. So low are the degeneration temperatures for other usual gases as well. They are considerably lower than the gas liquefaction temperatures. Therefore, the condition $T \gg T_0$ is always satisfied for usual gases, and the Boltzmann (Maxwell) distribution is always applicable to them.

As another example, let us consider free electrons in metals. To a certain approximation, they can also be considered as an ideal gas. The electron mass $m = 9.1 \times 10^{-31}$ kg and the spin $l_s = \frac{1}{2}$. The electron concentration equals $n_0 = 8{\times}10^{28}$ m^{-3} for copper and $n_0 = 5.8{\times}10^{28}$ m^{-3} for silver. Then, making use of formula (24.21), we obtain $T_0\,(\mathrm{Cu}) = 6 \times 10^4$ K and $T_0\,(\mathrm{Ag}) = 4.7 \times 10^4$ K. Hence, at temperatures lower than the melting one in metals, the condition $T \ll T_0$ is satisfied. This means that the electron gas in metals is degenerate, so that classical statistics is inapplicable to it.

24.3. Quantum ideal Bose gas at low temperatures. Bose condensation

Let us consider a quantum ideal gas of bosons (the Bose gas) consisting of N identical particles, each of them having the rest mass m different from zero and the spin l_s. The number of particle microstates in the energy interval from ε to , in accordance with (21.25) and (21.26), is determined by the formula

$$dg = aV\sqrt{\varepsilon}d\varepsilon, \tag{24.22}$$

where

$$a = \frac{m^{1/2}\,(2l_s + 1)}{\sqrt{2}\pi^2\hbar^3}.$$

The total number of Bose-particles should be calculated as

$$N = \int \bar{n} dg, \tag{24.23}$$

which, taking into account (24.13) and (24.22), transforms into

$$N = aV \int_0^\infty \frac{\sqrt{\varepsilon}d\varepsilon}{\mathrm{e}^{\dfrac{\varepsilon - \mu}{k_0 T}} - 1}, \quad \mu = \mu\,(T)\,. \tag{24.24}$$

If the rest mass of Bose-particles differs from zero, the condition $N = \mathrm{const}$ must be satisfied. Then, the integral in (24.24) and, hence, the power exponent $(\varepsilon - \mu\,(T))\,/\,(k_0 T)$ must be independent of the temperature.

Therefore, as the temperature decreases, the value of $-\mu(T)$ must also go down. Since, according to (24.11), the chemical potential for bosons $\mu \leq 0$, the chemical potential has to grow and, at the temperature $T = T_0$, reach the maximum value $\mu(T_0) = 0$. In this case,

$$N = aV \int_0^{\infty} \frac{\sqrt{\varepsilon} d\varepsilon}{\mathrm{e}^{\frac{\varepsilon - \mu}{k_0 T}} - 1}. \tag{24.25}$$

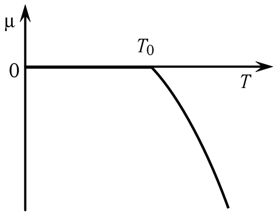

Fig. 24.1.

As the temperature decreases further, the chemical potential remains equal to zero (Fig. 24.1). Therefore, the exponential function in (24.25) will grow, and the number of particles N will decrease, which contradicts the condition $N = \text{const}$. This means that, at $T < T_0$, formula (24.25) does not take into account all particles, but only those with the energy $\varepsilon > 0$. In other words, this formula determines the number of particles N' that are not occupy the level $\varepsilon = 0$.

Changing the variable $x = \dfrac{\varepsilon}{k_0 T}$ in integral (24.25), we obtain

$$N' = (k_0 T)^{3/2} aV \int_0^{\infty} \frac{x^{1/2} dx}{\mathrm{e}^x - 1}. \tag{24.26}$$

The integral in (24.26) can by expressed only in terms of special functions (see Appendix N). Its numerical value approximately equals 2.315. Therefore,

$$N' \approx 2,315 aV (k_0 T)^{3/2}. \tag{24.27}$$

According to this formula, $N' = 0$ at $T = 0$. This means that, at $T = 0$, all particles in the Bose gas occupy the ground level with the energy $\varepsilon = 0$. This phenomenon is called the **Bose-Einstein condensation** (**Bose condensation**). It begins at the temperature $T = T_0$, which can be determined with the help of (24.25) and taking into account (24.27),

$$N = 2.315 aV (k_0 T_0)^{3/2}.$$

Making allowance for (24.22), we obtain

$$T_0 = \frac{3.31 \hbar^2}{m k_0} \left(\frac{n_0}{2 l_s + 1} \right)^{2/3}. \tag{24.28}$$

The temperature T_0, below which the Bose condensation tales place, is called the **temperature of Bose condensation** or the **temperature of Bose gas degeneration** (cf. (24.21) and (24.28)). The system of Bose gas particles on the ground energy level at $T < T_0$ is called the **Bose condensate**. This term is conditional, because it has nothing in common with the condensation of gas (vapor) into liquid and means only a transition of Bose particles into the ground energy state.

Since liquid has specific properties in comparison with gas, the Bose condensate also has its specific properties. At the energy of particle $\varepsilon = 0$, its momentum $p = 0$, i.e. the momentum has an exact value. In this case, however, according to the uncertainty relation $\Delta x \Delta p_x = \hbar$, the particle coordinate is completely indefinite (to within the volume V). As a result, the concepts of a separate particle in the Bose condensate loses any sense, and the Bose condensate has to be considered as a whole, in particular, while analyzing its energy spectrum.

The first excited energy level in the Bose condensate corresponds to the energy required for a particle to escape from it. However, at $T < T_0$, this energy exceeds the average thermal energy of particles $k_0 T$. Therefore, taking into account that the variation of Bose condensate energy can occur owing to the interaction with particles that do not belong to it, we come to a conclusion that the Bose condensate moves as a whole without any resistance and “friction” (superconductivity, superfluidity).

The observation of Bose condensation in usual gases with integer-spin molecules is complicated due to crystallization at temperatures much higher than T_0. An exclusion is liquid helium-4 (the ${}_2\mathrm{He}^4$ isotope), which remains in the liquid state at atmospheric pressure and temperatures close the absolute zero. Helium-4 is liquefied under atmospheric pressure at the temperature $T = 4.2$ K. It has the density $\rho = 130$ kg/m^3 and the concentration of atoms $n_0 = \rho N_A / M = 1.95 \times 10^{28}$ m^{-3}. Using those data and formula (24.28), we obtain the temperature of Bose condensation for helium-4 atoms $T_0 = 2.6$ K.

Liquid helium-4 has a phase transition of the second kind at a temperature of $T_c = 2.18$ K, which results in the emergence of a superfluid component (helium-II). The motion of the latter is not accompanied by friction, and this component is identified as the Bose condensate. At $T \to 0$, liquid helium becomes a superfluid liquid in whole. The temperature difference between $T_0 = 2.6$ K and $T_c = 2.18$ K is explained by the fact that liquid helium is not an ideal gas.

24.4. Equilibrium radiation as a Bose gas of photons

In section 24.3, the ideal Bose gas of particles with the rest mass different from zero was considered. Here, using the equilibrium electromagnetic radiation as an example, let us study the ideal gas of zero-mass bosons.

From the quantum-mechanical viewpoint, electromagnetic irradiation is a set of photons, the rest mass of which equals zero. They have the energy $\hbar\omega$ and the momentum $\hbar\vec{k}$, where ω is the cyclic frequency, and $\vec{k}$ the wave vector ($\left|\vec{k}\right| = 2\pi/\lambda$). The spin of photons $l_s = 1$, i.e. photons belong to bosons.

The photon spin can have two orientations: along and against the photon momentum vector. In classical electrodynamics, those orientations correspond to two mutually perpendicular orientations of the electric field strength vector, both being perpendicular to the propagation direction of electromagnetic wave

Hence, equilibrium radiation can be considered as a Bose gas of photons. However, according to the superposition principle, photons do not interact with one another, but interact with charged particles. Therefore, the equilibrium photon gas can exist in a system consisting of radiation and a substance, the atoms of which can emit and absorb photons.

Photon is a completely quantum-mechanical concept. Therefore, the photon gas has to be considered using quantum statistics, i.e. the Bose–Einstein distribution. In section 24, it was shown that the chemical potential of bosons satisfies the condition $\mu \leq 0$. Let us prove that it equals zero for photons.

A system consisting of photons and a substance is in the state of thermodynamic equilibrium provided the constant temperature and pressure over it. Therefore, the thermodynamic condition of its equilibrium is (see section 20.4) is the equality $\delta\Phi = 0$. However, the thermodynamic Gibbs potential Φ is an additive quantity and equals a sum of the thermodynamic potentials of photons, Φ_{ph}, and the substance, Φ_s. Therefore,

$$\delta\Phi = \delta\left(\Phi_{ph} + \Phi_s\right) = \delta\Phi_{ph} + \delta\Phi_s = 0. \tag{24.29}$$

Taking into account that, according to (20.41),

$$d\Phi = -SdT + Vdp + \mu dN$$

in the system with the variable number of particles, equilibrium condition (24.29) is written as follows:

$$-S_{ph}\delta T + V_{ph}\delta p + \mu_{ph}\delta N_{ph} - S_s\delta T + V_s\delta p + \mu_s\delta N_s.$$

However, in our case, $T = \text{const}$, $p = \text{const}$, and $N_s = \text{const}$, i.e. $\delta T = \delta p = \delta N_s = 0$, whereas $N_{ph} \neq \text{const}$, i.e. $\delta N_{ph} \neq 0$. Therefore,

$$\mu_{ph}\delta N_{ph} = 0$$

or $\mu_{ph} = 0$.Hence, the Bose-Einstein distribution with respect to the photons of equilibrium radiation acquires the form

$$\bar{n} = \frac{1}{\mathrm{e}^{\frac{\hbar\omega}{k_0 T}} - 1}. \tag{24.30}$$

Since photons are zero-mass particles, the number of quantum states of the photon with the momentum within the interval from p to $p + dp$, in accordance with (21.27), equals

$$dg = \frac{V\chi}{2\pi^2\hbar^3}p^2 dp,$$

where χ is the number of independent polarisations (in our case, it equals two). Taking into account that $p = \hbar k = \hbar\omega/c$, we find

$$dg = \frac{\omega^2 V d\omega}{\pi^2 c^3}. \tag{24.31}$$

On the basis of (24.30) and (24.31), the energy of the gas of photons with frequencies within the interval from ω to $\omega + d\omega$ per unit volume equals

$$\rho\left(\omega, T\right) d\omega = \hbar\omega\bar{n}dg = \frac{\hbar\omega^3}{\pi^2 c^3}\frac{d\omega}{\mathrm{e}^{\frac{\hbar\omega}{k_0 T}} - 1}, \tag{24.32}$$

so that the spectral density of radiation is

$$\rho\left(\omega, T\right) = \frac{\hbar\omega^3}{\pi^2 c^3}\frac{1}{\mathrm{e}^{\frac{\hbar\omega}{k_0 T}} - 1}. \tag{24.33}$$

Formula (24.33) was derived for the first time by M. Planck in 1900, and it is called the **Planck formula**.

In the long-wave approximation, $\hbar\omega/\left(k_0T\right) \ll 1$, we have

$$e^{\dfrac{\hbar\omega}{k_0T}} \approx 1 + e^{\dfrac{\hbar\omega}{k_0T}}.$$

In this case, the Planck formula transforms into the **Rayleigh-Jeans** one

$$\rho\left(\omega, T\right) = \frac{\omega^2}{\pi^2 c^3} k_0 T. \tag{24.34}$$

In the short-wave approximation, $\hbar\omega/\left(k_0T\right) \ll 1$, the unity in the denominator in the right hand side of formula (24.33) can be neglected. As a result, we obtain the **Wien formula**

$$\rho\left(\omega, T\right) = \frac{\hbar\omega^3}{\pi^2 c^3} e^{\dfrac{\hbar\omega}{k_0T}}. \tag{24.35}$$

On the basis of (24.33), the spectral radiant emittance of blackbody equals

$$\varepsilon\left(\omega, T\right) = \frac{c}{4}\rho\left(\omega, T\right) = \frac{\hbar\omega^3}{4\pi^2 c^3} \frac{1}{e^{\dfrac{\hbar\omega}{k_0T}} - 1}. \tag{24.36}$$

Function (24.36) describes the energy distribution over the frequencies in the blackbody radiation spectrum. To find the energy distribution over the wavelengths, it is necessary to use the relationship

$$\varepsilon\left(\omega, T\right)\left|d\omega\right| = \varepsilon\left(\lambda, T\right)\left(d\lambda\right),$$

i.e.

$$\varepsilon\left(\lambda, T\right) = \varepsilon\left(\omega, T\right)\left|\frac{d\omega}{d\lambda}\right| = \frac{2\pi c}{\lambda^2}\varepsilon\left(\omega, T\right).$$

Then, we obtain

$$\varepsilon\left(\lambda, T\right) = \frac{4\pi\hbar c^2}{\lambda^5} \frac{1}{e^{\dfrac{2\pi\hbar c}{\lambda k_0 T}} - 1}. \tag{24.37}$$

The plot of function (24.37) is depicted in Fig. 24.2.

Let us determine the wavelength λ_m at the maximum of the spectral radiant emittance of blackbody. For this purpose, we equate the derivative $d\varepsilon(\lambda, T)/d\lambda$ to zero and obtain the equation

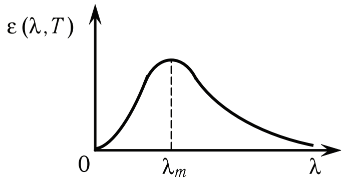

Fig. 24.2.

$$\frac{4\pi\hbar c^2}{\lambda^6}\frac{1}{\mathrm{e}^{\dfrac{2\pi\hbar c}{\lambda k_0 T}}-1}\left\{-5+\frac{\dfrac{2\pi\hbar c}{\lambda k_0 T}\mathrm{e}^{\dfrac{2\pi\hbar c}{\lambda k_0 T}}}{\mathrm{e}^{\dfrac{2\pi\hbar c}{\lambda k_0 T}}-1}\right\}=0. \tag{24.38}$$

It has three roots. Two of them, $\lambda_1 = \infty$ and $\lambda_2 = 0$, do not correspond to the maximum of the function $\varepsilon(\lambda, T)$ (see Fig. 24.2). The required root satisfies the equation

$$-5+\frac{xe^x}{\mathrm{e}^x-1}=0, \tag{24.39}$$

where the notation

$$x=\frac{2\pi\hbar c}{\lambda k_0 T} \tag{24.40}$$

was used. Equation (24.39) can be rewritten in the form of transcendental equation,

$$\mathrm{e}^x=\frac{5}{5-x}. \tag{24.41}$$

Its root $x \approx 4.965$ can be found graphically. Then, on the basis of (24.40), we obtain **Wien's displacement law**

$$\lambda_m=\frac{2\pi\hbar c}{4,96 k_0 T}=\frac{b}{T}, \tag{24.42}$$

in which **Wien's displacement constant** b equals

$$b=\frac{2\pi\hbar c}{4.96 k_0}=2.89\times10^{-3}\ \mathrm{m}\times\mathrm{K}. \tag{24.43}$$

According to (24.33), the total energy density of photon gas is calculated as

$$u\left(T\right)=\int_{0}^{\infty}\rho\left(\omega,\ T\right)d\omega=\frac{\hbar}{\pi^{2}c^{3}}\int_{0}^{\infty}\frac{\omega^{3}d\omega}{\mathrm{e}^{\frac{\hbar\omega}{k_{0}T}}-1}. \qquad (24.44)$$

Introducing the dimensionless variable $x=\hbar\omega/\left(k_{0}T\right)$, we obtain

$$u\left(T\right)=\frac{\hbar}{\pi^{2}c^{3}}\left(\frac{k_{0}T}{\hbar}\right)^{4}\int_{0}^{\infty}\frac{x^{3}dx}{\mathrm{e}^{x}-1}. \qquad (24.45)$$

The integral in (24.45) is calculate in terms of special functions (see Appendix N). It equals $\pi^{4}/15$. Therefore,

$$u\left(T\right)=\frac{\pi^{2}k_{0}^{4}}{15c^{3}\hbar^{3}}T^{4}=\sigma_{0}T^{4}, \qquad (24.46)$$

where

$$\sigma_{0}=\frac{\pi^{2}k_{0}^{4}}{15c^{3}\hbar^{3}}=7.6\times10^{-16}\ \frac{\mathrm{J}}{\mathrm{m}\times\mathrm{K}}.$$

The total radiant emittance of blackbody is determined using the **Stefan–Boltzmann law**

$$\varepsilon\left(T\right)=\frac{A}{4}u\left(T\right)=\sigma T^{4}, \qquad (24.47)$$

in which the coefficient

$$\sigma=\frac{\pi^{2}k_{0}^{4}}{60c^{2}\hbar^{3}}=5.67\times10^{-8}\ \frac{\mathrm{W}}{\mathrm{m}^{2}\times\mathrm{K}^{4}}$$

is called the **Stefan–Boltzmann constant**.

24.5. Free electrons in metals as a degenerate Fermi gas

The availability of free elements in metals follows from experimental facts (electroconductivity of metals, the Stewart–Tolman experiments, and so forth) and is proved theoretically, in particular, in the framework of the band theory (see, e.g., [22]). Experimental data testify that, in general, metals contain one free electron per one metal atom. Hence, their concentration should be $n_0 \sim 10^{28} \div 10^{29}\ \mathrm{m}^{-3}$.

In effect, free electrons in metals are not absolutely free, because they interact both with one another and with the positive ions of crystal lattice. Of special importance is their interaction with lattice ions that exert thermally induced vibrations around their equilibrium positions, which gives rise to the electron scattering at those lattice vibrations. Hence, free electrons in metals can be considered as the ideal gas only approximately. However, even if the model of ideal gas is applied, some obtained results agree, at least at the qualitative level, with experimental data.

In section 24.3, it was shown that the gas of free electrons in metals is so degenerate that the methods of classical statistics are inapplicable. Hence, since electrons are fermions, the Fermi–Dirac distribution (24.8),

$$\bar{n} = \frac{1}{\mathrm{e}^{\frac{\varepsilon - \mu}{k_0 T}} + 1}, \tag{24.48}$$

where $\mu = \mu(T)$, has to be applied.

First, let us consider the limiting case $T = 0$. From the viewpoint of band theory, free electrons in metals are electrons in the partially filled conduction band or electrons in a completely filled valence band that overlaps with the higher located free (conduction) band. The conduction band contains $2N(2l+1)$ electron energy sublevels, where N is the number of metal atoms, and l is the orbital quantum number for the atomic energy level, which transforms into the band in the crystal.

At $T = 0$, the energy of the electron Fermi gas has to be minimal. Besides, according to the Pauli principle, every quantum state can be occupied by only one electron. Therefore, at $T = 0$, free electrons fill the lowest N energy levels in the conduction band from the level with the energy $\varepsilon = 0$ (the conduction band bottom) to the level with the energy $\varepsilon = \varepsilon_{\max}$, which is called the **Fermi level** and which energy is designated as ε_{F} ($\varepsilon_{\mathrm{F}} = \varepsilon_{\max}$). Besides, at $T = 0$, the chemical potential $\mu(T = 0) = \mu_0$. Therefore, taking into account that

$$\lim_{T \to 0} \mathrm{e}^{\frac{\varepsilon - \mu_0}{k_0 T}} = \begin{cases} 0, & \varepsilon < \mu_0, \\ \infty, & \varepsilon > \mu_0, \end{cases}$$

we obtain for the Fermi–Dirac distribution (24.48) that

$$\bar{n}\,(T=0)=\lim_{T\to 0}\frac{1}{\mathrm{e}^{\frac{\varepsilon-\mu_0}{k_0T}}+1}=\begin{cases}1, \ \varepsilon<\mu_0,\\ 0, \ \varepsilon>\mu_0.\end{cases} \tag{24.49}$$

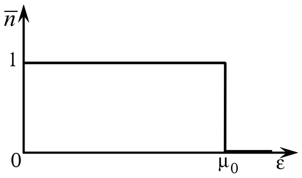

Fig. 24.3.

The plot of distribution (24.49) is schematically shown in Fig. 24.3.

In order to calculate the Fermi energy $\varepsilon_{\mathrm{F}}=\mu_0$, we can use the formula

$$N=\int \bar{n}dg,$$

in which $0\le\varepsilon\le\mu_0$, $\bar{n}=1$, and $dg=aV\sqrt{\varepsilon}d\varepsilon$ (see (24.22)). Then,

$$N=aV\int_0^{\mu_0}\varepsilon^{1/2}d\varepsilon=\frac{2}{3}aV\mu_0^{3/2}. \tag{24.50}$$

Substituting the value of the parameter a (see (21.26) or (24.22)) into (24.50) and taking into account the electron spin $l_s=\frac{1}{2}$, we obtain

$$\mu_0=\left(\frac{3N}{2aV}\right)^{2/3}=\frac{\hbar}{2m}\left(3\pi^2n_0\right)^{2/3}. \tag{24.51}$$

Analogously, we determine the total energy of the free electron gas at $T=0$,

$$U_0\int_0^{\mu_0}\varepsilon\bar{n}dg=aV\int_0^{\mu_0}\varepsilon^{3/2}d\varepsilon=\frac{2}{5}aV\mu_0^{5/2}=\frac{3}{5}N\mu_0. \tag{24.52}$$

The Fermi energy $\varepsilon_{\mathrm{F}}=\mu_0$ is the maximum kinetic energy of free electrons in the metal at $T=0$. Since, the electron mass $m=9.1\times10^{-31}$ kg, we obtain $\varepsilon_{\mathrm{F}}\,(\mathrm{Ag})=5$ eV for silver ($n_0=5.5\times10^{28}$ m^{-3}) and $\varepsilon_{\mathrm{F}}\,(\mathrm{Cu})=6.5$ eV for copper ($n_0=8.8\times10^{28}$ m^{-3}). Then, in accordance with (24.52), the total energy of 1 mol of free electrons in silver equals $U_0\approx0.3$ MJ.

Now, let us consider the gas of free electrons in the metal at temperatures different from zero ($T\neq0$) provided that $T\ll T_0$. At $T\neq0$, free electrons in the metal can exchange the energy with the ions of crystal lattice. The average energy of thermal vibrations of the latter equals k_0T. For $T=300_k$, we obtain $k_0T\approx0.02$ eV; hence, $T\ll T_0$.

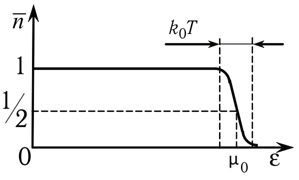

Fig. 24.4.

Therefore we may put $\mu(T) \approx \mu_0$ and, which is important, assume that only those electrons can exchange the energy with the crystal lattice that occupy the energy levels located within the interval of about k_0T near the Fermi level. For other electrons, this exchange is forbidden by the Pauli principle. Therefore, the distribution function at $T \neq 0$ looks like the plot schematically depicted in Fig. 24.4.

Detailed calculations demonstrate (see Appendix P) that, at $T \neq 0$,

$$\mu(T) = \mu_0 \left\{ 1 - \frac{1}{12} \left(\frac{\pi k_0 T}{\mu_0} \right)^2 \right\}; \qquad (24.53)$$

$$U(T) = U_0 \left\{ 1 + \frac{5}{12} \left(\frac{\pi k_0 T}{\mu_0} \right)^2 \right\}. \qquad (24.54)$$

From (24.54), it follows that the molar heat capacity of the free electron gas in the metal equals

$$C_V \frac{\partial U}{\partial T} = \frac{5}{6} U_0 \left(\frac{\pi k_0}{\mu_0} \right)^2 T = \frac{\pi^2 k_0 T}{3\mu_0} \frac{3}{2} R = \frac{\pi^2 k_0 T}{3\mu_0} C_{V,\text{class}}.$$

Hence, at $T = 300_k$, $C_V \approx 10^{-2} C_{V,\text{class}}$.

EXERCISES TO PART III

1. Find the work done by 1 mol of ideal gas in an isothermal process.

   - Solution: $A = RT \ln \dfrac{V_2}{V_1}$.

2. Calculate the entropy of ideal gas.

   - Solution: $S = C_V \ln T + R \ln V$.

3. Calculate the work of evaporation for 1 mol of water at a temperature of 100 °C and normal atmospheric pressure if the specific volumes of water and its vapor equal 1 and 1674 cm^3/g, respectively.

   - Solution: $A = 3000$ J.

4. What fraction of the heat spent to produce water vapor at a temperature of 100 °C and normal atmospheric pressure is spent to do the work? The specific volumes of water and its vapor equal 1 and 1674 cm^3/g, respectively, and specific heat of evaporation is $\lambda = 2.26 \times 10^6$ J/kg.

   - Solution: $A/Q = 0.074$.

5. Prove that $\lim\limits_{T \to 0} C(T) = 0$.

6. Find the free energy of 1 mol of single-atomic ideal gas. Instruction: use the results of exercise 2.

   - Solution: $F = C_V T(1 - \ln T) - RT \ln V + U_0$.

7. Calculate the changes of the free energy and the thermodynamic Gibbs potential at the evaporation of 1 mol of liquid chlorine at the normal boiling point ($t = -34.7$ °C). The density of liquid chlorine $\rho = 1570\ \mathrm{kg/m}^3$, and the molar mass $M = 0.03546$ kg/mol. Instructions: use the formulas $dF = -SdT - pdV$ and $d\Phi = -SdT + VdP$.

   - Solution: $\Delta F = RT - \dfrac{pm}{\rho} = -1978$ J, $\Delta\Phi = 0$.

8. Find the shift of the ice melting point when the pressure changes by $\Delta p = 1$ atm (101.325 kPa). The specific volume of ice $V_1 = 1.09 \times 10^{-3}$ m^3/kg, the specific volume of water $V_2 = 1 \times 10^3$ m^3/kg, the specific heat of ice melting $\Lambda = 3.4 \times 10^5$ J/kg. Instruction: use the Clapeyron–Clausius equation.

    - Solution: $\Delta T/\Delta p = 7.2 \times 10^{-4}$ K/atm.

9. A soap bubble with an initial diameter of 2 cm is blown up isothermally to a diameter of 10 cm. Find the change of its free energy if the surface tension coefficient for the soap solution equals $\sigma = 45$ mN/m.

    - Solution: $\Delta F = 2.7$ mJ.

10. Find the equation for the phase trajectory of a linear harmonic oscillator.

    - Solution: $\dfrac{x^2}{A^2} + \dfrac{p^2}{m^2\omega^2A^2} = 1$, where A is the amplitude.

11. Find the equation for the phase trajectory of a particle with the mass m that moves along the axis x under the action of the constant force F if, at the initial moment ($t = 0$), $x_0 = 0$ and $v = 0$.

    - Solution: $p^2 = 2mFx$.

12. Find the number of quantum states for a photon with the frequency in the interval from ω to $\omega + d\omega$ in the volume V.

    - Solution: $dg = \dfrac{\omega^2 V d\omega}{\pi^2 c^3}$.

13. Substantiate the number of accessible microstates for a classical single-atomic ideal gas in the volume V and with the energy in the interval ΔE. Instructions: use the solution of the problem in section 4.14 in [22].

    - Solution: $\Omega = \dfrac{3NA_N V^N E^{\frac{3}{2}N-1}\Delta E}{2N!(2\pi\hbar)^{3N}}$,
      where $A_N = \left(\dfrac{4\pi me}{3N}\right)^{3N/2}$.

14. Two bodies with the constant temperatures $t_1 = 27$ °C and $t_2 = 28$ °C exchange the heat $\delta Q = 10^{-7}$ J. Find the change of the system entropy.

    - Solution: $\Delta S = \delta Q \left(\frac{1}{T_1} - \frac{1}{T_2} \right) = 1.1 \times 10^{-12}\ \frac{\mathrm{J}}{\mathrm{K}}$.

15. According to the condition and the result of previous exercise, find the change of the thermodynamic probability of the system state (the number of accessible microstates). Instructions: use the Boltzmann principle.

    - Solution: $\frac{W_2}{W_1} = \exp \frac{\Delta S}{k_0} \approx e^{10^{11}}$.

16. Two bodies with the constant temperatures $t_1 = 27$ °C and $t_2 = 28$ °C exchange the heat. Find the probability that a heat quantity of (a) 10^{-7} J and (b) 1.2×10^{-18} J will be transferred from the cooler body to the warmer one. Draw a conclusion.

    - Solution: (a) $\frac{W_1}{W_2} = \exp \left(-\frac{\Delta S_1}{k_0} \right) \approx e^{-10^{11}}$,

      (b) $\frac{W_1}{W_2} = \exp \left(-\frac{\Delta S_2}{k_0} \right) = e^{-1} \approx 0.37$.

17. Find the number of oxygen molecules that are contained in a volume of 1 cm^3 under normal conditions and have the velocities in the interval from 600 to 610 km/s. Instruction: take into account that $\Delta v \ll v$; hence, $\Delta N \approx dN$.

    - Solution: $\Delta N = 3.2 \times 10^{13}\ \mathrm{cm}^{-3}$.

18. Find the most probable kinetic energy ε_m for a molecule of ideal gas.

    - Solution:.$\varepsilon_m = k_0 T/2$.

19. Calculate the altitude h, at which the air pressure is three times as lower as at $h = 0$. The air temperature equals 0 °C.

- Solution: $h = \dfrac{RT}{Mg} \ln \dfrac{p_0}{p} = 8.76$ km.

20. What fraction of oxygen molecules in the Earth atmosphere can overcome the Earth gravitational field at a temperature of 300 K? Draw a conclusion.

    - Solution: $n/n_0 = e^{-802} \approx 10^{-348} \to 0$.

21. A quantum ideal gas consists of particles, every of which can be in two nondegenerate quantum states with the energies ε_0 and ε_1. Find the internal energy of the gas.

    - Solution: $U = k_0 T^2 \dfrac{\partial \ln z}{\partial T} = N\varepsilon_0 T \dfrac{\mathrm{e}^{T_c/T}}{1+\mathrm{e}^{T_c/T}}$, where $T_c = \dfrac{\varepsilon_1 - \varepsilon_0}{k_0}$.

22. At what temperature does the energy density of equilibrium radiation equal the energy density of the classical ideal gas under normal conditions?

    - Solution: $T = \left(\dfrac{3n_0k_0}{2\sigma_0}\right)^{1/3} \approx 10^6$ K, where $n_0 = 2.7 \times 10^{25}\ \mathrm{m}^{-3}$ and $\sigma_0 = 7.6 \times 10^{-16}\ \dfrac{\mathrm{J}}{\mathrm{m}^3\mathrm{K}^4}$.

23. Calculate the light pressure at the centre of atomic bomb explosion, where the temperature equals $T = 10^8$ K. The radiation is assumed to be at equilibrium.

    - Solution: $p = 2.3 \times 10^{11}$ atm.

24. Calculate the number of photons in a unit volume of equilibrium thermal radiation at the temperature $T = 300$ K. Instruction: use the Planck formula and Appendix N.

    - Solution: $N = \dfrac{2.404}{\pi^2 c^3}\left(\dfrac{k_0 T}{\hbar}\right)^3 = 5.5 \times 10^{14}\ \mathrm{m}^{-3}$.

Appendices

APPENDIX A

DIRAC DELTA-FUNCTION

There are physical quantities (functions) that differ from zero in a narrow interval of their argument (the coordinate, time, frequency, and so on). At the same time, the corresponding function can be considered to differ from zero only at a single value of the variable x, e.g., at $x = x_0$. However, all physical quantities are spread either in space, or in time, or over the frequencies, and so on. Therefore, P. Dirac introduced a special function, $\delta(x-x_0)$, which is now called the Dirac delta-function. It has the following properties:

$$\delta(x - x_0) = \begin{cases} \infty & \text{if } x = x_0, \\ 0 & \text{if } x \neq x_0, \end{cases} \tag{A.1}$$

$$\int_{-\infty}^{\infty} \delta(x - x_0) dx = 1, \tag{A.2}$$

$$\int_{-\infty}^{\infty} F(x)\delta(x - x_0) dx = F(x_0). \tag{A.3}$$

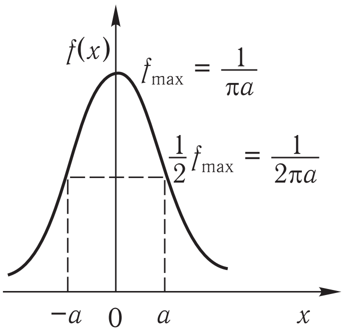

Fig. A.1.

The Dirac delta-function can be considered as the limiting case for various functions. For instance, this is the function

$$f(x) = \frac{1}{\pi} \frac{a}{x^2 + a^2}, \tag{A.4}$$

where a is a constant. The plot of function (A.4) is shown in Fig. A.1. It is easy to find that

$$f_{\max} = f(0) = \frac{1}{\pi a}, \quad f(\pm a) = \frac{1}{2} f_{\max} = \frac{1}{2\pi a}.$$

Besides,

$$\int_{-\infty}^{\infty} f(x)dx = \frac{a}{\pi}\int_{-\infty}^{\infty}\frac{dx}{x^2+a^2} = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{d\left(\dfrac{x}{a}\right)}{\left(\dfrac{x}{a}\right)^2+1}$$

$$= \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{dz}{z^2+1} = \frac{1}{\pi}\left[\arctan z\right]_{-\infty}^{\infty} = 1, \quad z = \frac{x}{a}. \qquad \text{(A.5)}$$

Notice that the value integral (A.5) does not depend on the parameter a. As follows from Fig. A.1, if the parameter a decreases, the peak in the plot becomes narrower, whereas the value of $f_{\max}$ grows. At $a \to 0$,

$$\lim_{a\to 0}\frac{1}{\pi}\frac{a}{x^2+a^2} = \begin{cases} \infty & \text{if } x = 0, \\ 0 & \text{if } x \neq 0. \end{cases} \qquad \text{(A.6)}$$

On the basis of (A.5) and (A.6), we may write

$$\lim_{a\to 0}\frac{1}{\pi}\frac{a}{x^2+a^2} = \delta(x-0) = \delta(x) \quad (x_0 = 0). \qquad \text{(A.7)}$$

It is clear that

$$\lim_{a\to 0}\frac{1}{\pi}\frac{a}{(x-x_0)^2+a^2} = \delta(x-x_0). \qquad \text{(A.8)}$$

Let us consider property (A.3) in more details. Since the delta-function $\delta(x-x_0)$ differs from zero only at $x = x_0$, integral (A.3) can be written as

$$\int_{-\infty}^{\infty} F(x)\delta(x-x_0)dx = \int_{x_0-\varepsilon}^{x_0+\varepsilon} F(x)\delta(x-x_0)dx,$$

where ε is an infinitesimally small number. The interval 2ε that contains the point $x = x_0$ is small, so that the function $F(x)$ can be considered constant within its limits and equal to $F(x_0)$. Therefore,

$$\int_{x_0-\varepsilon}^{x_0+\varepsilon} F(x)\delta(x-x_0)dx = F(x_0)\int_{x_0-\varepsilon}^{x_0+\varepsilon}\delta(x-x_0)dx$$

$$= F(x_0)\int_{-\infty}^{\infty}\delta(x-x_0)dx = F(x_0).$$

Hence,

$$\int_{-\infty}^{\infty} F(x)\delta(x-x_0)dx = F(x_0).$$

Another important representation of the delta-function is

$$\delta(x) = \int_{-\infty}^{\infty} \mathrm{e}^{ikx} dk. \tag{A.9}$$

The direct integration gives rise to the uncertainty. Therefore, let us write the right hand side of (A.9) in the form

$$I(x) = \lim_{N\to\infty} \frac{1}{2\pi} \int_{-N}^{N} \mathrm{e}^{ikx} dk = \lim_{N\to\infty} \frac{N}{\pi} \frac{\sin Nx}{Nx} = \lim_{N\to\infty} \varphi(N,x). \tag{A.10}$$

The plot of the function

$$\varphi(N,x) = \frac{N}{\pi} \frac{\sin Nx}{Nx} \tag{A.11}$$

is shown in Fig A.2. One can easily see that

$$\lim_{N\to\infty} \varphi(N,x) = \lim_{N\to\infty} \frac{N}{\pi} \frac{\sin Nx}{Nx} = \begin{cases} \infty & \text{if } x = 0, \\ 0 & \text{if } x \neq 0. \end{cases} \tag{A.12}$$

Moreover,

$$\int_{-\infty}^{\infty} \varphi(N,x)dx = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\sin Nx}{Nx} d(Nx) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\sin t}{t} dt = 1, \tag{A.13}$$

and the integral does not depend on N. From (A.12) and (A.13), we obtain (A.9). The general form of (A.9) looks like

$$\delta(x-x_0) = \int_{-\infty}^{\infty} \mathrm{e}^{ik(x-x_0)} dk. \tag{A.14}$$

Hence,

$$\delta(k-k_0) = \int_{-\infty}^{\infty} \mathrm{e}^{ix(k-k_0)} dx. \tag{A.15}$$

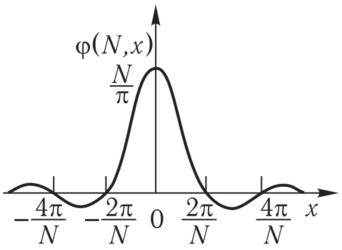

Fig. A.2.

Below, we present, without the proof, the relationships that are often used while solving problems:

$$\frac{1}{(2\pi)^3}\int \mathrm{e}^{i(\vec{k}-\vec{k}_0)\vec{r}} d^3r = \delta\left(\vec{k}-\vec{k}_0\right)$$

$$= \delta\left(k_x - k_{0x}\right)\delta\left(k_y - k_{0y}\right)\delta\left(k_z - k_{0z}\right), \tag{A.16}$$

$$\frac{1}{(2\pi)^3}\int \mathrm{e}^{i(\vec{r}-\vec{r}_0)\vec{k}} d^3k = \delta\left(\vec{r}-\vec{r}_0\right) = \delta\left(x - x_0\right)\delta\left(y - y_0\right)\delta\left(z - z_0\right), \tag{A.17}$$

$$\delta(ax) = \frac{1}{|a|}\delta(x). \tag{A.18}$$

APPENDIX B

ELEMENTS OF VECTOR ANALYSIS

Vectors

All physical quantities have a tensor character, i.e. they are tensors of some rank: zeroth (scalars), first (vectors), second, third, and so on. This classification is considered in Appendix E in more details. Here, we confine the consideration to scalars and vectors only.

The scalar is a quantity characterized by a single number, e.g., the length, mass, and so forth). The vector is a quantity that is characterized by a numerical value (the absolute value) and a direction. In the three-dimensional space, the vector is given by three numbers: these are either its absolute value and the two directional cosines or three components (A_x, A_y, A_z):

$$\vec{A} = A_x\vec{i} + A_y\vec{j} + A_z\vec{k}, \tag{B.1}$$

$$\vec{A} \equiv \{A_x, A_y, A_z\} \equiv \{A_i\}. \tag{B.2}$$

Operations with vectors:

(a) addition of two vectors (Fig. B.1,a and b),

$$\vec{A} + \vec{B} = \vec{C};$$

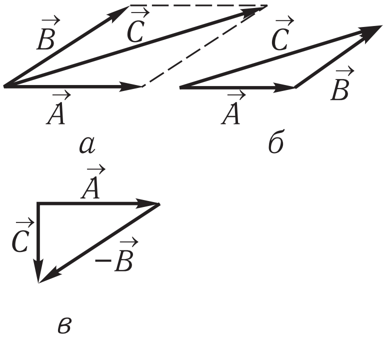

Fig. B.1.

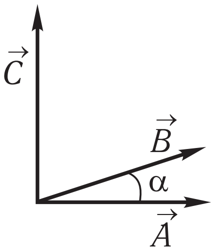

Fig. B.2.

(b) subtraction of two vectors (Fig. B.1,c),

$$\vec{A} - \vec{B} = \vec{A} + \left(-\vec{B}\right) = \vec{C};$$

(c) product of a vector and a scalar,

$$\lambda\vec{A} = \overrightarrow{\lambda A};$$

(d) scalar product (dot product) of two vectors,

$$\vec{A}\cdot\vec{B} = A\cdot B\cdot\cos\alpha = A_B\cdot B = A\cdot B_A$$

$$= A_xB_x + A_yB_y + A_zB_z; \tag{B.3}$$

(e) vector product (cross product) of two vectors (Fig. B.2),

$$\begin{aligned} \left[\vec{A}\cdot\vec{B}\right] &\equiv \vec{A}\times\vec{B}, \\ \left|\left[\vec{A}\cdot\vec{B}\right]\right| &= A\cdot B\cdot\sin\alpha, \end{aligned}$$

$$\left[\vec{A}\cdot\vec{B}\right] = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{vmatrix}$$

$$= \left(A_yB_z - A_zB_y\right)\vec{i} + \left(A_zB_x - A_xB_z\right)\vec{j} + \left(A_xB_y - A_yB_x\right)\vec{k}; \tag{B.4}$$

(f) scalar triple product (mixed triple product) of three vectors,

$$\vec{A}\cdot\left[\vec{B}\cdot\vec{C}\right] = \left[\vec{A}\cdot\vec{B}\right]\cdot\vec{C} = \left[\vec{C}\cdot\vec{A}\right]\cdot\vec{B}; \tag{B.5}$$

(g) vector triple product of three vectors,

$$\left[\vec{A}\cdot\left[\vec{B}\cdot\vec{C}\right]\right] = \vec{B}\cdot\left(\vec{A}\cdot\vec{C}\right) - \vec{C}\cdot\left(\vec{A}\cdot\vec{B}\right). \tag{B.6}$$

Gradient

Let us consider a scalar field $\varphi(x, y, z)$, i.e. a space, at every point of which a scalar value φ depending on coordinates (x, y, z) is given. We select a direction determined by the vector $\vec{l}$ and try to find the derivative of function φ along this direction,

$$\varphi' = \frac{d\varphi}{dl}. \tag{B.7}$$

Supposing that $\vec{l} = \vec{l}(x, y, z)$, we may write

$$\begin{aligned}\frac{d\varphi}{dl} &= \frac{\partial\varphi}{\partial x}\frac{dx}{dl} + \frac{\partial\varphi}{\partial y}\frac{dy}{dl} + \frac{\partial\varphi}{\partial z}\frac{dz}{dl} \\ &= \left(\frac{\partial\varphi}{\partial x}\vec{i} + \frac{\partial\varphi}{\partial y}\vec{j} + \frac{\partial\varphi}{\partial z}\vec{k}\right)\left(\frac{dx}{dl}\vec{i} + \frac{dy}{dl}\vec{j} + \frac{dz}{dl}\vec{k}\right) \\ &= \left(\frac{\partial\varphi}{\partial x}\vec{i} + \frac{\partial\varphi}{\partial y}\vec{j} + \frac{\partial\varphi}{\partial z}\vec{k}\right)\frac{d}{dl}\left(x\vec{i} + y\vec{j} + z\vec{k}\right) \\ &= \left(\frac{\partial\varphi}{\partial x}\vec{i} + \frac{\partial\varphi}{\partial y}\vec{j} + \frac{\partial\varphi}{\partial z}\vec{k}\right)\frac{d\vec{r}}{dl}. \end{aligned} \tag{B.8}$$

From Fig. B.3, it follows that

$$\frac{d\vec{r}}{dl} = \frac{d\vec{l}}{dl} = \vec{l}_0. \tag{B.9}$$

Therefore,

$$\frac{d\varphi}{dl} = \left(\frac{\partial\varphi}{\partial x}\vec{i} + \frac{\partial\varphi}{\partial y}\vec{j} + \frac{\partial\varphi}{\partial z}\vec{k}\right)\vec{l}_0. \tag{B.10}$$

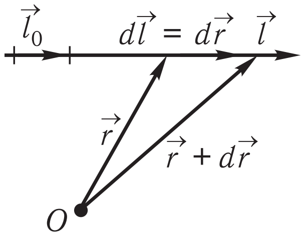

Fig. B.3.

The vector

$$\operatorname{grad}\varphi = \frac{\partial\varphi}{\partial x}\vec{i} + \frac{\partial\varphi}{\partial y}\vec{j} + \frac{\partial\varphi}{\partial z}\vec{k} \tag{B.11}$$

is called the **gradient of scalar function** φ. Hence,

$$\frac{d\varphi}{dl} = \operatorname{grad}\varphi \cdot \vec{l}_0 = \operatorname{grad}_l \varphi, \tag{B.12}$$

where $\mathrm{grad}_l\,\varphi$ is the projection of vector $\mathrm{grad}\,\varphi$ on the direction $\vec{l}$.

The projection of a vector has the maximum value if it is a projection on the own vector direction. Therefore, the vector $\mathrm{grad}\,\varphi$ points in the direction of the steepest slope of function φ, and its absolute value equals the rate of φ-growth along this direction.

The vector $\mathrm{grad}\,\varphi$ is written in the form

$$\mathrm{grad}\,\varphi = \left(\vec{i}\frac{\partial}{\partial x} + \vec{j}\frac{\partial}{\partial y} + \vec{k}\frac{\partial}{\partial z}\right)\varphi = \vec{\nabla}\varphi, \tag{B.13}$$

where $\vec{\nabla}$ is the "nabla" operator,

$$\vec{\nabla} = \vec{i}\frac{\partial}{\partial x} + \vec{j}\frac{\partial}{\partial y} + \vec{k}\frac{\partial}{\partial z}. \tag{B.14}$$

If $\varphi = \varphi(\vec{r})$, the gradient of this function is convenient to be calculated by the formula

$$\mathrm{grad}\,\varphi = \frac{\partial\varphi}{\partial\vec{r}}. \tag{B.15}$$

Really,

$$\frac{\partial\varphi}{\partial\vec{r}} = \frac{\partial\varphi}{\partial x}\frac{dx}{d\vec{r}} + \frac{\partial\varphi}{\partial y}\frac{dy}{d\vec{r}} + \frac{\partial\varphi}{\partial z}\frac{dz}{d\vec{r}}. \tag{B.16}$$

Taking into account that, e.g.,

$$\frac{dx}{d\vec{r}} = \left(\frac{d\vec{r}}{dx}\right)^{-1} = \frac{1}{\vec{i}} = \frac{\vec{i}}{\vec{i}\cdot\vec{i}} = \vec{i}, \tag{B.17}$$

we obtain

$$\frac{\partial\varphi}{\partial\vec{r}} = \frac{\partial\varphi}{\partial x}\vec{i} + \frac{\partial\varphi}{\partial y}\vec{j} + \frac{\partial\varphi}{\partial z}\vec{k} = \mathrm{grad}\,\varphi. \tag{B.18}$$

Example 1. Let us calculate the gradient of the absolute value of radius vector in two ways.

(i) $\mathrm{grad}\,|\vec{r}| = \mathrm{grad}\,\sqrt{x^2+y^2+z^2} = \frac{2x}{2\,|\vec{r}|}\vec{i} + \frac{2y}{2\,|\vec{r}|}\vec{j} + \frac{2z}{2\,|\vec{r}|}\vec{k} = \frac{\vec{r}}{|\vec{r}|} = \vec{r}_0.$

(ii) $\mathrm{grad}\,|\vec{r}| = \frac{\partial\,|\vec{r}|}{\partial\vec{r}} = \frac{\partial}{\partial\vec{r}}\sqrt{\vec{r}^2} = \frac{2\vec{r}}{2\,|\vec{r}|} = \vec{r}_0.$

Divergence

Let we have a vector field $\vec{A}(x, y, z)$, i.e. some vector $\vec{A}$ depending on coordinates (x, y, z) is given at every point of the space. The flux of the vector $\vec{A}$ (the number of force lines) through any surface S is determined by the formula

$$N = \int_S A_n dS \equiv \int_S \vec{A} \cdot d\vec{S}. \tag{B.19}$$

Let a closed surface S confines the volume ΔV. Then,

$$\Delta N = \oint_S A_n dS. \tag{B.20}$$

It is evident that

$$\lim_{\Delta V \to 0} \frac{\Delta N}{\Delta V} = \frac{dN}{dV} = \lim_{\Delta V \to 0} \frac{1}{\Delta V} \oint_S A_n dS. \tag{B.21}$$

This is a flux of the vector $\vec{A}$ from a unit volume taken at a definite point. This quantity is called the **divergence of vector** $\vec{A}$:

$$\operatorname{div} \vec{A} = \lim_{\Delta V \to 0} \frac{1}{\Delta V} \oint_S A_n dS = \frac{1}{dV} \oint_S A_n dS \tag{B.22}$$

Let us calculate the right hand side in (B.22) directly. For this purpose, let us tale the volume dV in the form of a cube with the edges dx, dy, and dz (Fig. B.4). The flow of vector $\vec{A}$ through the left edge $dy \cdot dz$ equals

$$dN_1 = A_{n1} dS_1 = -A_x dydz. \tag{B.23}$$

The flow through the right edge $dy \cdot dz$ equals

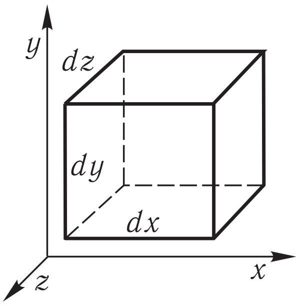

Fig. B.4.

$$dN_2 = A_{n2} dS_2 = \left(A_x + \frac{\partial A_x}{\partial x} dx \right) dydz. \tag{B.24}$$

The total flux trough both those edges amounts to

$$dN_x = \frac{\partial A_x}{\partial x} dxdydz = \frac{\partial A_x}{\partial x} dV. \tag{B.25}$$

By analogy, the total fluxes through the upper and lower edges, and through the front and back ones equal, respectively,

$$dN_y = \frac{\partial A_y}{\partial y} dV, \quad dN_z = \frac{\partial A_z}{\partial z} dV. \tag{B.26}$$

The total flux of vector $\vec{A}$ from the volume dV through all its edges equals

$$dN = \left(\frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} \right) dV. \tag{B.27}$$

Therefore,

$$\operatorname{div} \vec{A} = \frac{dN}{dV} = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}. \tag{B.28}$$

On a basis of (B.28), we obtain

$$dN = \operatorname{div} \vec{A} dV, \quad N = \oint_S A_n dS = \int_V \operatorname{div} \vec{A} dV. \tag{B.29}$$

This is the so-called **Gauss's theorem**,

$$\oint_S A_n dS = \int_V \operatorname{div} \vec{A} dV. \tag{B.30}$$

Using the operator "nabla" (B.14), the divergence of the vector can be written as

$$\operatorname{div} \vec{A} = \vec{\nabla} \cdot \vec{A}. \tag{B.31}$$

Example 2. Let us calculate the divergence of the radius vector $\vec{r}$:

$$\operatorname{div} \vec{r} = \frac{\partial r_x}{\partial x} + \frac{\partial r_y}{\partial y} + \frac{\partial r_z}{\partial z} = \frac{\partial x}{\partial x} + \frac{\partial y}{\partial y} + \frac{\partial z}{\partial z} = 3.$$

rot

Let us consider the circulation of vector $\vec{A}$ along a closed contour L,

$$\Gamma = \oint_L A_l dl = \oint_L \vec{A} \cdot d\vec{l}. \tag{B.32}$$

We suppose that the contour is infinitesimally small (dL), dS is the surface of any form confined by this contour, and $\vec{n}$ is the normal to the surface dS, the direction of which is connected with the direction of path-tracing in the right-handed system.. Then, the circulation along the contour equals

$$d\Gamma = \oint_{dL} A_l dl. \tag{B.33}$$

Hence,

$$\frac{d\Gamma}{dS} = \frac{1}{dS} \oint_{dL} A_l dl = \lim_{\Delta S \to 0} \frac{1}{\Delta S} \oint_L A_l dl$$

is the circulation of vector $\vec{A}$ related to a unity of surface area with the normal $\vec{n}$. This quantity is equal to the projections of a vector, which is called the **rotor of vector** $\vec{A}$, on the normal $\vec{n}$. This vector is designated as $\operatorname{rot} \vec{A}$ (Fig. B.5),

$$\operatorname{rot}_n \vec{A} = \lim_{\Delta S \to 0} \frac{1}{\Delta S} \oint_L A_l dl. \tag{B.34}$$

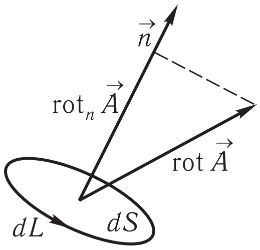

Fig. B.5.

Since

$$d\Gamma = \operatorname{rot}_n \vec{A} \cdot dS, \tag{B.35}$$

we may write

$$d\Gamma = \operatorname{rot} \vec{A} \cdot d\vec{S}, \quad d\vec{S} = dS \cdot \vec{n}. \tag{B.36}$$

In order to find a procedure to calculate $\operatorname{rot} \vec{A}$, first let us take an infinitesimally small contour in the form of a square with the edges dx and dy in the plane xy (Fig. B.6). The circulation of vector $\vec{A}$ along this contour equals

$$d\Gamma_z = d\Gamma_{12} + d\Gamma_{23} + d\Gamma_{34} + d\Gamma_{41}, \tag{B.37}$$

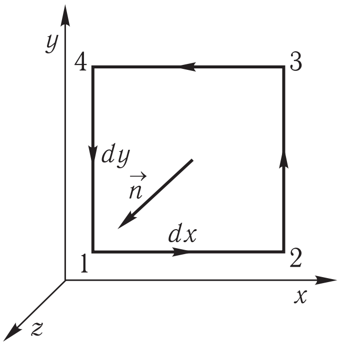

Fig. B.6.

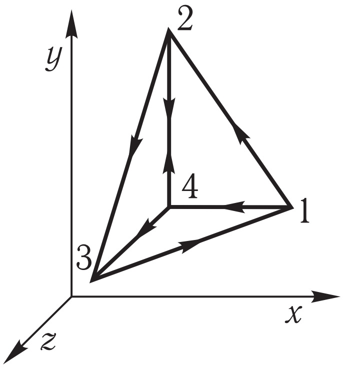

Fig. B.7.

where

$$d\Gamma_{12} = A_{l12} \cdot dl_{12} = A_x dx,$$

$$d\Gamma_{34} = A_{l34} \cdot dl_{34} = -\left(A_x + \frac{\partial A_x}{\partial y} dy\right) dx,$$

$$d\Gamma_{23} = A_{l23} \cdot dl_{23} = A_y dy,$$

$$d\Gamma_{41} = A_{l41} \cdot dl_{41} = -\left(A_y - \frac{\partial A_y}{\partial x} dx\right) dy. \tag{B.38}$$

Hence,

$$d\Gamma_z = \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right) dxdy = \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right) dS_z, \tag{B.39}$$

where $dS_z = |dS_z|$ is the area of contour surface perpendicular to the axis z. By analogy, we may write

$$d\Gamma_y = \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}\right) dS_y, \quad d\Gamma_x = \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}\right) dS_x. \tag{B.40}$$

Let us take three contours in the form of rectangular triangles in the planes xy, xz, and yz, as is illustrated in Fig. B.7. Then,

$$\begin{aligned} d\Gamma_z &= d\Gamma_{1241} = d\Gamma_{12} + d\Gamma_{24} + d\Gamma_{41}, \\ d\Gamma_y &= d\Gamma_{1431} = d\Gamma_{14} + d\Gamma_{43} + d\Gamma_{31}, \\ d\Gamma_x &= d\Gamma_{2342} = d\Gamma_{23} + d\Gamma_{34} + d\Gamma_{42}. \end{aligned} \tag{B.41}$$

Since $d\Gamma_{24} = -d\Gamma_{42}$, $d\Gamma_{34} = -d\Gamma_{43}$, and $d\Gamma_{14} = -d\Gamma_{41}$, then,

$$d\Gamma_x + d\Gamma_y + d\Gamma_z = d\Gamma_{12} + d\Gamma_{23} + d\Gamma_{31} = d\Gamma_{123} = d\Gamma, \tag{B.42}$$

i.e. $d\Gamma = d\Gamma_{123}$ is the circulation along an arbitrary oriented contour 123. Taking into account (B.39) and (B.40), we obtain

$$\begin{aligned} d\Gamma &= \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}\right) dS_x + \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}\right) dS_y + \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right) dS_z \\ &= \left[\left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}\right)\vec{i} + \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}\right)\vec{j} + \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right)\vec{k}\right] \\ &\qquad \cdot \left(\vec{i}dS_x + \vec{j}dS_y + \vec{k}dS_z\right) = \operatorname{rot}\vec{A} \cdot d\vec{S}. \end{aligned} \tag{B.43}$$

Hence,

$$\operatorname{rot}\vec{A} = \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}\right)\vec{i} + \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}\right)\vec{j} + \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right)\vec{k}. \tag{B.44}$$

On the basis of (B.35), we may write

$$\Gamma = \int d\Gamma = \oint_L A_l dl = \int_S \operatorname{rot}_n \vec{A} \cdot dS, \tag{B.45}$$

i.e.

$$\oint_L A_l dl = \int_S \operatorname{rot}_n \vec{A} \cdot dS. \tag{B.46}$$

This relation is called the **Stokes theorem**. It will be recalled that S in (B.46) is the surface confined by the contour L.

The rot of vector $\vec{A}$ can also be written in the form

$$\operatorname{rot}\vec{A} = \left[\vec{\nabla} \cdot \vec{B}\right] = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ A_x & A_y & A_z \end{vmatrix}. \tag{B.47}$$

Some important formulas

Here, we present the most important formulas of the vector analysis:

$$\operatorname{grad}(\varphi\Psi) = \nabla(\varphi\Psi) = \varphi\nabla\Psi + \Psi\nabla\varphi = \varphi\operatorname{grad}\Psi + \Psi\operatorname{grad}\varphi, \tag{B.48}$$

$$\operatorname{div}\left(\varphi\vec{A}\right) = \nabla\left(\varphi\vec{A}\right) = \varphi\nabla\vec{A} + \vec{A}\nabla\varphi = \varphi\operatorname{div}\vec{A} + \vec{A}\operatorname{grad}\varphi, \tag{B.49}$$

$$\operatorname{rot}\left(\varphi\vec{A}\right) = \left[\nabla\cdot\left(\varphi\vec{A}\right)\right] = \varphi\left[\nabla\vec{A}\right] + \left[\nabla\varphi\cdot\vec{A}\right] = \varphi\operatorname{rot}\vec{A} + \left[\operatorname{grad}\varphi\cdot\vec{A}\right], \tag{B.50}$$

$$\operatorname{div}\left[\vec{A}\cdot\vec{B}\right] = \nabla\left[\vec{A}\cdot\vec{B}\right] = \vec{B}\left[\nabla\vec{A}\right] - \vec{A}\left[\nabla\vec{B}\right] = \vec{B}\operatorname{rot}\vec{A} - \vec{A}\operatorname{rot}\vec{B} \tag{B.51}$$

$$\operatorname{rot}\left[\vec{A}\cdot\vec{B}\right] = \vec{B}\left[\nabla\vec{A}\right] - \left(\vec{A}\nabla\right)\vec{B} + \vec{A}\operatorname{div}\vec{B} - \vec{B}\operatorname{div}\vec{A}, \tag{B.52}$$

$$\operatorname{div}\operatorname{grad}\varphi = \nabla\nabla\varphi = \nabla^2\varphi = \Delta\varphi = \frac{\partial^2\varphi}{\partial x^2} + \frac{\partial^2\varphi}{\partial y^2} + \frac{\partial^2\varphi}{\partial z^2}, \tag{B.53}$$

$$\operatorname{rot}\operatorname{grad}\varphi = 0, \tag{B.54}$$

$$\operatorname{div}\operatorname{rot}\vec{A} = 0, \tag{B.55}$$

$$\operatorname{rot}\operatorname{rot}\vec{A} = \operatorname{grad}\operatorname{div}\vec{A} - \Delta\vec{A}, \tag{B.56}$$

$$\operatorname{div}\left[\vec{A}\cdot\vec{B}\right] = \left[\vec{A}\cdot\operatorname{rot}\vec{B}\right] + \left[\vec{B}\cdot\operatorname{rot}\vec{A}\right] + \left(\vec{A}\nabla\right)\vec{B} + \left(\vec{B}\nabla\right)\vec{A}. \tag{B.57}$$

APPENDIX C

OPERATOR ∇ IN SPHERICAL COORDINATES

By the gradient of scalar field $\Psi(\vec{r})$ is meant a vector $\operatorname{grad}\Psi = \nabla\Psi$, the direction of which coincides with the direction of the steepest slope of the function $\Psi(\vec{r})$ and the absolute value equals the derivative of the function $\Psi(\vec{r})$ along this direction. In the Cartesian coordinate system, the growth rates of function $\Psi(\vec{r})$ equal $\vec{i}\partial\Psi/\partial x$ along the axis x, $\vec{j}\partial\Psi/\partial y$ along the axis y, and $\vec{k}\partial\Psi/\partial z$ along the axis z. Therefore, the gradient of field $\Psi(\vec{r})$ is defined as follows:

$$\operatorname{grad}\Psi(\vec{r}) \equiv \nabla\Psi(\vec{r}) = \vec{i}\frac{\partial\Psi}{\partial x} + \vec{j}\frac{\partial\Psi}{\partial y} + \vec{k}\frac{\partial\Psi}{\partial z}, \tag{C.1}$$

i.e. the operator of gradient (the operator “nabla”) in the Cartesian coordinates looks like

$$\nabla = \vec{i}\frac{\partial}{\partial x} + \vec{j}\frac{\partial}{\partial y} + \vec{k}\frac{\partial}{\partial z}. \tag{C.2}$$

Let us find its expression in the spherical coordinate system.

The axes x, y, and z of the Cartesian coordinate system with the unit vectors $\vec{i}$, $\vec{j}$, and $\vec{k}$, as well as the axes ξ_r, ξ_θ, and ξ_φ of the orthogonal spherical system with the unit vectors $\vec{e}_r$, $\vec{e}_\theta$, and $\vec{e}_\varphi$, are exhibited in Fig. C.1. The vector $\vec{e}_r$ is directed along the radius vector $\vec{r}$, the vector $\vec{e}_\theta$ along the tangent to a circle with the large radius, and the vector e_φ along the tangent to a circle with the small radius that pass through the point with the axes ξ_r, ξ_θ, and ξ_φ.

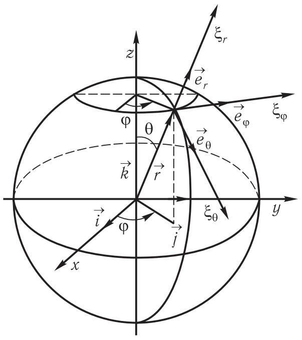

Fig. C.1.

By analogy with (C.1), in the coordinate system with the axes ξ_r, ξ_θ, and ξ_φ, the quantity grad Ψ has to look like

$$\operatorname{grad}\Psi = \vec{e}_r \frac{\partial \Psi}{\partial \xi_r} + \vec{e}_\theta \frac{\partial \Psi}{\partial \xi_\theta} + \vec{e}_\varphi \frac{\partial \Psi}{\partial \xi_\varphi}. \tag{C.3}$$

From Fig. C.1, it follows that $d\xi_\theta = rd\theta$, $d\xi_\varphi = r\sin\theta d\varphi$, and $\mathrm{d}\xi_r = dr$; therefore,

$$\operatorname{grad}\Psi = \vec{e}_r \frac{\partial \Psi}{\partial r} + \vec{e}_\theta \frac{1}{r}\frac{\partial \Psi}{\partial \theta} + \vec{e}_\varphi \frac{1}{r\sin\varphi}\frac{\partial \Psi}{\partial \varphi}. \tag{C.4}$$

Hence, in the spherical coordinates, the operator grad $= \nabla$ looks like

$$\nabla = \vec{e}_r \frac{\partial}{\partial r} + \vec{e}_\theta \frac{1}{r}\frac{\partial}{\partial \theta} + \vec{e}_\varphi \frac{1}{r\sin\varphi}\frac{\partial}{\partial \varphi}. \tag{C.5}$$

APPENDIX D

THE CONCEPT OF TENSORS

Every physical property is associated with ap hysical quantity. For instance, inertia is associated with the mass, conductivity is characterized by the specific electric conductance, and so forth. In turn, all physical quantities have a tensor character, i.e. they are tensors. Tensors are quantities that changes in a proper way in the case of coordinate transformation, namely, at the rotation of coordinate system.

Let us consider a point in a three-dimensional rectangular Cartesian coordinate system with the coordinates (x, y, z) or, for convenience, (x_1, x_2, x_3). Let us rotate the coordinate axes around the coordinate origin. Then the point coordinates are (x'_1, x'_2, x'_3). It will be recalled that

$$\begin{aligned} x'_1 &= \alpha_{11}x_1 + \alpha_{12}x_2 + \alpha_{13}x_3, \\ x'_2 &= \alpha_{21}x_1 + \alpha_{22}x_2 + \alpha_{23}x_3, \\ x'_3 &= \alpha_{31}x_1 + \alpha_{32}x_2 + \alpha_{33}x_3. \end{aligned}$$

These formulas describe the direct transformation of coordinates, and α_{ij} is the a cosine of the angle between the axis $\vec{j}$ (before the rotation) and the axis $\vec{i}'$ (after the rotation). In a compact form, those formulas look like

$$x'_i = \sum_{j=1}^{3} \alpha_{ij} x_j. \tag{D.1}$$

The formulas for the inverse transformation look like

$$x_j = \sum_{j=1}^{3} \alpha_{ij} x'_i.$$

The transformation coefficients α_{ij} have the following properties:

$$\sum_{j=1}^{3} \alpha_{ij}\alpha_{kj} = \delta_{ik}, \quad \sum_{j=1}^{3} \alpha_{ij}\alpha_{ik} = \delta_{jk}.$$

The product of coordinates $x_i x_j$ transforms as

$$x'_i x'_j = \sum_{k,l} \alpha_{ik}\alpha_{jl} x_k x_l, \tag{D.2}$$

the product $x_i x_j x_k$ as

$$x'_i x'_j x'_k = \sum_{l,m,n} \alpha_{il}\alpha_{jm}\alpha_{kn} x_l x_m x_n, \tag{D.3}$$

and so on.

Let us define tensors in the three-dimensional coordinate space.

As the **tensor of the zeroth rank** or the **scalar** is meant a quantity determined by a single ($3^0 = 1$) number that does not change at coordinate transformations. The corresponding examples include the mass, the length, and so forth.

The **tensor of the first rank** or the **vector** is a quantity determined by three ($3^1 = 3$) numbers (functions) or components A_i $(i = 1, 2, 3)$ that transforms as coordinates (see (D.1)) at the coordinate axis rotation, i.e.

$$A'_i = \sum_{j=1}^{3} \alpha_{ij} A_j. \tag{D.4}$$

The corresponding examples include the velocity $\vec{v} = \{v_x, v_y, v_z\}$, the force $\vec{F} = \{F_x, F_y, F_z\}$, and so on.

The **tensor of the second rank** is a quantity determined by 9 ($3^2 = 9$) numbers (functions) or components A_{ij} $(i, j = 1, 2, 3)$ that transforms as the coordinate product (see (D.2)) at the coordinate axis rotation, i.e.

$$A'_{ij} = \sum_{j=1}^{3} \alpha_{ik}\alpha_{jl}A_{kl}. \tag{D.5}$$

The **tensor of the third rank** is a quantity determined by 27 ($3^3 = 27$) numbers (functions) or components A_{ijk} $(i, j, k = 1, 2, 3)$ that transforms as the product of three coordinates (see (D.3)) at the coordinate axis rotation, i.e.

$$A'_{ijk} = \sum_{j=1}^{3} \alpha_{il}\alpha_{jm}\alpha_{kn}A_{lmn}.$$

The tensors of higher ranks are defined analogously.

Example: the inertia tensor.

Let us consider an object rotating around a certain axis with the angular velocity $\vec{\omega}$. We consider this object to consist of N material points, with the n-th point having the mass m_n. The angular momentum of the n-th material point is

$$\vec{L}_n = \left[\vec{r}_n \cdot \vec{p}_n\right].$$

Therefore, the total angular momentum of the object equals

$$\vec{L} = \sum_n \left[\vec{r}_n \cdot \vec{p}_n\right].$$

Taking into account that

$$\vec{p}_n = m_n\vec{v}_n, \quad \vec{v}_n = \left[\vec{\omega} \cdot \vec{r}_n\right],$$

we obtain

$$\vec{L} = \sum_n m_n \left[\vec{r}_n \cdot \left[\vec{\omega} \cdot \vec{r}_n\right]\right] = \sum_n m_n \left\{\vec{\omega}(\vec{r}_n \cdot \vec{r}_n) - \vec{r}_n(\vec{\omega} \cdot \vec{r}_n)\right\}$$

$$= \sum_n m_n \left\{\vec{\omega} \sum_l x_l^{(n)} x_l^{(n)} - \vec{r}_n \sum_j \omega_j x_j^{(n)})\right\}. \tag{D.6}$$

Then, the i-th projection of vector $\vec{L}$ looks like

$$L_i = \sum_n m_n \left\{\omega_i \sum_l x_l^{(n)} x_l^{(n)} - x_i^{(n)} \sum_j \omega_j x_j^{(n)})\right\}.$$

$$= \sum_{j,n} m_n \left\{ \omega_j \delta_{ij} \sum_l x_l^{(n)} x_l^{(n)} - \omega_j x_i^{(n)} x_j^{(n)} \right\} = \sum_j I_{ij} \omega_j,$$

where

$$I_{ij} = \sum_n m_n \left\{ \delta_{ij} \sum_l x_l^{(n)} x_l^{(n)} - x_i^{(n)} x_j^{(n)} \right\}.$$

The quantity I_{ij} is a tensor of the second rank, because it is composed of the coordinate products. It is called the **inertia tensor**.

According to the relativity principle, any physical law, i.e. the corresponding equation, must have the same form in all inertial reference frames, i.e. it must be covariant with respect to the coordinate transformation. This condition requires that all summands in the equation should be tensors of the same rank; for instance,

$$A_{ik} = \alpha B_{ik} + \beta C_{ik}, \tag{D.7}$$

where α and β are constants. Multiplying (D.7) by $\alpha_{li}\alpha_{mk}$ and summing up the result over i and k, on the basis of (D.5), we obtain

$$A'_{lm} = \alpha B'_{lm} + \beta C'_{lm}, \tag{D.8}$$

i.e. the equation keeps its form.

APPENDIX E

ANGULAR MOMENTUM OPERATOR IN SPHERICAL COORDINATES

In the Cartesian coordinates, the operators $\hat{L}_x$, $\hat{L}_y$, $\hat{L}_z$, and $\hat{L}$ look like

$$\hat{L}_x = -i\hbar \left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y} \right), \quad \hat{L}_y = -i\hbar \left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z} \right),$$

$$\hat{L}_z = -i\hbar \left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x} \right), \quad \hat{L}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2. \tag{E.1}$$

The Cartesian coordinates (x, y, z) transform into the spherical one (r, θ, φ) with the help of formulas

$$x = r\sin\theta\cos\varphi, \quad y = r\sin\theta\sin\varphi, \quad z = r\cos\theta. \tag{E.2}$$

In order to find the transformation formulas from the operators $\dfrac{\partial}{\partial x}$, $\dfrac{\partial}{\partial y}$, and $\dfrac{\partial}{\partial z}$ to the operators $\dfrac{\partial}{\partial r}$, $\dfrac{\partial}{\partial \theta}$, and $\dfrac{\partial}{\partial \varphi}$, we write the formulas for the inverse coordinate transformation, which follow from (E.2),

$$r = \sqrt{x^2 + y^2 + z^2}, \quad \theta = \arccos\frac{z}{r}, \quad \varphi = \arctan\frac{y}{x}. \tag{E.3}$$

As an example, let us consider the operator $\dfrac{\partial}{\partial x}$:

$$\frac{\partial\Psi}{\partial x} = \frac{\partial\Psi}{\partial r}\frac{\partial r}{\partial x} + \frac{\partial\Psi}{\partial\theta}\frac{\partial\theta}{\partial x} + \frac{\partial\Psi}{\partial\varphi}\frac{\partial\varphi}{\partial x}. \tag{E.4}$$

Making use of (E.2) (E.3), we find

$$\frac{\partial r}{\partial x} = \frac{r}{x} = \sin\theta\cos\varphi, \quad \frac{\partial\theta}{\partial x} = \frac{\partial}{\partial x}\arccos\frac{z}{r} = \frac{\cos\theta\cos\varphi}{r},$$

$$\frac{\partial\varphi}{\partial x} = \frac{\partial}{\partial x}\arctan\frac{y}{x} = -\frac{\sin\varphi}{r\sin\theta}. \tag{E.5}$$

Substituting (E.5) into (E.4) and omitting the function Ψ, we obtain

$$\frac{\partial}{\partial x} = \sin\theta\cos\varphi\frac{\partial}{\partial r} + \frac{\cos\theta\cos\varphi}{r}\frac{\partial}{\partial\theta} - \frac{\sin\varphi}{r\sin\theta}\frac{\partial}{\partial\varphi}. \tag{E.6}$$

Following the same procedure, we find

$$\frac{\partial}{\partial y} = \sin\theta\sin\varphi\frac{\partial}{\partial r} + \frac{\cos\theta\sin\varphi}{r}\frac{\partial}{\partial\theta} + \frac{\cos\varphi}{r\sin\theta}\frac{\partial}{\partial\varphi}, \tag{E.7}$$

$$\frac{\partial}{\partial z} = \cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial\theta}. \tag{E.8}$$

Substituting (E.2) and (E.6)–(E.8) into (E.1), we obtain the formulas for the operators $\hat{L}_x$, $\hat{L}_y$, and $\hat{L}_z$ in the spherical coordinates,

$$\hat{L}_x = -i\hbar\left(\sin\varphi\frac{\partial}{\partial\theta} + \cot\theta\cos\varphi\frac{\partial}{\partial\varphi}\right), \tag{E.9}$$

$$\hat{L}_y = -i\hbar\left(\cos\varphi\frac{\partial}{\partial\theta} - \cot\theta\sin\varphi\frac{\partial}{\partial\varphi}\right), \tag{E.10}$$

$$\hat{L}_z = -i\hbar\frac{\partial}{\partial\varphi}. \tag{E.11}$$

To find the operator $\hat{L}^2$ in the spherical coordinates, we write

$$\hat{L}^2\Psi = \left(\hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2\right)\Psi = \hat{L}_x\hat{L}_x\Psi + \hat{L}_y\hat{L}_y\Psi + \hat{L}_z\hat{L}_z\Psi.$$

Substituting operators (E.9)–(E.11) into this relation, after simple calculations, we obtain

$$\hat{L}^2 = -\hbar^2\left\{\frac{\partial^2}{\partial\theta^2} + \cot\theta\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right\}$$

or

$$\hat{L}^2 = -\hbar^2\left\{\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right\}. \tag{E.12}$$

For the Laplace operator Δ, using (E.6)–(E.8), we obtain

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

$$= \frac{1}{r^2}\left\{\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right\}$$

$$= \frac{1}{r^2}\left\{\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) - \frac{1}{\hbar^2}\hat{L}^2\right\}. \tag{E.13}$$

Whence it follows that

$$\hat{H} = -\frac{\hbar^2}{2m}\Delta + U(\vec{r}) = \hat{T}_r + \frac{1}{2mr^2}\hat{L}^2 + U(\vec{r}).$$

APPENDIX F

EIGENVALUES AND EIGENFUNCTIONS OF OPERATOR $\hat{L}^2$

The eigenvalue-eigenfunction equation for the square of the orbital angular momentum operator, $\hat{L}^2$, looks like

$$-\hbar^2\left\{\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right\}Y=L^2Y. \tag{F.1}$$

Dividing it by $\hbar^2$ and introducing the notation $\lambda = L^2/\hbar^2$, we obtain the equation for the eigenvalues and eigenfunctions of Legendre operator $\hat{\Lambda}=\hat{L}^2/\hbar^2$,

$$\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2 Y}{\partial\varphi^2}+\lambda Y=0. \tag{F.2}$$

The functions $Y = Y(\theta,\varphi)$ are called the **spherical functions**.

The operators $\hat{L}^2$ and $\hat{L}_z$ commute with each other. Therefore, the function $Y(\theta,\varphi)$ must also by the eigenfunction of the operator $\hat{L}_z$, which has the eigenfunctions $U_m(\varphi)=\dfrac{1}{2\pi}e^{im\varphi}$. In this connection, we seek the spherical functions $Y(\theta,\varphi)$ in the form

$$Y(\theta,\varphi)=\Phi(\theta)\mathrm{e}^{im\varphi},\quad m\notin Z. \tag{F.3}$$

Substituting (F.3) into (F.2), we obtain the equation

$$\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Phi}{\partial\theta}\right)+\frac{m^2}{\sin^2\theta}\Phi+\lambda\Phi=0. \tag{F.4}$$

Changing to a new variable $\xi=\cos\theta$ and taking into account that

$$\frac{\partial\Phi}{\partial\theta}=\frac{\partial\Phi}{\partial\xi}\frac{\partial\xi}{\partial\theta}=\frac{\partial\Phi}{\partial\xi}(-\sin\theta)=-\sqrt{1-\xi^2}\frac{\partial\Phi}{\partial\xi},$$

instead of equation (F.4), we obtain

$$\left(1-\xi^2\right)\frac{\partial^2\Phi}{\partial\xi^2}-2\xi\frac{\partial\Phi}{\partial\xi}-\frac{m^2}{1-\xi^2}\Phi+\lambda\Phi=0. \tag{F.5}$$

Since $-1 \leq \xi \leq 1$, let us first find an asymptotic solution of equation (F.5) at $\xi \to \pm 1$. In this case, the summand $\lambda\Phi$ can be neglected,

$$\left(1-\xi^2\right)\frac{\partial^2\Phi_a}{\partial\xi^2} - 2\xi\frac{\partial\Phi_a}{\partial\xi} - \frac{m^2}{1-\xi^2}\Phi_a = 0. \tag{F.6}$$

The solution of this equation is tried in the form

$$\Phi_a = \left(1-\xi^2\right)^\gamma. \tag{F.7}$$

Substituting (F.7) into (F.6) and taking into account that $1-\xi^2 \to 0$, we find the equation

$$\gamma(\gamma-1) + \gamma - \frac{m^2}{4} = 0,$$

the roots of which are $\gamma = \pm\frac{m}{2}$. If $\gamma < 0$, $\Phi_a(\xi \to \pm 1) \to \infty$. Therefore, there must be $\gamma = \pm\frac{m}{2} > 0$, i.e.$\gamma = \frac{1}{2}|m|$, and solution (F.7) takes the form

$$\Phi_a = \left(1-\xi^2\right)^{|m|/2}. \tag{F.8}$$

Hence, the exact solution of equation (F.5) should be sought in the form

$$\Phi(\xi) = \left(1-\xi^2\right)^{|m|/2} f(\xi). \tag{F.9}$$

Substituting (F.9) into (F.5), we obtain the equation for the function $f(\xi)$,

$$\left(1-\xi^2\right)\frac{\partial^2 f}{\partial\xi^2} - 2\left(|m+1|\right)\xi\frac{\partial f}{\partial\xi} - \left(\lambda - |m| - m^2\right)f = 0, \tag{F.10}$$

which is sought in the form of a power series

$$f(\xi) = \sum_{i=0}^{\infty} a_i\xi^i. \tag{F.11}$$

Substituting series (F.11) into (F.10) and equating the coefficient a_k at ξ^k to zero, we obtain the recurrent formula for the determination of a_k,

$$a_{k+2} = \frac{k(k-1) + 2(|m|+1)k - \lambda + |m| + m^2}{(k+2)(k+1)}a_k. \tag{F.12}$$

The finiteness of function $\Phi(\xi)$ dictates that series (F.11) has to be a polynomial of the finite order n ($n = 0, 1, 2, \ldots$); in this case, $a_n \neq 0$, but

$$a_{n+2} = \frac{n(n-1) + 2(|m|+1)n - \lambda + |m| + m^2}{(n+2)(n+1)} a_n = 0. \tag{F.13}$$

From (F.13), we find

$$\lambda = (n + |m|)(n + |m| + 1)$$

or, using the notation

$$l = n + |m| \,, \tag{F.14}$$

$\lambda = l(l+1)$. Since n and $|m|$ can equal 0, 1, 2, …, the quantity l also can equal 0, 1, 2,…. Taking into account that $\lambda = L^2/\hbar^2$, we find the eigenvalues of operator $\hat{L}^2$:

$$L^2 = l(l+1)\hbar, \quad l = 0, 1, 2, \ldots. \tag{F.15}$$

The number l is called the **orbital quantum number**.

From (F.14), it follows that, for the given l, we have $|m|_{\min} = 0$ and $|m|_{\max} = l$. Hence, the magnetic quantum number m can acquire values 0, ±1, ±2,…,$\pm l$; i.e. $m = l, l-1, l-2, \ldots, -l$; a total of $2l + 1$ values.

For the given l and m, the number $n = l - |m|$ can be even or odd. Therefore, in accordance with (F.11), for the even n,

$$f_n(\xi) = a_0 + a_2\xi^2 + a_4\xi^4 + \ldots + a_n\xi^n, \tag{F.16a}$$

and for the odd n,

$$f_n(\xi) = a_1 + a_3\xi^3 + \ldots + a_n\xi^n. \tag{F.16b}$$

With the help of recurrent formula (F.13), for the given l and m—and, hence, for the known $n = l - |m|$—we can determine all coefficients a_i in (F.16) in terms of a_0 at the even n or in terms of a_1 at the odd n, i.e. we can find the polynomial $f_n(\xi) = f_{l-|m|}(\cos\theta)$. Substituting it into (F.9), we obtain the function $\Phi(\xi)$ in the polynomial form, for which the notation

$$\Phi(\xi) = P_l^{|m|}(\cos\theta)$$

is used, and spherical function (F.3), which is written in the form

$$Y_{lm}(\theta, \varphi) = A_l^{|m|} P_l^{|m|}(\cos\theta) \mathrm{e}^{im\varphi}. \tag{F.17}$$

The uncertainty of coefficient a_0 (or a_1) is eliminated by the condition of spherical function normalization,

$$\int \left|Y_{lm(\theta,\varphi)}\right|^2 d\Omega = 1, \tag{F.18}$$

where $d\Omega = \sin\theta d\theta d\varphi$. As a result, the normalizing coefficient $A_l^{|m|}$ in (F.17) equals

$$A_l^{|m|} = \sqrt{\frac{(l-|m|)!(2l+1)}{4\pi(l+|m|)!}}. \tag{F.19}$$

The polynomials $P_l^{|m|}(\cos\theta)$ are called the **associated Legendre polynomials**. At the given l and m, they can be determined without calculating the coefficients a_i, but with the help of formula

$$P_l^{|m|}(\cos\theta) = \left(1-\cos^2\theta\right)^{|m|/2} \frac{d^{|m|}}{d(\cos\theta)^{|m|}} P_l(\cos\theta). \tag{F.20}$$

Here,

$$P_l(\cos\theta) = \frac{1}{2^l l!} \frac{d^l}{d(\cos\theta)^l} \left(\cos^2\theta - 1\right)^l \tag{F.21}$$

are also polynomials, which are called the **Legendre polynomials**.

APPENDIX G

PAULI OPERATORS

The commutation relations for the spin operators look like

$$\begin{aligned} \hat{L}_{sx}\hat{L}_{sy} - \hat{L}_{sy}\hat{L}_{sx} &= i\hbar\hat{L}_{sz}, \\ \hat{L}_{se}\hat{L}_{sz} - \hat{L}_{sz}\hat{L}_{sy} &= i\hbar\hat{L}_{sx}, \\ \hat{L}_{sz}\hat{L}_{sx} - \hat{L}_{sx}\hat{L}_{sz} &= i\hbar\hat{L}_{sy}. \end{aligned} \tag{G.1}$$

The Pauli operators $(\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z)$ are introduced as follows:

$$\hat{L}_{sx} = \frac{\hbar}{2}\hat{\sigma}_x, \quad \hat{L}_{sy} = \frac{\hbar}{2}\hat{\sigma}_y, \quad \hat{L}_{sz} = \frac{\hbar}{2}\hat{\sigma}_z. \tag{G.2}$$

On the basis of (G.1), we obtain

$$\begin{aligned} \hat{\sigma}_x\hat{\sigma}_y - \hat{\sigma}_y\hat{\sigma}_x &= i\hbar\hat{\sigma}_z, \\ \hat{\sigma}_y\hat{\sigma}_z - \hat{\sigma}_z\hat{\sigma}_y &= i\hbar\hat{\sigma}_x, \\ \hat{\sigma}_z\hat{\sigma}_x - \hat{\sigma}_x\hat{\sigma}_z &= i\hbar\hat{\sigma}_y. \end{aligned} \tag{G.3}$$

The operators $\hat{L}_{si}$ $(i = x, y, z)$ and, hence, the Pauli operators $\hat{\sigma}_i$ must be 2 × 2-matrices. Since $L_{sz} = \pm\hbar/2$, the eigenvalues of the matrix $\hat{L}_{sz}$, i.e. its diagonal members must be equal to $\pm\hbar/2$, and the other members to zero,

$$L_{sz} = \begin{pmatrix} \hbar/2 & 0 \\ 0 & -\hbar/2 \end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{G.4}$$

It is easy to be convinced that

$$\hat{\sigma}_z^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{G.5}$$

Therefore, on the basis of equality of all three coordinate axes, we may write

$$\hat{\sigma}_x^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \hat{\sigma}_y^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{G.6}$$

The operators $\hat{\sigma}_y^2$ and $\hat{\sigma}_z$ commute, since $\hat{\sigma}_y^2\hat{\sigma}_z - \hat{\sigma}_z\hat{\sigma}_y^2 = 0$. The latter relation can be rewritten in the form

$$\hat{\sigma}_y^2\hat{\sigma}_z - \hat{\sigma}_y\hat{\sigma}_z\hat{\sigma}_y + \hat{\sigma}_y\hat{\sigma}_z\hat{\sigma}_y - \hat{\sigma}_z\hat{\sigma}_y^2 = 0$$

or

$$\hat{\sigma}_y\left(\hat{\sigma}_y\hat{\sigma}_z - \hat{\sigma}_z\hat{\sigma}_y\right) + \left(\hat{\sigma}_y\hat{\sigma}_z - \hat{\sigma}_z\hat{\sigma}_y\right)\hat{\sigma}_y = 0.$$

Making use of the second relation in (G.3), we obtain

$$2i\left(\hat{\sigma}_y\hat{\sigma}_x + \hat{\sigma}_x\hat{\sigma}_y\right) = 0.$$

Hence,

$$\hat{\sigma}_x\hat{\sigma}_y = -\hat{\sigma}_y\hat{\sigma}_x. \tag{G.7}$$

Relation (G.7) can be supplemented by two analogous ones:

$$\hat{\sigma}_y\hat{\sigma}_z = -\hat{\sigma}_z\hat{\sigma}_y, \tag{G.8}$$

$$\hat{\sigma}_z\hat{\sigma}_x = -\hat{\sigma}_x\hat{\sigma}_z. \tag{G.9}$$

Writing down the operator $\hat{\sigma}_x$ in the general form as

$$\hat{\sigma}_x = \begin{pmatrix} b_1 & b_2 \\ b_3 & b_4 \end{pmatrix} \tag{G.10}$$

and substituting (G.4) and (G.10) into (G.9), we obtain

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} b_1 & b_2 \\ b_3 & b_4 \end{pmatrix} = - \begin{pmatrix} b_1 & b_2 \\ b_3 & b_4 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Multiplying the matrices, we find

$$\begin{pmatrix} b_1 & b_2 \\ -b_3 & -b_4 \end{pmatrix} = - \begin{pmatrix} b_1 & -b_2 \\ b_3 & -b_4 \end{pmatrix},$$

i.e. $b_1 = -b_1 = 0$ and $b_4 = -b_4 = 0$. Therefore, the operator $\hat{\sigma}_x$ looks like

$$\hat{\sigma}_x = \begin{pmatrix} 0 & b_2 \\ b_3 & 0 \end{pmatrix}. \tag{G.11}$$

Since the operator $\hat{\sigma}_x$ is Hermitian, $b_3 = b_2^*$. Taking into account (G.6), we may write

$$\hat{\sigma}_x^2 = \begin{pmatrix} 0 & b_2 \\ b_3 & 0 \end{pmatrix} \begin{pmatrix} 0 & b_2 \\ b_3 & 0 \end{pmatrix} = \begin{pmatrix} b_2 b_3 & 0 \\ 0 & b_2 b_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Therefore, $b_2 b_3 = |b_2|^2 = 1$. Whence it follows that $b_2 = \mathrm{e}^{i\alpha}$ and $b_3 = \mathrm{e}^{-i\alpha}$ (or vice versa, $b_2 = \mathrm{e}^{-i\alpha}$ and $b_3 = \mathrm{e}^{i\alpha}$), where α is an arbitrary phase multiplier, which can be equated to zero. Then, $b_2 = b_3 = 1$ and

$$\hat{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{G.12}$$

In order to find the form of operator $\hat{\sigma}_t$, let us use formulas (G.3) and (G.7),

$$\hat{\sigma}_y \hat{\sigma}_x = -i\hat{\sigma}_z.$$

For $\hat{\sigma}_y$ in the form

$$\hat{\sigma}_y = \begin{pmatrix} C_1 & C_2 \\ C_3 & C_4 \end{pmatrix},$$

we obtain

$$\begin{pmatrix} C_1 & C_2 \\ C_3 & C_4 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = -i \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

or

$$\begin{pmatrix} C_2 & C_1 \\ C_4 & C_3 \end{pmatrix} = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}.$$

Therefore, $C_1 = C_4 = 0$, $C_2 = -i$, $C_3 = i$, and the operator $\hat{\sigma}_y$ takes the form

$$\hat{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \tag{G.13}$$

Summarizing the results of calculations, we write

$$\hat{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{G.14}$$

APPENDIX H

PROOF OF THE CARNOT THEOREM

The Carnot cycle is shown in Fig. 19.1 (see p. 329). Let us first consider the reversible Carnot cycle for the ideal gas. Taking into account that the work done by one mole of ideal gas, when its volume isothermally changes from V_1 to V_2, is equal to

$$A = RT \ln \frac{V_2}{V_1},$$

and, when the temperature adiabatically changes from T_1 to T_2, to

$$A = C_V (T_1 - T_2),$$

let us calculate this work at every stage of cycle 1-2-3-4-1:

$$\begin{aligned} A_{12} &= RT_1 \ln \frac{V_2}{V_1} = Q_1, \\ A_{23} &= C_V (T_1 - T_2), \\ A_{34} &= -RT_2 \ln \frac{V_4}{V_3} = Q_2, \\ A_{41} &= C_V (T_2 - T_1). \end{aligned}$$

The efficiency of the cycle equals

$$\eta = \frac{A}{Q_1} = \frac{A_{12} + A_{23} + A_{34} + A_{44}}{A_{12}} = \frac{T_1 \ln \dfrac{V_2}{V_1} - T_2 \ln \dfrac{V_3}{V_4}}{T_1 \ln \dfrac{V_2}{V_1}}. \tag{H.1}$$

Using the relation between the temperatures and the volumes at adiabatic process,

$$\frac{T_2}{T_1} = \left(\frac{V_1}{V_2}\right)^{\gamma-1},$$

where $\gamma = c_p/c_V$, we obtain for adiabatic curves 2-3 and 4-1 that

$$\frac{T_2}{T_1} = \left(\frac{V_2}{V_3}\right)^{\gamma-1}, \quad \frac{T_2}{T_1} = \left(\frac{V_1}{V_4}\right)^{\gamma-1},$$

whence it follows that $V_2/V_3 = V_1/V_4$ or

$$\frac{V_2}{V_1} = \frac{V_3}{V_4}. \tag{H.2}$$

From formulas (H.1) and (H.2), we obtain

$$\eta = \frac{T_1 - T_2}{T_1}.$$

Now, let us calculate the efficiency of the reversible Carnot cycle for an arbitrary working medium,

$$\eta' = \frac{A'}{Q_1'} = \frac{Q_1' - Q_2'}{Q_1'}. \tag{H.3}$$

From this equation, it is easy to obtain the relation

$$A' = \frac{\eta'}{1 - \eta'} Q_2'.$$

At the same time, for the Carnot cycle with the ideal gas,

$$A = \frac{\eta}{1 - \eta} Q_2.$$

Therefore,

$$\frac{A'}{A} = \frac{\eta'(1-\eta)}{\eta(1-\eta')}\frac{Q_2'}{Q_2}. \tag{H.4}$$

Suppose that $\eta' > \eta$. Then, let us consider an aggregate of two Carnot engines with the common heater and refrigerator. Engine 1 works following the direct reversible Carnot cycle and has an arbitrary working medium. Within a cycle, it obtains the heat Q_1' from the heater, does the work A', and transfers the heat Q_2' to the refrigerator. Engine 2 works following the inverse reversible Carnot cycle (as a refrigerator) and has the ideal gas as a working medium. Within a cycle, it obtains the heat Q_2 from the refrigerator, consumes the work A, and transfers the heat Q_1 to the heater. If $Q_2 = Q_2'$, the state of refrigerator does not change. However, according to (H.4),

$$\frac{A'}{A} = \frac{\eta'(1-\eta)}{\eta(1-\eta')} > 1,$$

i.e. $A' > A$. This means that the aggregate does a positive work at the expense of the heater and transfers no heat to the refrigerator, which is impossible (see section 19.2), because contradicts the second law of thermodynamics. Hence, the relation $\eta' > \eta$ is also impossible.

In a similar way, we can prove that the relation $\eta' < \eta$ is also impossible. In this case, the engine with the ideal gas should work following the direct reversible Carnot cycle, and the engine with an arbitrary working medium following the inverse reversible Carnot cycle. Hence, we obtain that $\eta' = \eta$, i.e. the efficiency of the reversible Carnot cycle does not depend on the choice of working medium and equals

$$\eta_{\text{rev}} = \frac{T_1 - T_2}{T_1},$$

being determined only by the heater and refrigerator temperatures.

In the case of irreversible Carnot cycle,

$$\eta_{\text{irr}} = \frac{A_{\text{irr}}}{Q_1}.$$

However, as was shown in section 19.2, $A_{\text{irr}} < A_{\text{rev}}$. Hence,

$$\eta_{\text{irr}} < \eta_{\text{rev}}.$$

APPENDIX I

EULER GAMMA FUNCTION AND RIEMANN ZETA-FUNCTION

The Euler gamma-function $\Gamma(\alpha)$ and the Riemann zeta-function $\zeta(n)$ are special functions widely used in theoretical physics.

The gamma-function is defined with the help of integral

$$\Gamma(\alpha) = \int_0^\infty \mathrm{e}^{-x} x^{\alpha-1} dx, \quad (\alpha > 0). \tag{I.1}$$

Integrating (I.1) by parts, we obtain the relation

$$\Gamma(\alpha+1) = \alpha\Gamma(\alpha). \tag{I.2}$$

For $\alpha = 1$ and $\alpha = \dfrac{1}{2}$, the gamma function is calculated easily:

$$\Gamma(1) = 1, \quad \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}. \tag{I.3}$$

Therefore, for integer and half-integer α's, the value of gamma-function can be determined with the help of relation (I.2); in particular,

$$\Gamma(1) = 1, \quad \Gamma(2) = 1, \quad \Gamma(3) = 2, \quad \Gamma(4) = 6, \ldots;$$
$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}, \quad \Gamma\left(\frac{3}{2}\right) = \frac{1}{2}\sqrt{\pi}, \quad \Gamma\left(\frac{5}{2}\right) = \frac{3}{4}\sqrt{\pi}, \ldots.$$

One can easily see that, for the positive integer n, $\Gamma(n) = (n-1)!$.

The Riemann zeta-function is defined as

$$\zeta(n) = \sum_{m=1}^{\infty} \frac{1}{m^n}. \tag{I.4}$$

The values of $\zeta(n)$ for some n's are given below:

$$\zeta\left(\frac{3}{2}\right) \approx 2.612, \quad \zeta(2) = \frac{\pi^2}{6}, \quad \zeta\left(\frac{5}{2}\right) \approx 1.341,$$
$$\zeta(3) \approx 1.202, \quad \zeta(4) = \frac{\pi^4}{90.}$$

APPENDIX J

LIOUVILLE THEOREM

A statistical ensemble in the phase space is represented by a set of phase points distributed with the density proportional to the statistical distribution function $\rho(q,p)$. Since the microstates change in time, the phase points move along phase trajectories. If all the systems in the ensemble are at equilibrium, the distribution function does not depend explicitly on the time, $\partial\rho/\partial t = 0$.. Therefore, the motion of phase points can formally be considered as a stationary flow of a "gas", to which we may apply the continuity equation corresponding to the conservation law for the number of "gas particles", i.e. phase points.

In the three-dimensional coordinate space, the continuity equation looks like

$$\frac{\partial\rho}{\partial t} + \mathrm{div}\ (\rho\vec{v}) = 0,$$

where ρ is the density of the medium (liquid, electric charge, and so forth), and $\vec{v}$ is the velocity of its motion:

$$\mathrm{div}\ (\rho\vec{v}) = \sum_{i=1}^{3} \frac{\partial\left(\rho\dot{x}_i\right)}{\partial x_i}.$$

The continuity equation for the motion of phase points in the $2fN$-dimension phase space looks like

$$\frac{\partial\rho}{\partial t} + \sum_{i=1}^{fN} \left(\frac{\partial\left(\rho\dot{q}_i\right)}{\partial q_i} + \frac{\partial\left(\rho\dot{p}_i\right)}{\partial p_i} \right) = 0, \tag{J.1}$$

where $\rho = \rho(q,p)$ is the statistical distribution function proportional to the density of phase points. Taking into account that $\partial\rho/\partial t = 0$, the equation (J.1) can be rewritten as

$$\sum_{i=1}^{fN} \left(\dot{q}_i \frac{\partial\rho}{\partial q_i} + \dot{p}_i \frac{\partial\rho}{\partial p_i} \right) + \rho \sum_{i=1}^{fN} \left(\frac{\partial\dot{q}_i}{\partial q_i} + \frac{\partial\dot{p}_i}{\partial p_i} \right) = 0. \tag{J.2}$$

Applying the Hamilton equations of classical mechanics,

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q_i},$$

where H is the Hamilton function, we obtain that the second summand in (J.2) equals zero. Hence,

$$\frac{\partial \rho}{\partial t} + \sum_{i=1}^{fN} \left(\dot{q}_i \frac{\partial \rho}{\partial q_i} + \dot{p}_i \frac{\partial \rho}{\partial p_i} \right) = 0. \tag{J.3}$$

From relation (J.3), it follows that the statistical distribution function remains constant along phase trajectories. This is the content of the Liouville theorem. One of its most important consequences consists in the preservation of the magnitude (not the form!) of the phase volume containing a definite number of phase points.

APPENDIX K

POISSON INTEGRALS

The first Poisson integral looks like

$$I_1 = \int_0^\infty \mathrm{e}^{-ax^2} dx, \tag{K.1}$$

where $a = \mathrm{const} > 0$. Since the magnitude of any integral does not depend on its variable notation, we may write

$$I_1 = \int_0^\infty \mathrm{e}^{-ay^2} dy. \tag{K.2}$$

Multiplying (K.1) and (K.2), we obtain

$$I_1^2 = \int_0^\infty \int_0^\infty \mathrm{e}^{-a\left(x^2+y^2\right)} dx. \tag{K.3}$$

In double integral (K.3), we change to the polar coordinates on the plane by means of the formulas

$$x^2 + y^2 = r^2, \quad dxdy = rdrd\varphi.$$

Taking also into account that $0 \leq x < \infty$ and $0 \leq y < \infty$, integration over r should be carried out from 0 to ∞, and over φ from 0 to $\pi/2$. Then,

$$I_1^2 = \int_0^\infty dr \int_0^{\pi/2} d\varphi \mathrm{e}^{-ar^2} r = \frac{\pi}{2} \int_0^\infty \mathrm{e}^{-ar^2} r dr. \tag{K.4}$$

The last integral equals

$$\int_0^\infty \mathrm{e}^{-ar^2} r dr = \frac{1}{2a} \int_0^\infty \mathrm{e}^{-ar^2} d\left(ar^2\right) = \frac{1}{2a}. \tag{K.5}$$

Substituting (K.5) into (K.4), we find the first Poisson integral,

$$I_1 = \int_0^\infty \mathrm{e}^{-ax^2} dx = \frac{1}{2}\sqrt{\frac{\pi}{a}}. \tag{K.6}$$

The second Poisson integral is integral (K.5),

$$I_2 = \int_0^\infty \mathrm{e}^{-ar^2} r dr = \frac{1}{2a}. \tag{K.7}$$

The other Poisson integrals can be calculated by differentiating the integrals I_1, I_2, and the following integrals with respect to the parameter a:

$$I_3 = \int_0^\infty \mathrm{e}^{-ax^2} x^2 dx = \frac{1}{4}\sqrt{\frac{\pi}{a^3}}, \tag{K.8}$$

$$I_4 = \int_0^\infty \mathrm{e}^{-ax^2} x^3 dx = \frac{1}{2a^2}, \tag{K.9}$$

$$I_5 = \int_0^\infty \mathrm{e}^{-ax^2} x^4 dx = \frac{3}{8}\sqrt{\frac{\pi}{a^5}}, \tag{K.10}$$

$$I_6 = \int_0^\infty \mathrm{e}^{-ax^2} x^5 dx = \frac{1}{a^3}, \tag{K.11}$$

and so on.

APPENDIX L

STIRLING FORMULA

Let N be a large integer positive number. Then,

$$\frac{\Delta(\ln N!)}{\Delta N} = \frac{\ln N! - \ln(N-1)!}{1} = \ln \frac{N!}{(N-1)!} = \ln N. \tag{L.1}$$

Since $\Delta N = 1 \ll N$, instead of (L.1), we may write

$$\frac{d(\ln N!)}{dN} = \ln N$$

or

$$d(\ln N!) = \ln N \; dN. \tag{L.2}$$

Integrating equation (L.2) from 1 to N, we obtain

$$\ln N! = \int_1^N \ln N \; dN.$$

This integral can be calculated by parts. Taking into account that $N \gg 1$, we have

$$\ln N! = N \ln N - N. \tag{L.3}$$

Bearing in mind that $N = \ln \mathrm{e}^N$, instead of (L.3), we may write

$$N! = \left(\frac{N}{\mathrm{e}}\right)^N. \tag{L.4}$$

Formula (L.3) (or the equivalent formula (L.4)) is called the **Stirling formula**. It is applicable at larges N-values.

APPENDIX M

ERROR INTEGRAL

The error integral is defined by the equality

$$\Phi(x) = \mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \mathrm{e}^{-t^2} \, dt.$$

It is easy to be convinced that

$$\int_{\pm x}^{\infty} \mathrm{e}^{-t^2} dt = \frac{\sqrt{\pi}}{2}\left[1 \mp \Phi(x)\right].$$

If $x \ll 1$,

$$\Phi(x) = \frac{2}{\sqrt{\pi}}\left(x - \frac{x^3}{3 \cdot 1!} + \frac{x^5}{5 \cdot 2!} - \dots\right),$$

and if $x = 1$,

$$\Phi(1) = \frac{2}{\sqrt{\pi}} \int_0^1 \mathrm{e}^{-t^2} dt \approx 0.8427.$$

The values of error integral for some arguments are quoted in the following table.

Error integral values

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0.0	0.0000	0.5	0.5205	1.2	0.9103	1.9	0.9928
0.1	0.0564	0.6	0.6039	1.3	0.9340	2.0	0.9953
0.2	0.1125	0.7	0.6778	1.4	0.9523	2.1	0.9970
0.3	0.1680	0.8	0.7421	1.5	0.9661	2.2	0.9981
0.4	0.2227	0.9	0.7969	1.6	0.9763	2.3	0.9989
0.5	0.3286	1.0	0.8427	1.7	0.9838	2.4	0.9993
0.6	0.4284	1.1	0.8802	1.8	0.9891	2.5	0.9996

APPENDIX N

INTEGRALS K_N

Let us calculate the integral

$$K_n = \int_0^{\infty} \frac{x^n dx}{\mathrm{e}^x - 1}, \tag{N.1}$$

where n is an integer or half-integer number. Since

$$\frac{1}{\mathrm{e}^x - 1} = \frac{\mathrm{e}^{-x}}{1 - \mathrm{e}^{-x}},$$

then, dividing e^{-x} by $1-\mathrm{e}^{-x}$, we find

$$\frac{1}{\mathrm{e}^x - 1} = \sum_{l=0}^{\infty} \mathrm{e}^{-(l+1)x},$$

so that

$$K_n = \int_0^{\infty} x^n \sum_{l=0}^{\infty} \mathrm{e}^{-(l+1)x} dx = \sum_{l=0}^{\infty} \int_0^{\infty} x^n \sum_{l=0}^{\infty} \mathrm{e}^{-(l+1)x} dx \tag{N.2}$$

Substituting $x = t/(t+1)$, we obtain

$$K_n = \sum_{l=0}^{\infty} \frac{1}{(l+1)^{n+1}} \int_0^{\infty} t^n \mathrm{e}^{-t} dt = \zeta(n+1)\Gamma(n+1). \tag{N.3}$$

The values of gamma- and zeta-functions for some values of their argument are quoted in Appendix I. Hence,

$$K_{1/2} \approx 2.315, \quad K_1 = \pi^2/6, \quad K_{3/2} \approx 1.783,$$
$$K_2 \approx 2.404, \quad K_3 = \pi^4/15.$$

APPENDIX O

INTEGRALS L_N

Let us consider the integral

$$L_n = \int_0^{\infty} \frac{x^n dx}{\mathrm{e}^x + 1}. \tag{O.1}$$

It is easy to verify that

$$\frac{1}{\mathrm{e}^x + 1} = \sum_{l=0}^{\infty} (-1)^l \mathrm{e}^{-(l+1)x}.$$

Therefore,

$$L_n = \sum_{l=0}^{\infty} (-1)^l \int_0^{\infty} x^n \mathrm{e}^{-(l+1)x} dx.$$

Substituting $x = t/(t+1)$, we obtain

$$L_n = \sum_{l=0}^{\infty} \frac{(-1)^l}{(l+1)^{n+1}} \int_0^{\infty} t^n \mathrm{e}^{-t} dt = \Gamma(n+1) \sum_{l=0}^{\infty} \frac{(-1)^l}{(l+1)^{n+1}}. \tag{O.2}$$

First, we sum up the sum in (O.2) with the analogous sum, but only with odd l-values. In this case, the added summands with odd l's and the initial summands with even l's give together the Riemann zeta-function,

$$\sum_{l=0}^{\infty} \frac{1}{(l+1)^{n+1}} = \zeta(n+1).$$

To compensate new summands, we must subtract the same sum, but with odd l-values. As a result, we obtain

$$\sum_{l=0}^{\infty} \frac{(-1)^l}{(l+1)^{n+1}} = \zeta(n+1) - 2 \sum_{m=0}^{\infty} \frac{1}{(2m+2)^{n+1}} = \zeta(n+1) - \frac{1}{2^n}\zeta(n+1). \tag{O.3}$$

Substituting (O.3) into (O.2), we obtain

$$L_n = \int_0^{\infty} \frac{x^n dx}{\mathrm{e}^x + 1} = \left(1 - \frac{1}{2^n}\right) \Gamma(n+1)\zeta(n+1). \tag{O.4}$$

Using the values of gamma- and zeta-functions quoted in Appendix I, we find

$$L_{1/2} \approx 0.673, \quad L_1 = \pi^2/12, \quad L_3 = 7\pi^4/120. \tag{O.5}$$

APPENDIX P

INTEGRAL M

Let us consider the integral

$$M = \int_0^{\infty} \frac{f(\varepsilon) d\varepsilon}{\mathrm{e}^{\frac{\varepsilon - \mu}{k_0 T}} + 1} \tag{P.1}$$

for the function $f(\varepsilon)$, at which the integral converges, and provided that $\frac{\mu}{k_0T} \gg 1$. Changing the integration variable, $\frac{\varepsilon - \mu}{k_0T} = z$, we obtain

$$M = k_0T \int_{-\frac{\mu}{k_0T}}^{\infty} \frac{f(\mu + k_0Tz)}{\mathrm{e}^z + 1}\, dz$$

$$= k_0T \int_{-\frac{\mu}{k_0T}}^{0} \frac{f(\mu + k_0Tz)}{\mathrm{e}^z + 1}\, dz + k_0T \int_0^{\infty} \frac{f(\mu + k_0Tz)}{\mathrm{e}^z + 1}\, dz. \tag{P.2}$$

In the first integral in (P.2), we make the transformation $z \to -z$ and use the identity

$$\frac{1}{\mathrm{e}^{-z} + 1} = 1 - \frac{1}{\mathrm{e}^z + 1}.$$

As a result, we have

$$M = k_0T \int_0^{\frac{\mu}{k_0T}} f(\mu - k_0Tz) dz$$

$$-k_0T \int_0^{\frac{\mu}{k_0T}} \frac{f(\mu - k_0Tz)}{\mathrm{e}^z + 1}\, dz + k_0T \int_0^{\infty} \frac{f(\mu + k_0Tz)}{\mathrm{e}^z + 1}\, dz. \tag{P.3}$$

Taking into account that

$$k_0T \int_0^{\frac{\mu}{k_0T}} f(\mu - k_0Tz) dz = \int_0^{\mu} f(\varepsilon) d\varepsilon,$$

we obtain

$$M = \int_0^{\mu} f(\varepsilon) d\varepsilon + k_0T \int_0^{\infty} \frac{f(\mu + k_0Tz)}{\mathrm{e}^z + 1}\, dz - k_0T \int_0^{\frac{\mu}{k_0T}} \frac{f(\mu - k_0Tz)}{\mathrm{e}^z + 1}\, dz. \tag{P.4}$$

If $\frac{\mu}{k_0T} \gg 1$, we may put $\frac{\mu}{k_0T} \to \infty$. Than,

$$M = \int_0^{\mu} f(\varepsilon) d\varepsilon + k_0T \int_0^{\infty} \frac{f(\mu + k_0Tz) - f(\mu - k_0Tz)}{\mathrm{e}^z + 1}\, dz. \tag{P.5}$$

According to the same condition $\dfrac{\mu}{k_0 T} \gg 1$, the functions $f(\mu \pm k_0 T z)$ can be expanded in power series of $k_0 T z$ up to first few terms,

$$f(\mu \pm k_0 T z) = f(\mu) \pm f'(\mu) k_0 T z + \frac{1}{2} f''(\mu) \left(k_0 T z\right)^2 \pm \frac{1}{6} f'''(\mu) \left(k_0 T z\right)^3 + \dots .$$

Then, instead of (P.5), we obtain

$$\begin{aligned} M &= \int_0^\mu f(\varepsilon) d\varepsilon \\ &\quad + 2 \left(k_0 T\right)^2 f''(\mu) \int_0^\infty \frac{z dz}{\mathrm{e}^z + 1} \\ &\quad + \frac{1}{3} \left(k_0 T\right)^4 f'''(\mu) \int_0^\infty \frac{z^3 dz}{\mathrm{e}^z + 1} + \dots . \end{aligned}$$

The last two integrals equal (see Appendix O) $L_1 = \pi^2/12$ and $L_3 = 7\pi^4/120$. Therefore,

$$\begin{aligned} M &= \int_0^\mu f(\varepsilon) d\varepsilon \\ &\quad + \frac{1}{6} \left(\pi k_0 T\right)^2 f''(\mu) + \frac{7}{360} \left(\pi k_0 T\right)^4 f'''(\mu) + \dots . \end{aligned} \tag{P.6}$$

In section 24.3, M-integrals with $f(\varepsilon) = \varepsilon^{1/2}$ and $\varepsilon^{3/2}$ were used. In those cases, retaining first two summands in (P.6), we find

$$\int_0^\infty \frac{\varepsilon^{1/2} d\varepsilon}{\mathrm{e}^{\frac{\varepsilon - \mu}{k_0 T}} + 1} \approx \frac{2}{3} \mu^{3/2} \left[1 + \frac{1}{8} \left(\frac{\pi k_0 T}{\mu} \right)^2 \right], \tag{P.7}$$

$$\int_0^\infty \frac{\varepsilon^{3/2} d\varepsilon}{\mathrm{e}^{\frac{\varepsilon - \mu}{k_0 T}} + 1} \approx \frac{2}{5} \mu^{5/2} \left[1 + \frac{5}{8} \left(\frac{\pi k_0 T}{\mu} \right)^2 \right]. \tag{P.8}$$

REFERENCES

[1] I.E. Tamm, *Fundamentals the Theory of Electricity* (Gostekhteorizdat, Moscow, 1957) (in Russian).

[2] V.G. Levich, *Course of Theoretical Physics:, Vols. 1–2* (Fizmatgiz, Moscow, 1962) (in Russian).

[3] A.M. Fedorchenko, *Theoretical Physics. Classical Mechanics and Electrodynamics* (Vyshcha Shkola, Kyiv, 1992) (in Ukrainian).

[4] V.Y. Sugakov, *Theoretical Physics. Electrodynamics* (Vyshcha Shkola, Kyiv, 1974) (in Ukrainian).

[5] V.A. Ugarov, *Special Theory of Relativity* (Nauka, Moscow, 1977) (in Russian).

[6] A.N. Matveev, *Electrodynamics and the Relativity Theory* (Vysshaya Shkola, Moscow, 1964) (in Russian).

[7] G.O. Bugaenko and M.E. Fonkych, *Electrodynamics. Theory of Relativity* (Radyanska Shkola, Kyiv,1965) (in Ukrainian).

[8] V.V. Multanovskii ans A.S. Vasilevskii, *Course of Theoretical Physics. Classical Electrodynamics* (Prosveshchenie, Moscow, 1990) (in Russian).

[9] A.I. Borisenko and I.E. Tarapov, *Vector Analysis and Basics of Tensor Calculus* (Vysshaya Shkola, Moscow, 1963) (in Russian).

[10] E.F. Venger, V.M. Gryban, and O.V. Melnichuk, *Fundamentals of Electrodynamics and Relativity Theory* (Vyshcha Shkola, Kyiv, 2009) (in Ukrainian).

[11] A.S. Davydov, *Quantum Mechanics* (Nauka, Moscow, 1973) (in Russian).

[12] D.I. Blokhintsev, *Basics of Quantum Mechanics* (Nauka, Moscow, 1976) (in Russian)..

[13] I.R. Yukhnovskyi, *Quantum Mechanics* (Lybid, Kyiv, 1995) (in Ukrainian).

[14] I.O. Vakarchuk, *Quantum Mechanics* (Lviv Strate Univ., Lviv, 1998) (in Ukrainian).

[15] V.G. Levich, Yu.A. Vdovin, and V.A. Myamlin, *Course of Theoretical Physics, Vols. 1–2* (Nauka, Moscow, 1973) (in Russian), Vol. 2.

[16] E.V. Shpolskii, *Atomic Physics, Vols. 1–2* (Nauka, Moscow, 1974).

[17] I.V. Savel'yev, *Basics of Theoretical Physics. Quantum Mechanics, Vols. 1–2* (Nauka, Moscow, 1977), Vol. 2.

[18] A.M. Fedorchenko, *Theoretical Physics, Vols. 1–2* (Vyshcha Shkola, Kyiv, 1993) (in Ukrainian), Vol. 2.

[19] V.V. Multanovskii and A.S. Vasilevskii, *Course of Theoretical Physics. Quantum Mechanics* (Prosveshchenie, Moscow, 1991) (in Russian).

[20] E.F. Venger, V.M. Gryban, and O.V. Melnichuk, *Fundamentals of Quantum Mechanic* (Vyshcha Shkola, Kyiv, 2002).

[21] A.S. Vasilevskii and V.V. Multanovskii, *Statistical Physics and Thermodynamics* (Prosveshchenie, Moscow, 1985) (in Russian).

[22] E.F. Venger, V.M. Gryban, and O.V. Melnichuk, *Fundamentals of Statistical Physics and Thermodynamics* (Vyshcha Shkola, Kyiv, 2004) (in Ukrainian).

[23] A.M. Fedorchenko, *Introduction to Statistical Physics and Thermodynamics* (Vyshcha Shkola, Kyiv, 1973) (in Ukrainian).

[24] N.P. Bazarov, *Thermodynamics* (Fizmatgiz, Moscow, 1961) (in Russian).

[25] V.F. Nozdrev, *Course of Thermodynamics* (Prosveshchenie, Moscow, 1967) (in Russian).

[26] F. Reif, *Fundamentals of Statistical and Thermal Physics* (McGraw-Hill, New York, 1965).

[27] Yu.B. Rumer and M.Sh. Ryvkin, *Thermodynamics. Statistical physics and Kinetics* (Nauka, Moscow, 1977) (in Russian).

Index

У короткому курсі теоретичної фізики викладено поняття, положення та закони класичної електродинаміки, теорії відносності, квантової механіки, термодинаміки і статистичної фізики.

Для студентів вищих навчальних закладів.

Навчальне видання

НАЦІОНАЛЬНА АКАДЕМІЯ НАУК УКРАЇНИ

ВЕНГЕР Євген Федорович
ГРИБАНЬ Володимир Миколайович
МЕЛЬНИЧУК Олександр Володимирович

ОСНОВИ ТЕОРЕТИЧНОЇ ФІЗИКИ

Англійською мовою

Видається в авторській редакції

Підп. до друку 07.06.2016. Формат 60 × 84/16.
Ум. друк. арк. 28,6. Обл.-вид. арк. 36,9.

Видавець і виготовлювач
Видавничий дім «Академперіодика» НАН України
01004, Київ-4, вул. Терещенківська, 4

Свідоцтво субєкта видавничої справи
ДК № 544 від 27.07.2001 р.