I. V. Savelyev

PHYSICS

A General Course

QUANTUM OPTICS

ATOMIC PHYSICS

SOLID STATE PHYSICS

PHYSICS OF THE ATOMIC NUCLEUS AND ELEMENTARY PARTICLES

Mir Publishers
Moscow

III

И. В. САВЕЛЬЕВ

КУРС ОБЩЕЙ ФИЗИКИ

ТОМ III

КВАНТОВАЯ ОПТИКА
АТОМНАЯ ФИЗИКА
ФИЗИКА ТВЕРДОГО ТЕЛА
ФИЗИКА АТОМНОГО ЯДРА
И ЭЛЕМЕНТАРНЫХ ЧАСТИЦ

ИЗДАТЕЛЬСТВО «НАУКА»
МОСКВА

I. V. SAVELYEV

PHYSICS

A GENERAL COURSE

(In three volumes)

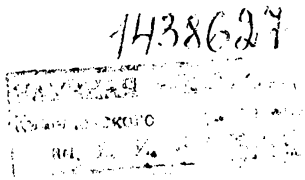
VOLUME III

QUANTUM OPTICS
ATOMIC PHYSICS
SOLID STATE PHYSICS
PHYSICS OF
THE ATOMIC NUCLEUS
AND ELEMENTARY
PARTICLES

MIR PUBLISHERS
MOSCOW

Translated from Russian by G. Leib

First published 1981
Revised from the 1979 Russian edition
Second printing 1985
Third printing 1989



На английском языке

Printed in the Union of Soviet Socialist Republics

ISBN 5-03-000903-5
ISBN 5-03-000900-0

PREFACE

This volume completes my work on a new version of a general course in physics for higher technical educational institutions (the first version was written by the author in the beginning of the 1960's). In this connection, I would like to note the following. Writing of this course required a fresh view on a number of questions, the rejection of obsolete traditions in the teaching of physics that were formed during many decades. This rejection was not at all simple for me because I myself was brought up on these traditions and for a number of years supported some of them (in particular, in the preceding version of the three-volume course). Speaking figuratively, I had to "reject myself". This difficult process was facilitated by daily contact with my young colleagues at the Department of General Physics of the Moscow Institute of Engineering Physics. Of the greatest importance was not so much the influence of these young people on the nature of the treatment of individual concrete questions of physics as the spirit of creative criticism and innovation that was established in the department after their joining it. Special mention must be made of the part played by associate professors N.B. Narozhny, V.I. Gervids, and V.N. Likhachev.

In addition to the influence of my young colleagues noted above, a decisive part in my work on this course was played by constant active contact with my students at lectures, exercises, consultations, and examinations. It is impossible to write a textbook without being in contact with whom it is intended for, associating with them only unilaterally at lectures. In instruction, as in any other vocation, experiments are needed. Among several possible ways of setting out a question, preference must be given to the one that produces the best result in the course of instruction. Such experiments were conducted quite broadly during my work on the new three-volume course.

I shall note in conclusion that the present course is intended above all for higher technical schools with an extended syllabus in physics. The material has been arranged, however, so that the book can be used as a teaching aid for higher technical schools with an ordinary syllabus simply by omitting some sections.

Igor Savelyev

Moscow, April, 1980

CONTENTS

Preface 5

PART I. QUANTUM OPTICS

CHAPTER 1. THERMAL RADIATION

1.1. Thermal Radiation and Luminescence 11
1.2. Kirchhoff's Law 12
1.3. Equilibrium Density of Radiant Energy 16
1.4. The Stefan-Boltzmann Law and Wien's Displacement Law 18
1.5. Standing Waves in Three-Dimensional Space 20
1.6. The Rayleigh-Jeans Formula 28
1.7. Planck's Formula 29

CHAPTER 2. PHOTONS

2.1. Bremsstrahlung 34
2.2. The Photoelectric Effect 36
2.3. Bothe's Experiment. Photons 40
2.4. The Compton Effect 44

PART II. ATOMIC PHYSICS

CHAPTER 3. THE BOHR THEORY OF THE ATOM

3.1. Regularities in Atomic Spectra 48
3.2. The Thomson Model of the Atom 50
3.3. Experiments in Scattering Alpha-Particles. The Nuclear Model of the Atom 52
3.4. Bohr's Postulates. The Franck-Hertz Experiment 57
3.5. Rule for Quantization of Circular Orbits 60
3.6. The Elementary Bohr Theory of the Hydrogen Atom 62

CHAPTER 4. ELEMENTS OF QUANTUM MECHANICS

4.1. De Broglie's Hypothesis. Wave Properties of Matter 65
4.2. The Unusual Properties of Microparticles 68
4.3. The Uncertainty Principle 70
4.4. The Schrödinger Equation 74
4.5. The Meaning of the Psi-Function 78
4.6. Quantization of Energy 79
4.7. Quantization of Angular Momentum 83

4.8. The Superposition Principle 87
4.9. Penetration of Particles Through a Potential Barrier 88
4.10. Harmonic Oscillator 93

CHAPTER 5. THE PHYSICS OF ATOMS AND MOLECULES

5.1. The Hydrogen Atom 95
5.2. Spectra of the Alkali Metals 101
5.3. Breadth of Spectral Lines 104
5.4. Multiplicity of Spectra and Spin of an Electron 109
5.5. Resultant Mechanical Angular Momentum of an Atom with Many Electrons 115
5.6. The Magnetic Moment of an Atom 117
5.7. The Zeeman Effect 122
5.8. Electron Paramagnetic Resonance 127
5.9. The Pauli Principle. Distribution of Electrons by Energy Levels of an Atom 129
5.10. Mendeleev's Periodic System of Elements 131
5.11. X-Ray Spectra 136
5.12. Energy of a Molecule 138
5.13. Molecular Spectra 142
5.14. Combination Scattering of Light 146
5.15. Stimulated Emission 148
5.16. Lasers 150
5.17. Non-Linear Optics 155

PART III. SOLID STATE PHYSICS

CHAPTER 6. OSCILLATIONS OF A CRYSTAL LATTICE

6.1. Crystal Lattice. Miller Indices 158
6.2. Heat Capacity of Crystals. Einstein's Theory 160
6.3. Oscillations of Systems with a Large Number of Degrees of Freedom 162
6.4. Debye's Theory 164
6.5. Phonons 168
6.6. The Mössbauer Effect 170

CHAPTER 7. THE BAND THEORY OF SOLIDS

7.1. The Quantum Theory of Free Electrons in a Metal 177
7.2. The Fermi-Dirac Distribution 182
7.3. Energy Bands in Crystals 185
7.4. Dynamics of Electrons in a Crystal Lattice 190

CHAPTER 8. THE ELECTRICAL CONDUCTANCE OF METALS AND SEMICONDUCTORS

8.1. The Electrical Conductance of Metals 194
8.2. Superconductivity 197
8.3. Semiconductors 200
8.4. Intrinsic Conductance of Semiconductors 202
8.5. Impurity Conductance of Semiconductors 205

CHAPTER 9. CONTACT AND THERMOELECTRIC PHENOMENA

9.1. Work Function 208
9.2. Thermionic Emission Electronic Tubes 210

9.3. Contact Potential Difference 215
9.4. Thermoelectric Effects 218
9.5. Semiconductor Diodes and Triodes 224
9.6. The Barrier-Layer Photoelectric Effect 229

PART IV. PHYSICS OF THE ATOMIC NUCLEUS AND ELEMENTARY PARTICLES

CHAPTER 10. THE ATOMIC NUCLEUS

10.1. Composition and Characteristic of the Atomic Nucleus 231
10.2. Mass and Binding Energy of a Nucleus 234
10.3. Models of the Atomic Nucleus 237
10.4. Nuclear Forces 238
10.5. Radioactivity 243
10.6. Nuclear Reactions 251
10.7. The Fission of Nuclei 256
10.8. Thermonuclear Reactions 262

CHAPTER 11. ELEMENTARY PARTICLES

11.1. Kinds of Interactions and Classes of Elementary Particles 266
11.2. Methods for Detecting Elementary Particles 267
11.3. Cosmic Rays 271
11.4. Particles and Antiparticles 272
11.5. Isotopic Spin 280
11.6. Strange Particles 283
11.7. Non-Conservation of Parity in Weak Interactions 286
11.8. The Neutrino 290
11.9. Systematization of Elementary Particles 294
11.10. Quarks 298
11.11. Conclusion 301

APPENDIX. LIST OF SYMBOLS 305

Name Index 495 307

Subject Index 497

PART I QUANTUM OPTICS

CHAPTER 1 THERMAL RADIATION

1.1. Thermal Radiation and Luminescence

Bodies can emit electromagnetic waves (glow) at the expense of various kinds of energy. The most widespread is **thermal radiation,** i.e. the emission of electromagnetic waves at the expense of the internal energy of bodies. All the other kinds of glow produced at the expense of any kind of energy except internal (thermal) energy are combined under the single term **"luminescence"**.

Phosphorus oxidizing in the air glows at the expense of the energy liberated upon the chemical transformation. This kind of glow is known as **chemiluminescence.** The glow produced in different kinds of self-sustained gas discharge is called **electroluminescence.** The glow of solid bodies due to their being bombarded by electrons is known as **cathodoluminescence.** The glow due to a body absorbing electromagnetic radiation is called **photoluminescence.**

Thermal radiation occurs at any temperature, but at low temperatures practically only long (infrared) electromagnetic waves are emitted.

Let us put an emitting body into an enclosure having an ideally reflecting surface (Fig. 1.1). We shall evacuate the air from the enclosure. The radiation reflected by the enclosure will fall on the body and be absorbed by it (partly or completely). Consequently, a continuous exchange of energy between the body and the radiation filling the enclosure will occur. If the distribution of energy between the body and the radiation remains constant for every wavelength, the state of the body-radiation system will be an equilibrium one. Experiments show that the only kind of radiation that can be in equilibrium with emitting bodies is thermal radiation. All other kinds of radiation are non-equilibrium ones.

The ability of thermal radiation to be in equilibrium with emitting bodies is due to the fact that its intensity grows with elevation

of the temperature. Assume that equilibrium between a body and radiation is violated, and the body emits more energy than it absorbs. The internal energy of the body will therefore diminish, which leads to lowering of the temperature. This, in turn, will result in a reduction in the amount of energy emitted by the body. The temperature of the body will lower until the amount of energy emitted by the body becomes equal to the amount of absorbed energy. If equilibrium is violated in the opposite direction, i.e. if less energy is emitted than absorbed, the temperature of the body will grow until equilibrium sets in again. Thus, the violation of equilibrium in the body-radiation system gives birth to processes restoring equilibrium.

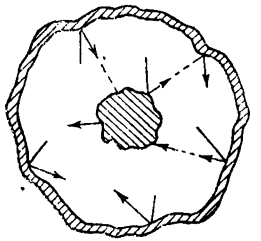


Fig. 1.1

Matters are different with luminescence. We shall show this using chemiluminescence as an example. Proceeding of the chemical reaction producing radiation causes the emitting body to become more and more remote from its initial state. The absorption of radiation by the body will not change the direction of the reaction, but, on the contrary, will result in the reaction proceeding at a faster rate (owing to heating) in the initial direction. Equilibrium will set in only when the reactants will be completely used up, and the glow due to chemical processes will be replaced by thermal radiation.

Thus, of all the kinds of radiation, only thermal radiation can be in equilibrium. The laws of thermodynamics can be applied to equilibrium states and processes. This is why thermal radiation must obey some general laws following from the principles of thermodynamics. We shall now pass over to a treatment of these laws.

1.2. Kirchhoff's Law

We shall characterize the intensity of thermal radiation by the magnitude of the energy flux measured in watts. The energy flux emitted by unit surface area of a radiating body in all directions (within the limits of a solid angle of 2π) is known as the **radiant emittance** of the body. We shall use the symbol R to designate this quantity. The radiant emittance is a function of the temperature.

Radiation consists of waves having different frequencies ω (or wavelengths λ). Let dR_ω be the energy flux emitted by unit surface area of a body within the frequency interval $d\omega$. When the interval

$d\omega$ is small, the flux dR_ω will be proportional to $d\omega$:

$$dR_\omega = r_\omega \, d\omega \tag{1.1}$$

The quantity r_ω is called the **emissivity** of a body. Like the radiant emittance, the emissivity depends greatly on the temperature of a body. Thus, r_ω is a function of the frequency and temperature.

The radiant emittance and the emissivity are related by the formula

$$R_T = \int dR_{\omega T} = \int_0^\infty r_{\omega T} \, d\omega \tag{1.2}$$

(to stress that the radiant emittance and the emissivity depend on the temperature, we have provided them with the subscript T).

Radiation can be characterized by its wavelength λ instead of its frequency ω. The wavelength interval $d\lambda$ will correspond to the spectrum portion $d\omega$. The quantities $d\omega$ and $d\lambda$ determining the same portion are related by a simple expression following from the formula $\lambda = 2\pi c/\omega$. Differentiation yields

$$d\lambda = -\frac{2\pi c}{\omega^2} \, d\omega = -\frac{\lambda^2}{2\pi c} \, d\omega \tag{1.3}$$

The minus sign in this expression is of no appreciable significance. It only indicates that when one of the quantities ω or λ grows, the other one diminishes. We shall therefore omit the minus sign in the following.

The fraction of the radiant emittance falling within the interval $d\lambda$, by analogy with Eq. (1.1), can be written in the form

$$dR_\lambda = r_\lambda \, d\lambda \tag{1.4}$$

If the intervals $d\omega$ and $d\lambda$ in expressions (1.1) and (1.4) are related by Eq. (1.3), i.e. belong to the same portion of the spectrum, then the quantities dR_ω and dR_λ must coincide:

$$r_\omega \, d\omega = r_\lambda \, d\lambda$$

Substituting for $d\lambda$ in this equation its value from Eq. (1.3), we get

$$r_\omega \, d\omega = r_\lambda \frac{2\pi c}{\omega^2} \, d\omega = r_\lambda \frac{\lambda^2}{2\pi c} \, d\omega$$

whence

$$r_\omega = r_\lambda \frac{2\pi c}{\omega^2} = r_\lambda \frac{\lambda^2}{2\pi c} \tag{1.5}$$

Equation (1.5) allows us to transfer from r_λ to r_ω and vice versa.

Assume that the flux of radiant energy $d\Phi_\omega$ due to electromagnetic waves whose frequency is within the interval $d\omega$ falls on an elementary area of a body's surface. A part of this flux $d\Phi'_\omega$ will be absorbed

by the body. The dimensionless quantity

$$a_{\omega T} = \frac{d\Phi'_{\omega}}{d\Phi_{\omega}} \tag{1.6}$$

is called the **absorptivity** of a body. The absorptivity of a body is a function of the frequency and temperature.

By definition, $a_{\omega T}$ cannot be greater than unity. For a body completely absorbing the radiation of all frequencies falling on it, $a_{\omega T} \equiv \equiv 1$. Such a body is known as a **blackbody.** A body for which $a_{\omega T} \equiv a_T = \text{const} < 1$ is called a **gray body.**

There is a definite relation between the emissivity and absorptivity of any body. We can convince ourselves that this is true by considering the following experiment. Assume that several bodies are confined in an enclosure maintained at a constant temperature T

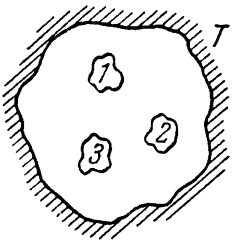


Fig. 1.2

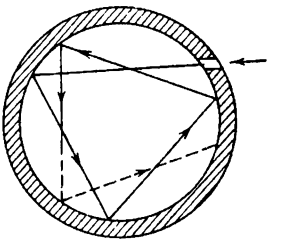


Fig. 1.3

(Fig. 1.2). The cavity inside the enclosure is evacuated so that the bodies can exchange energy with one another and with the enclosure only by emitting and absorbing electromagnetic waves. Experiments show that such a system will arrive at a state of thermal equilibrium after a certain time elapses—all the bodies will acquire the same temperature T equal to that of the enclosure. In this state, a body having a greater emissivity $r_{\omega T}$ loses more energy from unit surface area in unit time than a body whose emissivity $r_{\omega T}$ is lower. Since the temperature (and, consequently, the energy) of the bodies does not change, then the body emitting more energy must absorb more, i.e. have a greater $a_{\omega T}$. Thus, the greater the emissivity $r_{\omega T}$ of a body, the greater is its absorptivity $a_{\omega T}$. Hence follows the relation

$$\left(\frac{r_{\omega T}}{a_{\omega T}}\right)_1 = \left(\frac{r_{\omega T}}{a_{\omega T}}\right)_2 = \left(\frac{r_{\omega T}}{a_{\omega T}}\right)_3 = \ldots \tag{1.7}$$

where the subscripts 1, 2, 3, etc. relate to different bodies.

Relation (1.7) expresses the following law established by the German physicist Gustav Kirchhoff (1824-1887): *the ratio of the emissivity and the absorptivity does not depend on the nature of a body, it is the same (universal) function of the frequency (wavelength) and temperature for all bodies*:

$$\frac{r_{\omega T}}{a_{\omega T}} = f(\omega, T) \tag{1.8}$$

The quantities $r_{\omega T}$ and $a_{\omega T}$ can vary exceedingly greatly for different bodies. Their ratio, however, is identical for all bodies. This signifies that a body which absorbs certain rays to a greater extent will emit these rays to a greater extent too (do not confuse the emission of rays with their reflection).

For a blackbody, by definition, $a_{\omega T} \equiv 1$. It therefore follows from Eq. (1.8) that $r_{\omega T}$ for such a body equals $f(\omega, T)$. Thus, Kirchhoff's universal function $f(\omega, T)$ is nothing but the emissivity of a blackbody.

It is more convenient to use the function of the frequency $f(\omega, T)$ to characterize the spectral composition of equilibrium thermal radiation in theoretical investigations. The function of the wavelength $\varphi(\lambda, T)$ is more convenient in experimental studies. The two functions are related by the formula

$$f(\omega, T) = \frac{2\pi c}{\omega^2} \varphi(\lambda, T) = \frac{\lambda^2}{2\pi c} \varphi(\lambda, T) \tag{1.9}$$

similar to Eq. (1.5). According to Eq. (1.9), to find $\varphi(\lambda, T)$ from the known function $f(\omega, T)$, we must substitute $2\pi c/\lambda$ for the frequency ω in $f(\omega, T)$ and multiply the expression obtained by $2\pi c/\lambda^2$:

$$\varphi(\lambda, T) = \frac{2\pi c}{\lambda^2} f\left(\frac{2\pi c}{\lambda}, T\right) \tag{1.10}$$

To find $f(\omega, T)$ from the known function $\varphi(\lambda, T)$, we must use the relation

$$f(\omega, T) = \frac{2\pi c}{\omega^2} \varphi\left(\frac{2\pi c}{\omega}, T\right) \tag{1.11}$$

Blackbodies do not exist in nature. Carbon black and platinum black have an absorptivity $a_{\omega T}$ close to unity only within a limited range of frequencies. Their absorptivity is appreciably lower than unity in the far infrared region. It is possible to construct a device, however, whose properties are close to those of a blackbody as much as desired. Such a device is an almost completely enclosed cavity provided with a small hole (Fig. 1.3). The radiation penetrating into the cavity through the hole will undergo multifold reflections before emerging from it. Part of the energy is absorbed upon each reflection, and as a result virtually the entire radiation of any frequency is absorbed by such a cavity*. According to Kirchhoff's law, the emissivity of such a device is very close to $f(\omega, T)$, where T stands for the temperature of the cavity walls. Thus, if the cavity walls are maintained at the temperature T, then radiation will leak out through the hole very close in its spectral composition to the radiation of a blackbody at the same temperature. By obtaining the spectrum of this radiation with the aid of a diffraction grating and measur-

* For the same reason, the interior of a room seems dark when we look at it from a distance through an open window on a bright sunny day.

ing the intensity of different portions of the spectrum, we can find the form of the function $f(\omega, T)$ or $\varphi(\lambda, T)$ experimentally. The results of such experiments are shown in Fig. 1.4. Each curve is for

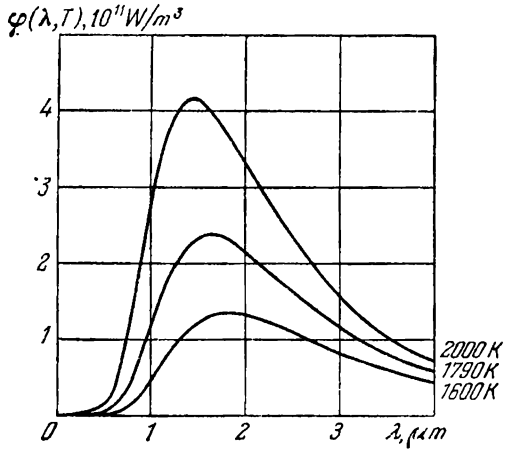


Fig. 1.4

a definite value of the temperature T of our blackbody. The area enclosed by the curve gives the radiant emittance of the blackbody at the corresponding temperature.

A glance at Fig. 1.4 shows that the radiant emittance of a blackbody grows greatly with the temperature. The maximum of the emissivity shifts toward shorter waves with elevation of the temperature.

1.3. Equilibrium Density of Radiant Energy

Let us consider radiation that is in equilibrium with a substance For this purpose, let us imagine an evacuated cavity whose walls are maintained at a constant temperature T. In the equilibrium state, the radiant energy will be distributed throughout the volume of the cavity with a definite density $u = u(T)$. The spectral distribution of this energy can be characterized by the function $u(\omega, T)$ determined by the condition $du_\omega = u(\omega, T)\, d\omega$, where du_ω is the fraction of the energy density falling within the interval of frequencies $d\omega$. The total energy density $u(T)$ is related to the function $u(\omega, T)$ by the formula

$$u(T) = \int_0^\infty u(\omega, T)\, d\omega \qquad (1.12)$$

It follows from thermodynamic considerations that the equilibrium radiant energy density $u\ (T)$ depends only on the temperature and does not depend on the properties of the cavity walls. Let us consider two cavities whose walls are made from different materials and initially have the same temperature. Let the equilibrium energy density in the two cavities be different and, say, $u_1\ (T) > u_2\ (T)$. We shall connect the cavities by means of a small hole (Fig. 1.5) and thus permit the walls of the cavities to exchange heat by radiation. Since we have assumed that $u_1 > u_2$, the energy flux from the first cavity into the second one must be greater than the flux in the opposite direction. The walls of the second cavity, as a result, will absorb more energy than they emit, and their temperature will start growing. The walls of the first cavity, on the other hand, will absorb less energy than they emit, and they will cool. But two bodies having the same initial temperature cannot acquire different temperatures as a result of heat exchange with each other—this is forbidden by the second law of thermodynamics. We must therefore acknowledge that our assumption on u_1 and u_2 being different is not lawful. The conclusion on the equality of $u_1\ (T)$ and $u_2\ (T)$ covers each spectral component $u\ (\omega,\ T)$.

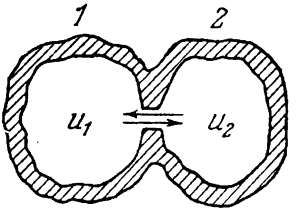


Fig. 1.5

That the equilibrium radiation does not depend on the nature of the cavity walls can be explained by the following considerations. Blackbody walls would absorb all the energy Φ_e falling on them and would emit the same energy flux Φ_e. Walls with the absorptivity a will absorb the fraction $a\Phi_e$ of the flux Φ_e falling on them and will reflect a flux equal to $(1 - a)\ \Phi_e$. In addition, they will emit the flux $a\Phi_e$ (equal to the absorbed flux). As a result, the walls of the cavity will return the same energy flux $\Phi_e = (1 - a)\ \Phi_e + a\Phi_e$ to the radiation that blackbody walls would return to it.

The equilibrium radiant energy density u is related to the radiant emittance of a blackbody R^* by a simple expression which we shall now proceed to derive*.

Let us consider an evacuated cavity with blackbody walls. In equilibrium, a radiant flux of the same density will pass through every point inside the cavity in any direction. If the radiation were to propagate in one given direction (i.e. if only one ray were to pass through a given point), the density of the energy flux at the point being considered would equal the product of the energy density u and the speed of an electromagnetic wave c. But a multitude of rays

* We have used the symbol R^* to stress that we are dealing with the radiant emittance of a blackbody.

whose directions are uniformly distributed within the limits of the solid angle 4π pass through every point. The energy flux cu is also distributed uniformly within the limits of this solid angle. Consequently, an energy flux whose density is

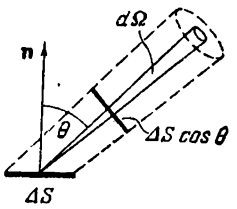


Fig. 1.6

$$dj = \frac{cu}{4\pi} d\Omega$$

will flow at every point within the limits of the solid angle $d\Omega$. Let us take the elementary area ΔS on the surface of the cavity (Fig. 1.6). This area emits the following energy flux within the limits of the solid angle $d\Omega = \sin\theta\, d\theta\, d\varphi$ in the direction making the angle θ with the normal **n**:

$$d\Phi_e = dj\, \Delta S \cos\theta = \frac{cu}{4\pi} d\Omega\, \Delta S \cos\theta =$$

$$= \frac{cu}{4\pi} \Delta S \cos\theta \sin\theta\, d\theta\, d\varphi$$

The area ΔS emits the energy flux

$$\Delta\Phi_e = \int d\Phi_e = \frac{cu}{4\pi} \Delta S \int_0^{\pi/2} \cos\theta \sin\theta\, d\theta \int_0^{2\pi} d\varphi = \frac{c}{4} u\, \Delta S \qquad (1.13)$$

in all the directions confined within the limits of the solid angle 2π.

At the same time, the energy flux emitted by the area ΔS can be found by multiplying the radiant emittance R^* by ΔS, i.e. $\Delta\Phi_e = R^* \Delta S$. A comparison with Eq. (1.13) shows that

$$R^* = \frac{c}{4} u \qquad (1.14)$$

Equation (1.14) must be satisfied for every spectral component of the radiation. It thus follows that

$$f(\omega, T) = \frac{c}{4} u(\omega, T) \qquad (1.15)$$

This formula relates the radiant emittance of a blackbody and the equilibrium energy density of thermal radiation.

1.4. The Stefan-Boltzmann Law and Wien's Displacement Law

The theoretical explanation of the laws of blackbody radiation had a tremendous significance in the history of physics—it led to the concept of energy quanta.

For a long time, attempts to obtain the form of the function $f(\omega, T)$* theoretically did not provide a general solution of the problem. In 1879, the Austrian physicist Joseph Stefan (1835-1893), analysing experimental data, arrived at the conclusion that the radiant emittance R of any body is proportional to the fourth power of the absolute temperature. But subsequent more accurate measurements, however, showed that his conclusions were erroneous. In 1884, the Austrian physicist Ludwig Boltzmann (1844-1906), on the basis of thermodynamic considerations, obtained theoretically the following value for the radiant emittance of a blackbody:

$$R^* = \int_0^\infty f(\omega, T)\, d\omega = \sigma T^4 \tag{1.16}$$

where σ is a constant quantity, and T is the absolute temperature. Thus, the conclusion which Stefan arrived at for gray bodies (he ran no experiments with blackbodies) was found to be true only for blackbodies.

Relation (1.16) between the radiant emittance of a blackbody and its absolute temperature was named the **Stefan-Boltzmann law.** The constant σ is called the **Stefan-Boltzmann constant.** Its experimental value is

$$\sigma = 5.7 \times 10^{-8} \text{ W/(m}^2 \cdot \text{K}^4) \tag{1.17}$$

In 1893, the German physicist Wilhelm Wien (1864-1928), using the electromagnetic theory in addition to thermodynamics, showed that the function of the spectral distribution must have the form

$$f(\omega, T) = \omega^3 F\left(\frac{\omega}{T}\right) \tag{1.18}$$

where F is a function of the ratio of the frequency to the temperature.

According to Eq. (1.10), the following expression is obtained for the function $\varphi(\lambda, T)$:

$$\varphi(\lambda, T) = \frac{2\pi c}{\lambda^2}\left(\frac{2\pi c}{\lambda}\right)^3 F\left(\frac{2\pi c}{\lambda T}\right) = \frac{1}{\lambda^5}\psi(\lambda T) \tag{1.19}$$

where $\psi(\lambda, T)$ is a function of the product λT.

Equation (1.19) makes it possible to establish the relation between the wavelength λ_m corresponding to the maximum of the function $\varphi(\lambda, T)$ and the temperature. Let us differentiate this expression with respect to λ:

$$\frac{d\varphi}{d\lambda} = \frac{1}{\lambda^5} T\psi'(\lambda T) - \frac{5}{\lambda^6}\psi(\lambda T) = \frac{1}{\lambda^6}[\lambda T\psi'(\lambda T) - 5\psi(\lambda T)] \tag{1.20}$$

The expression in brackets is a certain function $\Psi(\lambda T)$. At the wavelength λ_m corresponding to the maximum of the function $\varphi(\lambda, T)$

* Or, which is the same, the function $u(\omega, T)$ [see Eq. (1.15)].

Eq. (1.20) must become equal to zero:

$$\left(\frac{d\varphi}{d\lambda}\right)_{\lambda=\lambda_m} = \frac{1}{\lambda_m^6}\Psi(\lambda_m T) = 0 \tag{1.21}$$

It is known from experiments that λ_m is finite (i.e. $\lambda_m \neq \infty$). Therefore, the condition $\Psi(\lambda_m T) = 0$ must be satisfied. By solving Eq. (1.21) relative to the unknown quantity $\lambda_m T$, we get a certain value for this quantity which we shall denote by the symbol b. We thus obtain the relation

$$\lambda_m T = b \tag{1.22}$$

called Wien's **displacement law.** The experimental value of the constant b is

$$b = 2.90 \times 10^{-3}\ \text{m}\cdot\text{K} = 2.90 \times 10^{7}\ \text{Å}\cdot\text{K} \tag{1.23}$$

1.5. Standing Waves in Three-Dimensional Space

In finding the function $f(\omega, T)$, and also in calculating the heat capacities of solids (see Sec. 6.4), it becomes necessary to calculate the number of standing waves that can be produced in a volume of finite dimensions. We shall treat this question in the present section.

Assume that two plane waves produced as a result of reflection from walls at the points $x = 0$ and $x = a$ (Fig. 1.7) travel along the x-axis in opposite directions. The equations of the waves have the form

$$\begin{aligned} \xi_1 &= A\cos(\omega t - kx) \\ \xi_2 &= A\cos(\omega t + kx + \alpha) \end{aligned} \tag{1.24}$$

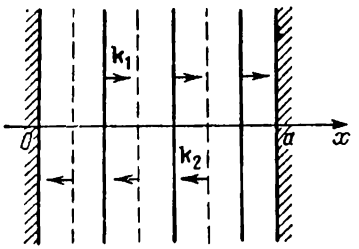


Fig. 1.7

(the initial phase of the first wave has been made to vanish by properly choosing the initial moment of counting the time). We know that in this case a standing wave is set up in the region $0 \leqslant x \leqslant a$, there being either nodes or antinodes at the boundaries of the region depending on the real conditions. Thus, nodes are observed at the ends of a string, and antinodes at the ends of a bar fixed at its middle.

Examination of Eqs. (1.24) reveals that for an antinode to appear at the boundary $x = 0$, the phase α must be zero (therefore at points with $x = 0$ the oscillations will occur in the same phase). In this case upon reflection from the boundary, the phase of the wave does not change*. For a node to appear at the boundary $x = 0$, the phase α

* This follows from the fact that in direct proximity to a wall (at $x \approx 0$) the phases of the oscillations ξ_1 and ξ_2 coincide.

must be π (consequently, at points with $x = 0$, the oscillations ξ_1 and ξ_2 will occur in counterphase). In this case upon reflection from the boundary, the phase of the wave undergoes a jump through π.

Thus, when antinodes are observed at the boundaries of the region, Eqs. (1.24) have the form

$$\xi_1 = A \cos(\omega t - kx),$$
$$\xi_2 = A \cos(\omega t + kx)$$

When nodes are observed at the boundaries of the region, Eqs. (1.24) have the following form:

$$\xi_1 = A \cos(\omega t - kx),$$
$$\xi_2 = A \cos(\omega t + kx + \pi)$$

Addition of the oscillations ξ_1 and ξ_2 for antinodes at the boundaries leads to the equation

$$\xi = \xi_1 + \xi_2 = 2A \cos kx \cdot \cos \omega t \tag{1.25}$$

and for nodes at the boundaries, to the equation

$$\xi = \xi_1 + \xi_2 = 2A \cos\left(kx + \frac{\pi}{2}\right) \cos\left(\omega t + \frac{\pi}{2}\right) \tag{1.26}$$

It is easy to see that when $x = 0$, the amplitude is maximum in the first case and equals zero in the second one.

To observe an antinode at the other boundary (i.e. when $x = a$) in the case described by Eq. (1.25) or a node in the case described by Eq. (1.26), the product ka must be an integral multiple of π, that is $ka = n\pi$. Thus, regardless of what is observed at the boundaries of the region (antinodes or nodes), the magnitude of the wave vector must have the values

$$k = \frac{\pi}{a} n \quad (n = 1, 2, \ldots) \tag{1.27}$$

Assume that $k' = (\pi/a)\, n'$, $k'' = (\pi/a)\, n''$. The difference $n'' - n'$ gives the number of standing waves ΔN_k, the magnitudes of whose wave vectors are within the interval $\Delta k = k'' - k'$. Taking into account the values of k' and k'', we find that

$$\Delta N_k = \frac{a}{\pi} \Delta k \tag{1.28}$$

The values of N_k form a discrete sequence. Replacing this sequence with a continuous function, we can write that

$$dN_k = \frac{a}{\pi} dk \tag{1.29}$$

The magnitude of the wave vector is related to the frequency ω and the velocity v by the expression

$$k = \frac{\omega}{v} \tag{1.30}$$

Accordingly,

$$dk = \frac{d\omega}{v} \tag{1.31}$$

(we consider that there is no dispersion, i.e. v = const). Substituting $d\omega/v$ for dk in Eq. (1.29), we arrive at the formula

$$dN_\omega = \frac{a}{\pi v} d\omega \tag{1.32}$$

where dN_ω is the number of standing waves whose frequencies are within the interval from ω to $\omega + d\omega$.

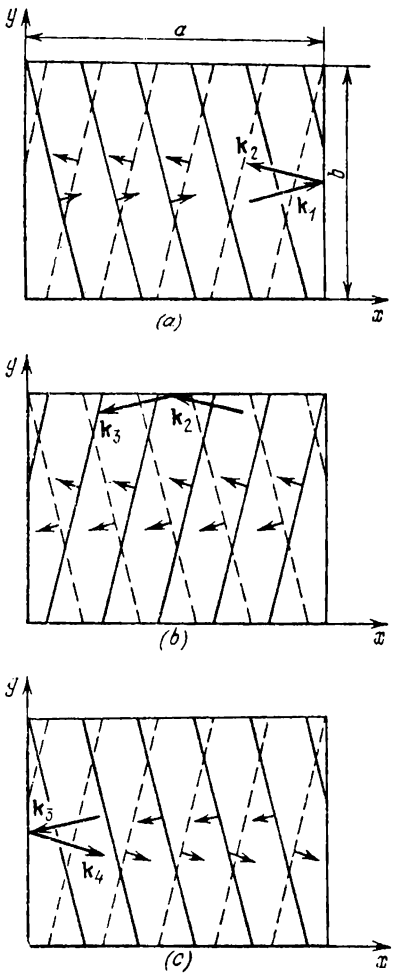


Fig. 1.8

Now let us consider the two-dimensional case. Assume that a plane wave (1) running in the direction of the wave vector $\mathbf{k}_1$ has been produced in a rectangular region with sides a and b (Fig. 1.8*a*). As a result of reflection from the right-hand boundary of the region, a running wave (2) will be produced with the wave vector $\mathbf{k}_2$. Reflection of the wave (2) from the top boundary (Fig. 1.8*b*) will produce a wave (3) with the wave vector $\mathbf{k}_3$. Finally, reflection of the wave (3) from the left-hand boundary (Fig. 1.8*c*) will produce a wave (4) with the wave vector $\mathbf{k}_4$. No other waves will be produced. Indeed, reflection of the wave (1) from the top boundary produces the wave (4), reflection of the wave (2) from the left-hand boundary produces the wave (1), reflection of the wave (3) from the bottom boundary produces the wave (2), and, finally, reflection of the wave (4) from the bottom and the right-hand boundaries of the region produces the waves (1) and (3), respectively.

Thus, the two-dimensional region will be filled with four plane waves running in the directions of the wave vectors $\mathbf{k}_1$, $\mathbf{k}_2$, $\mathbf{k}_3$, and $\mathbf{k}_4$. If we denote the projections of the vector $\mathbf{k}_1$ onto the axes x and y (see Fig. 1.8) by k_x and k_y, then the projections of all four vectors will be (the number of the vector is indicated in parentheses)

(1) k_x, k_y; (2) $-k_x$, k_y; (3) $-k_x$, $-k_y$; (4) k_x, $-k_y$

We have established above that antinodes are obtained at boundaries if the phase of a wave does not change upon reflection from a wall. In this case, the equations of the running waves have the form

$$\xi_1 = A\cos(\omega t - k_x x - k_y y), \quad \xi_3 = A\cos(\omega t + k_x x + k_y y)$$
$$\xi_2 = A\cos(\omega t + k_x x - k_y y), \quad \xi_4 = A\cos(\omega t - k_x x + k_y y)$$

Adding these equations in pairs, we obtain

$$\xi_1 + \xi_2 = 2A\cos k_x x\cos(\omega t - k_y y)$$
$$\xi_3 + \xi_4 = 2A\cos k_x x\cos(\omega t + k_y y)$$

The sum of the expressions found gives an equation describing a two-dimensional standing wave obtained when reflection from a boundary occurs without a jump in the phase of the running wave:

$$\xi = \xi_1 + \xi_2 + \xi_3 + \xi_4 = 4A\cos k_x x\cos k_y y\cos\omega t \qquad (1.33)$$

It can be seen from Eq. (1.33) that the amplitude is maximum at the point (0, 0). The following conditions must be satisfied for it also to be maximum at the points (0, b), (a, 0), and (a, b), i.e. at the other three apices of the rectangle:

$$k_x a = n_1\pi, \quad k_y b = n_2\pi \quad (n_1, n_2 = 1, 2, \ldots) \qquad (1.34)$$

We must note that owing to the presence of the multiplier $\cos k_y y$ in Eq. (1.33), the amplitude reaches its maximum value not along the entire length of the sides $x = 0$ and $x = a$, but only at the ends of these sides (where $y = 0$ and $y = b$), and also at $n_2 - 1$ intermediate points [at these points $k_y y$ takes on the values of π, 2π, ..., $(n_2 - 1)\pi$]. In the spaces between these points, the amplitude varies according to a cosine law. Similarly, the amplitude reaches a maximum not along the entire length of the sides $y = 0$ and $y = b$, but only at the ends of these sides, and also at $n_1 - 1$ intermediate points.

Nodes are obtained at the boundaries if upon reflection from a wall the phase of a wave undergoes a jump through π. Each of the waves (2), (3), (4) can be considered as the result of reflection of the preceding wave from a wall (see Fig. 1.8). Accordingly, the equations of the waves must be written in the form

$$\left.\begin{aligned}
\xi_1 &= A\cos(\omega t - k_x x - k_y y),\\
\xi_2 &= A\cos(\omega t + k_x x - k_y y + \pi),\\
\xi_3 &= A\cos(\omega t + k_x x + k_y y + 2\pi)\\
\xi_4 &= A\cos(\omega t - k_x x + k_y y + 3\pi)
\end{aligned}\right\} \qquad (1.35)$$

The phase of an oscillation permits the addition to or subtraction from it of a whole number of 2π's. With this in view, we shall alter

Eqs. (1.35) as follows:

$$\xi_1 = A\cos(\omega t - k_x x - k_y y),$$
$$\xi_2 = A\cos(\omega t + k_x x - k_y y + \pi),$$
$$\xi_3 = A\cos(\omega t + k_x x + k_y y)$$
$$\xi_4 = A\cos(\omega t - k_x x + k_y y + \pi)$$

Adding these equation in pairs, we get

$$\xi_1 + \xi_2 = 2A\cos\left(k_x x + \frac{\pi}{2}\right)\cos\left(\omega t - k_y y + \frac{\pi}{2}\right) \qquad (1.36)$$

$$\xi_3 + \xi_4 = 2A\cos\left(k_x x - \frac{\pi}{2}\right)\cos\left(\omega t + k_y y + \frac{\pi}{2}\right) \qquad (1.37)$$

Let us reverse the signs of the two cosines in Eq. (1.37) by adding π to the argument of the first cosine and subtracting π from the argument of the second cosine (the expression itself retains its previous magnitude). As a result, the sum $\xi_3 + \xi_4$ acquires the form

$$\xi_3 + \xi_4 = 2A\cos\left(k_x x + \frac{\pi}{2}\right)\cos\left(\omega t + k_y y - \frac{\pi}{2}\right)$$

Adding this sum to Eq. (1.36), we get the equation of a standing wave observed when the phase of a running wave undergoes a jump through π upon reflection from a boundary:

$$\xi = \xi_1 + \xi_2 + \xi_3 + \xi_4 = 4A\cos\left(k_x x + \frac{\pi}{2}\right)\cos\left(k_y y - \frac{\pi}{2}\right)\cos\omega t \qquad (1.38)$$

We must note that by adding (or subtracting) π to the arguments of two of the last three multipliers, we can impart the following form to the equation of a standing wave:

$$\xi = 4A\cos\left(k_x x - \frac{\pi}{2}\right)\cos\left(k_y y + \frac{\pi}{2}\right)\cos\omega t$$

or

$$\xi = 4A\cos\left(k_x x + \frac{\pi}{2}\right)\cos\left(k_y y + \frac{\pi}{2}\right)\cos(\omega t + \pi)$$

It follows from Eq. (1.38) that the amplitude is zero at all points of the boundary $x = 0$ and of the boundary $y = 0$. Conditions (1.34) must be satisfied for it to be zero too at the points of the boundaries $x = a$ and $y = b$.

Thus, regardless of what is obtained at the boundaries of the region (antinodes at the corners and at certain intermediate points, or nodes along the entire boundary), the projections of the wave vector must have the values

$$k_x = \frac{\pi}{a}n_1, \quad k_y = \frac{\pi}{b}n_2 \quad (n_1\ n_2 = 1, 2, \ldots) \qquad (1.39)$$

[compare with Eq. (1.27)]

We shall point out that the magnitude of the wave vector of all four running waves whose superposition leads to the setting up of a standing wave is the same and is

$$k = \sqrt{k_x^2 + k_y^2} \tag{1.40}$$

We shall call quantity (1.40) the magnitude of the wave vector of a standing wave.

Let us take a coordinate system on the k-plane with the axes k_x and k_y (Fig. 1.9). The four symmetrical points depicted in the figure

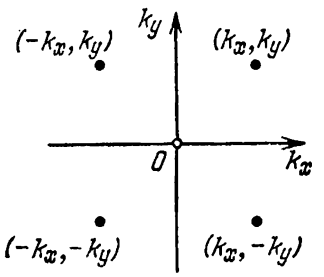


Fig. 1.9

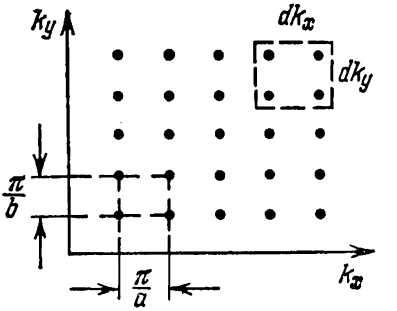


Fig. 1.10

correspond in the k-plane to the wave vectors of the four running waves forming a given standing wave. All these points correspond to the same standing wave. Therefore, when using the points to count the number of standing waves, we must take into account only the points in one of the quadrants of the k-plane. It is natural to consider the points in the first quadrant.

According to Eq. (1.39), the points corresponding to all possible standing waves are at the apices of rectangles with the sides π/a and π/b (Fig. 1.10). It is easy to see that an area equal to $\pi^2/ab = \pi^2/S$ (where S is the area of the two-dimensional region in whose limits a standing wave is produced) falls to the share of each standing wave on the k-plane. Hence, the density of the points on the k-plane is S/π^2.

Let us find the number of standing waves dN_{k_x, k_y} for which the projections of the wave vectors are within the limits from k_x to $k_x + dk_x$ and from k_y to $k_y + dk_y$. This number equals the density of the points multiplied by the area $dk_x\, dk_y$:

$$dN_{k_x, k_y} = \frac{S}{\pi^2}\, dk_x\, dk_y \tag{1.41}$$

Now let us find the number of standing waves dN_k for which the magnitude of the wave vector ranges from k to $k + dk$. This number

equals the number of points within the region confined between quarter-circles of radii k and $k + dk$ (Fig. 1.11). The area of this region is $\frac{1}{2}\pi k\, dk$. Multiplying the density of the points by the area of the region, we get

$$dN_k = \frac{S}{\pi^2}\frac{1}{2}\pi k\, dk = \frac{S}{2\pi} k\, dk \tag{1.42}$$

With a view to Eqs. (1.30) and (1.31), we can write that

$$dN_\omega = \frac{S}{2\pi v^2}\omega\, d\omega \tag{1.43}$$

where dN_ω is the number of standing waves whose frequencies are within the limits from ω to $\omega + d\omega$ [compare with Eq. (1.32)].

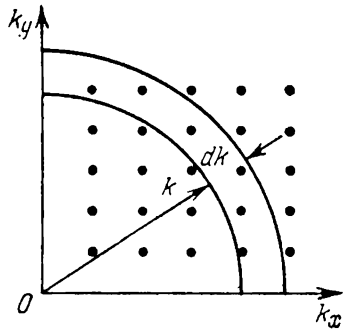


Fig. 1.11

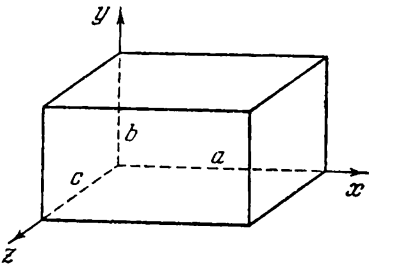


Fig. 1.12

It is simple to generalize the results obtained for the three-dimensional case. A standing wave produced within the limits of a rectangular region with sides a, b, and c parallel to the coordinate axes (Fig. 1.12) is formed by the superposition of eight running waves, the projections of whose wave vectors are

(1)	k_x,	k_y,	k_z;	(5)	k_x,	$-k_y$,	$-k_z$;
(2)	$-k_x$,	k_y,	k_z;	(6)	$-k_x$,	$-k_y$,	$-k_z$;
(3)	$-k_x$,	$-k_y$,	k_z;	(7)	$-k_x$,	k_y,	$-k_z$;
(4)	k_x,	$-k_y$,	k_z;	(8)	k_x,	k_y,	$-k_z$

We recommend our reader to write the equations of these waves and by performing the relevant calculations, to convince himself that the equation of a standing wave has the form

$$\xi = \xi_1 + \xi_2 + \ldots + \xi_8 = 8A \cos k_x x \cos k_y y \cos k_z z \cos \omega t \tag{1.44}$$

when a wave is reflected from the walls of a cavity without a change in phase, and

$$\xi = \xi_1 + \xi_2 + \ldots + \xi_8 =$$

$$= 8A \cos\left(k_x x + \frac{\pi}{2}\right) \cos\left(k_y y - \frac{\pi}{2}\right) \cos\left(k_z z + \frac{\pi}{2}\right) \cos\left(\omega t + \frac{\pi}{2}\right) \tag{1.45}$$

when the phase of a wave undergoes a jump through π upon reflection* [compare with Eqs. (1.33) and (1.38)].

We must note that in Eq. (1.45) we may simultaneously reverse the sign of $\pi/2$ in any two multipliers without changing the sign of ξ.

It can be seen from Eqs. (1.44) and (1.45) that for the amplitude of a standing wave to have the same value at all eight apices of the region in which the standing wave has been produced, the following conditions must be satisfied:

$$k_x = \frac{\pi}{a} n_1, \quad k_y = \frac{\pi}{b} n_2, \quad k_z = \frac{\pi}{c} n_3$$

$$(n_1, n_2, n_3 = 1, 2, \ldots) \tag{1.46}$$

[compare with Eq. (1.39)].

According to Eq. (1.45), the amplitude is zero everywhere at a boundary of the region. In the case described by Eq. (1.44), on the other hand, the maximum amplitude is obtained at the apices of the region, and also at separate points on the planes enclosing the region.

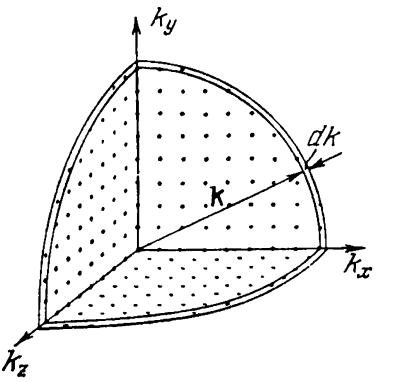


Fig. 1.13

A point in the first octant corresponds to every standing wave in k-space with the axes k_x, k_y, k_z (Fig. 1.13). The volume $\pi^3/abc = \pi^3/V$ (V is the volume of the region) falls to the share of each point. Hence, the density of the points is V/π^3.

The number of standing waves for which the projections of the wave vectors are within the limits from k_x to $k_x + dk_x$, from k_y to $k_y + dk_y$, and from k_z to $k_z + dk_z$ is determined by the expression

$$dN_{k_x, k_y, k_z} = \frac{V}{\pi^3} dk_x\, dk_y\, dk_z \tag{1.47}$$

[compare with Eq. (1.41)].

The number of standing waves for which the magnitude of the wave vector ranges from k to $k + dk$ equals the number of points

* In this case, the initial phase of odd-numbered waves can be taken equal to zero, and the phase of waves with even numbers taken equal to π.

getting within the confines of one-eighth of a spherical layer of radius k and thickness dk (see Fig. 1.13). Consequently,

$$dN_k = \frac{V}{\pi^3} \cdot \frac{1}{8} 4\pi k^2 \, dk = V \frac{k^2 \, dk}{2\pi^3} \tag{1.48}$$

[compare with Eq. (1.42)].

Taking into account Eqs. (1.30) and (1.31), we get the number of standing waves whose frequencies are in the interval from ω to $\omega + d\omega$:

$$dN_\omega = V \frac{\omega^2 \, d\omega}{2\pi^2 v^3} \tag{1.49}$$

Equation (1.49) is proportional to the volume of the space V. We can therefore speak of the number of standing waves dn_ω per unit volume of the space. This number is

$$dn_\omega = \frac{\omega^2 \, d\omega}{2\pi^2 v^3} \tag{1.50}$$

In the following, we shall introduce into this expression a refinement due to the need of taking the possible kinds of polarization of the waves into account.

1.6. The Rayleigh-Jeans Formula

The British physicists Lord Rayleigh (John William Strutt, 1842-1919) and James Jeans (1877-1946) made an attempt to determine the equilibrium density of radiation $u(\omega, T)$ on the basis of the theorem of classical statistics on the uniform distribution of energy among degrees of freedom. They assumed that an energy equal to two halves of kT falls on an average to each electromagnetic oscillation—one half to the electrical and the other to the magnetic energy of the wave (we remind our reader that according to classical notions an energy equal to two halves of kT falls on an average to each vibrational degree of freedom).

Equilibrium radiation in a cavity is a system of standing waves. With no account taken of the possible kinds of polarization, the number of standing waves related to unit volume of a cavity is determined by Eq. (1.50), in which the velocity v must be assumed equal to c. Two electromagnetic waves of the same frequency differing in their direction of polarization (polarized in mutually perpendicular directions) can propagate in a given direction. To take this circumstance into account, we have to multiply Eq. (1.50) by two. The result is

$$dn_\omega = \frac{\omega^2 \, d\omega}{\pi^2 c^3} \tag{1.51}$$

As we have already noted, Rayleigh and Jeans, proceeding from the law of equal distribution of energy among degrees of freedom, ascribed an energy of $\langle \varepsilon \rangle$ equal to kT to each oscillation. Multiplying Eq. (1.51) by $\langle \varepsilon \rangle$, we get the energy density falling to the frequency interval $d\omega$:

$$u(\omega, T)\, d\omega = \langle \varepsilon \rangle\, dn_\omega = kT\, \frac{\omega^2}{\pi^2 c^3}\, d\omega$$

whence

$$u(\omega, T) = \frac{\omega^2}{\pi^2 c^3}\, kT \qquad (1.52)$$

Passing over from $u(\omega, T)$ to $f(\omega, T)$ according to Eq. (1.15), we get an expression for the emissivity of a blackbody:

$$f(\omega, T) = \frac{\omega^2}{4\pi^2 c^2}\, kT \qquad (1.53)$$

We must note that function (1.53) satisfies condition (1.18) obtained by Wien.

Expressions (1.52) and (1.53) are known as the **Rayleigh-Jeans formula.** It agrees with experimental data satisfactorily only for large wavelengths and sharply diverges from these data for small wavelengths (see Fig. 1.14 in which the solid line depicts an experimentally obtained curve, and the dash line depicts a curve constructed according to the Rayleigh-Jeans formula).

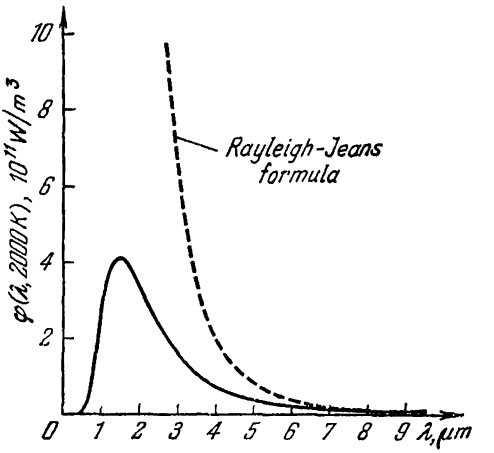


Fig. 1.14

Integration of Eq. (1.52) with respect to ω within the limits from 0 to ∞ gives an infinitely great value for the equilibrium energy density $u(T)$. This result, which has been named the **ultraviolet catastrophe**, also contradicts experimental data. Equilibrium between radiation and the body emitting it sets in at finite values of $u(T)$.

1.7. Planck's Formula

The derivation of the Rayleigh-Jeans formula is faultless from the classical viewpoint. Therefore, the failure of this formula to agree with experimental data pointed to the existence of laws that are incompatible with the notions of classical physics.

In 1900, the German physicist Max Planck (1858-1947) succeeded in finding a form of the function $u(\omega, T)$ that exactly corresponded to experimental results. For this purpose, he had to make an assumption absolutely alien to classical notions, namely, to assume that electromagnetic radiation is emitted in the form of separate portions of energy (quanta) whose magnitude is proportional to the frequency of radiation:

$$\varepsilon = \hbar\omega \tag{1.54}$$

The constant of proportionality $\hbar$ was subsequently named Planck's constant*. Its value determined experimentally** is

$$\hbar = 1.054 \times 10^{-34}\ \text{J}\cdot\text{s} = 1.054 \times 10^{-27}\ \text{erg}\cdot\text{s} = 0.659 \times 10^{-15}\ \text{eV}\cdot\text{s} \tag{1.55}$$

In mechanics, there is a quantity having the dimension "energy × × time" that is called **action**. Planck's constant is therefore sometimes called a **quantum of action**. It must be noted that the dimension of $\hbar$ coincides with that of the angular momentum.

If radiation is emitted in bundles or packets of $\hbar\omega$, then its energy ε_n must be a multiple of this quantity:

$$\varepsilon_n = n\hbar\omega \quad (n = 0, 1, 2, \ldots) \tag{1.56}$$

In a state of equilibrium, the distribution of oscillations by values of the energy must obey Boltzmann's law. According to Eq. (11.82) of Vol. I, p. 328, the probability P_n of the fact that the energy of oscillation of the frequency ω has the value ε_n is determined by the expression

$$P_n = \frac{N_n}{N} = \frac{\exp(-\varepsilon_n/kT)}{\sum\limits_n \exp(-\varepsilon_n/kT)} \tag{1.57}$$

(we have substituted N_n for N_i and ε_n for E_i).

Knowing the probability of various values of the oscillation energy, we can find the mean value of this energy $\langle\varepsilon\rangle$. According to Eq. (11.5) of Vol. I, p. 296,

$$\langle\varepsilon\rangle = \sum_n P_n\varepsilon_n$$

---

* Strictly speaking, the constant of proportionality h between ε and the frequency, $\varepsilon = h\nu$, is called Planck's constant. The constant $\hbar$ is Planck's constant h divided by 2π. The numerical value of Planck's constant is $h = 6.62 \times 10^{-34}\ \text{J}\cdot\text{s} = 6.62 \times 10^{-27}\ \text{erg}\cdot\text{s}$.

** Planck's constant is present in many physical relations, and in this connection it can be determined in various ways. The most accurate value is obtained from measurements of the short-wave boundary of the braking radiation (bremsstrahlung) X-ray spectrum (see Sec. 2.1).

Using Eqs. (1.56) and (1.57) for ε_n and P_n in this expression, we shall get the following formula for the mean value of the energy of radiation of frequency ω:

$$\langle \varepsilon \rangle = \frac{\sum_{n=0}^{\infty} n\hbar\omega \exp(-n\hbar\omega/kT)}{\sum_{n=0}^{\infty} \exp(-n\hbar\omega/kT)} \tag{1.58}$$

To perform the calculations, let us introduce the notation $\hbar\omega/kT = x$ and assume that x can change by taking on a continuous series of values. Equation (1.58) can therefore be written in the form

$$\langle \varepsilon \rangle = \hbar\omega \frac{\sum_{n=0}^{\infty} ne^{-nx}}{\sum_{n=0}^{\infty} e^{-nx}} = -\hbar\omega \frac{d}{dx} \ln \sum_{n=0}^{\infty} e^{-nx} \tag{1.59}$$

Inside the logarithm in Eq. (1.59) is the sum of the terms of an infinite geometrical progression with the first term equal to unity and the common ratio equal to e^{-x}. Since the denominator is less than unity, the progression will be a diminishing one, and according to the formula known from algebra

$$\sum_{n=0}^{\infty} e^{-nx} = \frac{1}{1-e^{-x}}$$

Introducing this value of the sum into Eq. (1.59) and differentiating, we obtain

$$\langle \varepsilon \rangle = -\hbar\omega \frac{d}{dx} \ln \frac{1}{1-e^{-x}} = \hbar\omega \frac{e^{-x}}{1-e^{-x}} = \frac{\hbar\omega}{e^{x}-1}$$

Now, replacing x with its value $\hbar\omega/kT$, we get a final expression for the mean energy of radiation of the frequency ω:

$$\langle \varepsilon \rangle = \frac{\hbar\omega}{\exp(\hbar\omega/kT)-1} \tag{1.60}$$

We must note that when $\hbar$ tends to zero, Eq. (1.60) transforms into the classical expression $\langle \varepsilon \rangle = kT$. We can convince ourselves in the truth of this statement by assuming that $\exp(\hbar\omega/kT) \approx \approx 1 + \hbar\omega/kT$, which is observed the more accurately, the smaller is $\hbar$. Thus, if the energy could take on a continuous series of values, its average value would equal kT.

Multiplying Eqs. (1.51) and (1.60), we find the density of the energy falling within the frequency interval $d\omega$:

$$u(\omega, T)\, d\omega = \frac{\hbar\omega}{\exp(\hbar\omega/kT)-1} \frac{\omega^2 d\omega}{\pi^2 c^3}$$

whence

$$u(\omega, T) = \frac{\hbar\omega^3}{\pi^2 c^3} \frac{1}{\exp(\hbar\omega/kT) - 1} \tag{1.61}$$

Using Eq. (1.15), we arrive at the formula

$$f(\omega, T) = \frac{\hbar\omega^3}{4\pi^2 c^2} \frac{1}{\exp(\hbar\omega/kT) - 1} \tag{1.62}$$

Equations (1.61) and (1.62) are called **Planck's formula.** This formula accurately agrees with experimental data throughout the entire interval of frequencies from 0 to ∞. Function (1.62) satisfies Wien's criterion (1.18). Provided that $\hbar\omega/kT \ll 1$ (small frequencies

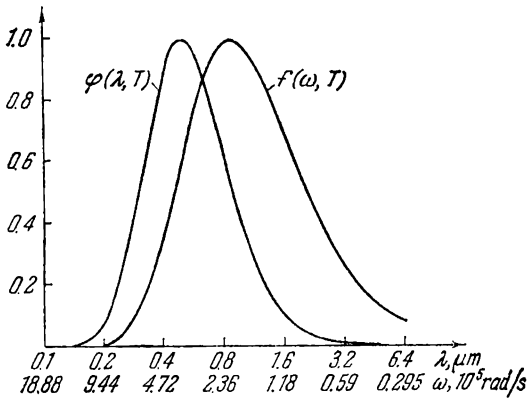


Fig. 1.15

or large wavelengths), the exponent $\exp(\hbar\omega/kT)$ may be assumed approximately equal to $1 + \hbar\omega/kT$. As a result, Planck's formula (1.61) or (1.62) transforms into the Rayleigh-Jeans formula [(1.52) or (1.53)]. This can also be seen from the fact that when this condition is satisfied, Eq. (1.60) approximately equals kT.

Transforming Eq. (1.62) in accordance with formula (1.10), we get

$$\varphi(\lambda, T) = \frac{4\pi^2\hbar c^2}{\lambda^5} \frac{1}{\exp(2\pi\hbar c/kT\lambda) - 1} \tag{1.63}$$

Figure 1.15 compares graphs of functions (1.62) and (1.63) plotted for the same temperature (5000 K). The logarithmic scales along the axis of abscissas have been chosen so that the values of λ and ω related by the expression $\lambda = 2\pi c/\omega$ have been superposed on one another. Examination of the figure reveals that the frequency ω_m corresponding to a maximum of $f(\omega, T)$ does not coincide with $2\pi c/\lambda_m$, where λ_m is the wavelength corresponding to the maximum of $\varphi(\lambda, T)$.

For the radiant emittance of a blackbody, we get the expression

$$R^* = \int_0^\infty f(\omega, T)\, d\omega = \int_0^\infty \frac{\hbar\omega^3}{4\pi^2c^2} \frac{d\omega}{\exp(\hbar\omega/kT) - 1}$$

Let us substitute the dimensionless variable $x = \hbar\omega/kT$ for ω. The substitution $\omega = (kT/\hbar)\, x$, $d\omega = (kT/\hbar)\, dx$ transforms the formula for R^* as follows:

$$R^* = \frac{\hbar}{4\pi^2c^2} \left(\frac{kT}{\hbar}\right)^4 \int_0^\infty \frac{x^3\, dx}{e^x - 1}$$

The definite integral in this expression can be calculated. It equals $\pi^4/15 \approx 6.5$. Introducing its value, we arrive at the Stefan-Boltzmann law:

$$R^* = \frac{\pi^2k^4}{60c^2\hbar^3} T^4 = \sigma T^4 \tag{1.64}$$

Substitution in this formula of the numerical values for k, c, and $\hbar$ gives the value of 5.6696×10^{-8} W/(m^2·K^4) for the Stefan-Boltzmann constant that agrees very well with the experimental value (1.17).

In concluding, let us find the value of the constant in Wien's displacement law (1.22). For this purpose, we shall differentiate function (1.63) with respect to λ and equate the expression obtained to zero:

$$\frac{d\varphi(\lambda, T)}{d\lambda} = \frac{4\pi^2\hbar c^2 [(2\pi\hbar c/kT\lambda)\, e^{2\pi\hbar c/kT\lambda} - 5\,(e^{2\pi\hbar c/kT\lambda} - 1)]}{\lambda^6\,(e^{2\pi\hbar c/kT\lambda} - 1)^2} = 0$$

The values of $\lambda = 0$ and $\lambda = \infty$ satisfying this equation correspond to minima of the function $\varphi(\lambda, T)$. The value of λ_m at which the function reaches a maximum converts the expression in brackets in the numerator to zero. Introducing the notation $2\pi\hbar c/kT\lambda_m = x$, we get the equation

$$xe^x - 5\,(e^x - 1) = 0$$

The solution* of this transcendental equation gives $x = 4.965$. Hence, $2\pi\hbar c/kT\lambda_m = 4.965$, whence

$$T\lambda_m = \frac{2\pi\hbar c}{4.965k} = b \tag{1.65}$$

Substitution of numerical values for $\hbar$, c, and k gives a value for b that coincides with the experimentally obtained one (1.23).

Thus, Planck's formula gives an exhaustive description of equilibrium thermal radiation.

* The solution can be found by the method of consecutive approximations. Noting that $e^5 \gg 1$, we can in the first approximation write the equation in the form $xe^x - 5e^x \approx 0$, whence $x \approx 5$. We get the second approximation from the equation $xe^5 - 5\,(e^5 - 1) = 0$, etc.

CHAPTER 2 PHOTONS

2.1. Bremsstrahlung

We learned in the preceding chapter that to explain the properties of thermal radiation, it was necessary to introduce the notion of electromagnetic radiation being emitted in portions of $\hbar\omega$. The quantum nature of radiation is also confirmed by the existence of a **short wavelength limit** of the bremsstrahlung X-ray spectrum.

X-rays are produced when solid targets are bombarded with fast electrons. An X-ray tube (Fig. 2.1) is an evacuated bulb with several electrodes. Cathode C heated by a current is the source of free electrons produced owing to thermoelectronic emission (see Sec. 9.2). Cylindrical electrode E is intended for focussing the electron beam. Anode A, also called an anticathode, is the target. It is made from heavy metals (W, Cu, Pt, etc.). The electrons are accelerated by the high voltage set up between the cathode and the anticathode. Virtually the entire energy of the electrons is liberated on the anticathode in the form of heat (only from 1 to 3% of the energy is transformed into radiation). This is why the anticathode has to be intensively cooled in powerful tubes. For this purpose, channels are made in the body of the anticathode for the circulation of a cooling liquid (water or oil).

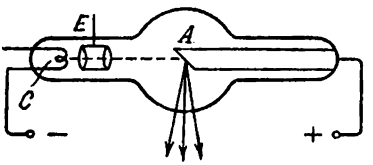


Fig. 2.1

If the voltage U is applied between the cathode and the anticathode, the electrons are accelerated to the energy eU. Upon getting into the substance of the anticathode, the electrons experience strong deceleration and become a source of electromagnetic waves. The radiant power P is proportional to the square of the charge of an electron and the square of its acceleration:

$$P \propto e^2 a^2$$

[see Eq. (15.47) of Vol. II, p. 317].

Let us assume that the acceleration of an electron a remains constant during the entire duration of deceleration τ. The radiant power will therefore also be constant, and during the deceleration

period an electron emits the energy

$$E = P\tau \propto e^2 a^2 \tau = \frac{e^2 v_0^2}{\tau}$$

where v_0 is the initial velocity of the electron.

The result obtained shows that appreciable radiation can be observed only upon sharp deceleration of the fast electrons. A voltage up to 50 kV is fed to X-ray tubes. Upon passing through such a potential difference, an electron acquires a velocity of $0.4c$. In a betatron (see Sec. 10.5 of Vol. II, p. 224 et seq.) electrons can be accelerated to an energy of 50 MeV. The velocity of the electrons at such an energy is $0.99995c$. By directing a beam of electrons accelerated in a betatron onto a solid target, we can get X-rays of a very small wavelength. The smaller the wavelength, the less are the rays absorbed in a substance. For this reason, X-rays obtained in a betatron have an especially high penetrability.

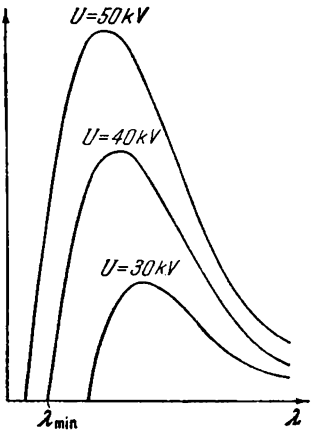


Fig. 2.2

At a sufficiently high velocity of the electrons, in addition to **bremsstrahlung—braking radiation** (i.e. radiation produced by deceleration of the electrons), there is also produced **characteristic radiation** (due to excitation of the internal electron shells of the anticathode atoms). This radiation is treated in Sec. 5.11. Now we shall be interested only in bremsstrahlung. According to classical electrodynamics, when an electron is decelerated, waves of all lengths—from zero to infinity—should be produced. The wavelength corresponding to the maximum radiant power should diminish with an increasing velocity of the electrons, i.e. with an increasing voltage U across the tube. Figure 2.2 gives experimental curves showing how the power of bremsstrahlung is distributed by wavelengths and obtained for different values of U. Inspection of the figure shows that the conclusions of theory are mainly confirmed experimentally. There is a fundamental deviation, however, from the requirements of classical electrodynamics. It consists in that the curves of power distribution do not pass to the origin of coordinates, but terminate at finite values of the wavelength λ_{min}. It has been established experimentally that the short wavelength limit of the bremsstrahlung spectrum λ_{min} is associated with the accelerating voltage U by the

relation

$$\lambda_{\min} = \frac{12\,390}{U} \tag{2.1}$$

where $\lambda_{\min}$ is in angstroms, and U in volts.

The existence of the short wavelength limit directly follows from the quantum nature of radiation. Indeed, if radiation is produced at the expense of the energy lost by an electron when it decelerates, then the magnitude of a quantum $\hbar\omega$ cannot exceed the energy of an electron eU:

$$\hbar\omega \leqslant eU$$

Hence, we find that the frequency of radiation cannot exceed the value $\omega_{\max} = eU/\hbar$ and, consequently, the wavelength cannot be smaller than the value

$$\lambda_{\min} = \frac{2\pi c}{\omega_{\max}} = \frac{(2\pi\hbar c/e)}{U} \tag{2.2}$$

We have thus arrived at empirical equation (2.1). The value of $\hbar$ found by comparing Eqs. (2.1) and (2.2) agrees quite well with the values determined by other methods. Of all the ways of finding $\hbar$, the one based on measuring the short-wave boundary of the bremsstrahlung spectrum is considered to be the most accurate.

2.2. The Photoelectric Effect

The **photoelectric effect** is the name given to the emission of electrons by a substance under the action of light. This phenomenon was discovered in 1887 by the German physicist Heinrich Hertz. He noted that the jumping of a spark between the electrodes of a discharger is considerably facilitated when one of the electrodes is illuminated with ultraviolet rays.

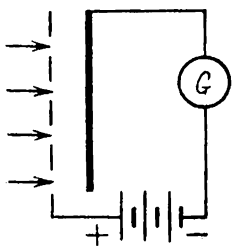


Fig. 2.3

In 1888-1889, the Russian physicist Aleksandr Stoletov systematically studied the photoelectric effect with the aid of the arrangement shown schematically in Fig. 2.3. The capacitor formed by a wire screen and a solid plate was connected in series with galvanometer G in the circuit of a battery. The light passing through a screen fell on the solid plate. As a result, a current was set up in the circuit that was registered by the galvanometer. Stoletov arrived at the following conclusions as a result of his experiments: (1) ultraviolet rays have the greatest action; (2) the current grows with increasing illumination of the plate; and (3) the charges emitted under the action of light have a negative sign.

Ten years later (in 1898), the German physicist Philipp Lenard (1862-1947) and the British physicist Joseph J. Thomson (1856-1940) measured the specific charge of the particles emitted under the action of light and found that these particles are electrons.

Lenard and other investigators improved Stoletov's arrangement by putting the electrodes into an evacuated bulb (Fig. 2.4). The light penetrating through quartz* window Q illuminates cathode C made from the substance being investigated. The electrons emitted as a result of the photoelectric effect move under the action of the electric field to anode A. As a result, a photocurrent measured by

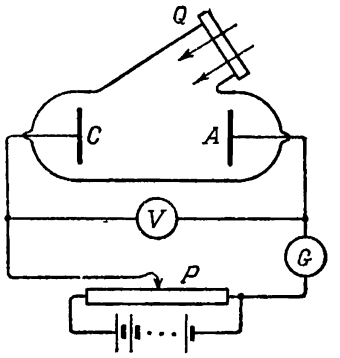


Fig. 2.4

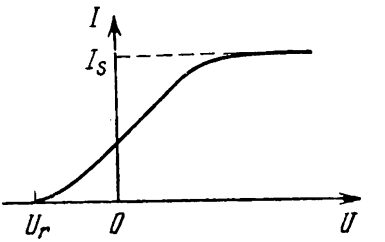


Fig. 2.5

means of galvanometer G flows through the circuit of the arrangement. The voltage between the anode and the cathode can be changed with the aid of potentiometer P.

The volt-ampere characteristic (i.e. the curve showing how the photocurrent I depends on the voltage U across the electrodes) obtained in such an arrangement is shown in Fig. 2.5. The characteristic is naturally read at a constant light flux Φ. A glance at the curve shows that at a certain not very high voltage the photocurrent reaches saturation—all the electrons emitted by the cathode reach the anode. Hence, the saturation current I_s is determined by the number of electrons emitted by the cathode in unit time under the action of light.

The gentle slope of the curve indicates that the electrons fly out of the cathode with velocities different in magnitude. A fraction of the electrons corresponding to the current when $U = 0$ have velocities sufficient for them to reach the anode "independently" without the aid of the accelerating field. For the current to vanish, the **retarding voltage** U_r must be applied. At this voltage, none of the

* Unlike ordinary glass, quartz transmits ultraviolet rays.

electrons, even those having the maximum value of the velocity v_m when flying out of the cathode, succeed in overcoming the retarding field and reaching the anode. We can therefore write that

$$\frac{1}{2}mv_m^2 = eU_r \tag{2.3}$$

where m is the mass of an electron. Thus, by measuring the retarding voltage U_r, we can find the maximum value of the velocity of photoelectrons.

By 1905, it was established that the maximum velocity of photoelectrons does not depend on the intensity of light, but depends only on its frequency—a growth in frequency leads to an increase in velocity. The experimentally established relations did not fit into the framework of the classical notions. For example, according to classical conceptions, the velocity of photoelectrons ought to grow with the amplitude, and, consequently, with the intensity of the electromagnetic wave.

In 1905, the German physicist Albert Einstein showed that all the laws of the photoelectric effect can readily be explained if it is assumed that light is absorbed in the same portions $\hbar\omega$ (quanta) in which, according to Planck's assumption, it is emitted. Einstein postulated that the energy received by an electron is supplied to it in the form of a quantum $\hbar\omega$, which it assimilates completely. Part of this energy, equal to the work function A*, goes to allow the electron to leave the body. If an electron is freed not at the very surface, but at a certain depth, then part of the energy equal to E' may be lost owing to chance collisions in the substance. The remaining energy is the kinetic energy E_k of the electron leaving the substance. The energy E_k will be maximum when $E' = 0$. In this case the equation

$$\hbar\omega = \frac{1}{2}mv_m^2 + A \tag{2.4}$$

known as **Einstein's formula** must be obeyed.

The photoelectric effect and the work function greatly depend on the state of the surface of a metal (in particular, on the oxides and adsorbed substances on it). Therefore, for a long time, Einstein's formula could not be verified with sufficient accuracy. In 1916, the American scientist Robert Millikan designed an apparatus in which the surfaces being studied were cleaned in a vacuum, after which the work function was measured and the dependence of the maximum kinetic energy of photoelectrons on the frequency of the incident light was determined (this energy was found by measuring the retarding potential U_r). The results agreed completely with formula (2.4).

* The smallest amount of energy that must be imparted to an electron in order to remove it from inside a solid or liquid body into a vacuum is known as the **work function of the body** (see Sec. 9.1).

Introducing the measured values of A and $\frac{1}{2}mv_m^2$ (for a given ω) into formula (2.4), Millikan determined the value of Planck's constant $\hbar$. It was found to coincide with the values determined from the spectral distribution of equilibrium thermal radiation and from the short-wave limit of the bremsstrahlung spectrum.

The method of studying the photoelectric effect was further improved in 1928 by the Soviet physicists P. Lukirsky and S. Prilezhaev, who designed an apparatus in the form of a spherical capacitor. The silver-coated walls of a spherical glass bulb were the anode in their apparatus. The cathode in the form of a sphere was placed at the centre of the bulb. Such a shape of the electrodes gives a steeper volt-ampere characteristic, which makes it possible to improve the accuracy of determining the retarding potential.

Examination of formula (2.4) reveals that when the work function A exceeds the energy of a quantum $\hbar\omega$, the electrons cannot leave the metal. Hence, for the photoelectric effect to appear, the condition $\hbar\omega \geqslant A$ or

$$\omega \geqslant \omega_0 = \frac{A}{\hbar} \tag{2.5}$$

must be satisfied. The condition for the wavelength, accordingly, is

$$\lambda \leqslant \lambda_0 = \frac{2\pi\hbar c}{A} \tag{2.6}$$

The frequency ω_0 or the wavelength λ_0 is called the **photoelectric threshold.**

The number of electrons freed owing to the photoelectric effect should be proportional to the number of light quanta falling on the relevant surface. At the same time, the light flux Φ is determined by the number of light quanta falling on the surface in unit time. Accordingly, the saturation current I_s must be proportional to the incident light flux:

$$I_s \propto \Phi \tag{2.7}$$

This relation is also confirmed experimentally. It must be noted that only a small part of the quanta transmit their energy to the photoelectrons. The energy of the remaining quanta goes to heat the substance absorbing the light.

In the phenomenon of the photoelectric effect considered above, an electron receives energy from only a single photon. Such processes are called **single-photon** ones. The invention of lasers was attended by the obtaining of light beam powers unachievable before that time. This made it possible to carry out **multiple-photon** processes. In particular, the **multiple-photon photoelectric effect** was observed. In this process, an electron flying out from a metal receives energy not from one, but from N photons ($N = 2, 3, 4, 5$).

Einstein's formula can be written as follows for the multiple-photon photoelectric effect:

$$N\hbar\omega = \frac{1}{2}mv_m^2 + A \tag{2.8}$$

The photoelectric threshold shifts accordingly in the direction of longer waves (λ_0 grows N times). Formula (2.7) for the N-photon effect has the form

$$I_N \propto \Phi^N \tag{2.9}$$

Apart from the **external photoelectric effect** (which is generally called simply the photoelectric effect) treated in this section, there also exists the **internal photoelectric effect** observed in dielectrics and semiconductors. It will be discussed in Sec. 9.6.

2.3. Bothe's Experiment. Photons

To explain the distribution of energy in the spectrum of equilibrium thermal radiation, it is sufficient, as Planck showed, to assume that light is only emitted in portions of $\hbar\omega$. To explain the photoelectric effect, it is sufficient to assume that light is absorbed in the same portions. Einstein, however, went considerably further. He advanced the hypothesis that light also propagates in the form of discrete particles initially called **light quanta**. These particles were later named **photons** (this term was introduced in 1926).

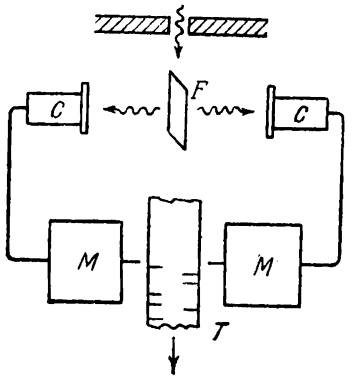


Fig. 2.6

The most direct confirmation of Einstein's hypothesis was given by an experiment run by the German physicist Walther Bothe (1891-1957). Thin metal foil F (Fig. 2.6) was placed between two gas-discharge counters C (see Sec. 12.3 of Vol.II, p. 240 et seq.). The foil was illuminated with a weak beam of X-rays under whose action it itself became a source of X-rays (this phenomenon is known as **X-ray fluorescence**). Owing to the low intensity of the primary beam, the number of quanta emitted by the foil was not great. When struck by X-rays, the counter operated and actuated special mechanism M that made a mark on moving tape T. If the emitted energy propagated uniformly in all directions as follows from wave notions, both counters ought to operate simultaneously, and the marks on the tape would be opposite one another. Actually, however, an absolutely chaotic arrangement of the marks is observed.

The only explanation is that in individual emission events, particles of light appear that fly first in one, and then in another direction.

Thus, the existence of special particles of light—**photons**—was proved experimentally. The energy of a photon is determined by its frequency:

$$E = \hbar\omega \tag{2.10}$$

We invite our reader to convince himself that the energy of a photon of $\hbar\omega = 2.5$ eV corresponds to a wavelength of $\lambda = 5000$ Å (the green part of the spectrum); when $\lambda = 1$ Å, we have $\hbar\omega = 12.5$ keV.

An electromagnetic wave has a momentum (see Sec. 15.5 of Vol. II, p. 312 et seq.). Accordingly, a photon must also have a momentum. To find the momentum of a photon, let us use the relations of the theory of relativity. We shall consider two reference frames K and K' moving relative to each other with the velocity v_0. We shall direct the axes x and x' along v_0. Assume that a photon flies in the direction of these axes. The energy of the photon in the frames K and K' is $\hbar\omega$ and $\hbar\omega'$, respectively. The frequencies ω and ω' are related by the expression

$$\omega' = \omega \frac{1 - v_0/c}{\sqrt{1 - v_0^2/c^2}}$$

(see Sec. 21.4 of Vol. II, p. 481 et seq.). Hence,

$$E' = E \frac{1 - v_0/c}{\sqrt{1 - v_0^2/c^2}} \tag{2.11}$$

Let $\mathbf{p}$ stand for the momentum of a photon in the frame K, and $\mathbf{p}'$ for the momentum of a photon in the frame K'. It follows from considerations of symmetry that the momentum of a photon must be directed along the x-axis. Therefore, $p_x = p$, and $p'_x = p'$. In passing from one reference frame to another, the energy and momentum are transformed by means of the formula

$$E' = \frac{E - v_0 p_x}{\sqrt{1 - v_0^2/c^2}} \tag{2.12}$$

[see the last of formulas (8.49) of Vol. I, p. 244; we have replaced β with its value v_0/c and written the formula for the reverse transformation; in this connection we have changed the sign of $v_0 p_x$]. In the case we are considering, we may substitute p for p_x in Eq. (2.12).

A comparison of Eqs. (2.11) and (2.12) shows that

$$E\left(1 - \frac{v_0}{c}\right) = E - v_0 p$$

(we have written p instead of p_x). Hence

$$p = \frac{E}{c} = \frac{\hbar\omega}{c} \tag{2.13}$$

We showed in Sec. 8.10 of Vol. I, p. 247, that such a relation between the momentum and the energy is possible only for particles having a zero rest mass and travelling with the speed c. It thus follows from the quantum relation $E = \hbar\omega$ and the general principles of the theory of relativity that (1) the rest mass of a photon is zero, and (2) a photon always travels with the speed c. This signifies that a photon is a special kind of particle differing from particles such as an electron and a proton that can exist when travelling at speeds less than c and even when at rest.

Replacing the frequency ω in Eq. (2.13) with the wavelength λ, we get the following expression for the momentum of a photon:

$$p = \frac{\hbar 2\pi}{\lambda} = \hbar k \tag{2.14}$$

(k is the wave number). A photon flies in the direction of propagation of the relevant electromagnetic wave. Therefore, the directions of the momentum $\mathbf{p}$ and the wave vector $\mathbf{k}$ coincide. Equation (2.14) can therefore be written in the vector form:

$$\mathbf{p} = \hbar\mathbf{k} \tag{2.15}$$

Assume that a flux of photons falls on a light-absorbing surface of a wall and that the photons are flying along a normal to the surface. If the density of the photons is n, then nc photons fall on unit surface area in unit time. Each photon when absorbed imparts the momentum $p = E/c$ to the wall. Multiplying p by nc, we get the momentum imparted to unit surface in unit time, i.e. the pressure $\mathcal{P}$ of the light on the wall

$$\mathcal{P} = \frac{E}{c} \cdot nc = En$$

The product En equals the energy of the photons confined in unit volume, i.e. the density w of electromagnetic energy. We have thus arrived at the formula $\mathcal{P} = w$, which coincides with the expression for the pressure obtained from the electromagnetic theory [see Eq. (15.41) of Vol. II, p. 314]. Upon reflection from a wall, a photon imparts the momentum $2p$ to it. Therefore, the pressure for a reflecting surface will be $2w$.

On the basis of the notion of an electromagnetic field as a collection of photons, it is a simple matter to obtain a relation between the emissivity of a blackbody and the equilibrium density of radiation. Assume that a unit volume of a cavity filled with equilibrium radiation contains dn_ω photons whose frequency ranges from ω to $\omega + d\omega$. The density of the energy falling to the same interval of frequencies will therefore be

$$du_\omega = u(\omega, T)\, d\omega = \hbar\omega\, dn_\omega \tag{2.16}$$

Like the molecules of a gas, the photons fly inside the cavity in all directions. Using Eq. (11.23) of Vol. I, p. 303, we get the value $\frac{1}{4} c\, dn_\omega$ for the number of photons colliding with a unit surface area in unit time. If the wall is a blackbody, it will absorb all these photons and, consequently, will receive energy equal to $\frac{1}{4} \hbar\omega c\, dn_\omega$. In equilibrium, the blackbody wall will emit the same energy. Thus,

$$f(\omega, T)\, d\omega = \frac{1}{4} \hbar\omega c\, dn_\omega \tag{2.17}$$

A comparison of Eqs. (2.16) and (2.17) shows that

$$f(\omega, T) = \frac{c}{4} u(\omega, T) \tag{2.18}$$

[compare with Eq. (1.15)].

We have treated a number of phenomena in this chapter in which light behaves like a flux of particles (photons). One must never forget, however, that phenomena such as the interference and diffraction of light can be explained only on the basis of wave notions. Thus, light displays **corpuscular-wave duality**: in some phenomena its wave nature manifests itself, and it behaves like an electromagnetic wave, whereas in other phenomena the corpuscular nature of light manifests itself, and it behaves like a flux of photons. We shall see in Sec. 4.1 that not only light particles, but also the particles of a substance (electrons, protons, atoms, etc.) have corpuscular-wave duality.

Let us find the relation between the wave and the corpuscular pictures. We can obtain an answer to this question by considering the illumination of a surface from both viewpoints. According to wave notions, the illumination at a point of a surface is proportional to the square of the amplitude of the light wave. From the corpuscular viewpoint, the illumination is proportional to the density of the photon flux. Consequently, direct proportionality exists between the square of the amplitude of a light wave and the density of a photon flux. Energy and momentum are carried by photons. Energy is liberated at the point of a surface onto which a photon falls. The square of the amplitude of a wave determines the probability of a photon falling on a given point of a surface. More exactly, the probability of the fact that a photon will be detected within the limits of the volume dV containing the point of space being considered is determined by the expression

$$dP = \chi A^2\, dV$$

where χ = constant of proportionality
A = amplitude of a light wave.

It follows from the above that the distribution of photons over a surface on which light is falling must have a statistical nature. The uniform illumination observed experimentally is due to the fact that the density of a photon flux is usually very high. For example, at an illumination of 50 lx (such an illumination is needed for the eyes not to become tired when reading) and a wavelength of 5500 Å, about 2×10^{13} photons fall on one square centimetre of a surface in one second. The relative fluctuation* is inversely proportional to the square root of the number of particles [see formula (11.89) of Vol. I, p. 335]. Hence, at this value of the photon flux, the fluctuations are negligible, and the surface appears to be illuminated uniformly.

Fluctuations of weak light fluxes were detected by the Soviet physicist Sergei Vavilov (1891-1951) and his collaborators. They found that in the region of its greatest sensitivity ($\lambda = 5550$ Å) the human eye begins to react to light when about 200 photons fall on the pupil a second. At such an intensity, Vavilov observed fluctuations of the light flux having a clearly expressed statistical nature. True, it must be borne in mind that the fluctuations of the perception of light observed in Vavilov's experiments were due not only to fluctuations of the light flux, but also to the fluctuations associated with the physiological processes occuring in the eye.

2.4. The Compton Effect

The corpuscular properties of light manifest themselves especially clearly in a phenomenon that was named the **Compton effect.** In 1923, the American physicist Arthur Compton (1892-1962), investigating the scattering of X-rays by different substances, discovered that the scattered rays in addition to radiation of the initial wavelength λ contain also rays of a greater wavelength λ'. The difference $\Delta\lambda = \lambda' - \lambda$ was found to depend only on the angle θ made by the direction of the scattered radiation with that of the initial beam. The value of $\Delta\lambda$ does not depend on the wavelength λ and on the nature of the scattering material.

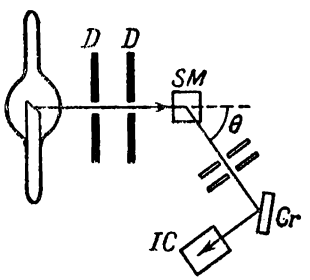


Fig. 2.7

Compton's experiment is shown schematically in Fig. 2.7. A narrow beam of monochromatic (characteristic) X-ray radiation separated by diaphragms *D* was directed onto scattering material *SM*.

* We remind our reader that by relative fluctuations are meant the relative deviations of statistical quantities from their mean value.

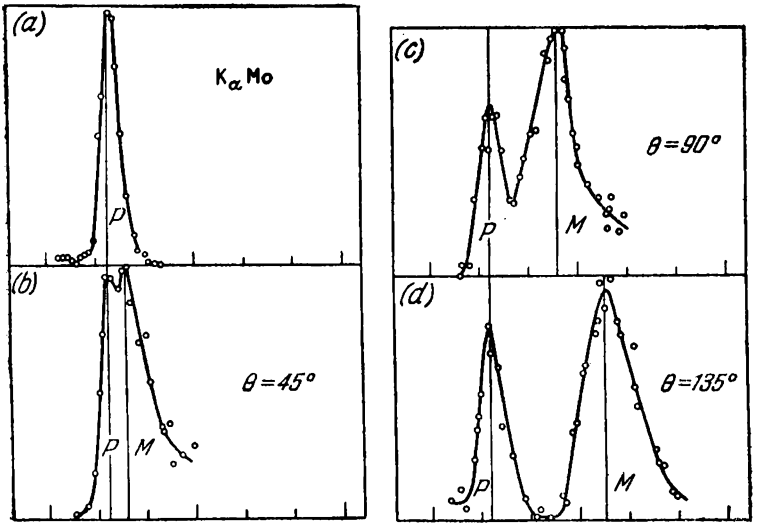

(a)
K_α Mo
P
(b)
θ=45°
P
M
(c)
θ=90°
P
M
(d)
θ=135°
P
M

Fig. 2.8

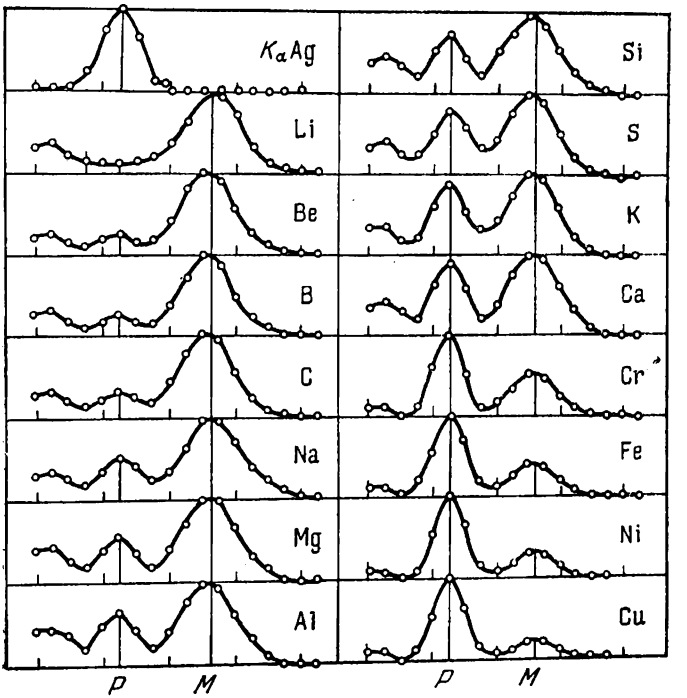

K_α Ag
Li
Be
B
C
Na
Mg
Al
Si
S
K
Ca
Cr
Fe
Ni
Cu
P
M
P
M

Fig. 2.9

The spectral composition of the scattered radiation was studied with the aid of an X-ray spectrograph consisting of crystal *Cr* and ionization chamber *IC*.

Figure 2.8 gives the results of studying the scattering of monochromatic X-rays (the line K_α* of molybdenum) on graphite. Curve *a* characterizes the primary radiation. The remaining curves relate to different scattering angles θ whose values are indicated in the figure. The intensity of radiation is laid off along the axis of ordinates, and the wavelength along the axis of abscissas.

Figure 2.9 shows how the relation between the intensities of the shifted *M* and unshifted *P* components depends on the atomic number of the scattering substance. The top curve in the left-hand column characterizes the primary radiation (the line K_α of silver). In scattering by substances with a low atomic number (Li, Be, B), virtually all the scattered radiation has a shifted wavelength. With an increase in the atomic number, a greater and greater part of the radiation is scattered without a change in the wavelength.

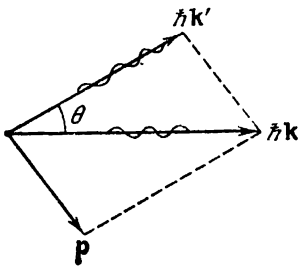


Fig. 2.10

All the features of the Compton effect can be explained by considering scattering as a process of elastic collision of the X-ray photons with practically free electrons. Those electrons may be considered free that are bound weakest to their atoms and whose binding energy is appreciably smaller than the energy which a photon can transmit to an electron when they collide**.

Assume that a photon having the energy $\hbar\omega$ and the momentum $\hbar\mathbf{k}$ falls on a free electron initially at rest (Fig. 2.10). The energy of the electron before the collision was mc^2 (here m is the rest mass of an electron), and its momentum was zero. After the collision, the electron will have the momentum $\mathbf{p}$ and an energy equal to $c\sqrt{p^2 + m^2c^2}$ [see Eq. (8.42) of Vol. I, p. 242]. The energy and momentum of the photon will also change and become equal to $\hbar\omega'$ and $\hbar\mathbf{k}'$. Two equations follow from the laws of energy and momentum conservation, namely,

$$\hbar\omega + mc^2 = \hbar\omega' + c\sqrt{p^2 + m^2c^2} \tag{2.19}$$

$$\hbar\mathbf{k} = \mathbf{p} + \hbar\mathbf{k}' \tag{2.20}$$

---

* See Sec. 5.11.

** In an elastic collision, a photon cannot transmit all its energy to an electron (or another particle). Such a process would violate the laws of conservation of energy and momentum.

Let us divide the first equation by c and write it in the form

$$\sqrt{p^2+m^2c^2} = \hbar(k-k')+mc$$

($\omega/c = k$). Squaring yields

$$p^2 = \hbar^2(k^2+k'^2-2kk')+2\hbar mc(k-k') \tag{2.21}$$

It can be seen from Eq. (2.20) that

$$p^2 = \hbar^2(\mathbf{k}-\mathbf{k}')^2 = \hbar^2(k^2+k'^2-2kk'\cos\theta) \tag{2.22}$$

(θ is the angle between the vectors $\mathbf{k}$ and $\mathbf{k}'$; see Fig. 2.10).

We find from a comparison of Eqs. (2.21) and (2.22) that

$$mc(k-k') = \hbar kk'(1-\cos\theta)$$

Multiplication of this equation by 2π and division by $mckk'$ yield

$$\frac{2\pi}{k'}-\frac{2\pi}{k} = \frac{2\pi\hbar}{mc}(1-\cos\theta)$$

Finally, taking into account that $2\pi/k = \lambda$, we arrive at the formula

$$\Delta\lambda = \lambda'-\lambda = \lambda_C(1-\cos\theta) \tag{2.23}$$

where

$$\lambda_C = \frac{2\pi\hbar}{mc} \tag{2.24}$$

The quantity λ_C determined by Eq. (2.24) is called the **Compton wavelength*** of the particle whose mass m we have in mind. In the case we are considering, λ_C is the Compton wavelength of an electron. Substituting for $\hbar$, m, and c in Eq. (2.24) their values, we get the following value for λ_C of an electron:

$$\lambda_C = 0.0243 \text{ Å} \tag{2.25}$$

($\lambda\!\!\!^{-}_C = 0.003\,86$ Å).

The results of measurements by Compton and of subsequent measurements are in complete agreement with Eq. (2.23) if we use in it the value of λ_C given by Eq. (2.25).

When photons are scattered on electrons whose bond to the atom is strong, the energy and momentum are exchanged with the atom as a whole. Since the mass of an atom is much greater than that of an electron, the Compton shift in this case is negligible, and λ' practically coincides with λ. An increase in the atomic number is attended by a growth in the relative number of electrons with a strong bond, and this is why the shifted line is weaker (see Fig. 2.9).

* The quantity

$$\lambda\!\!\!^{-}_C = \frac{\hbar}{mc}$$

is also known as the Compton wavelength.

PART II ATOMIC PHYSICS

CHAPTER 3 THE BOHR THEORY OF THE ATOM

3.1. Regularities in Atomic Spectra

The radiation of atoms that do not interact with one another consists of separate spectral lines. The emission spectrum of atoms is accordingly called a **line spectrum.** Figure 3.1 shows an emission spectrum of mercury vapour. The spectra of other atoms have the same nature.

The studying of atomic spectra served as a key to cognition of the structure of atoms. It was noted first of all that the lines in the spectra of atoms are arranged not chaotically, but are combined into groups or, as they are called, **series of lines.** This is revealed most clearly in the spectrum of the simplest atom—hydrogen. Figure 3.2 shows a part of the spectrum of atomic hydrogen in the visible and near ultraviolet region. The symbols H_α, H_β, H_γ, and H_δ designate the visible lines, and H_∞ shows the limit of the series (see below). The lines are evidently arranged in a definite order. The distance between the lines regularly diminishes upon passing from longer waves to shorter ones.

In 1885, the Swiss physicist Johann Balmer (1825-1898) discovered that the wavelengths of this series of hydrogen lines can be accurately represented by the formula

$$\lambda = \lambda_0 \frac{n^2}{n^2 - 4} \tag{3.1}$$

where λ_0 = constant

n = integer taking on values of 3, 4, 5, etc.

If we pass over from the wavelength to the frequency in Eq. (3.1), we get the formula

$$\omega = R\left(\frac{1}{2^2} - \frac{1}{n^2}\right) \qquad (n = 3, 4, 5, \ldots) \tag{3.2}$$

where R is a constant called the **Rydberg constant** in honour of the Swedish spectroscopist Johannes Rydberg (1854-1919). It equals

$$R = 2.07 \times 10^{16} \text{ rad/s} \tag{3.3}$$

Formula (3.2) is known as the **Balmer formula***, and the corresponding series of spectral lines of the hydrogen atom is known as the **Balmer series.** Further investigations showed that there are some other series in the hydrogen spectrum. The extreme ultraviolet part of the spectrum contains the Lyman series. The remaining series are in the infrared region. The lines of these series can be represented in the form of formulas similar to formula (3.2):

Lyman series $\omega = R\left(\frac{1}{1^2} - \frac{1}{n^2}\right)$

$(n = 2, 3, 4, \ldots)$

Paschen series $\omega = R\left(\frac{1}{3^2} - \frac{1}{n^2}\right)$

$(n = 4, 5, 6, \ldots)$

Brackett series $\omega = R\left(\frac{1}{4^2} - \frac{1}{n^2}\right)$

$(n = 5, 6, 7, \ldots)$

Pfund series $\omega = R\left(\frac{1}{5^2} - \frac{1}{n^2}\right)$

$(n = 6, 7, 8, \ldots)$

The frequencies of all the hydrogen atom spectrum lines can be represented by a single formula:

$$\omega = R\left(\frac{1}{m^2} - \frac{1}{n^2}\right) \qquad (3.4)$$

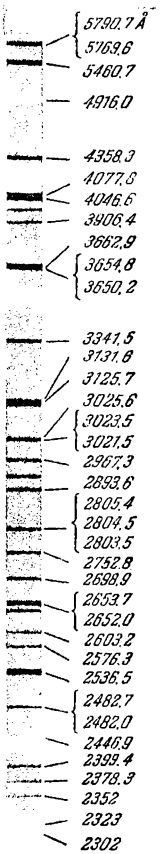


Fig. 3.1

* It is customary practice in spectroscopy to characterize spectrum lines not by the frequency, but by the quantity

$$\nu' = \frac{1}{\lambda} = \frac{\omega}{2\pi c}$$

that is the reciprocal of the wavelength and is called the wave number (do not confuse it with the wave number $k = 2\pi/\lambda = \omega/c$). The Balmer formula written for the wave number has the same form as Eq. (3.2):

$$\nu' = R\left(\frac{1}{2^2} - \frac{1}{n^2}\right) \quad (n = 3, 4, 5, \ldots)$$

The Rydberg constant in this case has the value

$$R = 109\ 737.309 \pm 0.012\ \text{cm}^{-1}$$

The number of authentic significant digits characterizes the accuracy of measurements achieved in spectroscopy. The value of the constant in Eq. (3.3) has been rounded off to the third digit.

where m has the value of 1 for the Lyman series, of 2 for the Balmer series, etc. At a given m, the number n takes on all integral values beginning from $m + 1$. Equation (3.4) is called the **generalized Balmer formula.**

When n grows, the frequency of the lines in each series tends to a limit value R/m^2 called **the series limit** (in Fig. 3.2 the symbol H_∞ indicates the limit of the Balmer series).

Let us take a number of values of the expression $T(n) = R/n^2$:

$$\frac{R}{1^2}, \frac{R}{2^2}, \frac{R}{3^2}, \ldots \tag{3.5}$$

The frequency of any hydrogen spectrum line can be represented in the form of the difference between two numbers of series (3.5). These

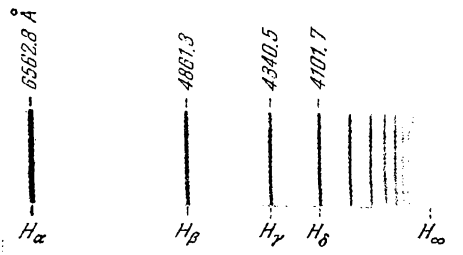


Fig. 3.2

numbers are called **spectral terms** or simply **terms.** For example, the frequency of the first line of the Balmer series is $T(2) - T(3)$, and of the second line of the Pfund series is $T(5) - T(7)$.

Studying of the spectra of other atoms showed that the frequencies of the lines in this case too can be represented as the differences between two terms:

$$\omega = T_1(m) - T_2(n) \tag{3.6}$$

But the term $T(n)$ usually has a more complicated form than for the hydrogen atom. In addition, the first and second terms of formula (3.6) are taken from different series of spectral terms.

3.2. The Thomson Model of the Atom

According to classical notions, an atom could emit a monochromatic wave (i.e. a spectral line) when an electron in the emitting atom performs harmonic oscillations, and, consequently, is retained near

its equilibrium position by a quasi-elastic force of the kind $F = -kr$, where r is the deviation of the electron from its equilibrium position. In 1903, J. J. Thomson suggested a model of an atom according to which an atom is a sphere uniformly filled with positive electricity, and there is an electron inside the sphere (Fig. 3.3). The total positive charge of the sphere equals the charge of an electron, so that the atom as a whole is neutral.

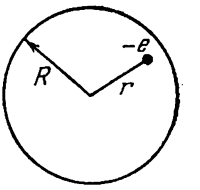


Fig. 3.3

The strength of the field inside a uniformly charged sphere is determined by the expression*

$$E(r) = \frac{e}{R^3} r \quad (0 \leqslant r \leqslant R)$$

where e is the charge of the sphere and R is its radius [see Eq. (1.125) of Vol. II, p. 60]. Hence, the following force will be exerted on an electron at the distance r from its equilibrium position (from the centre of the sphere):

$$F = (-e) E = -\frac{e^2}{R^3} r = -kr$$

In such conditions, the electron, brought out of its equilibrium position in some way or other, will oscillate with the frequency

$$\omega = \sqrt{\frac{k}{m}} = \sqrt{\frac{e^2}{mR^3}} \tag{3.7}$$

(e is the charge of an electron, m is the mass of an electron, and R is the radius of the atom). This equation can be used to assess the size of an atom. By Eq. (3.7)

$$R = \left(\frac{e^2}{m\omega^2}\right)^{1/3}$$

A frequency of $\omega \approx 3 \times 10^{15}$ rad/s corresponds to a wavelength of $\lambda = 6000$ Å (the visible part of the spectrum). Therefore,

$$R = \left(\frac{4.8^2 \times 10^{-20}}{0.91 \times 10^{-27} \times 3^2 \times 10^{30}}\right)^{1/3} \approx 3 \times 10^{-8} \text{ cm}$$

The value obtained coincides in the order of its magnitude with the gas-kinetic dimensions of atoms, which could have been considered as a confirmation of the Thomson model. Later, however, the unfoundedness of this model was established, and at present it is only of historical interest as one of the links in the chain of development of our notions on the structure of atoms.

---

* We shall use the Gaussian system of units here and further in this volume.

3.3. Experiments in Scattering Alpha-Particles. The Nuclear Model of the Atom

The distribution of positive and negative charges in an atom can be revealed by direct experimental "sounding" of its internal regions. Such sounding was performed by the British physicist Ernest Rutherford (1871-1937) and his collaborators with the aid of alpha-particles by watching the change in the direction of their flight (scattering) when passing through thin metal foils.

We remind our reader that alpha-particles are particles emitted by some substances in radioactive decay. The speeds of these particles are of the order of 10^9 cm/s. When Rutherford began to run his experiments, it was known that alpha-particles have a positive charge equal to twice the elementary charge, and that upon losing this charge (with the attachment of two electrons) an alpha-particle transforms into a helium atom.

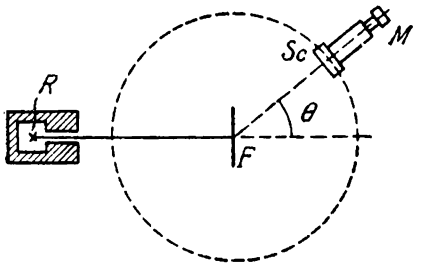


Fig. 3.4

The experiments were conducted as follows (Fig. 3.4). A narrow beam of alpha-particles emitted by radioactive substance R and separated by an aperture fell on thin metal foil F. In passing through the foil, the particles were deflected from their initial direction of motion through various angles θ. The scattered particles struck screen Sc coated with zinc sulphide, and the scintillations* they produced were observed in microscope M. The microscope and the screen could be rotated about an axis passing through the centre of the scattering foil and could thus be positioned at any angle θ. The entire apparatus was placed in an evacuated housing to exclude scattering of the alpha-particles due to collisions with air molecules.

Some of the alpha-particles were found to become scattered through very great angles (almost up to 180 degrees). Upon analysing the results of the experiments, Rutherford arrived at the conclusion that such a large deflection of the alpha-particles is possible only if there is an exceedingly strong electric field inside the atom that is produced by a charge associated with a large mass and concentrated in a very small volume. On the basis of this conclusion, Rutherford in 1911 proposed a nuclear model of the atom. According to Rutherford, an atom is a system of charges at whose centre there is a heavy positive nucleus of charge Ze having dimensions not exceeding 10^{-12} cm,

* By a scintillation is meant a flash of light produced by charged particles when they collide with a substance capable of luminescence.

while around the nucleus there are Z electrons distributed throughout the entire volume occupied by the atom. Almost the entire mass of the atom is concentrated in its nucleus.

On the basis of these assumptions, Rutherford developed a quantitative theory of alpha-particle scattering and derived a formula for the distribution of the scattered particles by the values of the angle θ. In deriving this formula, he reasoned as follows. The deflections of the alpha-particles are due to the action of the atomic nuclei on them. There cannot be a noticeable deflection because of interaction

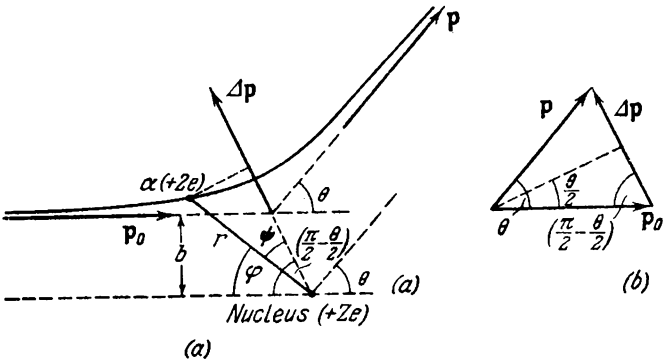


Fig. 3.5

with electrons since the mass of an electron is four orders of magnitude smaller than that of an alpha-particle. When a particle flies near a nucleus, it experiences the Coulomb force of repulsion

$$F = \frac{2Ze^2}{r^2} \tag{3.8}$$

In this case, the trajectory of the particle is a hyperbola (see Vol. I, pp. 116-7). Let θ stand for the angle between the asymptotes of the hyperbola (Fig. 3.5). This angle characterizes the deflection of the particle from its initial direction. The distance b from the nucleus to the initial direction of flight of an alpha-particle is called the **impact parameter.** The closer the trajectory of a particle approaches the nucleus (the smaller is b), the more, naturally, it is deflected (the greater is θ). There is a simple relation between b and θ which we shall now establish.

It follows from the law of energy conservation that at a considerable distance from a nucleus the magnitude of the momentum $\mathbf{p}$ of a scattered particle will be the same as the magnitude of the momentum $\mathbf{p}_0$ before scattering: $p = p_0$. Consequently (see Fig. 3.5*b*), we can write the following expression for the magnitude of the increment of a particle's momentum vector produced as a result of scatter-

ing:

$$|\Delta \mathbf{p}| = 2p_0 \sin\frac{\theta}{2} = 2m_\alpha v \sin\frac{\theta}{2} \tag{3.9}$$

where m_α = mass of an alpha-particle

v = its initial velocity.

At the same time according to Newton's second law, we have

$$\Delta \mathbf{p} = \int \mathbf{F}\, dt$$

Projecting the vectors in this equation onto the direction of $\Delta \mathbf{p}$, we get

$$|\Delta \mathbf{p}| = \int F_{\Delta p}\, dt \tag{3.10}$$

A glance at Fig. 3.5*a* shows that the projection of the force **F** onto the direction of the vector $\Delta \mathbf{p}$ is $F \cos\psi$. The angle ψ can be replaced by the polar angle φ and the angle of deflection θ:

$$\psi = \frac{\pi}{2} - \frac{\theta}{2} - \varphi$$

Hence,

$$F_{\Delta p} = F\cos\psi = F\sin\left(\varphi + \frac{\theta}{2}\right) = \frac{2Ze^2}{r^2}\sin\left(\varphi + \frac{\theta}{2}\right)$$

Using this expression in Eq. (3.10) and simultaneously substituting $d\varphi/\dot{\varphi}$ for dt, we obtain

$$|\Delta \mathbf{p}| = 2Ze^2 \int_0^{\pi-\theta} \frac{\sin(\varphi + \theta/2)\, d\varphi}{r^2\dot{\varphi}} \tag{3.11}$$

The expression $r^2\dot{\varphi}$ equals M/m_α, where M is the magnitude of the angular momentum of the alpha-particle taken relative to the scattering nucleus [see Eq. (3.123) of Vol. I, p. 114; in the present volume we have denoted the angular momentum by **M** instead of **L** for convenience]. The force experienced by the alpha-particle is a central one. The angular momentum M therefore remains constant all the time and equal to its initial value $M_0 = m_\alpha vb$. After replacing $r^2\dot{\varphi}$ with vb, the integral in Eq. (3.11) is calculated quite easily:

$$|\Delta \mathbf{p}| = \frac{2Ze^2}{vb}\int_0^{\pi-\theta} \sin\left(\varphi + \frac{\theta}{2}\right) d\varphi = \frac{2Ze^2}{vb}\, 2\cos\frac{\theta}{2} \tag{3.12}$$

A comparison of Eqs. (3.9) and (3.12) shows that

$$2m_\alpha v \sin\frac{\theta}{2} = \frac{2Ze^2}{vb}\, 2\cos\frac{\theta}{2}$$

Hence*

$$\cot\frac{\theta}{2}=\frac{m_\alpha v^2}{2Ze^2}b \tag{3.13}$$

Let us consider a thin layer of the scattering substance such that each particle when passing through it would fly near only one nucleus, i.e. that each particle will be scattered only once. To experience scattering through an angle within the limits from θ to $\theta + d\theta$, a particle must fly near a nucleus along a trajectory whose impact parameter is within the limits from b to $b + db$ (Fig. 3.6), $d\theta$ and db, as can be seen from Eq. (3.13), being related by the expression

$$-\frac{1}{\sin^2(\theta/2)}\frac{d\theta}{2}=\frac{m_\alpha v^2}{2Ze^2}db \tag{3.14}$$

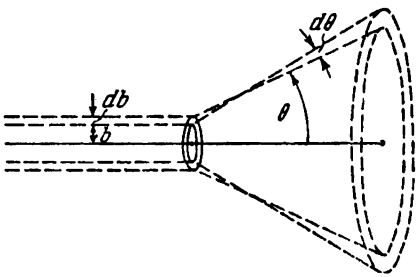


Fig. 3.6

The minus sign in this equation is due to the fact that the angle of deflection diminishes ($d\theta < 0$) with increasing b (i.e. at $db > 0$). In the following, we shall be interested only in the absolute value of db as a function of θ and $d\theta$, and we shall therefore omit the minus sign.

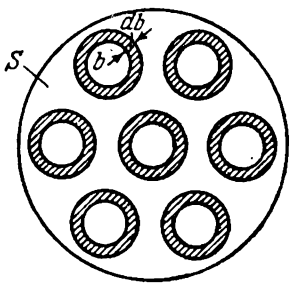


Fig. 3.7

Let us denote the cross-sectional area of a beam of alpha-particles by S. Hence, the number of atoms of the scattering foil in the path of the beam can be represented in the form nSa, where n is the number of atoms in unit volume, and a is the thickness of the foil. If the alpha-particles are distributed uniformly over the cross section of the beam and their number is very great (which is actually the case), then the relative number of alpha-particles flying near one of the nuclei along a trajectory with an impact parameter from b to $b + db$ (and, consequently, deflected within the limits of angles from θ to $\theta + d\theta$) will be (see Fig. 3.7):

$$\frac{dN_\theta}{N}=\frac{nSa\cdot 2\pi b\,db}{S}=na2\pi b\,db \tag{3.15}$$

In this expression, dN_θ is the flux of particles scattered within the limits of angles from θ to $\theta + d\theta$, and N is the total flux of particles in the beam.

* The derivation of formula (3.13) given above belongs to I. E. Irodov.

Using θ and $d\theta$ in Eq. (3.15) instead of b and db in accordance with Eqs. (3.13) and (3.14), we get

$$\frac{dN_\theta}{N} = na\left(\frac{2Ze^2}{m_\alpha v^2}\right)^2 2\pi \cot\frac{\theta}{2}\frac{1}{\sin^2(\theta/2)}\frac{d\theta}{2}$$

We transform the multipliers containing the angle θ:

$$\frac{\cot(\theta/2)}{\sin^2(\theta/2)} = \frac{\cos(\theta/2)\sin(\theta/2)}{\sin^4(\theta/2)} = \frac{\sin\theta}{2\sin^4(\theta/2)}$$

With account of this transformation

$$\frac{dN_\theta}{N} = na\left(\frac{2Ze^2}{m_\alpha v^2}\right)^2 \frac{2\pi\sin\theta\, d\theta}{4\sin^4(\theta/2)}$$

The expression $2\pi \sin\theta\, d\theta$ gives the solid angle $d\Omega$ confining the directions corresponding to scattering angles from θ to $\theta + d\theta$. We can therefore write:

$$\frac{dN_\theta}{N} = na\left(\frac{Ze^2}{m_\alpha v^2}\right)^2 \frac{d\Omega}{\sin^4(\theta/2)} \tag{3.16}$$

We have obtained the **Rutherford formula** for the scattering of alpha-particles. In 1913, Rutherford's collaborators verified this formula by counting the scintillations observed at different angles θ during identical time intervals. In the conditions of the experiment (see Fig. 3.4), the alpha-particles confined within the same solid angle were counted (this angle was determined by the area of screen *Sc* and its distance from the foil). Hence, the number of scintillations observed at different angles should be, according to the Rutherford formula, proportional to $1/[\sin^4(\theta/2)]$. This result of theory was confirmed quite well experimentally. The dependence of the scattering on the thickness of the foil and the speed of the alpha-particles was also found to agree with formula (3.16).

The truth of the theory proceeding from the Coulomb interaction between an alpha-particle and the nucleus of an atom indicates that even an alpha-particle thrown back in the opposite direction does not penetrate into the region occupied by the positive charge of the atom. At the same time, an alpha-particle flying exactly in the direction of a nucleus would approach its centre up to a distance that can be determined by equating the kinetic energy of the alpha-particle to the potential energy of interaction of the particle with the nucleus at the moment when the particle comes to a full stop:

$$\frac{m_\alpha v^2}{2} = \frac{2Ze^2}{r_{\min}}$$

($r_{\min}$ is the minimum distance between the centres of the alpha-particle and of the nucleus). Assuming that $Z = 47$, $v = 10^9$ cm/s,

and $m_\alpha = 4 \times 1.66 \times 10^{-24} = 6.6 \times 10^{-24}$ g, we get

$$r_{\min} = \frac{4Ze^2}{m_\alpha v^2} = \frac{4 \times 47 \times 4.8^2 \times 10^{-20}}{6.6 \times 10^{-24} \times 10^{18}} \approx 6 \times 10^{-12} \text{ cm}$$

Thus, the results of experiments involving the scattering of alpha-particles are witnesses in favour of the nuclear model of the atom presented by Rutherford. But the nuclear model contradicted the laws of classical mechanics and electrodynamics. Since a system of stationary charges cannot be in a stable state, Rutherford had to renounce the static model of the atom and assume that the electrons travel about the nucleus along curved trajectories. In this case, however, an electron would travel with acceleration. Consequently, according to classical electrodynamics, it must continuously emit electromagnetic (light) waves. The process of emission is attended by the loss of energy, so that the electron in hte long run should fall onto the nucleus (Fig. 3.8).

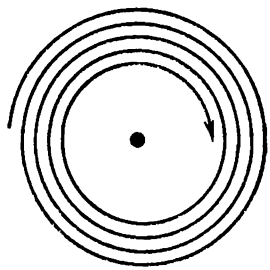


Fig. 3.8

3.4. Bohr's Postulates. The Franck-Hertz Experiment

We found in the preceding section that the nuclear model of the atom in combination with classical mechanics and electrodynamics was incapable of explaining the stability of an atom and the nature of an atomic spectrum. A way out of this impasse was found in 1913 by the Danish physicist Niels Bohr (1885-1962), true, at the price of introducing assumptions that contradicted the classical notions. The assumptions made by Bohr are contained in his following two postulates.

1. Among the infinite multitude of electron orbits possible from the viewpoint of classical mechanics, only several discrete orbits satisfying definite quantum conditions are actually encountered. An electron in one of these orbits does not emit electromagnetic waves (light) although it travels with acceleration.

2. Radiation is emitted or absorbed in the form of a quantum of light energy $\hbar\omega$ when an electron transfers from one stationary (stable) state to another. The magnitude of a light quantum equals the difference between the energies of the stationary states between which the quantum jump of the electron is performed:

$$\hbar\omega = E_n - E_m \tag{3.17}$$

The existence of discrete energy levels of an atom was confirmed by experiments run in 1914 by the German physicists James Franck

(1882-1964) and Gustav Hertz (born 1887). A schematic view of their apparatus is shown in Fig. 3.9*a*. A tube filled with mercury vapour at a low pressure (about 1 mmHg) contained three electrodes, name-

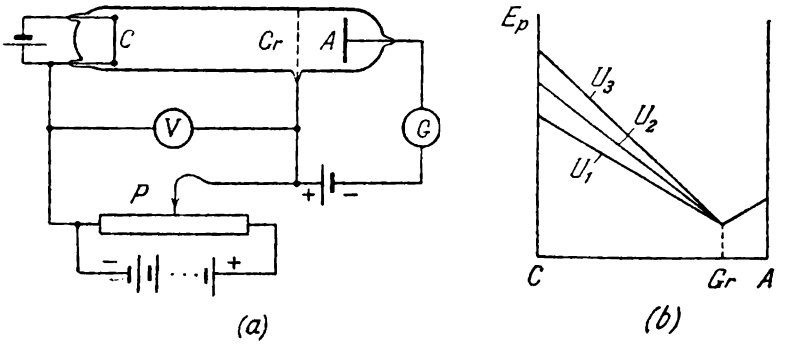


Fig. 3.9

ly, cathode *C*, grid *Gr* and anode *A*. The electrons flying out of the cathode owing to thermionic emission were accelerated by the potential difference *U* applied between the cathode and the grid. The potential difference could be smoothly varied with the aid of potentiometer *P*. A weak electric field (a potential difference of the order of 0.5 V) was set up between the grid and the anode that retarded the motion of the electrons to the anode. Figure 3.9*b* shows the change in the potential energy of an electron $E_p = -e\varphi$ in the space between the electrodes at different values of the voltage *U* between the cathode and the grid (φ is the potential at the corresponding point of the field).

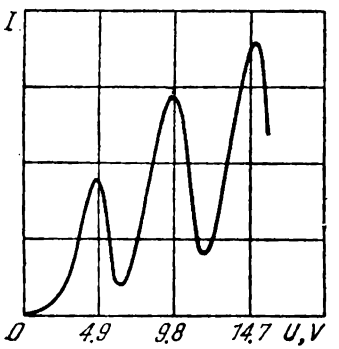


Fig. 3.10

The relation between the current *I* in the anode circuit and the voltage *U* between the cathode and the grid was studied. The current and the voltage were measured by galvanometer *G* and voltmeter *V*, respectively. The results obtained are shown in Fig. 3.10. It can be seen that the current first monotonously increased, reaching a maximum at $U = 4.9$ V, after which it sharply dropped with a further growth in *U*, reaching a minimum, and then again began to increase. Maxima of the current repeated at *U* equal to 9.8, 14.7 V, etc.*.

* Maxima were actually obtained at voltages of 4.1, 9.0, 13.9 V, etc., which is explained by the presence of a contact potential difference of the order of 0.8 V between the electrodes.

Such a shape of the curve is explained by the fact that owing to the discrete nature of their energy levels, atoms can absorb energy only in portions of

$$\Delta E_1 = E_2 - E_1, \quad \text{or} \quad \Delta E_2 = E_3 - E_1, \text{ etc.}$$

where E_1, E_2, E_3, . . . are the energies of the 1st, 2nd, 3rd, etc. stationary states.

As long as the energy of an electron is smaller than ΔE_1, the collisions between an electron and a mercury atom are of an elastic nature; since the mass of an electron is many times smaller than that of a mercury atom, the energy of an electron does not virtually change in the collisions. A part of the electrons get caught on the grid, while the remaining ones pass through the grid and reach the anode, producing a current in the circuit of galvanometer G. The greater the velocity with which the electrons reach the grid (the higher is U), the larger will be the fraction of the electrons passing through the grid and, consequently, the higher will be the current I.

When the energy accumulated by an electron in the space between the cathode and the grid reaches the value ΔE_1, the collisions stop being elastic—the electrons when they collide with the atoms transfer the energy ΔE_1 to them and then continue their motion with a lower velocity. Therefore, the number of electrons reaching the anode diminishes. For example, at $U = 5.3$ V, an electron transfers to an atom an energy corresponding to 4.9 V (the first excitation potential of a mercury atom) and continues to travel with an energy of 0.4 eV. Even if such an electron does get between the grid and the anode, it will not be able to overcome the retarding voltage of 0.5 V and will be returned to the grid.

The atoms that upon colliding with electrons receive an energy of ΔE_1 pass over into an excited state, from which after a time of the order of 10^{-8} s elapses they return to their ground state, emitting a photon having the frequency $\omega = \Delta E_1/\hbar$.

At a voltage exceeding 9.8 V, an electron along its path from the cathode to the anode may undergo an elastic collision with mercury atoms twice, losing an energy of 9.8 eV. As a result, the current I will again begin to fall. At a still higher voltage, three inelastic collisions of an electron with atoms are possible, which leads to the appearance of a maximum at $U = 14.7$ V, and so on.

At a sufficiently high rarefaction of the mercury vapour and the corresponding magnitude of the accelerating voltage, the electrons during the time before they collide with atoms may acquire a velocity high enough to transfer an atom to a state with the energy E_3. In this case, maxima are observed on the curve $I = f(U)$ at voltages that are multiples of the second excitation potential of an atom (this potential is 6.7 V for mercury), or at voltages equal to the sum of the first and second excitation potentials, etc.

Thus, the Franck-Hertz experiments directly detect the existence of discrete energy levels in atoms.

In spectroscopy, the frequencies of spectral lines are customarily represented in the form of the difference between positive numbers $T(n)$ called terms (see the next to last paragraph of Sec. 3.1). For example, for hydrogen, $T(n) = R/n^2$. Accordingly, the frequency of the photon emitted in a transition from state n to state m is determined by the formula

$$\omega_{nm} = T(m) - T(n) = \frac{R}{m^2} - \frac{R}{n^2} \quad (n > m) \tag{3.18}$$

[see formula (3.4)].

According to Bohr's second postulate

$$\omega_{nm} = \frac{E_n - E_m}{\hbar} = \left(-\frac{E_m}{\hbar}\right) - \left(-\frac{E_n}{\hbar}\right)$$

(we remind our reader that the energies of bound states of an electron are negative, so that the expressions in parentheses are greater than zero). Comparison with formula (3.18) shows that

$$T(n) = -\frac{E_n}{\hbar} \tag{3.19}$$

Thus, the term is closely associated with the energy of a stationary state of an atom, differing from it only in the factor $(-1/\hbar)$.

3.5. Rule for Quantization of Circular Orbits

Bohr obtained the condition for stationary orbits proceeding from Planck's postulate according to which only such states of a harmonic oscillator are possible whose energy is

$$E_n = n\hbar\omega \quad (n \text{ is an integer}) \tag{3.20}$$

Let us denote the coordinate of the oscillator by q and its momentum by p. The total energy of an oscillator is determined by the expression

$$E_n = \frac{p^2}{2m} + \frac{m\omega^2 q^2}{2} = n\hbar\omega$$

Hence,

$$\frac{q^2}{2n\hbar/m\omega} + \frac{p^2}{2mn\hbar\omega} = 1 \tag{3.21}$$

The coordinate plane q, p is called a **phase plane**, and a curve in this plane determining p as a function of q for a given motion is called a **phase trajectory**. It can be seen from Eq. (3.21) that the phase trajectory of a harmonic oscillator is an ellipse (Fig. 3.11).

The semiaxes of the ellipse are

$$a = \sqrt{\frac{2n\hbar}{m\omega}}, \qquad b = \sqrt{2mn\hbar\omega}$$

The area of the ellipse equals the product of the semiaxes multiplied by π:

$$S_n = \pi ab = 2\pi\hbar n \tag{3.22}$$

The area can also be represented in the form

$$S_n = \oint p\, dq \tag{3.23}$$

(in integration, the entire ellipse is circumvented; see Fig. 3.11).

The rule for quantization follows from a comparison of Eqs. (3.22) and (3.23):

$$\oint p\, dq = 2\pi\hbar n \tag{3.24}$$

Bohr extended rule (3.24) obtained for a harmonic oscillator to other mechanical systems. For an oscillator, $q = x$, $p = m\dot{x}$. For

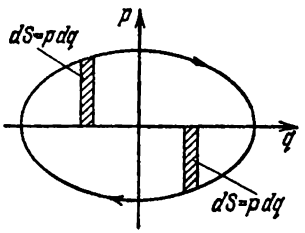


Fig. 3.11

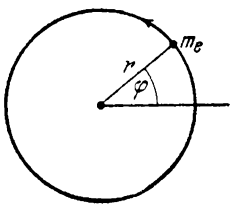


Fig. 3.12

other systems, q is meant to be a generalized coordinate*, and p the generalized momentum.

For an electron travelling around a nucleus in a circular orbit, it is natural to take the azimuthal angle φ (Fig. 3.12) as the generalized coordinate. Here, the generalized velocity will be $\dot{\varphi}$. We know that in rotation the part of the linear velocity passes over to the angular velocity $\dot{\varphi}$, and the part of the mass to the moment of inertia $m_e r^2$ (where m_e is the mass of an electron). The generalized momentum is accordingly $m_e r^2 \dot{\varphi} = m_e v r$. The latter expression determines the conventional angular momentum M taken relative to the nucleus. Thus, for an electron travelling in a circular orbit, condition (3.24)

* By generalized coordinates are meant any quantities with whose aid it is possible to set the position of a system in space.

has the form

$$\oint M\, d\varphi = 2\pi\hbar n \tag{3.25}$$

The force which the nucleus exerts on the electron is a central one. Consequently, $M = \text{const}$, and the left-hand side of Eq. (3.25) is $2\pi M$. We therefore arrive at the condition

$$M = n\hbar \tag{3.26}$$

Thus, according to Bohr's condition, of all the orbits of an electron possible from the viewpoint of classical mechanics, only those are actually encountered for which the angular momentum equals an integral multiple of Planck's constant $\hbar$.

3.6. The Elementary Bohr Theory of the Hydrogen Atom

According to Eq. (3.26), only such orbits are possible for which the angular momentum of an electron $m_e v r$ satisfies the condition

$$m_e v r = n\hbar \quad (n = 1, 2, 3, \ldots) \tag{3.27}$$

The number n is called the **principal quantum number.**

Let us consider an electron moving in the field of an atomic nucleus with the charge Ze. When $Z = 1$, such a system corresponds to a hydrogen atom, at other values of Z, to a hydrogen-like ion, i.e. to an atom with the atomic number Z from which all the electrons except one have been removed. The equation of motion of the electron has the form

$$m_e \frac{v^2}{r} = \frac{Ze^2}{r^2} \tag{3.28}$$

Deleting v from Eqs. (3.27) and (3.28), we get an expression for the radii of the allowed orbits:

$$r_n = \frac{\hbar^2}{m_e Z e^2} n^2 \quad (n = 1, 2, 3, \ldots) \tag{3.29}$$

The radius of the first orbit of the hydrogen atom is known as the **Bohr radius** (it is customarily designated by the symbol r_0 or a_0 instead of r_1). Its value is

$$r_0 = \frac{\hbar^2}{m_e e^2} = 0.529 \text{ Å} \tag{3.30}$$

We shall note that the Bohr radius has a value of the order of the gas-kinetic dimensions of an atom.

The internal energy of an atom consists of the kinetic energy of the electron (the nucleus is stationary) and of the energy of inter-

action of the electron with the nucleus

$$E = \frac{m_e v^2}{2} - \frac{Ze^2}{r}$$

It follows from Eq. (3.28) that

$$\frac{m_e v^2}{2} = \frac{Ze^2}{2r}$$

Hence,

$$E = \frac{Ze^2}{2r} - \frac{Ze^2}{r} = -\frac{Ze^2}{2r}$$

Using Eq. (3.29) for r in this expression, we shall find the allowed values of the internal energy of an atom:

$$E_n = -\frac{m_e e^4}{2\hbar^2}\frac{Z^2}{n^2} \qquad (n = 1, 2, 3, \ldots) \qquad (3.31)$$

The energy levels determined by formula (3.31) are shown schematically in Fig. 3.13.

When a hydrogen atom ($Z = 1$) passes from the state n to the state m, a photon is emitted

$$\hbar\omega = E_n - E_m = -\frac{m_e e^4}{2\hbar^2}\left(\frac{1}{n^2} - \frac{1}{m^2}\right)$$

The frequency of the emitted light is

$$\omega = \frac{m_e e^4}{2\hbar^3}\left(\frac{1}{m^2} - \frac{1}{n^2}\right)$$

We have arrived at the generalized Balmer formula [see Eq. (3.4)], the following value being obtained for the Rydberg constant:

$$R = \frac{m_e e^4}{2\hbar^3} \qquad (3.32)$$

When we introduce the numerical values of m_e, e, and $\hbar$ into Eq. (3.32), we get a quantity that strikingly well agrees with the experimental value of the Rydberg constant.

Fig. 3.13

Bohr's theory was a major step in the development of the theory of the atom. It showed very clearly the impossibility of applying classical physics to intra-atomic phenomena and the predominate significance of the quantum laws in the microworld.

The elementary theory which we have treated was subjected to further development and clarifications with which we shall not acquaint our reader because at present Bohr's theory has mainly a historical significance. After the first successes of the theory, its shortcomings began to stand out more and more. Especially distress-

ing was the failure of all attempts to construct a theory of the helium atom—one of the simplest atoms directly following the hydrogen atom in Mendeleev's periodic table of elements.

The weakest aspect of the Bohr theory underlying its subsequent failures was its internal logical contradiction: it was neither a consistent classical theory nor a consistent quantum one. After the discovery of the wave properties of matter, it became absolutely clear that the Bohr theory, based on classical mechanics, could be only a transition step on the path to the creation of a consistent theory of atomic phenomena.

CHAPTER 4 ELEMENTS OF QUANTUM MECHANICS

4.1. De Broglie's Hypothesis. Wave Properties of Matter

The inadequacy of Bohr's theory pointed to the necessity of revising the fundamentals of the quantum theory and the notions on the nature of microparticles (electrons, protons, etc.). The question arose as to how exhaustive is our notion of an electron as of a tiny mechanical particle characterized by definite coordinates and a definite velocity.

As a result of the broadening of our notions on the nature of light, it was found that a peculiar dual nature was detected in optical phenomena. In addition to such properties of light that in the most direct way point to its wave nature (interference, diffraction), there are other properties that just as directly reveal its corpuscular nature (the photoelectric effect, the Compton effect).

In 1924, the French physicist Louis de Broglie (born 1892) put forth a bold hypothesis that duality is not a feature of only optical phenomena, but has a universal significance. "In optics", he wrote, "the corpuscular way of treatment was neglected too much during a whole century; wasn't the opposite error made in the theory of matter?" Assuming that particles of matter have wave properties in addition to corpuscular ones, de Broglie transferred to the case of matter particles the same rules of transition from one picture to another that hold for light. A photon has the energy

$$E = \hbar\omega$$

and the momentum

$$p = \frac{2\pi\hbar}{\lambda}$$

De Broglie assumed that the motion of an electron or some other particle is associated with a wave process whose wavelength is

$$\lambda = \frac{2\pi\hbar}{p} = \frac{2\pi\hbar}{mv} \tag{4.1}$$

and whose frequency is

$$\omega = \frac{E}{\hbar} \tag{4.2}$$

De Broglie's hypothesis was soon confirmed experimentally. In 1927, the American physicists Clinton Davisson (1881-1958) and Lester Germer (born 1896) studied the reflection of electrons

from a monocrystal of nickel belonging to the cubic system. A narrow beam of monoenergetic electrons was directed onto the surface of the monocrystal polished at right angles to the major diagonal of a crystal cell [the crystal planes parallel to this surface are designated by the indices (111) in crystallography; see Sec. 6.1]. The reflected electrons were trapped by a cylindrical electrode connected to a galvanometer (Fig. 4.1). The intensity of the reflected beam was assessed according to the current flowing through the galvanometer. The velocity of the electrons and the angle φ were varied. Figure 4.2 shows how the current measured by the galvanometer depends on the angle φ at different energies of the electrons. The vertical axis in the graphs determines the direction of the incident beam. The current in a given direction is represented by the length of a line drawn from the origin of coordinates to its intersection with the curve. A glance at the figure shows that scattering is especially

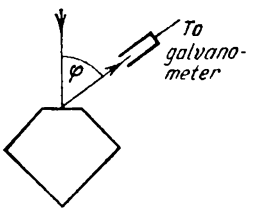


Fig. 4.1

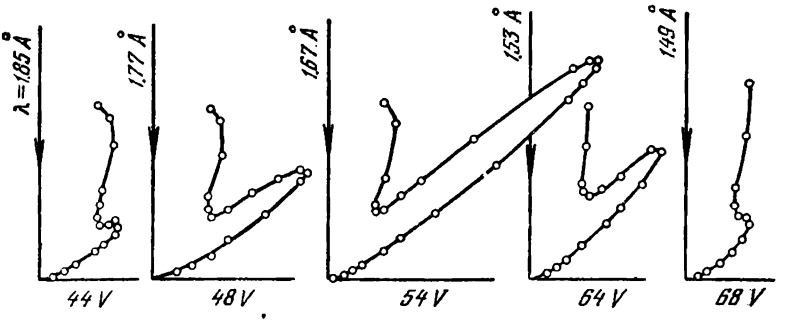


Fig. 4.2

intensive at a definite value of the angle φ. This angle corresponds to reflection from the atomic planes whose spacing d was known from X-ray investigations. At a given value of φ, the current was especially strong at an accelerating voltage of 54 V. The wavelength corresponding to this voltage and calculated by Eq. (4.1) is 1.67 Å. The Bragg wavelength corresponding to the condition*

$$2d \sin \theta = n\lambda$$

was 1.65 Å. The coincidence is so striking that the Davisson-Germer experiments must be acknowledged as a brilliant confirmation of de Broglie's idea.

* The slip angle θ is related to the angle φ by the expression

$$\theta = \frac{\pi}{2} - \frac{\varphi}{2}$$

In 1927, the British physicist George Thomson (born 1892) and independently of him the Soviet physicist Pyotr Tartakovsky obtained a diffraction pattern when an electron beam was passed through a metal foil. The experiment was run as follows (Fig. 4.3). A beam of electrons accelerated by a potential difference of the order of several scores of kilovolts was passed through a thin metal foil and impinged on a photographic plate. An electron colliding with the photographic plate has the same action on it as a photon. The electron-diffraction pattern of gold obtained in this way (Fig. 4.4*a*) is compared with an *X*-ray diffraction pattern of aluminium (Fig. 4.4*b*) obtained in similar conditions. The similarity of the two patterns is staggering. The German physicist Otto Stern (1888-1969) and his collaborators showed that diffraction phenomena are also detected in atomic and molecular beams. In all the cases listed above, the diffraction pattern corresponds to the wavelength determined by Eq. (4.1).

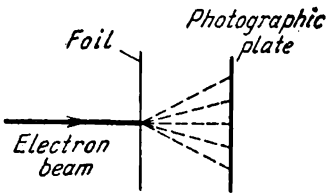


Fig. 4.3

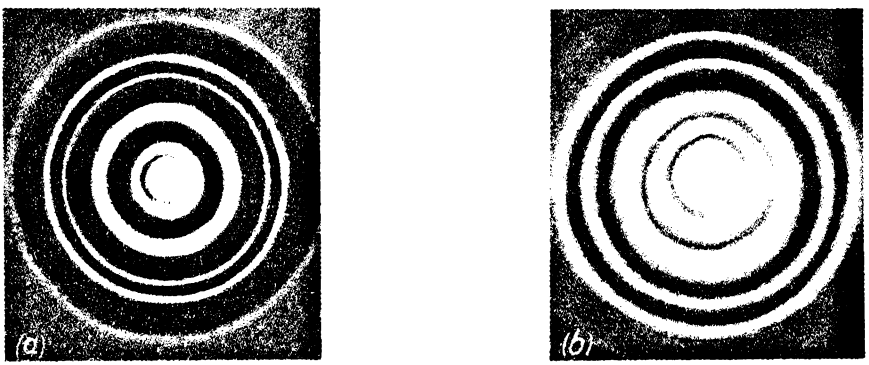


Fig.4.4

In the Davisson-Germer experiments, and also in G. Thomson's experiments, the intensity of the electron beams was so high that a large number of electrons passed through the crystal simultaneously. It was therefore possible to assume that the diffraction pattern observed was due to the simultaneous participation of a large number of electrons in the process, while a single electron passing through

the crystal does not display diffraction. To clarify this circumstance, the Soviet physicists Leon Biberman, Nikolai Sushkin, and Valentin Fabrikant in 1949 ran an experiment in which the intensity of the electron beam was so low that the electrons certainly passed through the instrument one at a time. The interval between the passage of two consecutive electrons through the crystal was about 30 000 times greater than the time needed for an electron to pass through the entire apparatus. With a sufficient exposure, a diffraction pattern was obtained that differed in no way from the one observed at the ordinary intensity of the beam. It was thus proved that a single electron has wave properties.

4.2. The Unusual Properties of Microparticles

Microparticles are defined as elementary particles (electrons, protons, neutrons, photons, and other simple particles), and also as complex particles formed from a comparatively small number of elementary particles (molecules, atoms, atomic nuclei, etc.).

The term "microparticle" reflects only one aspect of the object it is applied to. Any microobject (a molecule, atom, electron, photon, etc.) is a special kind of formation combining in itself the properties of both a particle and a wave. Perhaps it would be more correct to call it a "particle-wave".

A microobject is not capable of acting directly on our organs of sense—it can neither be seen nor felt. Nothing like microobjects exists in the world we perceive. Microbodies "do not behave like anything you have ever seen".*

"Because atomic behaviour is so unlike ordinary experience, it is very difficult to get used to and it appears peculiar and mysterious to everyone, both to the novice and to the experienced physicist. Even the experts do not understand it the way they would like to, and it is perfectly reasonable that they should not, because all of direct human experience and of human intuition applies to large objects. We know how large objects will act, but things on a small scale just do not act that way. So we have to learn about them in a sort of abstract or imaginary fashion, and not by connection with our direct experience".

In prequantum physics, to "understand" meant to form a visual image of an object or process. Quantum physics cannot be understood in this meaning of the word. Any visual model will inevitably function according to classical laws and will therefore not be suitable for representing quantum processes. Therefore, the best that we can

* This and the following passages in the given section in quotation marks have been taken from Feynman, R. P., Leighton, R. B., Sands, M. *The Feynman Lectures on Physics*. Reading, Mass., Addison-Wesley (1963), Chap. 37.

do is to discard all attempts to construct visual models of the behaviour of quantum objects. The absence of visualization may first give rise to a feeling of dissatisfaction, but this feeling passes with time, and everything takes its usual place.

Combining the properties of a particle and a wave, microbodies "do not behave like waves, they do not behave like particles". A microparticle differs from a wave in that it is always detected as an indivisible whole. Nobody ever observed, for example, a half of an electron. At the same time, a wave can be split into parts (for example, by directing a light wave onto a half-silvered mirror) and each part then perceived separately. A difference of a microparticle from a macroparticle which we are accustomed to is that it does not have definite values of a coordinate and momentum simultaneously, owing to which the concept of trajectory as applied to a microparticle loses its meaning.

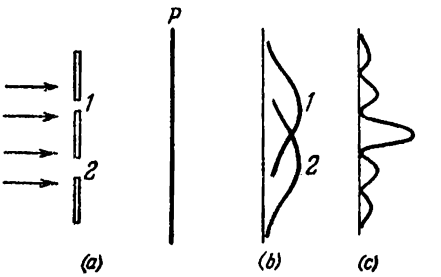


Fig. 4.5.

The peculiar nature of the properties of microparticles reveals itself with the greatest clarity in the following mental experiment*. Let us direct a parallel beam of monoenergetic (i.e. having the same kinetic energy) electrons onto a barrier with two narrow slits (Fig. 4.5). We shall place photographic plate P after the barrier. We shall first close the second slit and make an exposure during the time τ. Blackening on the processed plate will be characterized by curve *1* in Fig. 4.5*b*. We shall expose a second plate during the same time τ with the first slit closed. The nature of blackening of the plate is shown in this case by curve *2* in Fig. 4.5*b*. Finally, we shall open both slits and expose a third plate during the time τ. The pattern of the blackening obtained in the last case is shown in Fig. 4.5*c*. This pattern is not at all equivalent to the superposition of the first two patterns. It is similar to the pattern obtained upon the interference of two coherent light waves. The nature of the pattern shows that the motion of each electron is affected by both slits. This conclusion is incompatible with our notion of trajectories. If an electron at each moment of time were at a definite point in space and travelled along a trajectory, it would pass through a defi-

* In a mental (thought) experiment, the aspect of a phenomenon being studied is revealed in the simplest and clearest form. The authenticity of the effect observed in a mental experiment follows from observations obtained in a number of real experiments. In the given case, the experiments involving the diffraction of electrons described in the preceding section are such experiments.

nite slit—either the first or the second one. The phenomenon of diffraction proves, however, that both slits—the first and the second—participate in the passage of each electron.

Matters, however, should not be represented as if a part of an electron passes through one slit and its other part through the second one. We have already noted that an electron, like other microparticles, is always detected as an entirety, with its inherent mass, charge, and other characteristic quantities. Thus, an electron, proton, atomic nucleus are particles with very peculiar properties. A conventional sphere, even a very tiny one (a macroscopic particle) cannot be the prototype of a microparticle. A reduction in size is attended by the gradual appearance of qualitatively new properties not found in macroparticles.

In a number of cases, the statement that microparticles have no trajectories would seem to contradict experimental facts. For example, the path along which a microparticle travels in a Wilson chamber is detected in the form of narrow tracks produced by droplets of mist; the motion of electrons in a cathode-ray tube is calculated excellently according to classical laws, etc. This seeming contradiction is explained by the fact that in known conditions the concept of a trajectory may be applied to microparticles, but only with a certain degree of accuracy. Matters are exactly the same as in optics. If the dimensions of barriers or holes are great in comparison with the wavelength, the propagation of light takes place, as it were, along definite rays (trajectories). In definite conditions, the concept of a trajectory can also be approximately applied to the motion of microparticles, in the same way as the law of the rectilinear propagation of light is true.

4.3. The Uncertainty Principle

In classical mechanics, the state of a point particle (a classical particle) is set by giving the values of its coordinates, momentum, energy, etc. These quantities are known as **dynamic variables.** Strictly speaking, the above dynamic variables cannot be ascribed to a microobject. We obtain information on microparticles, however, by observing their interaction with instruments that are macroscopic bodies. Therefore, the results of such measurements are willy-nilly expressed in terms developed to characterize macrobodies, i.e. through the values of the dynamic variables. Accordingly, the measured values of the dynamic variables are ascribed to microparticles. For example, we speak of the state of an electron in which it has a certain value of the energy, and so on.

The peculiar nature of the properties of microparticles manifests itself in that measurements do not always give definite values for

all the variables. For example, an electron (and any other microparticle) cannot simultaneously have accurate values of its coordinate x and its momentum component p_x. The uncertainties in the values of x and p_x satisfy the expression

$$\Delta x \cdot \Delta p_x \geqslant \frac{\hbar}{2} \tag{4.3}$$

($\hbar$ is Planck's constant). It can be seen from expression (4.3) that the smaller the uncertainty of one of the variables (x or p_x), the greater is the uncertainty of the other one. A state is possible in which one of the variables has an accurate value, while the other one is absolutely uncertain (its uncertainty equals infinity).

A relation similar to expression (4.3) holds for y and p_y, for z and p_z, and also for a number of other pairs of quantities (in classical mechanics such pairs of quantities are called **canonically conjugate**). Using the symbols A and B to denote canonically conjugate quantities, we can write

$$\Delta A \cdot \Delta B \geqslant \frac{\hbar}{2} \tag{4.4}$$

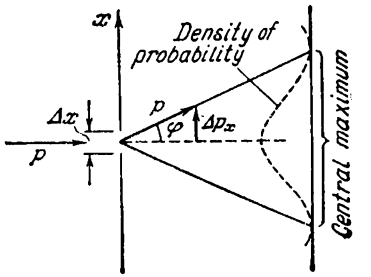


Fig. 4.6

Expression (4.4) is known as the **uncertainty relation** for the quantities A and B. This relation was discovered by the German physicist Werner Heisenberg (1901-1976) in 1927.

The statement that *the product of the uncertainties in the values of two conjugate variables cannot be less than Planck's constant $\hbar$ in the order of magnitude* is called the **Heisenberg uncertainty principle.**

Energy and time are canonically conjugate quantities. Therefore, the uncertainty relation also holds for them:

$$\Delta E \cdot \Delta t \geqslant \frac{\hbar}{2} \tag{4.5}$$

This relation signifies that the determination of the energy with an accuracy of ΔE must occupy an interval of time equal at least to $\Delta t \sim \hbar/\Delta E$.

The uncertainty relation was established when considering, in particular, the following example. Let us attempt to find the value of the coordinate x of a freely flying microparticle by placing in its path a slit of width Δx at right angles to the direction of motion of the particle (Fig. 4.6). Before the particle passes through the slit, its momentum component p_x has an accurate value equal to zero (the slit in accordance with our conditions is perpendicular to the

momentum), so that $\Delta p_x = 0$, but to make up for it the coordinate x of the particle is absolutely uncertain. At the moment when the particle passes through the slit, matters change. Instead of complete uncertainty in the coordinate x, the uncertainty Δx appears, but this is achieved at the price of a loss in the certainty in the value of p_x. Indeed, owing to diffraction, there is a certain probability of the fact that the particle will move within the limits of the angle 2φ, where φ is the angle corresponding to the first diffraction minimum (the higher order maxima may be ignored because their intensity is low in comparison with that of the central maximum). Thus, the following uncertainty appears:

$$\Delta p_x = p \sin \varphi$$

The angle φ for which

$$\sin \varphi = \frac{\lambda}{\Delta x}$$

corresponds to the edge of the central diffraction maximum (the first minimum) obtained from a slit of width Δx [see Eq. (18.25) of Vol. II, p. 406]. Consequently,

$$\Delta p_x \sim p \frac{\lambda}{\Delta x}$$

Hence with account taken of Eq. (4.1), we get the expression

$$\Delta x \cdot \Delta p_x \sim p\lambda = 2\pi\hbar$$

that agrees with expression (4.3).

The uncertainty relation is sometimes interpreted as follows: a microparticle actually does have accurate values of its coordinates and momenta, but the action of the measuring instrument perceptible for such a particle does not make exact determination of these values possible. Such an interpretation is absolutely wrong. It contradicts the phenomena of microparticle diffraction observed experimentally.

The uncertainty relation indicates to what extent we can apply the concepts of classical mechanics to microparticles, in particular to what degree of accuracy we can speak of the trajectories of microparticles. Motion along a trajectory is characterized by quite definite values of the coordinates and velocity at each moment of time. Substituting mv_x for p_x in expression (4.3), we get the relation

$$\Delta x \cdot \Delta v_x \geqslant \hbar/2m$$

We see that the greater the mass of a particle, the smaller is the uncertainty of its coordinate and velocity and, consequently, the greater is the accuracy with which we can apply the concept of trajectory. Already for a macroparticle only one micrometre in size, the uncertainties in the values of x and v_x are beyond the limits

of the accuracy of measuring these quantities so that its motion will virtually be indistinguishable from motion along a trajectory.

In definite conditions, even the motion of a microparticle may approximately be considered as occurring along a trajectory. We shall take as an example the motion of an electron in a cathode-ray tube. Let us assess the uncertainty in the coordinate and momentum of an electron for this case. Assume that the trace of the electron beam on the screen has a radius r of the order of 10^{-3} cm, and the length l of the tube is of the order of 10 cm (Fig. 4.7). Hence $\Delta p_x/p_x \sim 10^{-4}$. The momentum of an electron is related to the accelerating voltage U by the expression

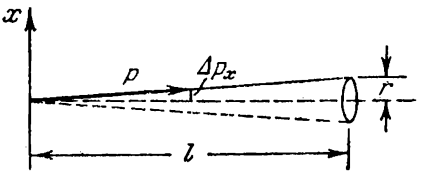


Fig. 4.7

$$\frac{p^2}{2m} = eU$$

Hence $p = \sqrt{2meU}$. At a voltage of $U \sim 10^4$ V, the energy of an electron is 10^4 eV $= 1.6 \times 10^{-8}$ erg. Let us assess the magnitude of the momentum:

$$p = \sqrt{2 \times 0.91 \times 10^{-27} \times 1.6 \times 10^{-8}} \approx 5 \times 10^{-18}$$

Consequently, $\Delta p_x \approx 5 \times 10^{-18} \times 10^{-4} = 5 \times 10^{-22}$. And, finally, according to expression (4.3):

$$\Delta x = \frac{\hbar/2}{\Delta p_x} \approx \frac{1.05 \times 10^{-27}/2}{5 \times 10^{-22}} \sim 10^{-6} \text{ cm}$$

The result obtained indicates that the motion of an electron in a cathode-ray tube is virtually indistinguishable from motion along a trajectory.

The uncertainty relation is one of the fundamental principles of quantum mechanics. This relation alone is sufficient to obtain a number of important results. In particular, it allows us to explain why an electron does not fall onto the nucleus of an atom, and also to assess the dimensions of the simplest atom and the minimum possible energy of an electron in such an atom.

If an electron were to fall onto a point nucleus, its coordinates and momentum would take on definite (zero) values, which is incompatible with the uncertainty principle. This principle requires that the uncertainty in the coordinate of the electron Δr and the uncertainty in the momentum Δp be related by condition (4.3). The energy would formally be minimum at $r = 0$ and $p = 0$. Therefore, in assessing the smallest possible energy, we must assume that $\Delta r \approx r$ and $\Delta p \approx p$. Using these values in expression (4.3), we get

$$rp = \hbar \qquad (4.6)$$

(since our calculations can only claim to give the orders of magnitude of the quantities being calculated, we have omitted the one-half in the right-hand side).

The energy of the electron in a hydrogen atom is

$$E = \frac{p^2}{2m} - \frac{e^2}{r}$$

Substituting $\hbar/r$ for p in accordance with Eq. (4.6), we find that

$$E = \frac{\hbar^2}{2mr^2} - \frac{e^2}{r} \tag{4.7}$$

Let us find the value of r at which E is minimum. Differentiating Eq. (4.7) with respect to r and equating the derivative to zero, we arrive at the equation

$$-\frac{\hbar^2}{mr^3} + \frac{e^2}{r^2} = 0$$

from which it follows that

$$r = \frac{\hbar^2}{me^2} \tag{4.8}$$

The value we have obtained coincides with the radius of the first Bohr orbit of the hydrogen atom [see Eq. (3.30)].

Introduction of Eq. (4.8) into Eq. (4.7) gives the energy of the ground state:

$$E_{\min} = \frac{\hbar^2}{2m}\left(\frac{me^2}{\hbar^2}\right)^2 - e^2\frac{me^2}{\hbar^2} = -\frac{me^4}{2\hbar^2}$$

The found value also coincides with the energy of the first Bohr level for $Z = 1$ [see Eq. (3.31)].

The circumstance that we have obtained accurate values of r and E is naturally simply good fortune. The calculations we have given above can only claim to give an assessment of the order of the quantities r and E.

4.4. The Schrödinger Equation

In 1926, the Austrian physicist Erwin Schrödinger (1887-1961) presented his famous equation as a development of de Broglie's ideas of the wave properties of matter. He associated with the motion of a microparticle a complex function of the coordinates and time which he called the wave function and designated by the Greek letter "psi" (ψ or Ψ). We shall call it the **psi-function.**

The psi-function characterizes the state of a microparticle. The form of the function is obtained from a solution of the Schrödinger

equation that appears as follows:

$$-\frac{\hbar^2}{2m}\nabla^2\Psi + U\Psi = i\hbar\frac{\partial\Psi}{\partial t} \tag{4.9}$$

Here m = mass of the particle
U = potential energy of the particle
i = imaginary unity
∇^2 = Laplacian operator.

The result of the action of this operator on a function is the sum of the second partial derivatives of this function with respect to the coordinates:

$$\nabla^2\Psi = \frac{\partial^2\Psi}{\partial x^2} + \frac{\partial^2\Psi}{\partial y^2} + \frac{\partial^2\Psi}{\partial z^2} \tag{4.10}$$

Inspection of Eq. (4.9) reveals that the form of the psi-function is determined by the function U, i.e. in the long run by the nature of the forces exerted on a particle.

The Schrödinger equation is a fundamental equation of non-relativistic quantum mechanics. It cannot be derived from other relations. It must be considered as a starting basic assumption whose truth is proved by the fact that all its corollaries agree with experimental data in the most accurate way.

Schrödinger derived his equation on the basis of an opticomechanical analogy. The latter consists in the similarity of the equations describing the path of light rays with the equations determining the trajectories of particles in analytical mechanics. In optics, the path of rays satisfies de Fermat's principle (see Sec. 16.6 of Vol.II, p. 334), in mechanics the form of a trajectory satisfies the so-called principle of least action.

If the force field in which a particle is travelling is stationary, then the function U does not depend explicitly on the time. In this case, the solution of the Schrödinger equation breaks up into two multipliers, one of which depends only on the coordinates, and the other only on the time:

$$\Psi(x,\ y,\ z,\ t) = \psi(x,\ y,\ z)\exp\left(-i\frac{E}{\hbar}t\right) \tag{4.11}$$

Here E is the total energy of a particle which in the case of a stationary field remains constant. To convince ourselves that Eq. (4.11) is true, let us introduce it into Eq. (4.9). As a result, we get the relation

$$-\frac{\hbar^2}{2m}\exp\left(-i\frac{E}{\hbar}t\right)\nabla^2\psi + U\psi\exp\left(-i\frac{E}{\hbar}t\right) = i\hbar\left(-i\frac{E}{\hbar}\right)\psi\exp\left(-i\frac{E}{\hbar}t\right)$$

Cancelling the common factor $\exp\left(-i\frac{E}{\hbar}t\right)$, we arrive at a differential equation determining the function ψ:

$$-\frac{\hbar^2}{2m}\nabla^2\psi+U\psi=E\psi \tag{4.12}$$

Equation (4.12) is known as the **Schrödinger equation for stationary states.** In the following, we shall have to do only with this equation and for brevity's sake we shall call it simply the Schrödinger equation. Equation (4.12) is often written in the form

$$\nabla^2\psi+\frac{2m}{\hbar^2}(E-U)\,\psi=0 \tag{4.13}$$

Let us explain how we can arrive at the Schrödinger equation. We shall limit ourselves to a one-dimensional case for simplicity. We shall consider a freely moving particle. According to de Broglie's idea, a plane wave must be compared with it:

$$\Psi = a\exp\left[-i\left(\omega t-kx\right)\right]$$

(in quantum mechanics, it is customary practice to take the exponent with the minus sign). Replacing ω and $k=2\pi/\lambda$ with E and p in accordance with Eqs. (4.1) and (4.2), we arrive at the expression

$$\Psi=a\exp\left[\frac{i}{\hbar}(px-Et)\right] \tag{4.14}$$

Differentiating this expression once with respect to t, and the second time twice with respect to x, we get

$$\frac{\partial\Psi}{\partial t}=-\frac{i}{\hbar}E\Psi, \qquad \frac{\partial^2\Psi}{\partial x^2}=\left(\frac{i}{\hbar}\right)^2p^2\Psi$$

Hence,

$$E=\frac{1}{\Psi}i\hbar\frac{\partial\Psi}{\partial t}, \qquad p^2=-\frac{1}{\Psi}\hbar^2\frac{\partial^2\Psi}{\partial x^2} \tag{4.15}$$

In non-relativistic mechanics, the energy E and momentum p of a free particle are related by the expression

$$E=\frac{p^2}{2m}$$

Using in this expression Eqs. (4.15) for E and p^2 and then cancelling Ψ, we get the equation

$$-\frac{\hbar^2}{2m}\frac{\partial^2\Psi}{\partial x^2}=i\hbar\frac{\partial\Psi}{\partial t}$$

that coincides with Eq. (4.9) if we assume that $U=0$ in the latter.

For a particle moving in a force field characterized by the potential energy U, the following relation exists between the energy E

and the momentum p:

$$\frac{p^2}{2m} = E - U$$

Extending Eqs. (4.15) for E and p^2 to this case too, we obtain

$$-\frac{1}{\Psi}\frac{\hbar^2}{2m}\frac{\partial^2\Psi}{\partial x^2} = \frac{1}{\Psi} i\hbar \frac{\partial \Psi}{\partial t} - U$$

Multiplying this equation by Ψ and transferring the term $U\Psi$ to the left-hand side, we arrive at the equation

$$-\frac{\hbar^2}{2m}\frac{\partial^2\Psi}{\partial x^2} + U\Psi = i\hbar\frac{\partial\Psi}{\partial t}$$

that coincides with Eq. (4.9).

The above reasoning does not have the validity of a proof and may not be considered as a derivation of the Schrödinger equation. Its object is to show how one could arrive at the establishing of this equation.

A great part is played in quantum mechanics by the concept of an operator. An **operator** is defined as a rule by means of which one function (we shall designate it by f) is correlated with another function (we shall designate it by φ). This is written symbolically as follows:

$$f = \hat{Q}\varphi \tag{4.16}$$

Here $\hat{Q}$ is the symbol of the operator (we could use any other letter with a "cap" over it, for example, $\hat{A}$, $\hat{U}$, or $\hat{M}$, with the same success). In Eq. (4.10), the part of $\hat{Q}$ is played by ∇^2, the part of φ by the function Ψ, and that of f by the right-hand side of the equation.

The symbol of an operator hides a complex of operations by means of which the initial function (φ) is transformed into another function (f). For example, the symbol ∇^2 hides double differentiation with respect to all three coordinates x, y, and z with the following summation of the expressions obtained. An operator may, in particular, represent the multiplication of the initial function φ by a certain function U. Thus, $f = \hat{U}\varphi = U\varphi$, and, consequently, $\hat{U} = U$.

If we consider the function U in Eq. (4.12) as an operator whose action on the psi-function consists in multiplying ψ by U, then Eq. (4.12) can be given the form

$$\hat{H}\psi = E\psi \tag{4.17}$$

In this equation, the symbol $\hat{H}$ stands for an operator equal to the sum of the operators $-(\hbar^2/2m)\nabla^2$ and U:

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + U \tag{4.18}$$

The operator $\hat{H}$ is called a **Hamiltonian.**

The Hamiltonian is an operator of the energy E. In quantum mechanics, operators are also correlated with the other dynamic variables. Accordingly, operators of the coordinates, momentum, angular momentum, etc. are considered. An equation similar to Eq. (4.17) is compiled for each dynamic variable q. It has the form

$$\hat{Q}\psi = q\psi \tag{4.19}$$

where $\hat{Q}$ is the operator being correlated with the dynamic variable q. The meaning of such equations will be revealed in Sec. 4.7.

4.5. The Meaning of the Psi-Function

A correct interpretation of the psi-function was given by the German physicist Max Born (1882-1970) in 1926. He postulated that the square of the magnitude of the psi-function determines the probability dP of the fact that a particle will be detected within the limits of the volume dV:

$$dP = A \mid \Psi \mid^2 dV = A\Psi^*\Psi \, dV \tag{4.20}$$

(A is a constant of proportionality).

The integral of Eq. (4.20) taken over the entire volume must equal unity:

$$\int dP = A \int \Psi^*\Psi \, dV = 1 \tag{4.21}$$

Indeed, this integral gives the probability of the fact that a particle is at some point in space, i.e. the probability of an authentic event, which is unity.

It is assumed in quantum mechanics that the psi-function allows multiplication by an arbitrary complex number C other than zero, Ψ and $C\Psi$ describing the same state of a particle. This circumstance makes it possible to select the psi-function so that it complies with the condition

$$\int \Psi^*\Psi \, dV = 1 \tag{4.22}$$

Condition (4.22) is known as the **normalization condition.** Functions satisfying this condition are called **normalized.** We shall always assume in the following that the psi-functions we are considering are normalized.

Equation (4.20) has the following form for a normalized function:

$$dP = \mid \Psi \mid^2 dV = \Psi^*\Psi \, dV \tag{4.23}$$

[this follows from a comparison of Eqs. (4.21) and (4.22)]. We conclude from Eq. (4.23) that the square of the magnitude of the psi-function gives the **density of the probability** (the probability related to unit volume) of a particle being in the relevant place in space.

The psi-function for a stationary force field has the form of Eq. (4.11). Accordingly,

$$\Psi^*\Psi = \exp\left(i\frac{E}{\hbar}t\right)\psi^* \exp\left(-i\frac{E}{\hbar}t\right)\psi = \psi^*\psi$$

so that the probability density is $\psi^*\psi$ and, consequently, is independent of time. This is why states described by psi-functions of the form given by Eq. (4.11) were called stationary ones.

It can be seen from the meaning of the psi-function that quantum mechanics has a statistical nature. It does not allow us to determine the whereabouts of a particle in space or the trajectory along which a particle is travelling. The psi-function only helps us to predict the probability of finding a particle in different points of space. It may seem at first sight that quantum mechanics provides a considerably less accurate and exhaustive description of the motion of a particle than classical mechanics, which determines the "exact" location and velocity of a particle at every moment of time. Actually, however, this is not true. Quantum mechanics reveals the true behaviour of microparticles to a much deeper extent. It only fails to determine what actually does not occur. As applied to microparticles, the concepts of a definite location and trajectory, as we have already noted, lose their meaning in general.

4.6. Quantization of Energy

The Schrödinger equation allows us to find the psi-function of a given state and, consequently, determine the probability of a particle being at different points in space. But this far from exhausts the significance of the equation. The rules for the quantization of energy directly follow from Eq. (4.17) and from the conditions imposed on the psi-function.

In accordance with its meaning, the psi-function must be single-valued, continuous, and finite (except, perhaps, for special points). In addition, it must have a continuous and finite derivative. The collection of the above requirements is called the **standard conditions.**

The Schrödinger equation includes the total energy E of a particle as a parameter. It is proved in the theory of differential equations that equations of the form of (4.17) have solutions satisfying the standard conditions not at any values of the parameter (i.e. of the energy E), but only at certain selected values. These selected values are known as the **eigenvalues** of the relevant quantity (in our case of the energy). The solutions corresponding to the eigenvalues of E are called the **eigenfunctions** of the problem.

A collection of eigenvalues is called a **spectrum** of a quantity. If this collection forms a discrete succession, the spectrum is called **discrete**. If the eigenvalues form a continuous succession, the spectrum is called **continuous**. In the following, we shall only consider problems in which the spectrum of the eigenvalues is discrete.

With a discrete spectrum, the eigenvalues and eigenfunctions can be numbered:

$$\begin{matrix} E_1, E_2, \ldots, E_n, \ldots; \\ \psi_1, \psi_2, \ldots, \psi_n, \ldots \end{matrix} \qquad (4.24)$$

The quantization of energy is thus obtained from the fundamental tenets of quantum mechanics without any additional assumptions.

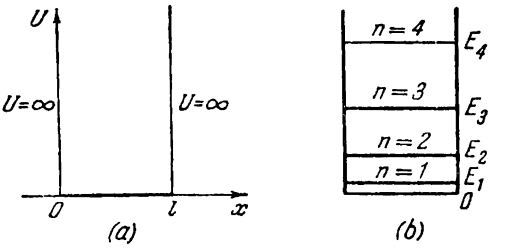


Fig. 4.8

The finding of the eigenvalues and eigenfunctions, as a rule, is a very difficult mathematical task. We shall consider an example that is simple enough to permit us to solve the Schrödinger equation without any appreciable difficulty.

Let us find the eigenvalues of the energy and the eigenfunctions corresponding to them for a particle in an infinitely deep one-dimensional potential well. We shall assume that the particle can move only along the x-axis. Let the motion be restricted by the walls $x = 0$ and $x = l$ that are impenetrable for the particle. The potential energy U has the following form in this case (Fig. 4.8*a*): it is zero at $0 \leqslant x \leqslant l$ and becomes equal to infinity at $x < 0$ and $x > l$.

Let us take the Schrödinger equation in its form (4.13). Since the psi-function depends only on the coordinate x, the equation is simplified as follows:

$$\frac{d^2\psi}{dx^2} + \frac{2m}{\hbar^2}(E-U)\,\psi = 0 \qquad (4.25)$$

The particle cannot get beyond the limits of the potential well. Therefore, the probability of detecting the particle and, consequently, the function ψ beyond the limits of the well are zero. It follows from the condition of continuity that ψ must also be zero at the

boundaries of the well, i.e. that

$$\psi(0) = \psi(l) = 0 \tag{4.26}$$

This is exactly the condition which the solutions of Eq. (4.25) must satisfy.

In the region where ψ does not identically equal zero, Eq. (4.25) has the form

$$\frac{d^2\psi}{dx^2} + \frac{2m}{\hbar^2} E\psi = 0 \tag{4.27}$$

(in this region, $U = 0$). Introducing the notation

$$\omega^2 = \frac{2m}{\hbar^2} E \tag{4.28}$$

we shall arrive at an equation that is well known from the theory of oscillations:

$$\psi'' + \omega^2\psi = 0$$

The solution of such an equation has the form*

$$\psi(x) = a \sin(\omega x + \alpha) \tag{4.29}$$

Conditions (4.26) can be satisfied by the corresponding choice of the constants ω and α. First of all from the condition $\psi(0) = 0$, we get

$$\psi(0) = a \sin \alpha = 0$$

whence we can see that α must equal zero. Further, the condition

$$\psi(l) = a \sin \omega l = 0$$

must be observed, which is possible only when

$$\omega l = \pm n\pi \quad (n = 1, 2, 3, \ldots) \tag{4.30}$$

($n = 0$ drops out because it yields $\psi \equiv 0$—the particle is nowhere).

Excluding ω from Eqs. (4.28) and (4.30), we find the eigenvalues of the energy of a particle:

$$E_n = \frac{\pi^2\hbar^2}{2ml^2} n^2 \quad (n = 1, 2, 3, \ldots) \tag{4.31}$$

The energy spectrum was found to be discrete. Figure 4.8*b* shows an energy level diagram.

Let us assess the spacings between two adjacent levels for different values of the mass of a particle m and the width of the well l. The difference between the energies of two adjacent levels is

$$\Delta E_n = E_{n+1} - E_n = \frac{\pi^2\hbar^2}{2ml^2}(2n+1) \approx \frac{\pi^2\hbar^2}{ml^2} n$$

* See Eq. (7.55) of Vol. I, p. 195. In the present case, it is more convenient to take the sine instead of the cosine.

If we take m of the order of the mass of a molecule ($\sim 10^{-23}$ g), and l of the order of 10 cm (the molecules of a gas in a vessel), we get

$$\Delta E_n \approx \frac{3.14^2 \times 1.05^2 \times 10^{-54}}{10^{-23} \times 10^2} n \approx 10^{-32} n \text{ erg}$$

Such densely arranged energy levels will practically be perceived like a continuous energy spectrum so that although energy quantization indeed occurs in principle, it will not affect the nature of motion of the molecules.

A similar result is obtained if we take m of the order of the mass of an electron ($\sim 10^{-27}$ g) at the same dimensions of the well (free electrons in a metal). In this case

$$\Delta E_n \approx 10^{-28} n \text{ erg} \approx 10^{-16} n \text{ eV}$$

An absolutely different result is obtained for an electron, however, if the region within which it is moving will be of the order of atomic dimensions ($\sim 10^{-8}$ cm). In this case,

$$\Delta E_n \approx \frac{3.14^2 \times 1.05^2 \times 10^{-54}}{10^{-27} \times 10^{-16}} n \approx 10^{-10} n \text{ erg} \approx 10^2 n \text{ eV}$$

so that the discreteness of the energy levels will be quite appreciable.

Introducing in Eq. (4.29) the value of ω obtained from condition (4.30), we shall find the eigenfunctions of the problem:

$$\psi_n(x) = a \sin \frac{n\pi x}{l}$$

(we remind our reader that $\alpha = 0$). To find the coefficient a, let us use normalization condition (4.22), which in the given case will be written as follows:

$$a^2 \int_0^l \sin^2 \frac{n\pi x}{l} dx = 1$$

At the ends of the integration interval, the integrand vanishes. Hence, the value of the integral can be obtained by multiplying the average value of $\sin^2 (n\pi x/l)$ $\left(\text{which, as is known, equals } \frac{1}{2}\right)$ by the length of the interval l. The result is $a^2 \left(\frac{1}{2}\right) l = 1$, whence $a = \sqrt{2/l}$. Thus, the eigenfunctions have the form

$$\psi_n(x) = \sqrt{\frac{2}{l}} \sin \frac{n\pi x}{l} \qquad (n = 1, 2, 3, \ldots) \qquad (4.32)$$

Graphs of the eigenfunctions are shown in Fig. 4.9*a*. Figure 4.9*b* gives the density of the probability of finding the particle at different distances from the walls of the well; it equals $\psi^*\psi$. Inspection of

the graphs reveals, for example, that in the state with $n = 2$ the particle cannot be found at the middle of the well. At the same time, it will be detected either in the left-hand or in the right-hand half of the well with equal frequency. This behaviour of the particle is

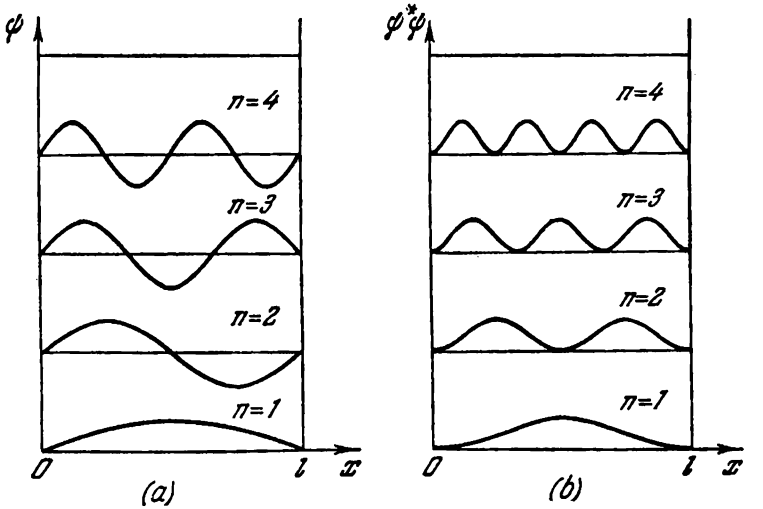


Fig. 4.9

evidently incompatible with our notion of trajectories. We must note that in accordance with classical notions, all the positions of the particle in the well are equally probable.

4.7. Quantization of Angular Momentum

We indicated in Sec. 4.4 that in quantum mechanics an operator $\hat{Q}$ is correlated with every physical quantity q (the operator has a different symbol for each quantity: $\hat{H}$ for energy, $\hat{\mathbf{p}}$ for momentum, etc.) By solving the equation

$$\hat{Q}\psi = q\psi$$

we find the eigenvalues $q_1, q_2, \ldots$ of the operator $\hat{Q}$. One of the postulates of quantum mechanics states that in measurements of the physical quantity q represented by the operator $\hat{Q}$, we can obtain only results coinciding with the eigenvalues of this operator.

States are possible for which measurements of a quantity q always give the same value q_n. Such states are said to be ones in which the quantity q has a definite value. States are also possible, however,

for which measurements give different eigenvalues of the operator $\hat{Q}$ with a different probability. Such states are said to be ones in which the quantity q does not have a definite value.

Four operators are introduced in quantum mechanics conformably to the angular momentum, namely, the operator of the square of the angular momentum $\hat{\mathbf{M}}^2$ and three operators of the projections of the angular momentum onto the coordinate axes: $\hat{M}_x$, $\hat{M}_y$, and $\hat{M}_z$. It was found that only the square of the angular momentum and one of the projections of the angular momentum onto the coordinate axes can simultaneously have definite values. The other two projections are absolutely indefinite*. This signifies that the "vector" of the angular momentum has no definite direction and, consequently, cannot be depicted with the aid of a directed length of a straight line as in classical mechanics.

The solution of the equation

$$\hat{\mathbf{M}}^2\psi = M^2\psi$$

is very difficult. We shall therefore only give the final results: the eigenvalues of the operator of the square of the angular momentum are

$$M^2 = l(l+1)\hbar^2 \quad (l = 0, 1, 2, \ldots) \tag{4.33}$$

Here l is a **quantum number** called the **azimuthal** (or **orbital**) one. Consequently, the magnitude of the angular momentum can have only discrete values determined by the formula

$$M = \hbar\sqrt{l(l+1)} \quad (l = 0, 1, 2, \ldots) \tag{4.34}$$

The operator $\hat{M}_z$ has a quite simple form. We can therefore consider the solution of the equation

$$\hat{M}_z\psi = M_z\psi \tag{4.35}$$

as another example of finding eigenvalues (the first example was treated in the preceding section, where we determined the eigenvalues of the energy for a particle in a potential well).

In spherical coordinates (r, θ, φ), the operator of the projection of the angular momentum onto the polar axis z (from which the polar angle θ is measured) has the form

$$\hat{M}_z = -i\hbar\frac{\partial}{\partial\varphi}$$

Hence, Eq. (4.35) appears as follows:

$$-i\hbar\frac{\partial\psi}{\partial\varphi} = M_z\psi \tag{4.36}$$

* An exception is the case $M = 0$, when all three projections of the angular momentum onto the axes x, y, z have a definite value equal to zero.

The introduction of $\psi = \exp(\alpha\varphi)$, after cancellation of the common factor $\exp(\alpha\varphi)$, leads to the algebraic equation

$$-i\hbar\alpha = M_z$$

from which we get the value $iM_z/\hbar$ for α. The solution of Eq. (4.36) thus has the form

$$\psi = C \exp\left(i \frac{M_z}{\hbar} \varphi\right)$$

For this function to be single-valued, it is necessary to satisfy the condition $\psi(\varphi + 2\pi) = \psi(\varphi)$ or

$$\exp\left[i \frac{M_z}{\hbar} (\varphi + 2\pi)\right] = \exp\left(i \frac{M_z}{\hbar} \varphi\right)$$

This condition will be satisfied if we assume that $M_z = m\hbar$, where m is a positive or negative integer or zero. Hence, the operator $\hat{M}_z$ has a discrete spectrum:

$$M_z = m\hbar \quad (m = 0, \pm 1, \pm 2, \ldots) \tag{4.37}$$

For reasons which will be revealed on a later page, m is called the **magnetic quantum number.** We remind our reader that quantization of the projection of the angular momentum was discovered experimentally by O. Stern and W. Gerlach (see Sec. 7.6 of Vol. II, p. 170).

Since the projection of a vector cannot exceed the magnitude of this vector, the following condition must be observed:

$$|m\hbar| \leqslant \hbar \sqrt{l(l+1)}$$

Hence, it follows that the maximum possible value of $|m|$ is l.

For convenience of reviewing, let us write the results obtained together:

$$\left.\begin{aligned} M &= \hbar \sqrt{l(l+1)} \quad (l = 0, 1, 2, \ldots) \\ M_z &= m\hbar \quad (m = 0, \pm 1, \pm 2, \ldots, \pm l) \end{aligned}\right\} \tag{4.38}$$

Inspection of these formulas shows that $|M_z|$ is always smaller than M. Consequently, the direction of the angular momentum cannot coincide with a direction earmarked in space. This agrees with the circumstance that the direction of the angular momentum in space is indefinite.

We must underline the fact that values of M and M_z differing from Eqs. (4.38) cannot be observed in any circumstances. Hence, the angular momenta of macroscopic bodies also obey rules (4.38). True, owing to the smallness of $\hbar$, the discreteness of the angular momenta of macroscopic bodies is virtually not detected, like the discreteness of macroscopic electric charges is not detected owing to the smallness of the elementary charge e.

We must note that it can be seen from the rules for the quantization of the angular momentum that the Planck constant $\hbar$ can be considered as a natural unit of angular momentum.

The angular momentum of a system consisting of several microparticles equals the sum of the momenta of the individual particles. The net angular momentum, like any angular momentum in general, is determined by the expression

$$M = \hbar \sqrt{L(L+1)} \tag{4.39}$$

where L is the azimuthal (orbital) quantum number of the resultant angular momentum. For a system consisting of two particles, the number L can have the values

$$L = l_1 + l_2,\ l_1 + l_2 - 1,\ \ldots,\ |l_1 - l_2| \tag{4.40}$$

where l_1 and l_2 are numbers determining the magnitudes of the angular momenta being summated according to the formula $M_i = \hbar \sqrt{l_i(l_i+1)}$.

It is a simple matter to see that the resultant angular momentum can have $2l_2 + 1$ or $2l_1 + 1$ different values (the smaller of the two l's must be taken).

For a system consisting of more than two particles, we must first add the angular momenta of any two particles. Next, we must add the result obtained to the angular momentum of a third particle, and so on. It is evident that the maximum value of the quantum number L equals the sum of the numbers l_i for the individual particles. The minimum value of L, for instance, for three particles is $|(|l_1 - l_2| - l_3)|$. If all the l_i's are the same and equal l, then the minimum value of L is zero with an even number of particles and l with an odd number of them.

The projection of the resultant angular momentum onto a certain direction z is determined, as for any angular momentum in general, by the expression

$$M_z = m_L \hbar \qquad (m_L = 0, \pm 1, \pm 2, \ldots, \pm L) \tag{4.41}$$

[see Eqs. (4.38)].

The mechanical angular momentum of a charged particle is inseparably associated with its magnetic moment (see Sec. 7.6 of Vol. II, p. 166 et seq.). The magnetic moments, as we know, interact with one another. A definite value of the interaction energy corresponds to each of the possible values of the resultant moment. When a system experiences a weak magnetic field, the coupling between the moments is not violated, and the resultant moment is projected onto the direction of **B**. When the magnetic field is strong enough, the moments no longer remain coupled, and each of these moments is projected onto the direction of **B** independently of the others.

4.8. The Superposition Principle

One of the main tenets of quantum mechanics is the **principle of superposition of states.** The essence of this principle consists in the following. Assume that a certain quantum-mechanical system can be both in state ψ' and in state ψ''. There is consequentlv a state of the system described by the function

$$\psi = c'\psi' + c''\psi''$$

(c' and c'' are arbitrary complex numbers).

Very important corollaries follow from the superposition principle. Let us consider the collection of eigenvalues of a physical quantity q and the eigenfunctions corresponding to them:

$$q_1, q_2, \ldots, q_n, \ldots;$$

$$\psi_1, \psi_2, \ldots, \psi_n, \ldots$$

In each of the states described by these functions, the quantity q has a definite value: the value q_1 in the state ψ_1, the value q_2 in the state ψ_2, etc. According to the superposition principle, a state described by the function

$$\psi = c_1\psi_1 + c_2\psi_2$$

is possible. In this state, the quantity q no longer has a definite value—measurements will give either the value q_1 or the value q_2. The probabilities of the appearance of these values equal the squares of the magnitudes of the coefficients c_1 and c_2, i.e. the probability of obtaining the result q_1 in measurements is $|c_1|^2$, and the probability of obtaining the result q_2 is $|c_2|^2$ (as we agreed on in Sec. 4.5, the functions ψ_1 and ψ_2 are assumed to be normalized).

It is assumed in quantum mechanics that a collection of eigenfunctions of any physical quantity q forms a **complete set.** This signifies that the psi-function of any state can be expanded by the eigenfunctions of this quantity, i.e. can be written in the form

$$\psi = \sum_n c_n\psi_n \tag{4.42}$$

where c_n are in the general case complex numbers not depending on the coordinates (for a time-varying state, the coefficients c_n do depend on t). The number of addends in the sum equals the number of different eigenfunctions of the quantity q (for different quantities, this number varies from 2 to ∞).

The squares of the magnitudes of the coefficients c_n give the probabilities of obtaining the corresponding values of the quantity q in measurements conducted on a system in the state ψ. Since the sum of all such probabilities must equal unity, the coefficients c_n

satisfy the condition

$$\sum_n |c_n|^2 = 1$$

This condition is always observed for normalized ψ_n's.

Knowing the probabilities of different values of the quantity q, we can find the average value of this quantity in the state ψ:

$$\langle q \rangle = \sum_n |c_n|^2 q_n \tag{4.43}$$

For non-stationary states, $c_n = c_n(t)$; hence, Eq. (4.43) shows how the average value of the quantity q varies with time.

4.9. Penetration of Particles Through a Potential Barrier

Assume that a particle moving from left to right encounters a potential barrier of height U_0 and of width l on its path (Fig. 4.10). According to classical notions, the particle will behave as follows. If the energy of the particle is higher than the height of the barrier ($E > U_0$), the particle passes over the latter without hindrance (on the section $0 \leqslant x \leqslant l$ only the speed of the particle diminishes, but then when $x > l$ it again acquires its initial value). If E is lower than U_0 (as is shown in the figure), then the particle is reflected from the barrier and flies in the reverse direction; the particle cannot penetrate through the barrier.

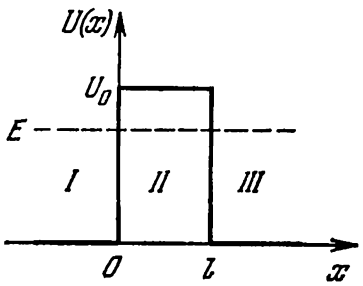


Fig. 4.10

The behaviour of the particle is absolutely different according to quantum mechanics. First, even when $E > U_0$, there is a probability other than zero that the particle will be reflected from the barrier and fly in the reverse direction. Second, when $E < U_0$, there is a probability other than zero that the particle will penetrate "through" the barrier and will be in the region where $x > l$. Such a behaviour of the particle, absolutely impossible from the classical viewpoint, follows directly from the Schrödinger equation.

Let us consider the case $E < U_0$. In this case, Eq. (4.13) has the form

$$\frac{d^2\psi}{dx^2} + \frac{2m}{\hbar^2} E\psi = 0 \tag{4.44}$$

for regions *I* and *III*, and

$$\frac{d^2\psi}{dx^2}+\frac{2m}{\hbar^2}(E-U_0)\,\psi=0 \tag{4.45}$$

for region *II*: as indicated above, $E-U_0<0$.

We shall seek the solution of Eq. (4.44) in the form $\psi=\exp(\lambda x)$ (see Sec. 7.4. of Vol. I, p. 192 et seq.). Introduction of this function into Eq. (4.44) leads to the characteristic equation

$$\lambda^2+\frac{2m}{\hbar^2}E=0$$

Hence, $\lambda=\pm i\alpha$, where

$$\alpha=\frac{1}{\hbar}\sqrt{2mE} \tag{4.46}$$

The general solution of Eq. (4.44) thus has the form

$$\left.\begin{aligned}\psi_1&=A_1e^{i\alpha x}+B_1e^{-i\alpha x}\ \text{for region } I\\ \psi_3&=A_3e^{i\alpha x}+B_3e^{-i\alpha x}\ \text{for region } III\end{aligned}\right\} \tag{4.47}$$

Solving Eq. (4.45) by introducing $\psi=\exp(\lambda x)$, we get a general solution of this equation in the form

$$\psi_2=A_2e^{\beta x}+B_1e^{-\beta x}\ \text{for region } II \tag{4.48}$$

Here

$$\beta=\frac{1}{\hbar}\sqrt{2m(U_0-E)} \tag{4.49}$$

We must note that a solution of the form $\exp(i\alpha x)$ corresponds to a wave propagating in the positive direction of the x-axis, and a solution of the form $\exp(-i\alpha x)$ to a wave propagating in the opposite direction. To understand this, we shall remember that an ordinary (sound, electromagnetic, etc.) plane wave propagating in the direction of a growth in x is described by the real part of the expression $\exp[i(\omega t-kx)]$, while a wave propagating in the direction of diminishing of x is described by the real part of the expression $\exp[i(\omega t+kx)]$. The function $\Psi=a\exp[(i/\hbar)(px-Et)]$ [see Eq. (4.14)] is compared with a particle moving in the positive direction of the x-axis. If we discard the temporal multiplier in this function, then we get the expression $a\exp[i(p/\hbar)x]$ for ψ. For a particle moving in the opposite direction, we get $\psi=a\exp[-i(p/\hbar)x]$.

In region *III*, there is only the wave that has penetrated through the barrier and is propagating from the left to the right. Consequently, in Eq. (4.47) for ψ_3, we must assume that the coefficient B_3 is zero. To find the other coefficients, we shall use the conditions which the function ψ must satisfy. For ψ to be continuous through the entire region of changes in x from $-\infty$ to $+\infty$, the conditions

$\psi_1(0) = \psi_2(0)$ and $\psi_2(l) = \psi_3(l)$ must be satisfied. For ψ to be smooth, i.e. for it to have no breaks, the conditions $\psi_1'(0) = \psi_2'(0)$ and $\psi_2'(l) = \psi_3'(l)$ must be satisfied. From these conditions, we get the relations

$$\left.\begin{aligned} A_1 + B_1 &= A_2 + B_2 \\ A_2 e^{\beta l} + B_2 e^{-\beta l} &= A_3 e^{i\alpha l} \\ i\alpha A_1 - i\alpha B_1 &= \beta A_2 - \beta B_2 \\ \beta A_2 e^{\beta l} - \beta B_2 e^{-\beta l} &= i\alpha A_3 e^{i\alpha l} \end{aligned}\right\}. \tag{4.50}$$

Let us divide all the equations by A_1 and introduce the notation

$$b_1 = \frac{B_1}{A_1}, \qquad a_2 = \frac{A_2}{A_1}, \qquad b_2 = \frac{B_2}{A_1}, \qquad a_3 = \frac{A_3}{A_1}$$

and also

$$n = \frac{\beta}{\alpha} = \sqrt{\frac{U_0 - E}{E}} \tag{4.51}$$

Equations (4.50) thus acquire the form

$$\left.\begin{aligned} 1 + b_1 &= a_2 + b_2 \\ a_2 e^{\beta l} + b_2 e^{-\beta l} &= a_3 e^{i\alpha l} \\ i - ib_1 &= na_2 - nb_2 \\ na_2 e^{\beta l} - nb_2 e^{-\beta l} &= ia_3 e^{i\alpha l} \end{aligned}\right\} \tag{4.52}$$

The ratio between the squares of the magnitudes of the reflected and incident wave amplitudes

$$R = \frac{|B_1|^2}{|A_1|^2} = |b_1|^2$$

determines the probability of a particle being reflected from the potential barrier and can be called the **reflection coefficient.**

The ratio between the squares of the magnitudes of the transmitted and incident wave amplitudes

$$T = \frac{|A_3|^2}{|A_1|^2} = |a_3|^2 \tag{4.53}$$

determines the probability of a particle penetrating through the barrier and can be called the **transmission coefficient.**

We shall be interested only in the penetration of particles through the barrier, and we shall limit ourselves to finding the quantity T. True, having found T, it is simple to find R because these two coefficients are related by the obvious expression $R + T = 1$.

Let us multiply the first of Eqs. (4.52) by i and add it to the third one. The result is

$$2i = (n + i)\, a_2 - (n - i)\, b_2 \tag{4.54}$$

Now let us multiply the second of Eqs. (4.52) by i and subtract it from the fourth one. We get

$$(n-i)\,e^{\beta l}\,a_2-(n+i)\,e^{-\beta l}\,b_2=0 \qquad (4.55)$$

Solving the simultaneous equations (4.54) and (4.55), we find that

$$a_2=\frac{2i\,(n+i)\,e^{-\beta l}}{(n+i)^2\,e^{-\beta l}-(n-i)^2\,e^{\beta l}}$$

$$b_2=\frac{2i\,(n-i)\,e^{\beta l}}{(n+i)^2\,e^{-\beta l}-(n-i)^2\,e^{\beta l}}$$

Finally, introducing the values of a_2 and b_2 which we have found into the second of Eqs. (4.52), we get an expression for a_3:

$$a_3=\frac{4nl}{(n+i)^2\,e^{-\beta l}-(n-i)^2\,e^{\beta l}}\,e^{-i\alpha l}$$

The quantity

$$\beta l=\frac{\sqrt{2m\,(U_0-E)}}{\hbar}\,l$$

is usually much greater than unity. For this reason, we may disregard the addend containing the multiplier $e^{-\beta l}$ in the denominator of the expression for a_3 in comparison with the addend containing the multiplier $e^{\beta l}$ (the complex numbers $n+i$ and $n-i$ have the same magnitude). We can thus assume that

$$a_3\approx-\frac{4nie^{-i\alpha l}}{(n-i)^2}\,e^{-\beta l}$$

According to Eq. (4.53), the square of the magnitude of this quantity gives the probability of the penetration of a particle through the potential barrier. Taking into account that $|n-i|=\sqrt{n^2+1}$, we get

$$T=|a_3|^2\approx\frac{16n^2}{(n^2+1)^2}\,e^{-2\beta l}$$

where

$$n^2=\frac{U_0-E}{E}=\frac{U_0}{E}-1$$

[see Eq. (4.51)].

The expression $16n^2/(n^2+1)^2$ has a magnitude of the order of unity*. We can therefore consider that

$$T\approx\exp(-2\beta l)=\exp\left[-\frac{2}{\hbar}\sqrt{2m\,(U_0-E)}\,l\right] \qquad (4.56)$$

* The function $16x/(x+1)^2$ has a maximum equal to 4 at $x=1$. When x ranges from 0.03 to 30, the values of the function range from 0.5 to 4.

It follows from the expression we have obtained that the probability of a particle penetrating through the potential barrier depends greatly on the width of the barrier l and on its superelevation above E, i.e. on $U_0 - E$. If at a certain width of the barrier the transmission coefficient T equals, say, 0.01, then when the width is doubled, T becomes equal to $0.01^2 = 0.0001$, i.e. diminishes to one-hundredth of its initial value. The same effect in this case would be caused by a four-fold growth in the quantity $U_0 - E$. The transmission coefficient decreases sharply when the mass m of a particle grows.

The relevant calculations show that when the potential barrier has an arbitrary shape (Fig. 4.11), formula (4.56) must be replaced with the more general formula

$$T \approx \exp\left[-\frac{2}{\hbar}\int_0^l \sqrt{2m(U-E)}\,dx\right] \qquad (4.57)$$

where $U = U(x)$.

When a particle overcomes a potential barrier, it passes, as it were, through a "tunnel" in this barrier (see the hatched region in Fig. 4.11), and in this connection the phenomenon we have considered is known as the **tunnel effect.**

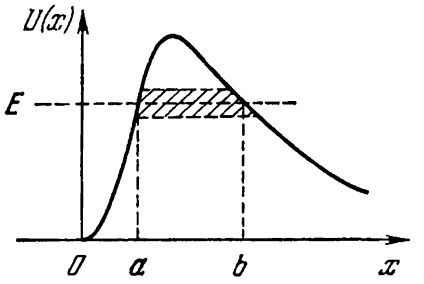


Fig.4.11

The tunnel effect is absurd from the classical viewpoint because a particle "in the tunnel" ought to have a negative kinetic energy (in the tunnel $E < U$). The tunnel effect, however, is aspecifically quantum phenomenon having no analogue in classical physics. In quantum mechanics, the division of the total energy into kinetic and potential energies has no sense because it contradicts the uncertainty principle. Indeed, the fact that a particle has a definite kinetic energy E_k would be equivalent to the particle having a definite momentum p. Similarly, the fact that a particle has a definite potential energy U would signify that the particle is in an exactly given place in space. Since the coordinate and the momentum of a particle cannot simultaneously have definite values, it is impossible to simultaneously find exact values of E_k and U. Thus, although the total energy of a particle E has a quite definite value, it cannot be represented in the form of the sum of the exactly determined energies E_k and U. It is clear that in this case the conclusion on E_k being negative "inside" the tunnel becomes groundless.

4.10. Harmonic Oscillator

A harmonic oscillator is defined as a particle performing one-dimensional motion under the action of the quasi-elastic force $F = -kx$. The potential energy of such a particle has the form

$$U = \frac{kx^2}{2} \tag{4.58}$$

The natural frequency of a classical harmonic oscillator is $\omega = \sqrt{k/m}$, where m is the mass of the particle (see Sec. 7.10 of Vol. I, p. 210). Expressing k through m and ω in Eq. (4.58), we have

$$U = \frac{m\omega^2 x^2}{2}$$

In the one-dimensional case, $\nabla^2\psi = d^2\psi/dx^2$. Therefore, the Schrödinger equation [see Eq. (4.13)] for an oscillator has the following form:

$$\frac{d^2\psi}{dx^2} + \frac{2m}{\hbar^2}\left(E - \frac{m\omega^2 x^2}{2}\right)\psi = 0 \tag{4.59}$$

(E is the total energy of the oscillator).

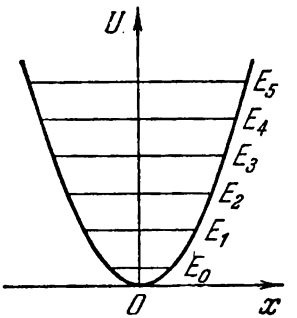


Fig. 4.12

It is proved in the theory of differential equations that Eq. (4.59) has finite, unambiguous, and continuous solutions at values of the parameter E equal to

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega \qquad (n = 0, 1, 2, \ldots) \tag{4.60}$$

Figure 4.12 shows schematically the energy levels of a harmonic oscillator. For purposes of illustration, the levels have been inscribed in the potential energy curve. It must be remembered, however, that in quantum mechanics the total energy cannot be represented in the form of the sum of exactly determined energies E_k and U (see the last paragraph of the preceding section).

The energy levels of a harmonic oscillator are equidistant, i.e. are equal distances apart. The smallest possible value of the energy is $E_0 = \frac{1}{2}\hbar\omega$. This value is called the **zero energy**. The existence of zero energy is confirmed by experiments studying the scattering of light by crystals at low temperatures. It was found that the intensity of scattered light with decreasing temperature tends not to zero, but to a certain finite value. This indicates that even at absolute zero, the oscillations of the atoms in a crystal lattice do not stop.

Quantum mechanics allows us to calculate the probability of various transitions of a quantum system from one state to another.

Such calculations show that for a harmonic oscillator only transitions between adjacent levels are possible. In such transitions, the quantum number n changes by unity:

$$\Delta n = \pm 1 \tag{4.61}$$

The conditions imposed on the changes in the quantum numbers upon transitions of a system from one state to another are known as the **selection rules.** Thus, a selection rule expressed by formula (4.61) exists for a harmonic oscillator.

It follows from rule (4.61) that the energy of a harmonic oscillator can change only in portions of $\hbar\omega$. This result, which is obtained naturally in quantum mechanics, coincides with the very alien assumption for classical physics which Planck had to make in order to calculate the emissivity of a blackbody (see Sec. 1.7). We must note that Planck assumed the energy of a harmonic oscillator to be only an integral multiplier of $\hbar\omega$. Actually, there is also a zero energy whose existence was established only after the appearance of quantum mechanics.

CHAPTER 5 THE PHYSICS OF ATOMS AND MOLECULES

5.1. The Hydrogen Atom

Let us consider a system formed by a stationary nucleus having the charge Ze (where Z is an integer) and an electron in motion around it. When $Z > 1$, such a system is known as a hydrogen-like ion; when $Z = 1$, it is a hydrogen atom.

The potential energy of an electron is

$$U = -\frac{Ze^2}{r}$$

(r is the distance to the electron from the nucleus). Hence, the Schrödinger equation has the form

$$\nabla^2\psi + \frac{2m_e}{\hbar^2}\left(E + \frac{Ze^2}{r}\right)\psi = 0 \qquad (5.1)$$

(m_e is the mass of an electron).

The field in which the electron travels is a centrally symmetrical one. It is therefore expedient to use a spherical coordinate system: r, θ, φ. Using the expression for the Laplacian operator in spherical coordinates in Eq. (5.1), we arrive at the equation

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right) + \\ + \frac{1}{r^2\sin^2\theta}\frac{\partial^2\psi}{\partial\varphi^2} + \frac{2m_e}{\hbar^2}\left(E + \frac{Ze^2}{r}\right)\psi = 0 \qquad (5.2)$$

It can be shown that Eq. (5.2) has the required (i.e. single-valued, finite, and continuous) solutions in the following cases: (1) at any positive values of E, and (2) at discrete negative values of the energy equal to

$$E_n = -\frac{m_e e^4}{2\hbar^2}\frac{Z^2}{n^2} \qquad (n = 1, 2, 3, \ldots) \qquad (5.3)$$

The case $E > 0$ corresponds to an electron flying near the nucleus and again travelling away to infinity. The case $E < 0$ corresponds to an electron bound to the nucleus. A comparison with Eq. (3.31) shows that quantum mechanics leads to the same values of the energy of a hydrogen atom that were obtained in Bohr's theory. In quantum mechanics, however, these values are obtained as a corollary of the fundamental tenets of this science. Bohr, on the

other hand, had to introduce special additional assumptions to obtain such a result.

The eigenfunctions of Eq. (5.2) contain three integral parameters n, l, and m:

$$\psi = \psi_{nlm}(r, \theta, \varphi) \tag{5.4}$$

The parameter n, called the **principal quantum number,** coincides with the number of the energy level [see Eq. (5.3)]. The parameters l and m are the azimuthal and magnetic quantum numbers determining by Eqs. (4.38) the magnitude of the angular momentum and the projection of the angular momentum onto a certain direction z.

Solutions satisfying the standard conditions are obtained only for values of l not exceeding $n - 1$. Hence, at a given n, the quantum number l can take on n different values:

$$l = 0, 1, 2, \ldots, n - 1$$

At a given l, the quantum number m can take on $2l + 1$ different values:

$$m = -l, -l + 1, \ldots, -1, 0, +1, \ldots, l - 1, l$$

[see Eqs. (4.38)].

According to Eq. (5.3), the energy of an electron depends only on the principal quantum number n. Hence, several eigenfunctions ψ_{nlm} differing in the values of the quantum numbers l and m correspond to each eigenvalue of the energy E_n (except for E_1). This signifies that a hydrogen atom can have the same value of the energy while being in several different states. Table 5.1 gives the states corresponding to the first three energy levels.

Table 5.1

Energy level E_n	Psi-function ψ_{nlm}	Value			Energy level E_n	Psi-function ψ_{nlm}	Value		
		n	l	m			n	l	m
E_1	ψ_{100}	1	0	0	E_3	ψ_{300}	3	0	0
						ψ_{31-1}	3	1	−1
						ψ_{310}	3	1	0
						ψ_{31+1}	3	1	+1
E_2	ψ_{200}	2	0	0		ψ_{32-2}	3	2	−2
	ψ_{21-1}	2	1	−1		ψ_{32-1}	3	2	−1
	ψ_{210}	2	1	0		ψ_{320}	3	2	0
	ψ_{21+1}	2	1	+1		ψ_{32+1}	3	2	+1
						ψ_{32+2}	3	2	+2

States having the same energy are called **degenerate,** and the number of different states with a certain value of the energy is known as the **degree of degeneracy** of the corresponding energy level.

It is simple to calculate the degree of degeneracy of hydrogen levels on the basis of the possible values for l and m. For each of the n values of the quantum number l there are $2l+1$ values of the quantum number m. Hence, the number of different states corresponding to the given n is

$$\sum_{l=0}^{n-1}(2l+1)=n^2$$

Thus, the degree of degeneracy of the energy levels in a hydrogen atom is n^2 (see Table 5.1), or, in other words, the energy levels of a hydrogen atom are n^2-fold degenerate.

States with different values of the azimuthal quantum number l differ in the magnitude of the angular momentum. In atomic physics, symbols are used for the states of an electron with different magnitudes of the angular momentum that have been borrowed from spectroscopy. An electron in a state with $l = 0$ is called an s-electron (the corresponding state is the s-state), with $l = 1$ is called a p-electron, with $l = 2$ a d-electron, with $l = 3$ an f-electron, then come g, h and so on according to the alphabet. The value of the principal quantum number is indicated before the symbol of the quantum number l. Thus, an electron in a state with $n = 3$ and $l = 1$ is designated by the symbol $3p$, etc.

Since l is always smaller than n, the following states of an electron are possible:

$$1s,$$
$$2s,\ 2p$$
$$3s,\ 3p,\ 3d$$
$$4s,\ 4p,\ 4d,\ 4f$$

and so on.

The energy levels could be depicted schematically as was done in Sec. 3.6 (see Fig. 3.13). But it is much more convenient to use the diagram shown in Fig. 5.1. This diagram reflects (true, only partly) the degeneracy of the levels. It also has a number of other appreciable advantages that will soon become evident.

We know that the emission and absorption of light occur upon transitions of an electron from one level to another. It is proved in quantum mechanics that the following selection rule holds for the azimuthal quantum number l:

$$\Delta l = \pm 1 \qquad (5.5)$$

This signifies that only such transitions are possible in which l changes by unity. Rule (5.5) is due to the fact that a photon has an intrinsic angular momentum (spin*) approximately equal to $\hbar$ (we

* Spin will be treated in Sec. 5.4.

shall define its value more precisely on a later page). When a photon is emitted, it carries this momentum along with it from the atom,

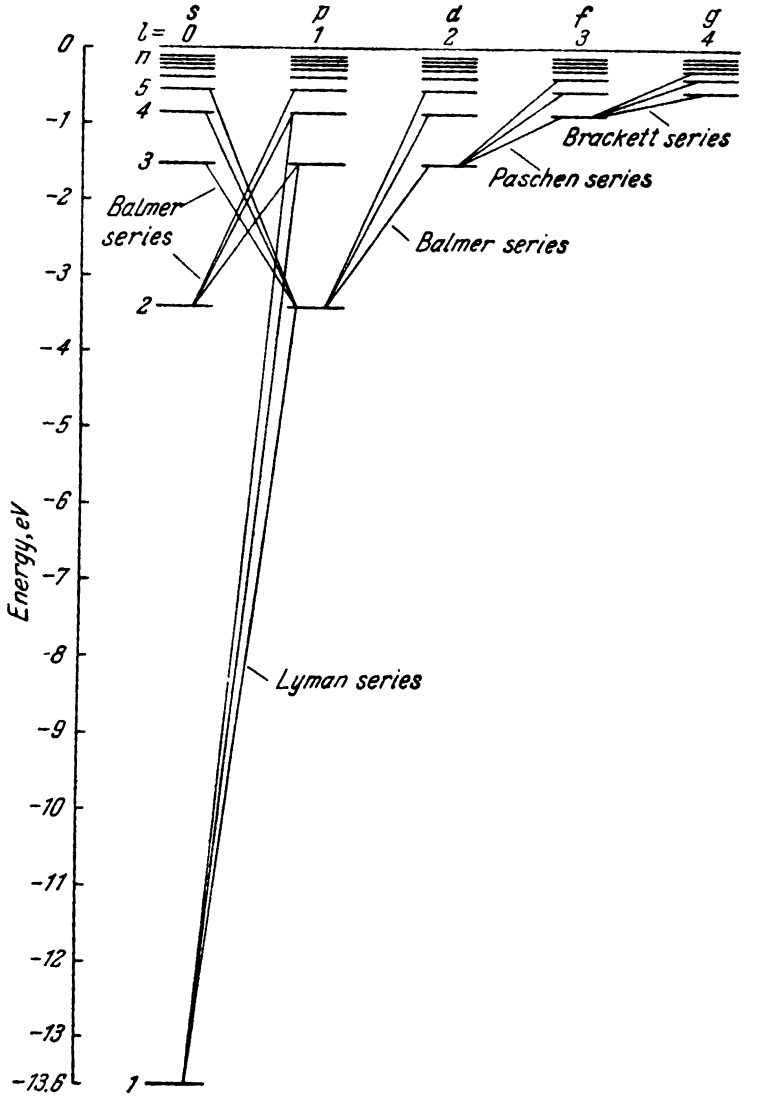


Fig. 5.1

while when it is absorbed it gives up this momentum, so that selection rule (5.5) is simply a corollary of the law of angular momentum conservation.

Figure 5.1 shows the transitions allowed by rule (5.5). Using the symbols of the electron states, we can write the transitions resulting in the production of the Lyman series in the form

$$np \to 1s \quad (n = 2, 3, \ldots)$$

the transitions

$$ns \to 2p \text{ and } nd \to 2p \quad (n = 3, 4, \ldots)$$

correspond to the Balmer series, etc.

The state $1s$ is the ground state of the hydrogen atom. The latter has the minimum energy in this state. To transfer an atom from its ground state to an excited one (i.e. to a state with a greater energy), it must be supplied with energy. This can be done at the expense of the thermal collision of atoms (this is why heated bodies glow—the atoms radiate upon returning from an excited state to the ground one), or at the expense of a collision of an atom with a sufficiently fast electron, or, finally, at the expense of the absorption of a photon by an atom.

A photon vanishes when it is absorbed by an atom, giving up to the latter all of its energy. An atom cannot absorb only a part of a photon because a photon, like an electron, and like other elementary particles, is indivisible. Consequently, in the absence of multiphoton processes (see Sec. 5.17), an atom can absorb only the photons whose energy corresponds exactly* to the difference between the energies of two of its levels. Since the absorbing atom is usually in the ground state, the absorption spectrum of the hydrogen atom must consist of lines corresponding to the transitions

$$1s \to np \quad (n = 2, 3, \ldots)$$

This result completely agrees with experimental data.

The eigenfunctions of Eq. (5.2) break up into two multipliers, one of which depends only on r, and the other only on the angles θ and φ:

$$\psi_{nlm} = R_{nl}(r)\, Y_{lm}(\theta, \varphi) \tag{5.6}$$

The multiplier $R_{nl}(r)$ is real and depends on the quantum numbers n and l, while the multiplier $Y_{lm}(\theta, \varphi)$ is complex and depends on the quantum numbers l and m.

The function $Y_{lm}(\theta, \varphi)$ is the eigenfunction of the operator of the square of the angular momentum. This function is constant for the s-states of an electron (i.e. for the states with an angular momentum equal to zero), so that a psi-function of the form ψ_{n00} depends only on r.

An element of volume in a spherical coordinate system equal to $dV = r^2 \sin\theta\, dr\, d\theta\, d\varphi$ can be represented in the form $dV = r^2\, dr\, d\Omega$,

* More correctly, with an accuracy up to a small correction that will be introduced in Sec. 5.3.

where $d\Omega = \sin\theta\, d\theta\, d\varphi$ is an element of a solid angle. Therefore, the condition of normalization of function (5.6) can be written as follows:

$$\int \psi^*_{nlm}\psi_{nlm}\, dV = \int_0^\infty R^2_{nl} r^2\, dr \int_{(4\pi)} Y^*_{lm} Y_{lm}\, d\Omega = 1 \qquad (5.7)$$

(the integral with respect to $d\Omega$ is taken over a complete solid angle equal to 4π). The eigenfunctions of the operator $\hat{M}^2$ are assumed to be normalized; this signifies that

$$\int_{(4\pi)} Y^*_{lm} Y_{lm}\, d\Omega = 1 \qquad (5.8)$$

Consequently, the condition of normalization of the function $R_{nl}(r)$ follows from Eq. (5.7):

$$\int_0^\infty R^2_{nl} r^2\, dr = 1 \qquad (5.9)$$

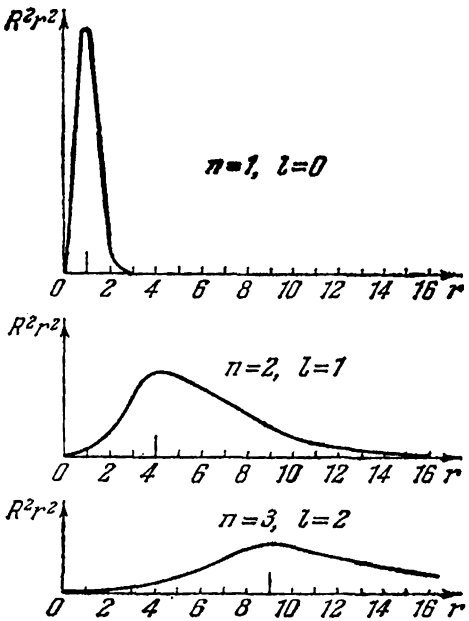


Fig. 5.2

The probability of an electron being in the volume element $dV = r^2 \sin\theta\, dr\, d\theta\, d\varphi = r^2\, dr\, d\Omega$ is determined by the expression

$$dP_{r,\theta,\varphi} = R^2_{nl} r^2\, dr\, Y^*_{lm} Y_{lm}\, d\Omega$$

Integrating this expression with respect to a complete solid angle 4π, we shall find the probability dP_r of an electron being in a thin spherical layer of radius r and thickness dr:

$$dP_r = R^2_{nl} r^2\, dr \int_{(4\pi)} Y^*_{lm} Y_{lm}\, d\Omega$$

Taking into account condition (5.8), we obtain

$$dP_r = R^2_{nl} r^2\, dr \qquad (5.10)$$

Examination of formula (5.10) shows that the expression $R^2_{nl} r^2$ is the density of the probability of an electron being at a distance r from the nucleus. Figure 5.2 gives graphs of the probability density for the hydrogen atom ($Z = 1$) for the states (1) $n = 1$, $l = 0$, (2) $n = 2$, $l = 1$, and (3) $n = 3$, $l = 2$. The scale unit for the r-axis is the Bohr radius r_0 [see Eq. (3.30)]. The long vertical strokes

on the r-axes of the graphs indicate the radii of the relevant Bohr orbits. A glance at the figure shows that these radii coincide with the most probable distances from the electron to the nucleus.

5.2. Spectra of the Alkali Metals

The emission spectra of alkali metal atoms, like the spectrum of hydrogen, consist of several series of lines. The most intensive of them have been named the **principal, sharp, diffuse,** and **fundamental** (or Bergmann) series. These names have the following origin. The principal series owes its name to its also being observed in absorption. Hence, it corresponds to transitions of an atom to its ground state. The sharp and diffuse series consist respectively of sharp and blurred (diffuse) lines. The Bergmann series was called fundamental because of its similarity with the hydrogen series.

The lines of the series of a sodium (Na) atom can be represented as transitions between the energy levels depicted in Fig. 5.3. This diagram differs from that of the hydrogen atom levels (see Fig. 5.1) in that similar levels in different sets are at different heights. Notwithstanding this distinction, both diagrams have a great similarity. This similarity gives us grounds to assume that the spectra of the alkali metals are emitted upon transitions of the outermost (the so-called **valence** or **outer**) electron from one level to another.

Inspection of Fig. 5.3 reveals that the energy of a state depends, apart from the quantum number n, also on the set which the given term is in, i.e. on the number of the set of terms. In the diagram of the hydrogen atom levels, the different sets of terms (with levels coinciding in height) differ in the values of the angular momentum of an electron. It is natural to assume that the different sets of terms of the alkali metals also differ in the values of the angular momentum of the valence electron. Since the levels of different sets in this case are at different heights, it should be assumed that the energy of a valence electron in an alkali metal atom depends on the magnitude of the angular momentum of the electron (which we did not observe for hydrogen).

The assumption that the energy of a valence electron of alkali metal atoms depends on the quantum number l (i.e. on the value of M) is confirmed by quantum-mechanical calculations. For atoms more complicated than the hydrogen atom, we may consider that each of the electrons moves in the averaged field of the nucleus and the remaining electrons. This field will no longer be a Coulomb one (i.e. be proportional to $1/r^2$), but has central symmetry (depends only on r). Indeed, depending on the degree of penetration of an electron into an atom, the charge of the nucleus will be screened for the given electron by the other electrons to some extent or other so

that the effective charge which the electron under consideration experiences will not be constant. At the same time, since electrons travel at tremendous speeds in an atom, the time-averaged field can be considered as a centrally symmetrical one.

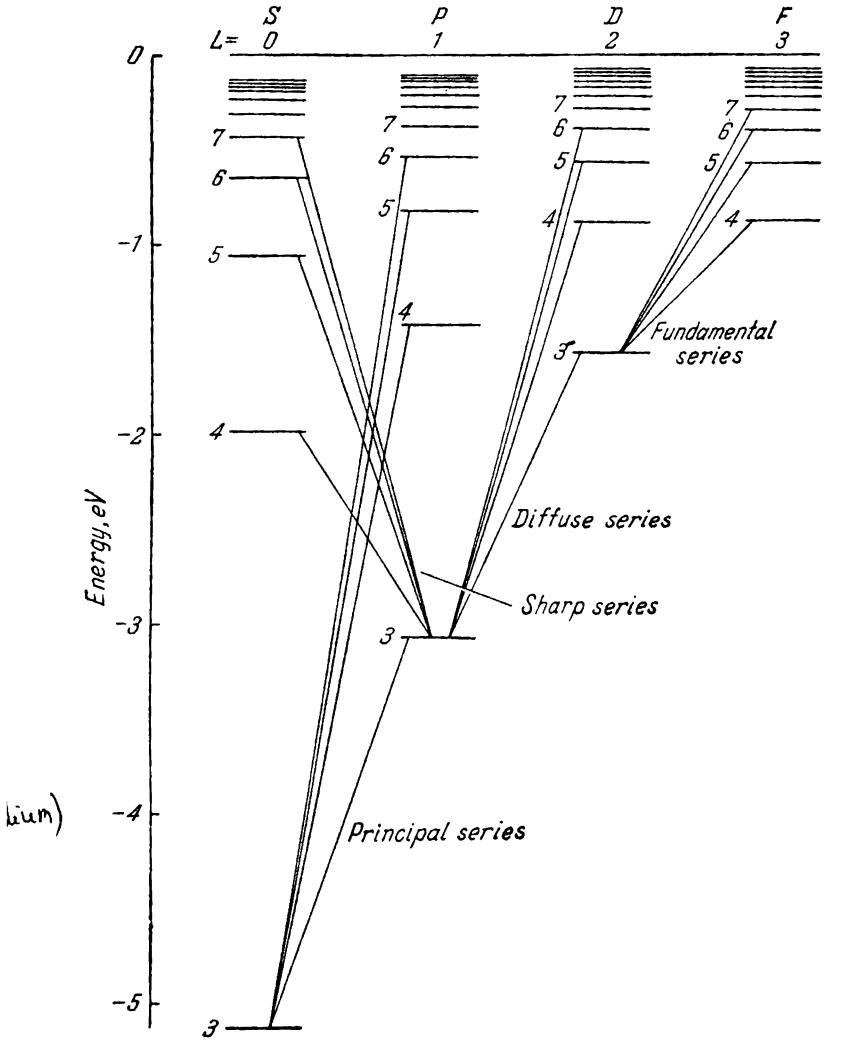


Fig. 5.3

The solution of the Schrödinger equation for an electron travelling in a centrally symmetrical non-Coulomb field gives a result similar to that for a hydrogen atom. A difference, however, is that the energy

levels depend not only on the quantum number n, but also on the quantum number l:

$$E = E_{nl}$$

Thus, in this case, the degeneracy with respect to l is removed. The difference in energy between states with different l's and identical n's is in general not so great as between states with different n's. With an increase in l, the energy of levels with identical n's grows.

The angular momentum of an atom as a whole is the sum of the angular momenta of all the electrons in the atom. The value of the resultant momentum is determined by the quantum number L (see Sec. 4.7). A different value of L corresponds to each column of levels in Fig. 5.3.

The symbols S, P, D, F used in the diagram in Fig. 5.3 are the initial letters of the series names: *s*harp, *p*rincipal, *d*iffuse, *f*undamental. Each of the series is produced at the expense of transitions from levels belonging to the corresponding set. After it had been established that different sets of levels differ in the value of the quantum number L, the symbols S, P, D, F (or s, p, d, f) were used to designate states with the corresponding values of L (or l).

Investigations of the optical spectra of alkali metal ions showed that the angular momentum of the atomic residue (i.e. of the nucleus and the remaining electrons except for the most loosely attached valence electron that leaves the atom in ionization) is zero. Hence, the angular momentum of an alkali metal atom equals that of its valence electron, and L of the atom coincides with l of this electron.

The same selection rule is in force for l of the valence electron of alkali metal atoms as for l of the hydrogen atom electron [see formula (5.5)].

When an alkali metal atom is excited and when it emits light, only the state of the valence electron changes. Therefore, the diagram of the levels of an alkali metal atom may be considered identical to that of the levels of the valence electron*.

Let us denote the terms corresponding to the columns of the levels labelled S, P, D, F in Fig. 5.3 by the symbols nS, nP, nD, and nF. According to Eq. (3.6), the frequency of a spectral line equals the difference between the terms of the final and the initial states. Hence, the spectral series of sodium can be represented in the following form:

$$\left.\begin{array}{lll} \text{sharp series:} & \omega = 3P - nS & (n = 4, 5, \ldots) \\ \text{principal series:} & \omega = 3S - nP & (n = 3, 4, \ldots) \\ \text{diffuse series} & \omega = 3P - nD & (n = 3, 4, \ldots) \\ \text{fundamental series} & \omega = 3D - nF & (n = 4, 5, \ldots) \end{array}\right\} \quad (5.11)$$

* The reasons why we ascribed a value of the principal quantum number equal to 3 (see Fig. 5.3) to the fundamental state of the valence electron of a sodium atom will be revealed on a later page.

Back at the end of the last century, Rydberg established that the terms of the alkali metals with a high degree of accuracy can be represented with the aid of the empirical formula

$$T(n) = \frac{R}{(n+\alpha)^2} \tag{5.12}$$

Here R is the Rydberg constant [see Eq. (3.3)], n is the principal quantum number, and α is a fractional number called the **Rydberg correction.** This correction has a constant value for a given set of terms. It is designated by the same letter used to denote the corresponding set of terms—the letter s for the S-terms, the letter p for the P-terms, etc. The values of the corrections are determined experimentally. They are different for different alkali metals. These values for sodium are

$$s = -1.35, \quad p = -0.87, \quad d = -0.01, \quad f = 0.00 \tag{5.13}$$

We must note that the term given by Eq. (5.12) differs from the term of the hydrogen atom [see Eq. (3.5)] only in the presence of the correction α. For F-terms, this correction is zero. Consequently, the fundamental series (appearing in transitions from the F-levels) is hydrogen-like.

Introducing the empirical expressions into Eqs. (5.11), we get the following formulas for the frequencies of the spectral series of nitrogen:

sharp series—

$$\omega = \frac{R}{(3+p)^2} - \frac{R}{(n+s)^2} \qquad (n = 4, 5, \ldots)$$

principal series—

$$\omega = \frac{R}{(3+s)^2} - \frac{R}{(n+p)^2} \qquad (n = 3, 4, \ldots)$$

diffuse series—

$$\omega = \frac{R}{(3+p)^2} - \frac{R}{(n+d)^2} \qquad (n = 3, 4, \ldots)$$

fundamental series—

$$\omega = \frac{R}{(3+d)^2} - \frac{R}{(n+f)^2} \qquad (n = 4, 5, \ldots)$$

The corrections s, p, d, f in the these formulas have the values given by Eqs. (5.13).

5.3. Breadth of Spectral Lines

An atom can transfer spontaneously from an excited state to a lower energy state. The time τ during which the number of atoms in a given excited state diminishes to $1/e$-th of its initial value is

called the **lifetime** of the excited state*. The lifetime of the excited states of atoms is of the order of 10^{-8} to 10^{-9} s. The lifetime of metastable states may reach tenths of a second.

The possibility of spontaneous transitions indicates that excited states cannot be considered as strictly stationary ones. Accordingly, the energy of an excited state is not exactly definite, and an excited

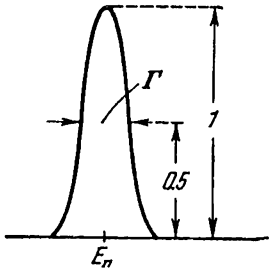


Fig. 5.4

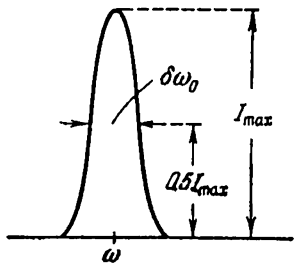


Fig. 5.5

energy level has the finite breadth Γ (Fig. 5.4). According to formula (4.5), the uncertainty of the energy Γ is associated with the lifetime of a state τ by the relation $\Gamma \cdot \tau \sim \hbar$. The breadth of a level is thus determined by the expression

$$\Gamma = \frac{\hbar}{\tau} \tag{5.14}$$

(we have written the equality sign for definiteness).

The ground state of an atom is stationary (a spontaneous transition from it to other states is impossible). Therefore, the energy of the ground state is determined quite accurately.

Owing to the finite breadth of the excited levels, the energy of the photons emitted by atoms is scattered as described by the curve depicted in Fig. 5.4. The spectral line (Fig. 5.5) accordingly has a finite breadth**:

$$\delta\omega_0 = \frac{\Gamma}{\hbar} = \frac{1}{\tau} \tag{5.15}$$

Taking $\tau \sim 10^{-8}$ s, we obtain a value of the order of 10^8 rad/s for $\delta\omega_0$.

The frequency interval $\delta\omega_0$ is related to the wavelength interval $\delta\lambda_0$ by the expression

$$\delta\lambda_0 = \frac{2\pi c}{\omega^2}\,\delta\omega_0 = \frac{\lambda}{\omega}\,\delta\omega_0 = \frac{\lambda^2}{2\pi c}\,\delta\omega_0 \tag{5.16}$$

* The lifetime determined in this way coincides with the average time spent by atoms in the excited state.

** The breadth of a spectral line $\delta\omega$ is determined as the difference between the frequencies which an intensity equal to half the intensity at the maximum corresponds to. In this connection, $\delta\omega$ is sometimes called the half-breadth of a spectral line. We shall use the term "breadth of a line".

(we have omitted the minus sign). Introducing $\lambda = 5000$ Å and $\delta\omega_0 = 10^8$ rad/s, we get a value of the order of 10^{-4} Å for $\delta\lambda_0$.

Expressions (5.15) and (5.16) determine the so-called **natural breadth** of a spectral line. The natural breadth is characterized by the values

$$\delta\omega_0 \sim 10^8 \text{ rad/s}; \quad \delta\lambda_0 \sim 10^{-4} \text{ Å} \tag{5.17}$$

The thermal motion of emitting atoms leads to the so-called **Doppler broadening** of spectral lines. Assume that at the moment of emission of a photon an atom has the momentum $\mathbf{p}_0$ and, accordingly, the energy of translational motion $p_0^2/2m_a$ (here m_a is the mass of an atom). The photon carries along with it the momentum $\hbar\mathbf{k}$ equal in magnitude to $\hbar\omega/c$. Hence, the momentum of the atom changes and becomes equal to $\mathbf{p} = \mathbf{p}_0 - \hbar\mathbf{k}$. Consequently, the energy of the translational motion of the atom changes too. The atom receives the recoil energy equal to

$$E_{\text{rec}} = \frac{(\mathbf{p}_0 - \hbar\mathbf{k})^2}{2m_a} - \frac{p_0^2}{2m_a} = \frac{(\hbar\mathbf{k})^2}{2m_a} - \frac{\mathbf{p}_0\hbar\mathbf{k}}{m_a} \tag{5.18}$$

Let us replace k with ω/c. In addition, we shall take into account that $\mathbf{p}_0/m_a$ is the velocity $\mathbf{v}_0$ of the atom prior to emission.

As a result, Eq. (5.18) acquires the form

$$E_{\text{rec}} = \frac{(\hbar\omega)^2}{2m_a c^2} - \frac{v_0}{c}\hbar\omega\cos\alpha \tag{5.19}$$

where α is the angle between the vectors $\mathbf{p}_0$ and $\mathbf{k}$, i.e. the angle between the direction of motion of the atom and the direction in which the photon is emitted.

Let ΔE_{nm} stand for the decrement of the internal energy of the atom, i.e. the difference $E_n - E_m$, where E_n and E_m are the values of the energies of the levels between which the transition occurs. On the basis of the law of energy conservation, ΔE_{nm} must equal the sum of the energy of the photon and the recoil energy acquired by the atom upon emission:

$$\Delta E_{nm} = \hbar\omega + E_{\text{rec}} \tag{5.20}$$

If atoms upon emission did not experience recoil, they would emit photons of the frequency ω_0. The value of this frequency is obtained from the condition

$$\hbar\omega_0 = \Delta E_{nm} \tag{5.21}$$

We must note that in the preceding sections by ω we meant ω_0.

The recoil energy for visible light is about 10^{-11} of the energy of an emitted photon. For gamma-quanta with $\hbar\omega = 100$ keV, the recoil energy is 10^{-6} of the energy of a photon. Therefore, we may substitute ω_0 for ω in Eq. (5.19). We assume that the velocity v_0

equals the average velocity v of the thermal motion of molecules. As a result, we get

$$E_{\text{rec}} = \frac{(\hbar\omega_0)^2}{2m_a c^2} - \frac{v}{c}\,\hbar\omega_0 \cos\alpha \tag{5.22}$$

The average value of this expression equals the first addend ($\cos\alpha$ takes on all the values from -1 to $+1$ with equal probability, owing to which the second addend is zero on an average).

Thus, denoting the average recoil energy acquired by an atom upon the emission of a photon by the symbol R, we can write:

$$R = \langle E_{\text{rec}} \rangle = \frac{(\hbar\omega_0)^2}{2m_a c^2} = \frac{\Delta E_{\hbar m}^2}{2m_a c^2} \tag{5.23}$$

With account taken of Eq. (5.23), we can represent Eq. (5.22) as follows:

$$E_{\text{rec}} = R - \frac{v}{c}\,\hbar\omega_0 \cos\alpha \tag{5.24}$$

It follows from Eqs. (5.20) and (5.21) that

$$\hbar\omega = \hbar\omega_0 - E_{\text{rec}}$$

Using Eq. (5.24) for E_{rec} in this expression and dividing the relation obtained by $\hbar$, we arrive at the formula

$$\omega = \omega_0 - \frac{R}{\hbar} + \frac{v}{c}\,\omega_0 \cos\alpha \tag{5.25}$$

Let us introduce the notation

$$\Delta\omega_R = \frac{R}{\hbar} = \frac{\hbar\omega_0^2}{2m_a c^2} \approx \frac{\hbar\omega^2}{2m_a c^2} \tag{5.26}$$

$$\delta\omega_D = 2\frac{v}{c}\,\omega_0 \approx 2\frac{v}{c}\,\omega \tag{5.27}$$

Using this notation, we can represent formula (5.25) in the form

$$\omega = \omega_0 - \Delta\omega_R + \frac{1}{2}\,\delta\omega_D \cos\alpha \tag{5.28}$$

In a source of radiation in which all the directions of thermal motion of the atoms are equally probable, the frequencies of the emitted photons will be confined within the limits of the interval $\delta\omega_D$. Consequently, Eq. (5.27) gives the **Doppler breadth** of a spectral line. A glance at Eq. (5.27) shows that the relative Doppler broadening of the lines $\delta\omega_D/\omega$ does not depend on the frequency and equals $2\,(v/c)$ [compare with Eq. (21.15) of Vol. II, p. 483, taking into account that $\Delta\omega$ in this equation corresponds to a half of $\delta\omega_D$].

According to Eq. (5.16), $\delta\lambda/\lambda = \delta\omega/\omega$. The average velocity of the atoms (with a relative atomic mass of about 100) at a temperature of the order of several thousand kelvins is approximately 10^3 m/s.

In these conditions, the Doppler breadth of a spectral line for $\lambda = 5000$ Å will be

$$\delta\lambda_D = 2\frac{v}{c}\lambda = 2 \times \frac{10^3}{3 \times 10^8} \times 5000 \approx 3 \times 10^{-2} \text{ Å}$$

[compare with expression (5.17)].

The actual breadth of a spectral line $\delta\omega$ is the sum of the natural breadth given by Eq. (5.15) and the Doppler breadth given by (5.27):

$$\delta\omega = \delta\omega_0 + \delta\omega_D$$

The middle of the line corresponds to the frequency $\omega_0 - \Delta\omega_R$ [see Eq. (5.28)]. The quantity ω_0 is the frequency which a photon would have provided that the energy ΔE_{nm} were completely used for radiation. The receiving by an atom of the recoil energy R in the emission of energy leads to shifting of the spectral line toward lower frequencies (i.e. larger wavelengths) by the amount $\Delta\omega_R$ determined by Eq. (5.26). It can be seen from this equation that the relative shift of the frequency $\Delta\omega_R/\omega$ is proportional to the frequency ω.

Let us assess $\Delta\omega_R$ for visible light ($\omega \sim 3 \times 10^{15}$ rad/s). We shall assume that the mass of an atom is 10^{-22} g (an atomic mass of the order of 100). By Eq. (5.26)

$$\Delta\omega_R = \frac{1.05 \times 10^{-27} \times 9 \times 10^{20}}{2 \times 10^{-22} \times 9 \times 10^{20}} \approx 5 \times 10^4 \text{ rad/s}$$

whence for $\Delta\lambda_R$ we get a value of the order of 10^{-7} Å, which we may disregard.

We must note that when an atom absorbs a photon $\hbar\omega$, the momentum of the photon $\hbar\mathbf{k}$ is communicated to the atom. As a result, the latter acquires translational motion. The average additional energy received by an atom when it absorbs a photon coincides with the average value of the recoil energy R. Consequently, to produce the transition $E_m \to E_n$ in an atom, a photon must have the energy

$$\hbar\omega' = \Delta E_{nm} + R$$

while the frequency of the photon must be $\omega' = \omega_0 + \Delta\omega_R$, where $\Delta\omega_R$ is determined by Eq. (5.26).

Thus, a spectral line whose middle is in the emission spectrum of the given atom at the frequency $\omega_0 - \Delta\omega_R$ will have the frequency $\omega_0 + \Delta\omega_R$ in the absorption spectrum of the same atom. For visible light, the shift of the emission and absorption lines relative to each other (which is $2\Delta\lambda_R \approx 10^{-7}$ Å) is five orders of magnitude less than the Doppler breadth of the line (equal to about 3×10^{-2} Å) and three orders of magnitude less than the natural breadth of the line (equal to about 10^{-4} Å). Consequently, for visible light, we may consider that the emission and absorption lines exactly coincide with one another.

5.4. Multiplicity of Spectra and Spin of an Electron

The investigation of alkali metal spectra with the aid of instruments having a high resolving power has shown that each line of these spectra is a double one (doublet). For example, the yellow line $3P \rightarrow 3S$ characteristic of sodium (see Fig. 5.3) consists of two lines with wavelengths of 5890 and 5896 Å. The same relates to the other lines of the principal series, and also to lines of other series.

The structure of a spectrum reflecting the splitting of the lines into their components is called **fine structure**. The complex lines consisting of several components are known as **multiplets**. A fine structure is a property of other elements in addition to the alkali metals. The number of components in a multiplet may be two (doublets), three (triplets), four (quartets), five (quintets), and so on. In a particular case, the spectral lines even with account of the fine structure may be single (singlets).

The splitting of spectral lines is evidently due to splitting of the energy levels. To explain the splitting of these levels, the Dutch physicists Samuel Goudsmit and George Uhlenbeck in 1925 advanced the hypothesis that an electron has an intrinsic angular momentum M_s not associated with the motion of the electron in space. This intrinsic angular momentum was called **spin**.

It was initially assumed that spin is due to rotation of an electron about its axis. According to these notions, an electron was considered similar to a top or spindle. This explains the origin of the term "spin". Very soon, however, it became necessary to reject such model ideas, in particular for the following reason. A spinning charged sphere must have a magnetic moment, and the ratio of the magnetic moment to the mechanical angular momentum must be

$$\frac{\mu}{M} = -\frac{e}{2m_e c} \qquad (5.29)$$

[see Eq. (7.41) of Vol. II, p. 167; we have used the symbol μ here instead of p_m for convenience].

Indeed, it was established that an electron in addition to its intrinsic mechanical angular momentum has an intrinsic magnetic moment μ_s. A number of experimental facts, however, in particular the complicated Zeeman effect, witness that the ratio between the intrinsic magnetic moment and intrinsic mechanical angular momentum is double that between the orbital magnetic moment and orbital mechanical angular momentum:

$$\frac{\mu_s}{M_s} = -\frac{e}{m_e c} \qquad (5.30)$$

Thus, the notion of an electron as of a spinning sphere was unfounded. Spin must be considered as an intrinsic property characterizing an electron in the same way as its charge and mass do.

The assumption on the spin of an electron was confirmed by a great number of experimental facts and should be considered as absolutely proved. It was also found that the presence of spin and all its properties automatically follow from the equation of quantum mechanics satisfying the requirements of the theory of relativity that was proposed by Paul Dirac. It was thus found that the spin of an electron is simultaneously a quantum and a relativistic property. Protons, neutrons, and other elementary particles (except mesons) also have a spin.

The magnitude of the intrinsic angular momentum of an electron is determined according to the general laws of quantum mechanics [see Eq. (4.34)] by the so-called **spin quantum number** s equal to 1/2*,

$$M_s = \hbar \sqrt{s(s+1)} = \hbar \sqrt{\frac{1}{2} \times \frac{3}{2}} = \frac{1}{2} \hbar \sqrt{3} \qquad (5.31)$$

The projection of the spin onto a given direction can take on quantized values differing from one another by $\hbar$:

$$M_{s,z} = m_s \hbar \qquad \left(m_s = \pm s = \pm \frac{1}{2}\right) \qquad (5.32)$$

To find the value of the intrinsic magnetic moment of an electron, we shall multiply M_s by the ratio of μ_s to M_s [see Eq. (5.30)]:

$$\mu_s = -\frac{e}{m_e c} M_s = -\frac{e\hbar}{m_e c} \sqrt{s(s+1)} =$$

$$= -2\mu_B \sqrt{s(s+1)} = -\mu_B \sqrt{3} \qquad (5.33)$$

[μ_B is the Bohr magneton; see Eq. (7.45) of Vol. II, p. 169]. The minus sign indicates that the mechanical angular momentum M_s and the magnetic moment μ_s of an electron are directed oppositely.

The projection of the intrinsic magnetic moment of an electron onto a given direction can have the following values:

$$\mu_{s,z} = -\frac{e}{m_e c} M_{s,z} = -\frac{e}{m_e c} \hbar m_s = -\frac{e\hbar}{m_e c} (\pm 1/2) = \mp \mu_B \qquad (5.34)$$

(the minus sign is obtained if $m_s = +\frac{1}{2}$, and the plus sign if $m_s = -\frac{1}{2}$).

Thus, the projection of the intrinsic angular momentum of an electron can taken on values of $+\frac{1}{2}\hbar$ and $-\frac{1}{2}\hbar$, and of the intrinsic

* For a proton and a neutron, s also equals one-half, for a photon, s equals unity.

magnetic moment—values of $+\mu_B$ and $-\mu_B$. A number of formulas, in particular the expression for energy, include not the angular momentum and magnetic moment themselves, but their projections. The intrinsic mechanical angular momentum (spin) of an electron is therefore customarily said to equal one-half (naturally, in units of $\hbar$), and the intrinsic magnetic moment to equal one Bohr magneton.

Let us now use the example of the sodium atom to show how the existence of the spin of an electron can explain the multiplet structure of its spectrum. Since the angular momentum of the atom residue is zero, the angular momentum of the sodium atom equals that of its valence electron. The angular momentum of the electron will consist, on the other hand, of two momenta: the orbital angular momentum M_l due to the motion of the electron in the atom and the spin angular momentum M_s not associated with the motion of the electron in space. The resultant of these two momenta gives the total angular momentum of the valence electron. Summation of the orbital and spin angular momenta to obtain the total momentum is performed according to the same quantum laws used to summate the orbital angular momenta of different electrons [see Eqs. (4.39) and (4.40)]. The magnitude of the total angular momentum M_j is determined by the quantum number j:

$$M_j = \hbar \sqrt{j(j+1)}$$

Here j can have the values

$$j = l + s, \quad |l - s|$$

where l and s are the azimuthal and spin quantum numbers, respectively. When $l = 0$, the quantum number j has only one value, namely, $j = s = \frac{1}{2}$. When l differs from zero, two values are possible: $j = l + \frac{1}{2}$ and $j = l - \frac{1}{2}$ that correspond to two possible mutual orientations of the angular momenta M_l and M_s—"parallel" and "antiparallel"*.

We shall now take into consideration that magnetic moments are associated with the mechanical angular momenta. The magnetic moments interact with each other like two currents or two magnetic pointers do. The energy of this interaction (called **spin-orbit interaction**) depends on the mutual orientation of the orbital and intrinsic angular momenta. Hence, states with different j's must have different energy.

Thus, each term of the set P ($l = 1$) splits into two terms corresponding to $j = 1/2$ and $j = 3/2$; each term of the set D ($l = 2$)

* The words "parallel" and "antiparallel" have been taken in quotation marks since two angular momenta being added are never directed along a single straight line.

splits into terms with $j = 3/2$ and $j = 5/2$, etc. Only one value of $j = 1/2$ corresponds to each term of the set S; therefore, the terms of the set S do not split.

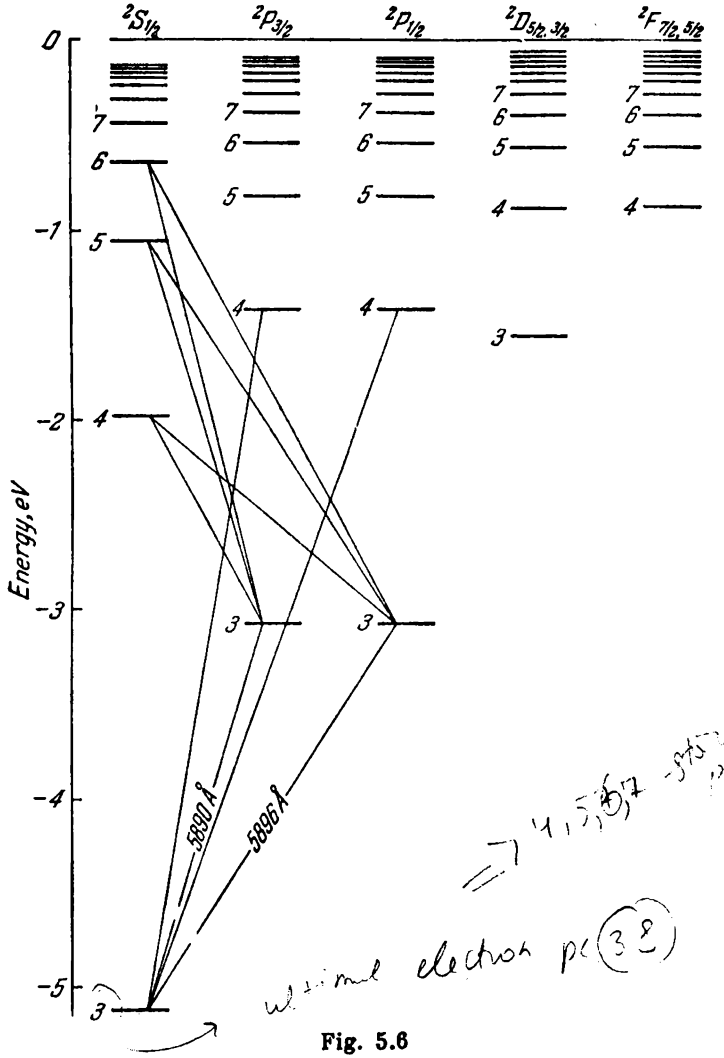


Fig. 5.6

Thus, each set of terms except for S splits into two sets—the terms have a doublet structure. It is customary practice to denote the terms by the symbols

$$^2S_{1/2},\ ^2P_{3/2},\ ^2P_{1/2},\ ^2D_{5/2},\ ^2D_{3/2},\ ^2F_{7/2},\ ^2F_{5/2},\ \ldots$$

The right-hand subscript gives the value of j. The left-hand superscript indicates the multiplicity of the terms. Although the set S is a single one, the superscript *2* is also used with its symbol to show that this set belongs to a system of terms that is a doublet one as a whole.

When the fine structure is taken into account, the diagram of the terms is more complicated, which is illustrated by the diagrams

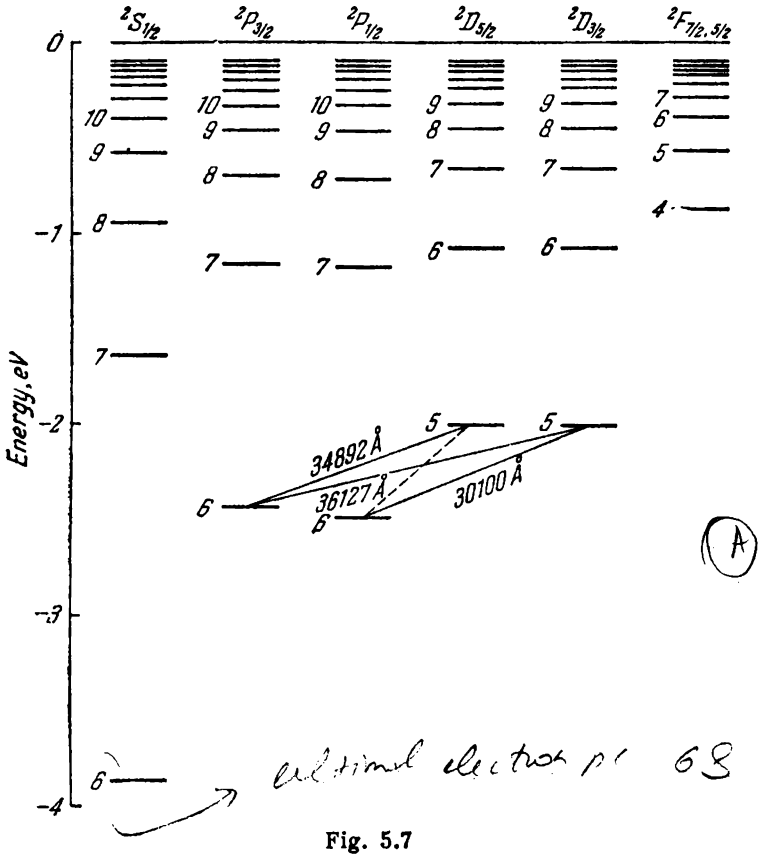


Fig. 5.7

of the levels of sodium (Fig. 5.6) and cesium (Fig. 5.7). The diagram for sodium should be compared with the one shown in Fig. 5.3. Since the multiplet splitting of the terms D and F for sodium is very small, the sublevels of D and F differing in their values of j are shown by single lines in the diagram.

The following selection rule exists for the quantum number of the total angular momentum of an atom:

$$\Delta j = 0, \pm 1 \qquad (5.35)$$

The multiplet splitting of the cesium atom is considerably greater than in the sodium one. It can be seen from the diagram for cesium in Fig. 5.7 that the fine structure of the diffuse series consists of three lines instead of two:

$$5^2D_{3/2} \rightarrow 6^2P_{3/2} \sim 36\,127 \text{ Å}$$
$$5^2D_{5/2} \rightarrow 6^2P_{3/2} \sim 34\,892 \text{ Å}$$
$$5^2D_{3/2} \rightarrow 6^2P_{1/2} \sim 30\,100 \text{ Å}$$

The appearance of these lines is explained additionally in Fig. 5.8. The transition $5^2D_{5/2} \rightarrow 6^2P_{1/2}$ depicted by the dash line is forbidden by selection rule (5.35). The lower part of the diagram shows what the multiplet itself looks like. The thickness of the lines in the diagram corresponds approximately to the intensity of the spectral lines. The collection of the lines obtained looks like a doublet in which one of the components, in turn, is double. Such a group of lines is called not a triplet, but a **complex doublet** because it is produced as a result of the combination of doublet terms.

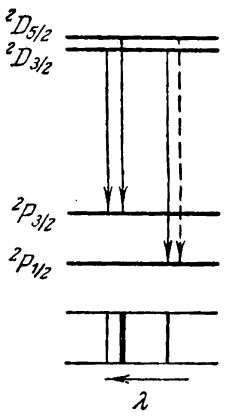


Fig. 5.8

We must note that in connection with an electron having a spin it is quite natural to presume that the levels with $l > 0$ must be double in the hydrogen atom too, and the spectral lines must be doublets. The fine structure of the hydrogen spectrum was really detected experimentally.

The splitting of the energy levels due to spin is a relativistic effect. The relativistic quantum theory gives the following value for the distance between the levels of the fine structure of the hydrogen atom:

$$\Delta E = \frac{\alpha^2}{16} E_1 \tag{5.36}$$

Here E_1 is the ionization energy of the hydrogen atom (calculated on the assumption that the mass of the nucleus is infinitely great), and α is a dimensionless quantity called the **fine structure constant.** It is determined by the expression

$$\alpha = \frac{e^2}{\hbar c} \approx \frac{1}{137} \tag{5.37}$$

We can use Eq. (5.36) to assess the magnitude of the multiplet splitting of levels. The distances between levels differing in the values of the principal quantum number have a magnitude of the order of E_1; the expression $\alpha^2/16$ has a value of the order of 10^{-5}. Consequently, the distance between the levels of the fine structure is about $1/10^5$ of the distance between the main levels.

The fine structure constant is one of the fundamental constants of nature. Its meaning becomes obvious when we pass over to the so-called **natural system of units** used in quantum electrodynamics. In this system, the unit of mass is the mass of an electron m_e, the unit of length is the Compton wavelength of an electron $\lambda_C = \hbar/m_e c$ (see Sec. 2.4), the unit of energy is the rest energy of an electron $m_e c^2$, etc. Let us calculate in these units the electric energy of interaction of two electrons at the distance of $\hbar/m_e c$ from each other. For this purpose, we must divide the expression $e^2/(\hbar/m_e c)$ by $m_e c^2$. As a result, we get a dimensionless quantity equal to

$$\frac{e^2/(\hbar/m_e c)}{m_e c^2} = \frac{e^2}{\hbar c} = \alpha \tag{5.38}$$

[see Eq. (5.37)]. If we expressed the charge of an electron q in natural units, then the formula for the interaction energy would have the form

$$\frac{q\,(\text{nat. un.})\,q\,(\text{nat. un.})}{1\text{ nat. un. of length}} = \alpha \text{ nat. un. of energy}$$

It thus follows that α is the square of the elementary charge expressed in natural units.

According to Eq. (5.38), the fine structure constant characterizes the energy of interaction of two electrons. We can say in other words that α determines how strong an electron is bound to an electromagnetic field. For this reason, the constant α is known as the **constant of electron coupling with an electromagnetic field.**

The mass of an electron is absent in Eq. (5.38) for α. Hence, α is a constant of coupling with an electromagnetic field for any elementary particle having the charge e.

5.5. Resultant Mechanical Angular Momentum of an Atom with Many Electrons

Every electron in an atom has an orbital angular momentum M_l and an intrinsic momentum M_s. The mechanical angular momenta are related to the relevant magnetic moments, owing to which there is interaction between all the M_l's and M_s's.

The angular momenta M_l and M_s add up to form the resultant angular momentum of the atom M_J. Here two cases are possible.

1. The angular momenta M_l have a stronger interaction with one another than with the M_s's which, in turn, are coupled more strongly to one another than to the M_l's. Consequently, all the M_l's add up to form the resultant M_L, the angular momenta M_s add up to form M_S, and only now do M_L and M_S give the total angular momentum of

the atom M_J. Such a kind of coupling is encountered most frequently and is known as the **Russell-Saunders or LS coupling.**

2. Each pair of the M_l's and M_s's displays a stronger interaction between the partners of the pair than between an individual partner and the other M_l's and M_s's. Consequently, resultant M_j's are formed for each electron separately, and they then combine into the M_J of the atom. This kind of coupling, called *jj* **coupling**, is observed in heavy atoms.

The angular momenta are summated with observance of the quantum laws (see Sec. 4.7). Let us consider in greater detail the summation of the angular momenta for a Russell-Saunders coupling.

The orbital quantum numbers l_i are always integers. Accordingly, the quantum number L of the total orbital angular momentum is also an integer (or zero).

The quantum number S of the resultant spin* angular momentum of an atom M_S may be an integer or half-integer depending on whether the number of electrons in the atom is even or odd. With an even number of electrons N, the quantum number S takes on all the integral values from $N \times 1/2$ (all the M_s's are "parallel" to one another) to zero (all the M_s's compensate one another in pairs). For example, when $N = 4$, the quantum number S can have values of 2, 1, 0. When N is an odd number, S takes on all the half-integral values from $N \times 1/2$ (all the M_s's are "parallel" to one another) to 1/2 (all the M_s's except one compensate one another in pairs). For example, when $N = 5$, the possible values of S are 5/2, 3/2, 1/2.

At given values of M_L and M_S, the quantum number J of the resultant angular momentum M_J can have one of the following values:

$$J = L + S, \quad L + S - 1, \ldots, |L - S|$$

Consequently, J will be an integer if S is an integer (i.e. with an even number of electrons in an atom), and a half-integer if S is a half-integer (i.e. with an odd number of electrons). For example,

(1) when $L = 2$ and $S = 1$, the possible values of J are 3, 2, 1;

(2) when $L = 2$ and $S = 3/2$, the possible values of J are 7/2, 5/2, 3/2, 1/2.

The energy of an atom depends on the mutual orientation of the angular momenta M_l (i.e. on the quantum number L), on the mutual orientation of the angular momenta M_s (i.e. on the quantum number S), and on the mutual orientation of M_L and M_S (on the quantum number J). The term of an atom is conventionally written as follows:

$$^{2S+1}L_J \qquad (5.39)$$

* Do not confuse the quantum number S with the symbol of a term S.

where L stands for one of the letters S, P, D, F, etc., depending on the value of the number L. For example, the terms

$$^3P_0, \quad ^3P_1, \quad ^3P_2 \tag{5.40}$$

relate to states with identical $L = 1$, identical $S = 1$, but different J equal to 0, 1, 2.

Symbol (5.39) contains information on the values of the three quantum numbers L, S, and J. When $S < L$, the left-hand superscript $2S + 1$ gives the multiplicity of the term, i.e. the number of sublevels differing in the value of the number J [see symbols (5.40)]. When $S > L$, the actual multiplicity is $2L + 1$. But the symbol of the term is written all the same in form (5.39) because otherwise it would contain no information on the value of the quantum number S.

We have already used symbols of type (5.39) in Sec. 5.2 for alkali metal atoms. It is characteristic of these elements, however, that S of an atom, coinciding with s of its valence electron, equals 1/2. Now we have acquainted ourselves with the symbols of terms for any cases.

5.6. The Magnetic Moment of an Atom

We have noted more than once that the magnetic moment μ is associated with the mechanical angular momentum of an atom M. The ratio μ/M is called the **gyromagnetic ratio.**

Although our notion of orbits, as in general our notion of the trajectories of microparticles, is illegitimate, the angular momentum due to the motion of the electrons in an atom is called **orbital.** The experimentally determined ratio of the orbital magnetic moment μ_L and the mechanical orbital angular momentum M_L coincides with the gyromagnetic ratio ensuing from the classical notions (see Sec. 7.6 of Vol. II, p. 167). This ratio is $-e/2m_ec$; accordingly,

$$\mu_L = -\frac{e}{2m_ec} M_L = -\frac{e\hbar}{2m_ec} \sqrt{L(L+1)} = -\mu_B \sqrt{L(L+1)} \tag{5.41}$$

The quantity

$$\mu_B = \frac{e\hbar}{2m_ec} = 0.927 \times 10^{-20} \text{ erg/Gs} \tag{5.42}$$

is called the **Bohr magneton** and is the natural unit of the magnetic moment. The minus sign in Eq. (5.41) indicates that the directions of the magnetic moment and the mechanical angular momentum are opposite (this is due to the fact that the charge of an electron is negative). The presence of the minus sign permits us to obtain the projection of μ_L onto the direction z by simply substituting the

quantum number m_L for $\sqrt{L(L+1)}$ in Eq. (5.41):

$$\mu_{L,z} = -\mu_B m_L \tag{5.43}$$

When $m_L > 0$, the projection of $\mathbf{M}_L$ is positive, while the projection of $\boldsymbol{\mu}_L$ is negative; when $m_L < 0$, the projection of $\mathbf{M}_L$ is negative, and that of $\boldsymbol{\mu}_L$ is positive.

A number of experimental facts indicate that the gyromagnetic ratio of the intrinsic (spin) magnetic moment and angular momentum is double the gyromagnetic ratio of the orbital magnetic moment and angular momentum. Thus,

$$\mu_S = -\frac{e}{m_e c} M_S = -2\mu_B \sqrt{S(S+1)} \tag{5.44}$$

In this connection, the spin is said to have a double magnetism.

The double magnetism of spin follows from the experiment of A. Einstein and W. de Haas, and from S. Barnett's experiment (see Sec. 7.6 of Vol. II, p. 167 et seq.). In addition, the notion of the double magnetism of spin makes it possible to give an exhaustive explanation of the complicated Zeeman effect (see the following section).

Owing to the double magnetism of spin, the gyromagnetic ratio of the total magnetic moment μ_J and total angular momentum M_J is a function of the quantum numbers L, S, and J. We must note that the numbers L and S characterize the ratio of the values of M_L and M_S, while the number J determines the mutual orientation of the orbital and spin angular momenta. The relevant quantum mechanical calculations give the following formula for the magnetic moment of an atom

$$\mu_J = -\mu_B g \sqrt{J(J+1)} \tag{5.45}$$

where

$$g = 1 + \frac{J(J+1) + S(S+1) - L(L+1)}{2J(J+1)} \tag{5.46}$$

Expression (5.46) is called the **Lande g factor.** When the total spin angular momentum of an atom is zero ($S = 0$), the total angular momentum coincides with the orbital one ($J = L$). Introducing $S = 0$ and $J = L$ into Eq. (5.46) yields $g = 1$, and we arrive at the value of the magnetic moment determined by Eq. (5.41). When the total orbital angular momentum of an atom is zero ($L = 0$), the total angular momentum coincides with the spin one ($J = S$). Introduction of these values of the quantum numbers into Eq. (5.46) yields $g = 2$, and we arrive at the value of the magnetic moment determined by Eq. (5.44). We must note that the Lande g factor can have values less than unity, and can even be zero (this is obtained, for example, when $L = 3$, $S = 2$, and $J = 1$). In the last

case, the magnetic moment of an atom is zero, although the mechanical angular momentum differs from zero.

We remind our reader that the presence of a minus sign in Eq. (5.45) makes it possible to obtain the projection of μ_J onto the z-axis by simply substituting m_J for $\sqrt{J(J+1)}$. Hence,

$$\mu_{J,z} = -\mu_B g m_J \qquad (m_J = 0, \pm 1, \ldots, \pm J) \qquad (5.47)$$

A number of questions of the physics of the atom can be treated with the aid of the so-called **vector model of an atom.** In the construction of such a model, the mechanical angular momenta and magnetic moments are depicted in the form of directed lengths of lines. Strictly speaking, owing to the uncertainty in the directions of the vectors **M** in space, such a procedure is not substantiated. Therefore, when working with a vector model, we must remember the conditional nature of the relevant constructions. A vector model must not be understood literally. It should be considered as a collection of rules permitting us to obtain results whose truth is confirmed by strict quantum mechanical calculations.

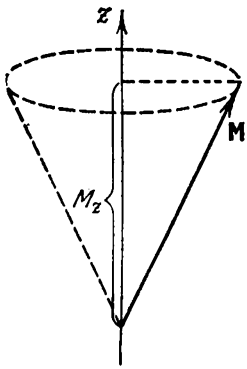


Fig. 5.9

A vector model is constructed according to the following rules. Let M and M_z have definite values (here M_x and M_y have not been determined). Consequently, the vector **M** can have the direction of one of the generatrices of the cone depicted in Fig. 5.9. We can imagine matters as if the vector **M** is uniformly rotating (precessing) about the direction z coinciding with the axis of the cone.

Assume that the magnetic field **B** has been set up in the direction z. The magnetic moment μ is associated with the mechanical angular momentum **M**. Therefore, the field is exerted on **M** (through μ). We assume that the velocity of precession of the momentum **M** about **B** is the higher, the stronger is the field acting on the angular momentum, i.e. the greater is **B**.

According to the rules for constructing a vector model, the angular momenta $\mathbf{M}_1$ and $\mathbf{M}_2$ being added precess about the direction of the resultant angular momentum **M** (Fig. 5.10). The angular momenta interact with each other (through the magnetic moments μ_1 and μ_2). The velocity of precession is assumed to be proportional to the intensity of interaction. In the state in which M and M_z have been determined, the vector **M** precesses in turn about the direction z. If we set up a magnetic field **B** along the z-axis, different phenomena will be observed depending on the relation between the interaction of the angular momenta with each other and with the magnetic

field. Let us consider two cases: (1) a weak field—the interaction of the angular momenta with each other is greater than the action of the magnetic field on each of them; (2) a strong field—the action of the field on each of the angular momenta exceeds their interaction with each other.

In the first case (Fig. 5.11*a*), the angular momenta add up to form the resultant angular momentum **M** that is projected onto the direc-

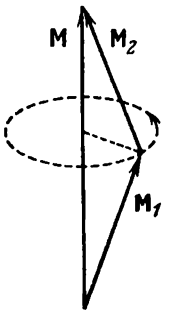


Fig. 5.10

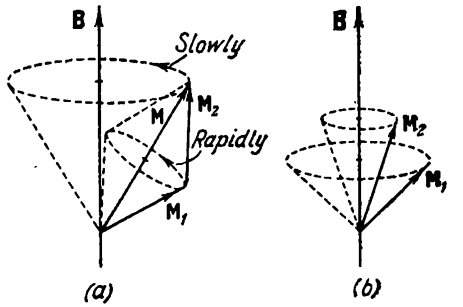


Fig. 5.11

tion of the field. Here two kinds of precession occur: precession of the angular momenta $\mathbf{M}_1$ and $\mathbf{M}_2$ about the direction of **M** and precession of the resultant vector **M** about the direction of **B**. The velocity of the first precession will be much higher because the interaction of the angular momenta with each other exceeds the action of the magnetic field on each of them.

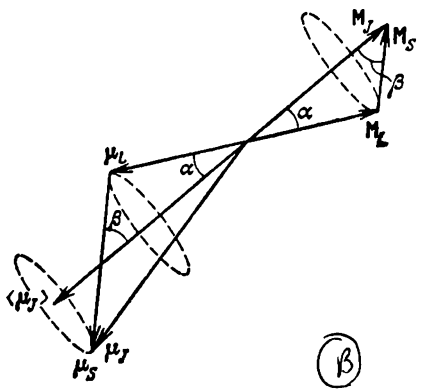


Fig.5.12

In the second case (Fig. 5.11*b*), the field breaks the coupling between the angular momenta $\mathbf{M}_1$ and $\mathbf{M}_2$, and each of them precesses about the direction of the field independently of the other one. Each of the vectors $\mathbf{M}_1$ and $\mathbf{M}_2$ will also be projected separately onto the direction of the field.

Let us obtain formula (5.45) with the aid of a vector model. Figure 5.12 depicts the vectors $\mathbf{M}_L$, $\mathbf{M}_S$, $\mathbf{M}_J$ and the vectors $\boldsymbol{\mu}_L$, $\boldsymbol{\mu}_S$, $\boldsymbol{\mu}_J$ corresponding to them. The scales have been selected so that the vectors $\mathbf{M}_L$ and $\boldsymbol{\mu}_L$ are depicted by arrows of the same length. In this condition, the vector $\boldsymbol{\mu}_S$ will be depicted by an arrow that is twice as long as the one depicting the vector $\mathbf{M}_S$.

Owing to the double magnetism of the spin, the vector μ_J is non-collinear with the vector $\mathbf{M}_J$. The vectors $\mathbf{M}_L$ and $\mathbf{M}_S$ precess about the direction of $\mathbf{M}_J$, also involving in this precession the resultant vector of the magnetic moment μ_J. During a sufficiently long observation time, the average value of the vector μ_J will be registered. It is designated in Fig. 5.12 by the symbol $\langle \mu_J \rangle$. Let us find the projection of this vector onto the direction of $\mathbf{M}_J$, which we shall denote simply by μ_J. A glance at the figure shows that

$$\mu_J = -|\mu_L| \cos\alpha - |\mu_S| \cos\beta \tag{5.48}$$

where $|\mu_L|$ and $|\mu_S|$ are the magnitudes of the relevant vectors. According to Eqs. (5.41) and (5.44)

$$|\mu_L| = \mu_B \sqrt{L(L+1)}; \qquad |\mu_S| = 2\mu_B \sqrt{S(S+1)} \tag{5.49}$$

To find the value of $\cos\alpha$, let us square the relation $\mathbf{M}_S = \mathbf{M}_J - \mathbf{M}_L$:

$$M_S^2 = M_J^2 + M_L^2 - 2M_J M_L \cos\alpha$$

whence

$$\cos\alpha = \frac{M_J^2 + M_L^2 - M_S^2}{2M_J M_L} = \frac{J(J+1) + L(L+1) - S(S+1)}{2\sqrt{J(J+1)}\sqrt{L(L+1)}} \tag{5.50}$$

To find the value of $\cos\beta$, let us square the relation $\mathbf{M}_L = \mathbf{M}_J - \mathbf{M}_S$:

$$M_L^2 = M_J^2 + M_S^2 - 2M_J M_S \cos\beta$$

whence

$$\cos\beta = \frac{M_J^2 + M_S^2 - M_L^2}{2M_J M_S} = \frac{J(J+1) + S(S+1) - L(L+1)}{2\sqrt{J(J+1)}\sqrt{S(S+1)}} \tag{5.51}$$

Introducing Eqs. (5.49), (5.50), and (5.51) into Eq. (5.48), we have

$$\mu_J = -\mu_B \sqrt{L(L+1)} \frac{J(J+1) + L(L+1) - S(S+1)}{2\sqrt{J(J+1)}\sqrt{L(L+1)}} - \\ -2\mu_B \sqrt{S(S+1)} \frac{J(J+1) + S(S+1) - L(L+1)}{2\sqrt{J(J+1)}\sqrt{S(S+1)}}$$

Let us perform cancellations, combine both addends and, in addition, multiply the numerator and the denominator by $\sqrt{J(J+1)}$. The result is the expression

$$\mu_J = -\mu_B \sqrt{J(J+1)} \frac{3J(J+1) + S(S+1) - L(L+1)}{2J(J+1)}$$

coinciding with Eq. (5.45).

5.7. The Zeeman Effect

The **Zeeman effect** is the name given to the splitting of the energy levels when a magnetic field acts on atoms*. Splitting of the levels leads to splitting of the spectral lines into several components. The splitting of the spectral lines when emitting atoms experience the action of a magnetic field is also known as the Zeeman effect.

The splitting of the lines was discovered in 1896 by the Dutch physicist Pieter Zeeman (1865-1943). The splitting is very slight—at B of the order of 10^4 Gs it is only several tenths of an angstrom.

The Zeeman splitting of the levels is explained by the fact that an atom having the magnetic moment μ_J acquires in the magnetic field the additional energy

$$\Delta E = -\mu_{J,B} B \tag{5.52}$$

where $\mu_{J,B}$ is the projection of the magnetic moment onto the direction of the field [see Eq. (6.76) of Vol. II, p. 136]. In accordance with Eq. (5.47), we have

$$\mu_{J,B} = -\mu_B g m_J$$

Introduction of this expression into Eq. (5.52) yields

$$\Delta E = \mu_B g B m_J \quad (m_J = 0, \pm 1, \ldots, \pm J) \tag{5.53}$$

It can be seen from this formula that the energy level corresponding to the term ${}^{2S+1}L_J$ splits into $2J + 1$ equally spaced sublevels, the amount of splitting depending on the Lande g factor, i.e. on the quantum numbers L, S, and J of the given level. Before the switching on of a field, the states differing in the values of the quantum number m_J had an identical energy, i.e. degeneracy was observed with respect to the quantum number m_J. The magnetic field removes the degeneracy with respect to m_J.

Let us first consider the Zeeman splitting of spectral lines having no fine structure (singlets). These lines appear in transitions between the levels corresponding to $S = 0$. For such levels, $g = 1$. Consequently, Eq. (5.53) has the form

$$\Delta E = \mu_B B m_J \quad (m_J = 0, \pm 1, \ldots, \pm L) \tag{5.54}$$

($J = L$, $m_J = m_L$).

Figure 5.13 shows the splitting of the levels and spectral lines for the transition between the states with $L = 1$ and $L = 0$ (for the $P \rightarrow S$-transition). In the absence of a field, one line is observed whose frequency is designated by ω_0. When a field is switched on,

* Splitting of the energy levels also occurs when an electric field acts on atoms. This phenomenon is called the Stark effect.

in addition to the line ω_0, two lines symmetrical relative to it appear having the frequencies $\omega_0 + \Delta\omega_0$ and $\omega_0 - \Delta\omega_0$.

Figure 5.14 gives a similar diagram for a more complicated case—for the transition $D \rightarrow P$. It may seem at first sight that in this case the initial line ought to split into seven components. Actually, however, only three components are obtained as in the preceding case: a line with the frequency ω_0 and two lines with the frequencies

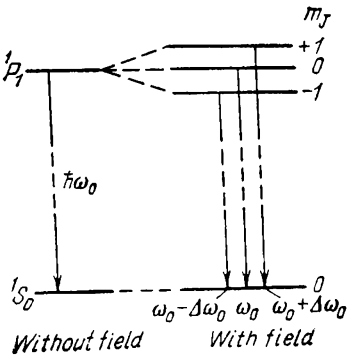


Fig. 5.13

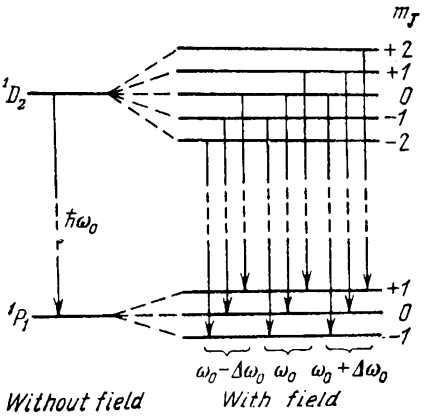


Fig. 5.14

$\omega_0 + \Delta\omega_0$ and $\omega_0 - \Delta\omega_0$ arranged symmetrically relative to it. The explanation is that there is a **selection rule** for the magnetic quantum number m_J according to which only those transitions are possible when m_J either remains constant or changes by unity:

$$\Delta m_J = 0, \pm 1 \tag{5.55}$$

Because of this rule, only the transitions indicated in Fig. 5.14 are possible. As a result, three components are obtained having the same frequencies as in the case depicted in Fig. 5.13.

The shift of the components $\Delta\omega_0$ obtained in the cases treated above is called the **normal** or the **Lorentz* shift**. According to Eq. (5.54), this shift is

$$\Delta\omega_0 = \frac{\mu_B B}{\hbar} = \frac{e\hbar}{2m_e c}\frac{B}{\hbar} = \frac{e}{2m_e c}B \tag{5.56}$$

The splitting into three lines considered above, with two of these lines at a distance equal to the normal shifting $\Delta\omega_0$ from the undisplaced line, is called the **simple** (or **normal**) **Zeeman effect**. Let us assess the magnitude of the simple Zeeman splitting for a field of

* H. Lorentz gave a classical explanation of the simple Zeeman effect and calculated the value of the normal shift. Give attention to the fact that $\Delta\omega_0$ coincides with the Larmor frequency [see Eq. (7.46) of Vol. II, p. 172].

the order of 10^4 Gs. Since $\lambda = 2\pi c/\omega$,

$$|\Delta\lambda| = \frac{2\pi c}{\omega^2}\Delta\omega_0 = \frac{\pi eB}{m_e\omega^2}$$

The frequency ω for visible light is about 3×10^{15} rad/s. Hence,

$$\Delta\lambda = \frac{3.14 \times 4.8 \times 10^{-10} \times 10^4}{0.91 \times 10^{-27} \times 9 \times 10^{30}} \approx 0.2 \times 10^{-8} \text{ cm} = 0.2 \text{ Å}$$

We have already noted that the simple Zeeman effect is observed when the initial lines have no fine structure, i.e. are singlets. For lines having a fine structure, the number of components may be greater than three, while the magnitude of the splitting is a rational fraction of the normal shifting $\Delta\omega_0$:

$$\Delta\omega = \Delta\omega_0 \frac{r}{q} \qquad (5.57)$$

where r and q are small integers. For example, the splitting of the yellow doublet of sodium appears as shown in Fig. 5.15. Such splitting of spectral lines is called the **complicated** (or **anomalous**) **Zeeman effect.**

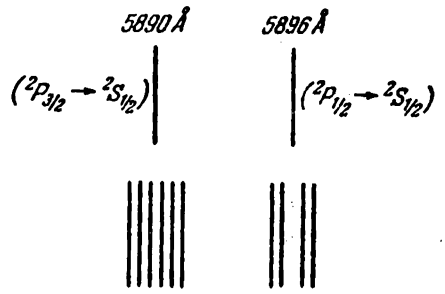


Fig. 5.15

The complicated Zeeman effect is explained by the dependence of the magnitude of splitting of levels on the Lande g factor, i.e. in the long run by the existence of spin of an electron and the double magnetism of spin. We shall explain this using the following example.

Let us consider the splitting of the sodium doublet formed by the transitions $3^2P_{1/2} \rightarrow 3^2S_{1/2}$ and $3^2P_{3/2} \rightarrow 3^2S_{1/2}$ (see Fig. 5.6). The Lande g factor has the values:

for the term $^2S_{1/2}$ ($L = 0$, $S = 1/2$, $J = 1/2$)

$$g = 1 + \frac{1/2 \times 3/2 + 1/2 \times 3/2 - 0 \times 1}{2 \times 1/2 \times 3/2} = 1 + 1 = 2$$

for the term $^2P_{1/2}$ ($L = 1$, $S = 1/2$, $J = 1/2$)

$$g = 1 + \frac{1/2 \times 3/2 + 1/2 \times 3/2 - 1 \times 2}{2 \times 1/2 \times 3/2} = 1 - 1/3 = 2/3$$

for the term $^2P_{3/2}$ ($L = 1$, $S = 1/2$, $J = 3/2$)

$$g = 1 + \frac{3/2 \times 5/2 + 1/2 \times 3/2 - 1 \times 2}{2 \times 3/2 \times 5/2} = 1 + 1/3 = 4/3$$

Figure 5.16*a* shows the splitting of the levels and the transitions allowed by rule (5.55) for the line $^2P_{1/2} \rightarrow {}^2S_{1/2}$. For the level $^2S_{1/2}$, the energy increment is

$$\Delta E' = \mu_B B g' m'_J$$

where $g' = 2 = 6/3$ [see Eq. (5.53)]. For the level $^2P_{1/2}$

$$\Delta E'' = \mu_B B g'' m_j''$$

where $g'' = 2/3$.

The shifting of the lines relative to the initial line is determined by the expression

$$\Delta\omega = \frac{\Delta E'' - \Delta E'}{\hbar} = \frac{\mu_B B}{\hbar}(g'' m_j'' - g' m_j') = \Delta\omega_0 (g'' m_j'' - g' m_j')$$

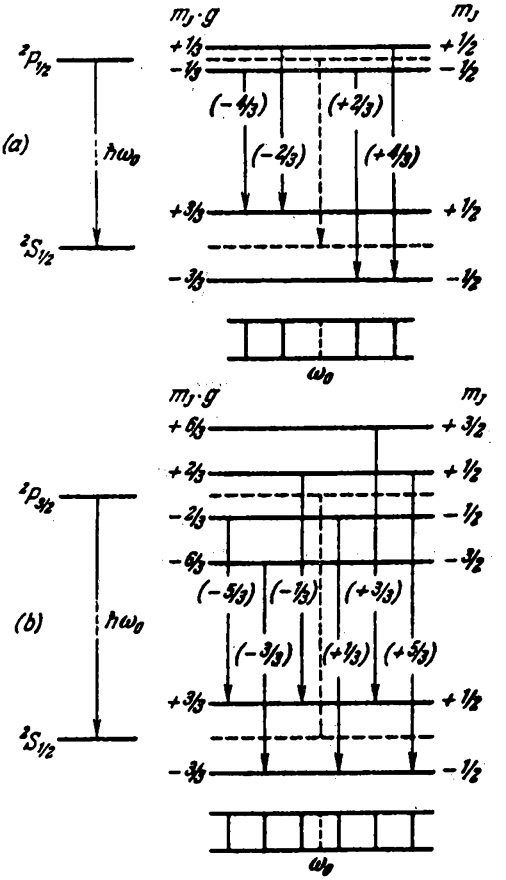


Fig. 5.16

The values of $(g'' m_j'' - g' m_j')$ for the relevant spectral lines are given in Fig. 5.16 in parentheses in gaps of the lines depicting the transitions between the levels.

Inspection of Fig. 5.16*a* reveals that when a field is switched on, the initial line is absent. Four lines appear instead of it. Their

shifts, expressed in units of normal shift, are $-4/3$, $-2/3$, $+2/3$, and $+4/3$, which can be written as follows:

$$\Delta\omega = \Delta\omega_0 \left[\pm \frac{2}{3}, \pm \frac{4}{3} \right]$$

The splitting of the levels and the allowed transitions for the line ${}^2P_{3/2} \rightarrow {}^2S_{1/2}$ are shown in Fig. 5.16*b*. It can be seen from the diagram

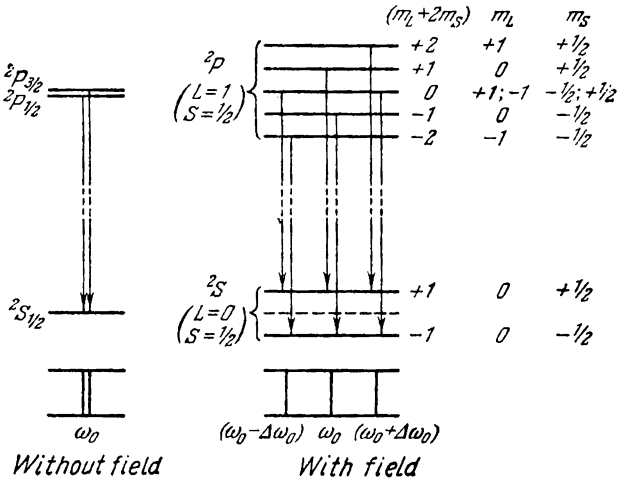


Fig. 5.17

that for such a transition the initial line is also absent when the field is switched on. The shifts of the six lines obtained are

$$\Delta\omega = \Delta\omega_0 \left[\pm \frac{1}{3}, \pm \frac{3}{3}, \pm \frac{5}{3} \right]$$

Everything said above holds for a weak magnetic field. A field is considered weak with respect to the Zeeman effect if the Zeeman splitting of the levels is less than the multiplet splitting.

The coupling between $\mathbf{M}_L$ and $\mathbf{M}_S$ is broken in a strong magnetic field, and these angular momenta are projected onto the direction of the field independently of each other. In this case

$$\Delta E = \mu_B B m_L + 2\mu_B B m_S = \mu_B B (m_L + 2m_S)$$

i.e. the splitting becomes an integral multiple of normal splitting. The following selection rules hold for the transitions:

$$\Delta m_L = 0, \pm 1, \qquad \Delta m_S = 0$$

As a result, we get a normal Zeeman triplet (Fig. 5.17). This phenomenon is known as the **Paschen-Back effect**. It is observed when the magnetic splitting of the lines becomes greater than the multiplet splitting.

5.8. Electron Paramagnetic Resonance

We established in the preceding section that when an atom having a magnetic moment other than zero is in a magnetic field, each level of the atom splits into $2J + 1$ Zeeman sublevels*. According to Eq. (5.53), the distance between the sublevels is

$$\delta E = \mu_B g B$$

Let us assume that an electromagnetic wave is incident on an atom in a constant magnetic field B, and that the frequency of the wave satisfies the condition

$$\hbar\omega = \delta E = \mu_B g B = \hbar \Delta\omega_0 g \tag{5.58}$$

where $\Delta\omega_0$ is the normal shift [see Eq. (5.56)]. We could expect that the action of the magnetic field of the incident wave would cause transitions of the atom between adjacent sublevels [rule (5.55) allows only transitions at which m_J changes by not more than unity]. Such a phenomenon is indeed observed. It was discovered by the Soviet physicist Yevgeni Zavoisky (1907-1976) in 1944 and was named **electron paramagnetic resonance.** This name is explained by the following circumstances. The phenomenon is of a resonance nature—transitions appear at a strictly definite frequency of the incident wave. The magnetic moment of the atom set up by the orbital and spin moments of the electrons is responsible for splitting of the levels (we must note that nuclear magnetic resonance due to the magnetic moment of the nucleus is observed in addition to electron paramagnetic resonance). The phenomenon occurs only in paramagnetic substances (in diamagnetics the magnetic moments of the atoms are zero).

It can be seen from Eq. (5.58) that the resonance frequencies are of the order of the normal shift $\Delta\omega_0$ (the factor g has a value of the order of unity). At $B = 10^4$ Gs

$$\omega \sim \Delta\omega_0 = \frac{\mu_B B}{\hbar} = \frac{0.927 \times 10^{-20} \times 10^4}{1.05 \times 10^{-27}} \sim 10^{11} \text{ rad/s}$$

[see Eq. (5.42)]. A wavelength of the order of a few centimetres corresponds to such a frequency. Consequently, the resonance frequencies are in the radio range.

An electromagnetic wave can cause an atom to pass over either to a higher energy state or to a lower one with equal probability (this will be discussed in detail in Sec. 5.15). In the first case, the wave will be weakened, and in the second amplified. If a paramagnet-

* If the magnetic moment of the atom is produced by an outer electron in the s-state, the number of sublevels is two—spin of the electron "with the field" and spin "against the field".

ic is in thermal equilibrium, the atoms are distributed by sublevels in accordance with the Boltzmann law [see Eq. (11.81) of Vol. I, p. 328]. Consequently, the number of atoms in a state with a lower energy exceeds their number in a state with a higher energy. Therefore, transitions occurring with an increase in energy will predominate over the transitions with a decrease in energy. As a result, the intensity of the wave will diminish—the paramagnetic absorbs electromagnetic radiation, and gets heated as a result.

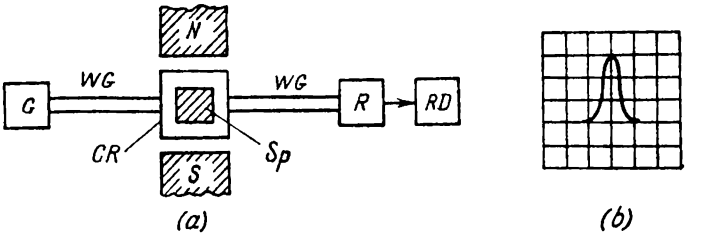


Fig. 5.18

It follows from the above that electron paramagnetic resonance is the selective absorption of the energy of a radio-frequency field in paramagnetic substances that are in a constant magnetic field.

We tacitly assumed in our reasoning that the atoms of a paramagnetic do not interact with one another. In practice, electron paramagnetic resonance is observed in crystalline or liquid paramagnetics (it was also observed in some gases). In condensed media, individual atoms, apart from the external magnetic field, also experience the action of chaotically oriented internal fields. For this reason, the resonance frequencies are slightly different for different atoms, and as a result the electron paramagnetic resonance lines have a finite breadth.

The instrument used to study electron paramagnetic resonance is called a **microwave spectroscope**. It consists (Fig. 5.18*a*) of electromagnetic wave generator *G*, wave guides *WG*, cavity resonator *CR** suspended between the poles of an electromagnet, receiver *R*, and registering device *RD*. The receiver is tuned to the frequency of the generator. An oscillograph or automatic recorder is used as the registering device. Paramagnetic specimen *Sp* is placed inside the cavity resonator. In the course of an experiment, the magnetic field produced by the electromagnet is smoothly changed. At a value of *B* corresponding to condition (5.58), intensive absorption of the wave by the specimen is observed. An absorption curve is shown in Fig. 5.18*b*. As noted above, it has a finite breadth.

* Wave guides are defined as tubes with conducting walls. A cavity resonator is a cavity with conducting walls.

Electron paramagnetic resonance is used to study the structure of crystals, the magnetic properties of atomic nuclei, and in a number of other cases.

5.9. The Pauli Principle. Distribution of Electrons by Energy Levels of an Atom

Every electron in an atom travels in a first approximation in a centrally symmetrical non-Coulomb field. The state of an electron in this case is determined by the three quantum numbers n, l, and m whose physical meaning was established in Sec. 5.1. In connection with the existence of spin of an electron, it is necessary to add to these quantum numbers the quantum number m_s that can take on values of $\pm 1/2$ and determines the projection of the spin onto the given direction. In the following, we shall use the symbol m_l instead of m for the magnetic quantum number to stress the circumstance that this number determines the projection of the orbital angular momentum whose value is given by the quantum number l.

Thus, the state of every electron in an atom is characterized by four quantum numbers:

principal	n ($n = 1, 2, 3, \ldots$)
azimuthal	l ($l = 0, 1, 2, \ldots, n-1$)
magnetic	m_l ($m_l = -l, \ldots, -1, 0, +1, \ldots, +l$)
spin	m_s ($m_s = +1/2, -1/2$)

The energy of a state mainly depends on the numbers n and l. In addition, there is a slight dependence of the energy on the numbers m_l and m_s because their values are associated with the mutual orientation of the angular momenta $\mathbf{M}_l$ and $\mathbf{M}_s$ on which the magnitude of the interaction between the orbital and intrinsic magnetic moments of an electron depends. The energy of a state grows at a greater rate with an increase in the number n than in the number l. Therefore, as a rule, a state with a greater value of n has a greater energy regardless of the value of l.

In the ground (unexcited) state of an atom, the electrons should be at the lowest energy levels available for them. It should therefore seem that in any atom in the ground state all the electrons ought to be in the state $1s$ ($n = 1$, $l = 0$), and the fundamental terms of all the atoms ought to be of the type of S-terms ($L = 0$). Experiments show, however, that this is not the case.

The explanation of the observed types of terms is as follows. According to one of the laws of quantum mechanics called the **Pauli prin-**

ciple* (named in honour of its discoverer, the Austrian physicist Wolfgang Pauli, 1900-1958), the same atom (or any other quantum system) cannot contain two electrons having the same set of the four quantum numbers, n, l, m_l, and m_s. In other words, two electrons cannot simultaneously be in the same state.

It was shown in Sec. 5.1 that n^2 states differing in the values of l and m_l correspond to a given n. The quantum number m_s can take on two values: $\pm 1/2$. Consequently, not more than $2n^2$ electrons can be in states with a given value of n in an atom:

Quantum number n	1	2	3	4	5	...
Maximum possible number of electrons in state	2	8	18	32	50	...

The set of electrons having identical values of the quantum number n forms a shell. The shells are further divided into subshells differing in the value of the quantum number l. In accordance with the

Table 5.2

Shell	n	l	m_l	m_s	Subshell	Shell	n	l	m_l	m_s	Subshell
K	1	0	0	↑↓	K (1s)			0	0	↑↓	N_1 (4s)
		0	0	↑↓	L_1 (2s)			1	−1	↑↓	N_2 (4p)
L	2		−1	↑↓					0	↑↓	
		1	0	↑↓	L_2 (2p)				+1	↑↓	
			+1	↑↓					−2	↑↓	
		0	0	↑↓	M_1 (3s)				−1	↑↓	
			−1	↑↓				2	0	↑↓	N_3 (4d)
		1	0	↑↓	M_2 (3p)	N	4		+1	↑↓	
M	3		+1	↑↓					+2	↑↓	
			−2	↑↓					−3	↑↓	
			−1	↑↓					−2	↑↓	
		2	0	↑↓	M_3 (3d)			3	−1	↑↓	N_4 (4f)
			+1	↑↓					0	↑↓	
			+2	↑↓					+1	↑↓	
									+2	↑↓	
									+3	↑↓	

* This principle is also known as the **Pauli exclusion principle** or simply the exclusion principle. It holds not only for electrons, but also for other particles with a half-integral spin.

value of n, the shells are given symbols taken from X-ray spectroscopy:

Quantum number n	1	2	3	4	5	6	7	...
Symbol of shell	*K*	*L*	*M*	*N*	*O*	*P*	*Q*	...

The division of the possible states of an electron in an atom into shells and subshells is shown in Table 5.2, in which the symbols $\uparrow\downarrow$ have been used instead of the designations $m_s = \pm 1/2$ for visualization. The subshells, as indicated in the table, can be designated in two ways (for example, L_1 or $2s$).

A completely filled subshell is characterized by the equality to zero of the total orbital and total spin angular momenta ($L = 0$, $S = 0$). Hence, the angular momentum of such a subshell equals zero ($J = 0$). Let us convince ourselves that this is true taking the $3d$-subshell as an example. The spins of all ten electrons in this subshell compensate one another in pairs, and as a result $S = 0$. The quantum number of the projection of the resultant orbital angular momentum M_L of this subshell onto the z-axis has the single value $m_L = \sum m_l = 0$. Consequently, L also equals zero.

Thus, when determining L and S of an atom, no attention may be given to filled subshells.

5.10. Mendeleev's Periodic System of Elements

The Pauli principle provides an explanation of the periodic repetition of the properties of atoms. Let us see how the **periodic system of elements** discovered by the Russian chemist Dmitri Mendeleev (1834-1907) is constructed. We shall begin with the hydrogen atom having one electron. Each following atom is obtained by increasing the charge of the preceding atom's nucleus by unity and adding one electron, which we shall place in the state with the smallest energy accessible for it in accordance with the Pauli principle.

The hydrogen atom has one $1s$-electron in the ground state with an arbitrary orientation of its spin. The quantum numbers of the atom have the values $L = 0$, $S = 1/2$, $J = 1/2$. Accordingly, the fundamental term of the hydrogen atom has the form $^2S_{1/2}$.

If we increase the charge of the hydrogen atom nucleus by unity and add another electron, we get the helium atom. Both electrons in this atom can be in the K-shell, but with an antiparallel orientation of their spins. The so-called **electron configuration** of the atom can be written as $1s^2$ (two $1s$-electrons). The fundamental term will be $^1S_0$ ($L = 0$, $S = 0$, $J = 0$).

Filling of the K-shell terminates in the helium atom. The third electron of the lithium atom can occupy only the level $2s$ (Fig. 5.19).

The electron configuration $1s^22s$ is obtained. The ground state is characterized by $L = 0$, $S = 1/2$, $J = 1/2$. Therefore, $^2S_{1/2}$ will be the fundamental term as in the hydrogen atom. The third electron of the lithium atom, occupying a higher energy level than the remaining two electrons, is bound to the nucleus of the atom more weakly than they are. As a result, it determines the optical and chemical properties of the atom.

In the fourth element, beryllium, the subshell $2s$ is completely filled. In the following six elements (B, C, N, O, F, and Ne), the subshell $2p$ is filled with electrons. As a result, the neon atom has completely filled shells K (with two electrons) and L (with eight electrons) forming a stable system like that of helium. This explains the specific properties of the inert (noble) gases.

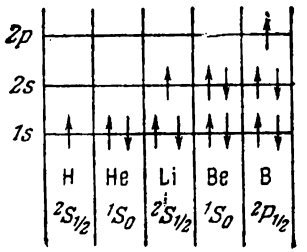


Fig. 5.19

The process of building up the electron shells of the first 36 elements of the periodic system is shown in Table 5.3. The eleventh element, sodium, in addition to the filled shells K and L, has one electron in the subshell $3s$. The electron configuration has the form $1s^22s^22p^63s$. Here, $^2S_{1/2}$ is the fundamental term. The electron $3s$ is bound to the nucleus most weakly of all the electrons and is the valence or optical electron. In this connection, the chemical and optical properties of sodium are similar to those of lithium. The ground state of the optical electron in the sodium atom is characterized by the value of $n = 3$. This is exactly what explains the circumstance that in the diagram of the sodium atom levels (Fig. 5.6) the ground level is indicated by the number 3. We shall note in passing that the cesium atom has the following electron configuration in the ground state

$$1s^22s^22p^63s^23p^63d^{10}4s^24p^64d^{10}5s^25p^66s$$

Consequently, its optical electron in the ground state has $n = 6$. The levels in Fig. 5.7 are marked accordingly.

In the elements following sodium, the subshells $3s$ and $3p$ are filled normally. The subshell $3d$ with the given general configuration is higher than the subshell $4s$ from the energy viewpoint. In this connection, with filling of the M shell not completed as a whole, filling of the N shell begins. The subshell $4p$ is already higher than $3d$, so that after $4s$ the subshell $3d$ is filled.

The electron levels of all the atoms are built up with similar deviations from the ordinary sequence repeating from time to time. Similar electron configurations (for example, $1s$, $2s$, and $3s$) periodically repeat above the completely filled subshells. This underlies

Table 5.3

Element	*K*	*L*		*M*			*N*		Fundamental term
	1*s*	2*s*	2*p*	3*s*	3*p*	3*d*	4*s*	4*p*	
1H	1	—	—	—	—	—	—	—	$^2S_{1/2}$
2He	2	—	—	—	—	—	—	—	$^1S_0$
3Li	2	1	—	—	—	—	—	—	$^2S_{1/2}$
4Be	2	2	—	—	—	—	—	—	$^1S_0$
5B	2	2	1	—	—	—	—	—	$^2P_{1/2}$
6C	2	2	2	—	—	—	—	—	$^3P_0$
7N	2	2	3	—	—	—	—	—	$^4S_{3/2}$
8O	2	2	4	—	—	—	—	—	$^3P_2$
9F	2	2	5	—	—	—	—	—	$^2P_{3/2}$
10Ne	2	2	6	—	—	—	—	—	$^1S_0$
11Na	2	8		1	—	—	—	—	$^2S_{1/2}$
12Mg	2	8		2	—	—	—	—	$^1S_0$
13Al	2	8		2	1	—	—	—	$^2P_{1/2}$
14Si	2	8		2	2	—	—	—	$^3P_0$
15P	2	8		2	3	—	—	—	$^4S_{3/2}$
16S	2	8		2	4	—	—	—	$^3P_2$
17Cl	2	8		2	5	—	—	—	$^2P_{3/2}$
18Ar	2	8		2	6	—	—	—	$^1S_0$
19K	2	8		8		—	1	—	$^2S_{1/2}$
20Ca	2	8		8		—	2	—	$^1S_0$
21Sc	2	8		8		1	2	—	$^2D_{3/2}$
22Ti	2	8		8		2	2	—	$^3F_2$
23V	2	8		8		3	2	—	$^4F_{3/2}$
24Cr	2	8		8		5	1	—	$^7S_3$
25Mn	2	8		8		5	2	—	$^6S_{5/2}$
26Fe	2	8		8		6	2	—	$^5D_4$
27Co	2	8		8		7	2	—	$^4F_{9/2}$
28Ni	2	8		8		8	2	—	$^3F_4$
29Cu	2	8		8		10	1	—	$^2S_{1/2}$
30Zn	2	8		8		10	2	—	$^1S_0$
31Ga	2	8		8		10	2	1	$^2P_{1/2}$
32Ge	2	8		8		10	2	2	$^3P_0$
33As	2	8		8		10	2	3	$^4S_{3/2}$
34Se	2	8		8		10	2	4	$^3P_2$
35Br	2	8		8		10	2	5	$^2P_{3/2}$
36Kr	2	8		8		10	2	6	$^1S_0$

the periodic repetition of the chemical and optical properties of atoms.

In establishing the kind of terms possible with a given electron configuration, we must bear in mind that the Pauli principle does not allow all the combinations of the values of L and S that follow from the configuration. For example, with the configuration np^2 (two electrons with the principal quantum number n and $l = 1$), the possible values of L are 0, 1, 2, while S can have the values 0 and 1. Accordingly, the following terms would seem to be possible:

$$^1S,\ ^1P,\ ^1D,\ ^3S,\ ^3P,\ ^3D \tag{5.59}$$

According to the Pauli principle, however, only such terms are possible for which the values of at least one of the quantum numbers m_l and m_s for equivalent electrons (i.e. electrons with the same n and l) do not coincide*. The term ^3D, for instance, does not comply with this requirement. Indeed, $L = 2$ signifies that the orbital angular momenta of the electrons are "parallel", consequently, the values of m_l for these electrons will coincide. Similarly, $S = 1$ signifies that the spins of the electrons are also "parallel", therefore, the values of m_s also coincide. As a result, all four quantum numbers (n, l, m_l, and m_s) are the same for both electrons, which contradicts the Pauli principle. Thus, the term ^3D in the system of two equivalent electrons cannot be realized.

Table 5.4

m_l			$m_L = \sum m_l$	$m_S = \sum m_s$	
+1	0	−1			
↑↓			+2	0	A
↑	↑		+1	+1	B
↑	↓		+1	0	A
↓	↑		+1	0	B
↓	↓		+1	−1	B
↑		↑	0	+1	B
↑		↓	0	0	A
↓		↑	0	0	B
↓		↓	0	−1	B
	↑↓		0	0	C
	↑	↑	−1	+1	B
	↑	↓	−1	0	A
	↓	↑	−1	0	B
	↓	↓	−1	−1	B
		↑↓	−2	0	A

* This requirement vanishes for non-equivalent electrons, i.e. electrons differing either in n, or in l, or in both of them.

The following procedure is employed to establish the terms of equivalent electrons allowed by the Pauli principle: the values of m_s are indicated in the form of arrows (an arrow pointing upward signifies $m_s = +1/2$, and one pointing downward signifies $m_s = -1/2$) in the columns of a table headed by the values of m_l for an individually taken electron (see Table 5.4 compiled for two equivalent p-electrons). The table contains all the combinations of the values of m_l and m_s for both electrons allowed by the Pauli principle. When both arrows get into one column (this signifies that m_l is the same for both electrons), they are directed oppositely (m_s must be different). In the next two columns of the table, we enter the values of the quantum numbers m_L and m_S equal to the algebraic sum of the numbers m_l and m_s and corresponding to the given combination. The set of allowable values of m_L and m_S permits us to establish the allowable combinations of the values of L and S. One of such sets, marked by the letter A in the last column of the table, corresponds to the combination $L = 2$, $S = 0$, i.e. to the term ^1D; the second set, marked by the letter B, corresponds to $L = 1$, $S = 1$, i.e. to the term ^3P, and, finally, the set marked by the letter C corresponds to $L = 0$, $S = 0$, i.e. to the term ^1S. Thus, of the six formally possible terms indicated in expression (5.59), only three do not contradict the Pauli principle, namely, ^1S, ^3P, ^1D, the term ^3P being a triplet—it splits up into the components $^3P_2$, $^3P_1$, $^3P_0$.

Now the question arises as to which of the terms

$$^1S_0,\ ^3P_2,\ ^3P_1,\ ^3P_0,\ ^1D_2 \tag{5.60}$$

corresponds to the ground state, i.e. to the state with the lowest energy. The answer to this question is given by two empirical **Hund's rules:**

1. Of the terms belonging to a given electron configuration, the term with the greatest possible value of S and with the greatest possible value of L at this S will have the lowest energy.

2. The multiplets formed by equivalent electrons are **normal** (this signifies that the energy of the state grows with an increase in J) if not more than half of the subshell is filled, and are **inverted** (the energy diminishes with an increase in J) if more than half of the subshell is filled.

It follows from Hund's second rule that when not more than half of a subshell is filled, the component of the multiplet with $J = |L - S|$ has the lowest energy, otherwise the component with $J = L + S$ has such an energy.

According to Hund's first rule, one of the P-terms of those given in (5.60) must have the least energy (S is the greatest for these terms). With the configuration np^2, the subshell p is filled only by one-third, i.e. less than half. Consequently, according to Hund's

second rule, the term with the smallest value of J, i.e. the term $^3P_0$, has the lowest energy. It is exactly this term that is the fundamental one for the configuration np^2 (see 6C, 14Si, and 32Ge in Table 5.3).

5.11. X-Ray Spectra

We already noted in Sec. 2.1 that there are two kinds of X-ray radiation—**bremsstrahlung (braking)** and **characteristic** radiation. At not too high energies of the electrons bombarding the anticathode, only bremsstrahlung is observed, which has a continuous spectrum

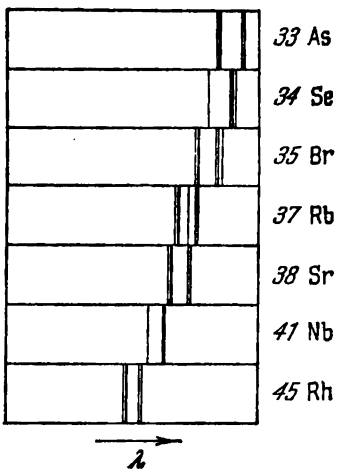


Fig. 5.20

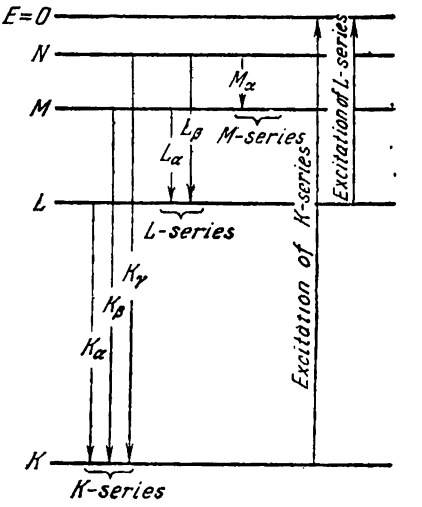


Fig. 5.21

and does not depend on the material of the anticathode. When the energy of the bombarding electron becomes sufficient to knock electrons out of the inner shells of an atom, sharp lines of characteristic radiation appear on the background of the bremsstrahlung. The frequencies of these lines depend on the nature of the substance which the anticathode is made of (this is exactly why the radiation is called characteristic).

X-ray spectra are distinguished by an appreciable simplicity. They consist of several series denoted by the letters K, L, M, N, and O. Each series contains a small number of lines designated in the order of growth of the frequency by the subscripts α, β, γ, ... (K_α, K_β, K_γ, . . .; L_α, L_β, L_γ, . . ., etc.). The spectra of different elements have a similar nature. With an increase in the atomic number Z, the entire X-ray spectrum only shifts to the short-

wave part without changing its structure (Fig. 5.20). The explanation is that the X-ray spectra are produced in transitions of electrons in the inner parts of atoms, and these parts have a similar structure.

A diagram showing how X-ray spectra are produced is given in Fig. 5.21. Excitation of an atom consists in removing one of the inner electrons. If one of the two electrons of the K-shell is knocked out, then the freed site can be occupied by an electron from an outer shell (L, M, N, etc.). Here a K-series is produced. Other series appear in a similar manner. The K-series is attended without fail by other series because when its lines are emitted, levels are freed in shells L, M, etc., which in turn will be filled by electrons from the higher layers.

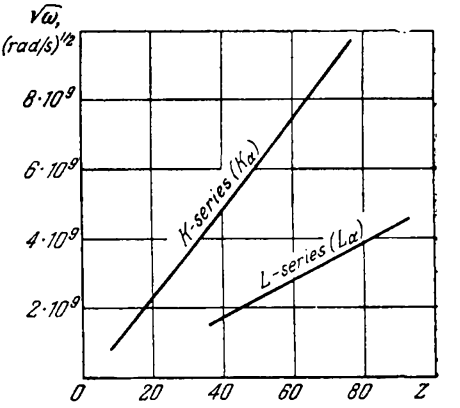


Fig. 5.22

The British physicist Henry Moseley (1887-1915) established a law in 1913 that relates the frequencies of the X-ray spectrum lines to the atomic number Z of the element emitting them. According to this law, the frequencies of the line K_α can be represented by the formula

$$\omega_{K_\alpha} = R(Z-1)^2\left(\frac{1}{1^2}-\frac{1}{2^2}\right)$$

(R is the Rydberg constant), of the line K_β by the formula

$$\omega_{K_\beta} = R(Z-1)^2\left(\frac{1}{1^2}-\frac{1}{3^2}\right)$$

of the line L_α by the formula

$$\omega_{L_\alpha} = R(Z-7.5)^2\left(\frac{1}{2^2}-\frac{1}{3^2}\right)$$

and so on. All these formulas have the form

$$\omega = R(Z-\sigma)^2\left(\frac{1}{n_1^2}-\frac{1}{n_2^2}\right) \qquad (5.61)$$

Moseley's law is usually expressed by the formula

$$\sqrt{\omega} = C(Z-\sigma) \qquad (5.62)$$

(C and σ are constants) and is formulated as follows: *the square root of the frequency is a linear function of the atomic number Z.*

Figure 5.22 shows graphs of $\sqrt{\omega}$ against Z constructed according to experimental points for the lines K_α and L_α. These graphs allow

us to assess the accuracy with which Moseley's law is obeyed. An attentive examination will show that the graph for the line K_α is not completely linear.

Moseley's law makes it possible to exactly establish the atomic number of a given element according to the measured wavelength of the X-ray lines; it played a great part in arranging the elements in the periodic table.

Moseley gave a simple theoretical explanation of the law he discovered. He noted that lines with frequencies determined by formula (5.61) coincide with the lines emitted upon the transition of an electron in the field of the charge $(Z - \sigma)\,e$ from the level numbered n_2 to the one numbered n_1. It is easy to understand the meaning of the constant σ: the electrons performing transitions upon the emission of X-rays are under the action of the nucleus whose attraction is weakened somewhat by the action of the other electrons surrounding it. It is exactly this so-called shielding (or screening) action that is expressed in the need to subtract a certain quantity σ, called the **shielding factor,** from Z.

We must note that Eq. (5.61) is based on the assumption that the shielding factor σ has the same value for both terms. Actually, however, the shielding, for example, for the K-term will be weaker than for the L-term because an electron in the L-shell is shielded by both electrons of the K-shell. In addition, the other electrons of the L-shell play a certain part in shielding, whereas an electron of the K-shell is shielded only by the second K-electron. Formula (5.61) ought to be written more strictly in the form

$$\omega = R\left\{\frac{(Z-\sigma_1)^2}{n_1^2} - \frac{(Z-\sigma_2)^2}{n_2^2}\right\}$$

5.12. Energy of a Molecule

Experiments show that the X-ray spectra of the heavy elements do not depend on what chemical compound the given element is in. It thus follows that the forces retaining atoms in a molecule are due to interaction of the outer electrons. The electrons of the inner shells remain in their previous states when atoms combine into molecules.

In the following, we shall limit ourselves to a consideration of only diatomic molecules. Two kinds of bond between the atoms in a molecule are distinguished. One of them is encountered when the electrons in a molecule can be divided into two groups each of which is constantly near one of the nuclei. The electrons are distributed so that a surplus of electrons is formed near one of the nuclei, and a shortage of them near the other one. Thus, the molecule, as it were, consists of two ions of opposite signs attracting each other. A bond

of this kind is called **heteropolar (or ionic)**. Examples of molecules with a heteropolar bond are NaCl, KBr, and HCl.

The second kind of bond is observed in molecules in which part of the electrons travel about both nuclei. Such a bond is called **homopolar (or covalent, or atomic)**. It is formed by pairs of electrons having oppositely directed spins. Among molecules of this kind, we must distinguish ones with identical nuclei (H_2, N_2, O_2) and ones with different nuclei (for example, CN). In the former molecules, the electrons are distributed symmetrically. In the latter ones, there is a certain asymmetry in the distribution of the electrons owing to which the molecules acquire an electric dipole moment.

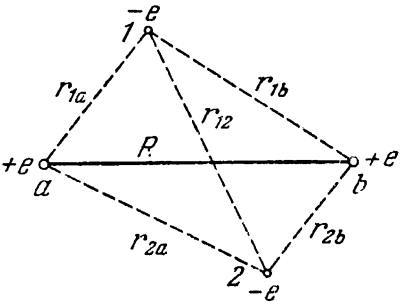


Fig. 5.23

The simplest molecule with a homopolar bond is the hydrogen molecule. Soon after the creation of quantum mechanics, W. Heitler and F. London (1927) successfully tried to perform a quantum-mechanical calculation of the ground state of the H_2 molecule. They succeeded in solving the Schrödinger equation for a system consisting of two protons (hydrogen atom nuclei) and two electrons (Fig. 5.23). The potential energy of such a system is

$$U = -\frac{e^2}{r_{1a}} - \frac{e^2}{r_{2a}} - \frac{e^2}{r_{1b}} - \frac{e^2}{r_{2b}} + \frac{e^2}{r_{12}} + \frac{e^2}{R}$$

The nuclei have a mass that is about 2000 times that of an electron. This is why they move much more slowly than electrons, and in a first approximation, they can be considered stationary. Hence, the Schrödinger equation has the form

$$\nabla_1^2\psi + \nabla_2^2\psi + \frac{2m_e}{\hbar^2}\left[E - e^2\left(\frac{1}{r_{12}} + \frac{1}{R} - \right.\right.$$
$$\left.\left. - \frac{1}{r_{1a}} - \frac{1}{r_{2a}} - \frac{1}{r_{1b}} - \frac{1}{r_{2b}}\right)\right]\psi = 0 \quad (5.63)$$

Here ∇_1^2 is the Laplacian operator containing the coordinates of one electron, and ∇_2^2 is the Laplacian operator containing the coordinates of the other electron.

The eigenvalues of the energy obtained from Eq. (5.63) are found to depend on the distance R between the nuclei, i.e. $E = E(R)$. The nature of this relation appreciably differs for parallel and antiparallel orientation of the spins of the electrons (Fig. 5.24). The formation of a molecule is possible only when atoms with antiparallel spins approach each other. The asymptotic value E_0 which the

energy of a molecule tends to at $R \to \infty$ for both curves shown in the figure is the same and equals the sum of the energies of the isolated atoms.

Matters are similar for other diatomic molecules. The energy due to the electron configuration (the electron energy) has a minimum at a certain value of R and is depicted by a curve of the same kind as that for the hydrogen molecule (see curve *1* in Fig. 5.25).

A change in the electron configuration of a molecule leads to a change in the curve showing how the electron energy depends on the distance R between the nuclei. The asymptotic value of the

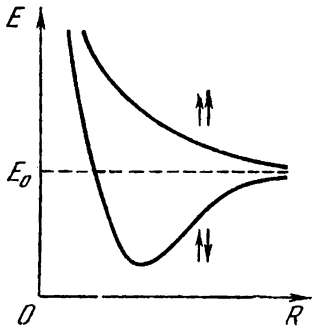


Fig. 5.24

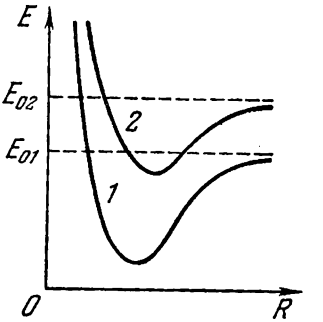


Fig. 5.25

energy also becomes different—equal to the total energy of the isolated atoms in the new quantum state (see curve *2* in Fig. 5.25).

The store of energy in a molecule mainly changes, as in an atom, as a result of changes in the electron configuration forming the peripheral part of the molecule. At a given electron configuration, however, the nuclei of the molecule may vibrate and rotate differently relative to the common centre of inertia. The stores of vibrational and rotational energy are associated with these kinds of motion, and they must be taken into consideration in the total balance. Let us introduce the following notation:

E_e = energy due to the electron configuration (electron energy)
E_v = energy corresponding to the vibrations of a molecule (vibrational energy)
E_r = energy associated with the rotation of a molecule (rotational energy).

In a first approximation, the separate kinds of molecular motions—motion of the electrons, vibration and rotation of a molecule—can be considered to be independent of one another. Hence, the total energy of a molecule can be represented in the form

$$E = E_e + E_v + E_r$$

According to Eq. (4.60), the energy of a harmonic oscillator is determined by the expression

$$E_v = (v + 1/2)\,\hbar\omega_v \quad (v = 0, 1, 2, \ldots) \tag{5.64}$$

where v is the vibrational quantum number, and ω_v is the classical frequency of the oscillator [in Eq. (4.60) these quantities were designated by the symbols n and ω]. We remind our reader that the following selection rule holds for the vibrational quantum number:

$$\Delta v = \pm 1 \tag{5.65}$$

[see Eq. (4.61)].

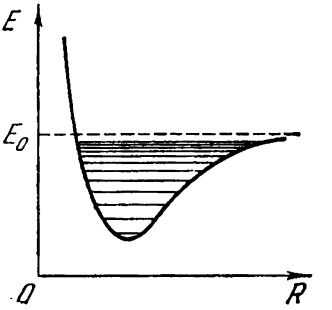


Fig. 5.26

The curve of the potential energy of a molecule (see Fig. 5.25) coincides with a parabola only at small vibrations. The anharmonicity (deviations from harmonicity) that sets in when the intensity of the vibrations grows results in the fact that an increase in the quantum number v is attended by crowding of the levels, their limit being the energy E_0 of a dissociated molecule (Fig. 5.26). At small values of v, however, we may consider with a sufficient degree of accuracy that the vibrational energy of a molecule is determined by Eq. (5.64).

Now let us turn to the rotational energy of a molecule. The energy of a system having the moment of inertia I and rotating with the angular velocity ω_r is

$$E_r = \frac{I\omega_r^2}{2} = \frac{(I\omega_r)^2}{2I} = \frac{M^2}{2I}$$

where $M = I\omega_r$ is the angular momentum of the system. According to Eq. (4.34), the angular momentum can take on only discrete values:

$$M = \hbar\sqrt{J(J+1)} \quad (J = 0, 1, 2, \ldots)$$

(J is the quantum number of the angular momentum). Hence, the rotational energy of a molecule can have only quantized values:

$$E_r = \frac{\hbar^2 J(J+1)}{2I} \tag{5.66}$$

where I is the moment of inertia of a molecule relative to the axis passing through its centre of inertia, and J is the **rotational quantum number** taking on values of 0, 1, 2, etc.

The following selection rule holds for the rotational quantum number:

$$\Delta J = \pm 1 \tag{5.67}$$

Thus, according to Eqs. (5.64) and (5.66), the total energy of a molecule is

$$E = E_e + \left(v + \frac{1}{2}\right)\hbar\omega_v + \frac{\hbar^2 J(J+1)}{2I} \qquad (5.68)$$

Experiments and theory show that the distance between the rotational levels ΔE_r is considerably smaller than the distance between the vibrational levels ΔE_v. The latter distance, in turn, is considerably smaller than that between the electron levels ΔE_e. Consequently, the diagram of the energy levels of a diatomic molecule has the appearance shown in Fig. 5.27 (only two electron levels are given). The collection of levels is contained in the right-hand column. The first two columns only explain the appearance of the levels.

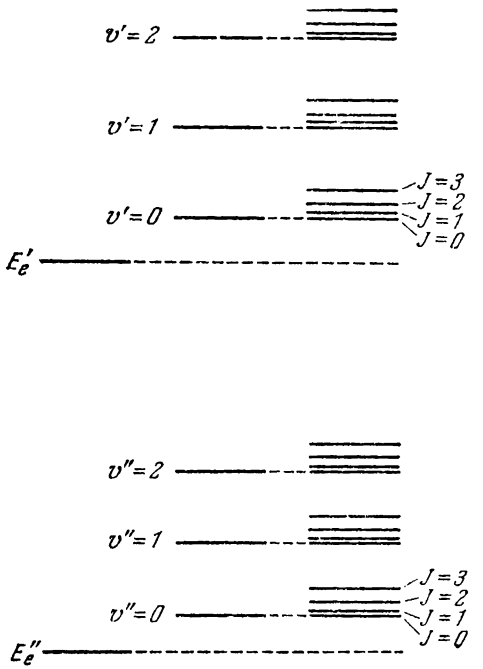


Fig. 5.27

5.13. Molecular Spectra

Whereas atomic spectra consist of separate lines, molecular spectra when observed in an instrument of medium resolving power are seen to consist of bands (see Fig. 5.28, which depicts a portion of the spectrum obtained for a glow discharge in air). When instruments having a high resolving power are used, the bands are found to consist of a great number of closely arranged lines (see Fig. 5.29, which depicts the fine structure of one of the bands of the nitrogen molecule spectrum).

In accordance with their nature, the spectra of molecules are known as **band spectra**. Three kinds of bands are distinguished depending on the kind of energy (electron, vibrational, or rotational) whose change results in a molecule emitting a photon. They are (1) rotational, (2) vibrational-rotational, and (3) electron-vibrational ones. The bands in Fig. 5.28 belong to the electron-vibrational type. Bands of this type are characterized by the presence of a sharp band edge. The other end of such a band is blurred. The edge is due to the

crowding of the lines forming the band. Rotational and vibrational-rotational bands have no edge.

We shall limit ourselves to a treatment of the rotational and vibrational-rotational spectra of diatomic molecules. The energy of such molecules consists of electron, vibrational, and rotational energies [see Eq. (5.68)]. All three kinds of energy have a minimum

Fig.5.28

Fig. 5.29

value in the ground state of a molecule. When a molecule receives a sufficient amount of energy, it transfers to an excited state, and then, performing a transition to one of the lower energy states allowed by the selection rules, it emits a photon:

$$\hbar\omega = \Delta E_e + \Delta E_v + \Delta E_r =$$
$$= E'_e - E''_e + (v' + 1/2)\,\hbar\omega'_v - (v'' + 1/2)\,\hbar\omega''_v +$$
$$+ \frac{\hbar^2 J'(J'+1)}{2I'} - \frac{\hbar^2 J''(J''+1)}{2I''}$$

(it must be borne in mind that both ω_v and I differ for different electron configurations of a molecule).

It was indicated in the preceding section that

$$\Delta E_e \gg \Delta E_v \gg \Delta E_r$$

Therefore, with weak excitations, only E_r changes, with stronger ones, E_v, and only with still stronger excitations does the electron configuration of the molecule, i.e. E_e, change.

Rotational Bands. Photons corresponding to transitions of a molecule from one rotational state to another have the smallest energy

(the electron configuration and the vibrational energy do not change in this case):

$$\hbar\omega = \Delta E_r = \frac{\hbar^2 J'(J'+1)}{2I} - \frac{\hbar^2 J''(J''+1)}{2I}$$

The possible changes in the quantum number J are restricted by selection rule (5.67). Hence, the frequencies of the lines emitted

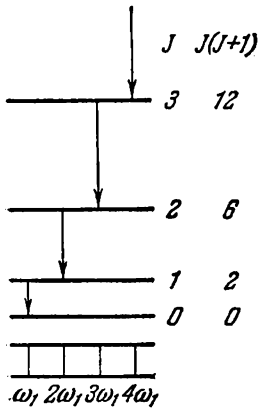


Fig. 5.30

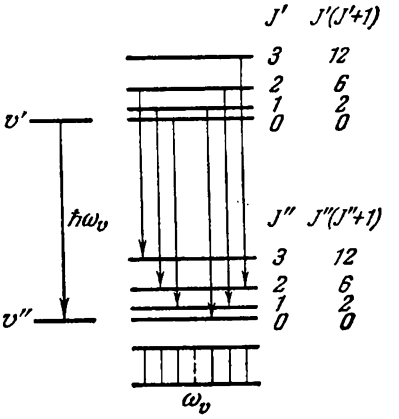


Fig. 5.31

in transitions between the rotational levels can have the values

$$\omega = \frac{\Delta E_r}{\hbar} = B[(J+1)(J+2) - J(J+1)] = 2B(J+1) = \omega_1(J+1)$$

where J is the quantum number of the level to which the transition occurs (it can have the values 0, 1, 2, . . .), and

$$B = \frac{\hbar}{2I} \tag{5.69}$$

Figure 5.30 shows schematically the appearance of a rotational band. A rotational spectrum consists of a number of equispaced lines in the very far infrared region. By measuring the distance between the lines $\Delta\omega = \omega_1$, we can find the constant B of Eq. (5.69) and calculate the moment of inertia of a molecule. Next, knowing the masses of the nuclei, we can calculate the equilibrium distance R_0 between them in a diatomic molecule.

The line spacing $\Delta\omega$ is of the order of 10^{13} rad/s, so that we get values of the order of 10^{-40} g·cm^2 for the moments of inertia of molecules. For example, for the molecule HCl, we have $I = 2.71 \times \times 10^{-40}$ g·cm^2, which corresponds to $R_0 = 1.29$ Å.

Vibrational-Rotational Bands. When both the vibrational and the rotational states of a molecule change in a transition (Fig. 5.31),

the energy of the emitted photon will be

$$\hbar\omega = \Delta E_v + \Delta E_r = \hbar\omega_v (v' + 1/2) - \hbar\omega_v (v'' + 1/2) + \\ + \frac{\hbar^2 J' (J'+1)}{2I} - \frac{\hbar^2 J'' (J''+1)}{2I}$$

The quantum number v obeys selection rule (5.65), and J obeys rule (5.67).

Since $\Delta E_v \gg \Delta E_r$, the emission of a photon can be observed not only when $J' > J''$, but also when $J' < J''$. If $J' > J''$, the frequencies of the photons are determined by the formula

$$\omega = \omega_v + B [(J + 1) (J + 2) - J (J + 1)] = \\ = \omega_v + 2B (J + 1) = \omega_v + 2Bk \quad (k = 1, 2, 3, \ldots)$$

where J is the rotational quantum number of the lower level that can take on the values 0, 1, 2, . . ., and B is the quantity given by Eq. (5.69). If $J' < J''$, the formula for the frequency of the photons has the form

$$\omega = \omega_v + B [(J - 1) J - J (J + 1)] = \\ = \omega_v - 2BJ = \omega_v - 2Bk \quad (k = 1, 2, 3, \ldots)$$

where J is the rotational quantum number of the lower level that can take on values of 1, 2, . . . (in this case $J'' = J$ cannot have the value of 0 because J' would be -1).

Both cases can be covered by the single formula

$$\omega = \omega_v \pm 2Bk = \omega_v \pm \omega_1 k \quad (k = 1, 2, 3, \ldots)$$

The collection of lines with frequencies determined by this formula is called the **vibrational-rotational band.** The vibrational part of the frequency ω_v determines the spectral region in which the band is; the rotational part $\pm\omega_1 k$ determines the fine structure of the band, i.e. the splitting of the individual lines. The region in which the vibrational-rotational bands are extends approximately from 8000 to 50 000 Å. A glance at Fig. 5.31 shows that the vibrational-rotational band consists of a collection of lines symmetrical relative to ω_v and spaced at a distance of $\Delta\omega = \omega_1$. Only at the middle of the band is the distance twice as great because no line having the frequency ω_v is produced.

The distance between the components of the vibrational-rotational band is related to the moment of inertia of a molecule by the same expression as holds for a rotational band. We can thus find the moment of inertia of a molecule by measuring this distance.

We must note that rotational and vibrational-rotational spectra completely corresponding to the conclusions of theory are observed experimentally only for non-symmetrical diatomic molecules (i.e.

for molecules formed by two different atoms). In symmetrical molecules, the dipole moment equals zero, which leads to the forbidding of rotational and vibrational-rotational transitions. Electron-vibrational spectra are observed both for non-symmetrical and for symmetrical molecules.

5.14. Combination Scattering of Light

In 1928, the Soviet scientists Grigori Landsberg (1890-1957) and Leonid Mandelshtam (1879-1944), and simultaneously the Indian physicist Chandrasekhara Raman (1888-1970) discovered a phenomenon consisting in that the scattered light spectrum produced when

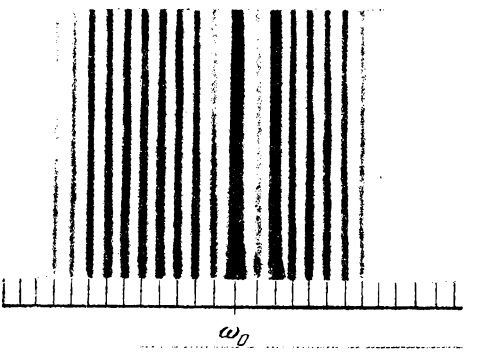


Fig. 5.32

light passes through gases, liquids, or transparent crystalline bodies apart from the unshifted line contains new lines whose frequencies ω are a combination of the frequency of the incident light ω_0 and the frequencies ω_i of the vibrational or rotational transitions of the molecules scattering the light:

$$\omega = \omega_0 \pm \omega_i \tag{5.70}$$

This phenomenon was named the **combination scattering of light.***

Figure 5.32 shows a spectrum of the combination scattering of oxygen excited by the line Hg 2536.5 Å. Onto the line of combination scattering to the right of the line of the source there was superposed

* This phenomenon is usually called the **Raman effect** in foreign publications.

the line Hg 2534.8 Å (less intensive than Hg 2536.5 Å) owing to which the intensity of this line obtained was greater than that of the others. Inspection of the figure shows that the combination scattering spectrum consists of the unshifted line ω_0 and a number of satellites arranged symmetrically relative to it. To each "red" satellite (i.e. a satellite shifted toward greater wavelengths) with the frequency $\omega_0 - \omega_i$ there corresponds a "violet" satellite* with the frequency $\omega_0 + \omega_i$. At ordinary temperatures, the intensity of the violet satellites is considerably lower than that of the red ones. The intensity of the violet satellites rapidly grows with elevation of the temperature.

According to the quantum theory, the process of light scattering can be considered as the inelastic collision of photons with molecules. In a collision, a photon can give up to a molecule or receive from it only such amounts of energy that equal the differences between two of its energy levels. If upon colliding with a photon, a molecule passes from a state with the energy E' to a state with the energy E'' ($E'' > E'$), then the energy of the photon after scattering will become equal to $\hbar\omega_0 - \Delta E$, where $\Delta E = E'' - E'$. Accordingly, the frequency of the photon will diminish by $\omega_1 = \Delta E/\hbar$— a red satellite appears. If a molecule was initially in a state with the energy E'', it may pass over into a state with the energy E' because of colliding with a photon and give up its surplus energy $\Delta E = E'' - E'$ to the photon. As a result, the energy of the photon will become equal to $\hbar\omega_0 + \Delta E$, and the frequency will grow by ω_1. The scattering of the photon $\hbar\omega_0$ may be attended by transitions of a molecule between different rotational or vibrational levels E', E'', E''', etc. The result is the appearance of a number of symmetrically arranged satellites.

At ordinary temperatures, the number of molecules in the ground state greatly exceeds the number of molecules in excited states. Hence, collisions attended by diminishing of the energy of a molecule occur much more rarely than collisions attended by an increase in the energy. This explains the low intensity of violet satellites in comparison with red ones. The number of excited molecules rapidly grows with elevation of the temperature, and the result is an increase in the intensity of the violet satellites.

The investigation of combination scattering gives a lot of information on the structure of molecules. The natural frequencies of vibrations of molecules are determined with the aid of this method. It also allows us to assess the nature of symmetry of a molecule. In crystals, the combination scattering of light is usually associated with the so-called optical branch of oscillations of a crystal lattice

* The red satellites are also called **Stokes lines**, and the violet ones, **anti-Stokes lines.**

(see Sec. 6.4). The spectra of combination scattering are so characteristic of molecules that they are used in the analysis of complicated molecular mixtures, especially of organic molecules whose analysis by chemical methods is very difficult or even impossible.

We shall note that combination scattering relates to the so-called non-linear effects (see Sec. 5.17).

5.15. Stimulated Emission

Up to now, we have considered only two kinds of transitions of atoms between energy levels—spontaneous ones from higher to lower levels, and transitions from lower to higher levels occurring under the action of radiation (stimulated transitions). Transitions of the first kind result in the spontaneous emission of photons by atoms, while transitions of the second kind result in the absorption of radiation by a substance.

In 1918, Albert Einstein gave attention to the circumstance that the two kinds of radiation indicated above are not sufficient for explaining the existence of states of equilibrium between radiation and a substance. Indeed, the probability of spontaneous transitions is determined only by the internal properties of atoms and, consequently, cannot depend on the intensity of the incident radiation, whereas the probability of "absorbing" transitions depends both on the properties of atoms and on the intensity of the incident radiation. To permit equilibrium to set in at an arbitrary intensity of the incident radiation, the existence of "emission" transitions is needed whose probability would grow with an increasing intensity of radiation, i.e. of "emission" transitions produced by radiation. The emission produced as a result of such transitions is called **stimulated or induced emission.**

Einstein proved on the basis of thermodynamic considerations that the probability of stimulated transitions attended by radiation must equal the probability of stimulated transitions attended by the absorption of light. Thus, stimulated transitions may occur with equal probability in either direction.

Stimulated emission has very important properties. The direction of its propagation exactly coincides with the direction of propagation of the stimulating radiation, i.e. of the external radiation producing a transition. The same relates to the frequency, phase, and polarization of the stimulated emission and stimulating radiation. Consequently, the stimulated emission and the radiation stimulating it are strictly coherent. This feature of stimulated emission underlies the action of light amplifiers and generators known as lasers (see the following section).

Assume that P_{nm} is the probability of a stimulated transition of an atom in unit time from the energy level E_n to the level E_m, and P_{mn} is the probability of the reverse transition. It was indicated above that at an identical intensity of radiation, $P_{nm} = P_{mn}$. The probability of stimulated transitions is proportional to the density of the energy u_ω of the electromagnetic field inducing the transition* falling to the frequency ω corresponding to the given transition $[\omega = (E_n - E_m)/\hbar]$. Letting B stand for the coefficient of proportionality, we get

$$P_{nm} = B_{nm}u_\omega, \quad P_{mn} = B_{mn}u_\omega \tag{5.71}$$

The quantities B_{nm} and B_{mn} are known as **Einstein's coefficients** According to what has been said above, $B_{nm} = B_{mn}$.

Einstein gave a very simple derivation of Planck's formula on the basis of the equal probability of the stimulated transitions $n \to m$ and $m \to n$. Equilibrium between a substance and radiation will be achieved provided that the number of atoms N_{nm} performing the transition from the state n to the state m in unit time will equal the number of atoms N_{mn} performing the transition in the opposite direction. Assume that $E_n > E_m$. Hence, the transitions $m \to n$ will be able to occur only under the action of radiation, whereas the transitions $n \to m$ will occur both under stimulation and spontaneously. Consequently,

$$N_{mn} = N_{mn}^{(\text{stim})}, \quad N_{nm} = N_{nm}^{(\text{stim})} + N_{nm}^{(\text{spont})}$$

The equilibrium condition has the form

$$N_{mn}^{(\text{stim})} = N_{nm}^{(\text{stim})} + N_{nm}^{(\text{spont})} \tag{5.72}$$

According to Eq. (5.71)

$$N_{mn}^{(\text{stim})} = P_{mn}N_m = B_{mn}u_\omega N_m \tag{5.73}$$

$$N_{nm}^{(\text{stim})} = P_{nm}N_n = B_{nm}u_\omega N_n \tag{5.74}$$

(N_m and N_n are the numbers of atoms in the states m and n).

Let us denote the probability of a spontaneous transition of an atom from the state n to the state m in unit time by A_{nm}. Hence, the number of atoms performing a spontaneous transition $n \to m$ in unit time is determined by the expression

$$N_{nm}^{(\text{spont})} = A_{nm}N_n \tag{5.75}$$

The introduction of Eqs. (5.73), (5.74), and (5.75) into Eq. (5.72) yields

$$B_{mn}u_\omega N_m = B_{nm}u_\omega N_n + A_{nm}N_n$$

* In Sec. 1.7, we denoted the equilibrium value of u_ω by $u(\omega, T)$.

The value of u_ω determined by this equation is the equilibrium value of this quantity, i.e. $u(\omega, T)$. Thus,

$$u(\omega, T) = \frac{A_{nm}N_n}{B_{mn}N_m - B_{nm}N_n} = \frac{A_{nm}}{B_{nm}} \frac{1}{N_m/N_n - 1}$$

(we have taken into account that $B_{mn} = B_{nm}$).

The equilibrium distribution of atoms among states with a different energy is determined by Boltzmann's law, according to which

$$\frac{N_m}{N_n} = \exp\left(\frac{E_n - E_m}{kT}\right) = \exp\left(\frac{\hbar\omega}{kT}\right)$$

Consequently, we arrive at the formula

$$u(\omega, T) = \frac{A_{nm}}{B_{nm}} \frac{1}{\exp(\hbar\omega/KT) - 1} \tag{5.76}$$

To determine the coefficient A_{nm}/B_{nm}, Einstein took advantage of the fact that at low frequencies, Eq. (5.76) must transform into the Rayleigh-Jeans formula. When $\hbar\omega \ll kT$, the substitution $\exp(\hbar\omega/kT) \approx 1 + \hbar\omega/kT$ can be made, as a result of which Eq. (5.76) acquires the form

$$u(\omega, T) = \frac{A_{nm}}{B_{nm}} \frac{kT}{\hbar\omega}$$

A comparison with Eq. (1.52) gives for A_{nm}/B_{nm} the value

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

Introduction of this value into Eq. (5.76) leads to Planck's formula [see Eq. (1.61)].

5.16. Lasers

In 1939, the Soviet physicist V. Fabrikant first indicated the possibility of obtaining media in which light will be amplified at the expense of stimulated emission (see the preceding section). In 1953, the first molecular generators operating in the range of centimetre waves and named **masers** were developed independently by the Soviet scientists N. Basov and A. Prokhorov and the American scientists C. Townes and J. Weber*. The word maser is an acronym for *m*icrowave *a*mplification by *s*timulated *e*mission of *r*adiation. In 1960, T. Meiman (USA) developed the first similar device operating in the optical range—the **laser** (*l*ight *a*mplification by *s*timulated *e*mission of *r*adiation). Lasers are sometimes known as **optical quantum generators.**

* In 1964, Basov, Prokhorov, and Townes were awarded the Nobel prize for this work.

We found out in the preceding section that when light of frequency ω acts on a substance and its frequency coincides with one of the frequencies $(E_n - E_m)/\hbar$ of the atoms of the substance $(E_n > E_m)$, it will set up two processes: (1) a stimulated transition $m \to n$, and (2) a stimulated transition $n \to m$. The first process leads to the absorption of light and attenuation of the incident beam, while the second one leads to an increase in the intensity of the incident beam. The resultant change in the intensity of the light beam depends on which of the two processes predominates.

For thermodynamic equilibrium, the distribution of the atoms by different energy states is determined by Boltzmann's law:

$$N_i = \frac{N \exp(-E_i/kT)}{\sum_j \exp(-E_j/kT)} = C \exp\left(-\frac{E_i}{kT}\right) \tag{5.77}$$

where N is the total number of atoms, and N_i is the number of atoms at the temperature T in a state with the energy E_i (we have assumed for simplicity that all the energy levels are not degenerate). It can be seen from this formula that the population of a level, i.e. the number of atoms in a given state, diminishes with an increase in the energy of the state. The number of transitions between two levels is proportional to the population of the initial level. Consequently, in a system of atoms in thermodynamic equilibrium, the absorption of the incident light wave will predominate over stimulated emission, so that the incident wave is attenuated when passing through the substance.

To obtain amplification of the incident wave, we must invert the population of the energy levels, i.e. ensure that there are more atoms in the state with the higher energy E_n than in the state with the lower energy E_m. In this case, the given collection of atoms is said to have an **inverse population**. According to Eq. (5.77)

$$\frac{N_n}{N_m} = \exp\left(-\frac{E_n - E_m}{kT}\right)$$

For an inverse population, $(N_n/N_m) > 1$ at $(E_n - E_m) > 0$. Formally extending distribution (5.77) to this case, we get a negative value for T. Therefore, states with an inverse population are sometimes called states with a negative temperature.

The change in the intensity of light when it passes through an absorbing medium is described by the formula

$$I = I_0 e^{-\varkappa l} \tag{5.78}$$

In a substance with inverse population of the energy levels, the stimulated emission may exceed the absorption of light by the atoms, and as a result the incident beam of light will be amplified when passing through the substance. The phenomenon of amplification of the incident beam proceeds as if the absorption coefficient $\varkappa$ in

formula (5.78) became negative. Accordingly, a collection of atoms with inverse population may be treated as a medium with a negative absorption coefficient.

The creation of the laser became possible after ways were found for inverting the population of the levels in certain substances. In the first laser constructed by Meiman, the working substance was a pink ruby cylindrical rod. The diameter of the rod was about 1 cm, and its length about 5 cm. The ends of the ruby rod were thoroughly polished to form mirrors strictly parallel to each other.

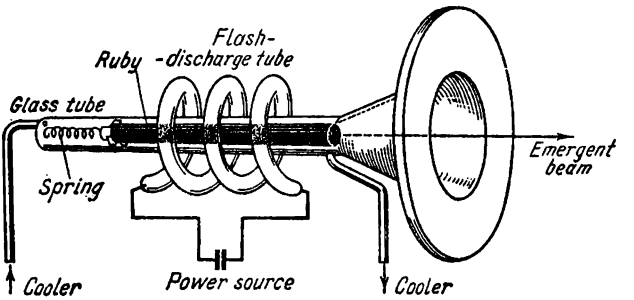


Fig. 5.33

One end was coated with a dense opaque layer of silver, and the other end was coated with a layer of silver that transmitted about eight per cent of the energy falling on it.

A ruby is aluminium oxide (Al_2O_3) in which some of the aluminium atoms are substituted by chromium atoms. When light is absorbed, the chromium ions Cr^{3+} (the chromium is in the ruby crystal in this form) become excited. The reverse transition to the ground state occurs in two stages. In the first of them, the excited ions give up part of their energy to the crystal lattice and pass into a metastable state. The transition from the metastable state to the ground one is forbidden by the selection rules. Therefore, the average lifetime of an ion in the metastable state ($\sim 10^{-3}$ s) is about 10^5 times greater than the lifetime in the ordinary excited state. In the second stage, the ions pass from the metastable state to the ground one* emitting a photon with $\lambda = 6943$ Å. Under the action of photons of the same wavelength, i.e. in stimulated emission, the chromium ions pass from the metastable state to the ground one much more rapidly than in spontaneous emission.

The ruby in a laser is illuminated by a flash-discharge xenon tube (Fig. 5.33) that produces light with a broad band of frequencies.

* The selection rules are not absolutely strict. The probability of the forbidden transitions is considerably smaller than of the allowed ones, but nevertheless differs from zero.

When the power of the tube is adequate, most of the chromium ions pass into the excited state. The process of imparting energy to the working substance of a laser to transfer the atoms into the excited state is called **pumping.** Figure 5.34 gives a diagram of the levels of the chromium ion Cr^{3+} (level 3 is a band formed by a collection of closely arranged levels).

The excitation of the ions as a result of pumping is depicted by arrow W_{13}. The lifetime of level *3* is very small ($\sim 10^{-8}$ s). During this time, some ions pass spontaneously from band *3* to ground level *1*. Such transitions are depicted by arrow A_{31}. Most of the ions, however, will pass to metastable level *2* (the probability of the transition depicted by arrow S_{32} is much greater than that of transition A_{31}). When the pumping power is adequate, the number of chromium ions at level *2* becomes greater than their number at level *1*. Consequently, levels *1* and *2* become inverted.

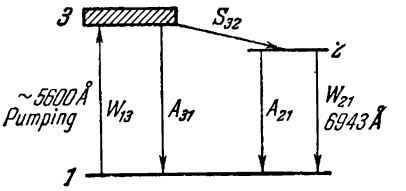


Fig.5.34

Arrow A_{21} depicts a spontaneous transition from the metastable level to the ground one. The emitted photon may produce stimulated emission of additional photons (transition W_{21}) which, in turn, will produce stimulated emission, etc. A cascade of photons is formed as a result. We remind our reader that the photons produced in stimulated emission fly in the same direction as the incident photons. The photons whose directions of motion form small angles with the axis of the crystal rod experience multifold reflection from its ends. Therefore, their path in the crystal will be very long, so that the cascades of photons in the direction of the axis will receive special development. The photons emitted spontaneously in other directions emerge from the crystal through its side surface.

The process of formation of a cascade is shown schematically in Fig. 5.35. Before the beginning of a pulse, the chromium ions are in the ground state (the black circles in Fig. 5.35*a*). The pumping light (the solid arrows in Fig. 5.35*b*) transfers most of the ions to the excited state (white circles). A cascade begins to develop when the excited ions spontaneously emit photons (the dash arrows in Fig. 5.35*c*) in a direction parallel to the axis of the crystal (the photons emitted in other directions emerge from the crystal). The photons multiply at the expense of the stimulated emission. This process develops (Fig. 5.35*d* and *e*) because the photons repeatedly pass along the crystal, being reflected from its ends. When the beam becomes sufficiently intense, part of it emerges through the half-silvered end of the crystal (Fig. 5.35*f*).

Ruby lasers are pulsed ones (with a frequency of the order of several pulses a minute). A large amount of heat is liberated inside the crystal. It therefore has to be intensively cooled, which is done with the aid of liquid air.

In 1961, A. Javan developed the first gas laser operating on a mixture of helium and neon. In 1963, the first semiconductor lasers

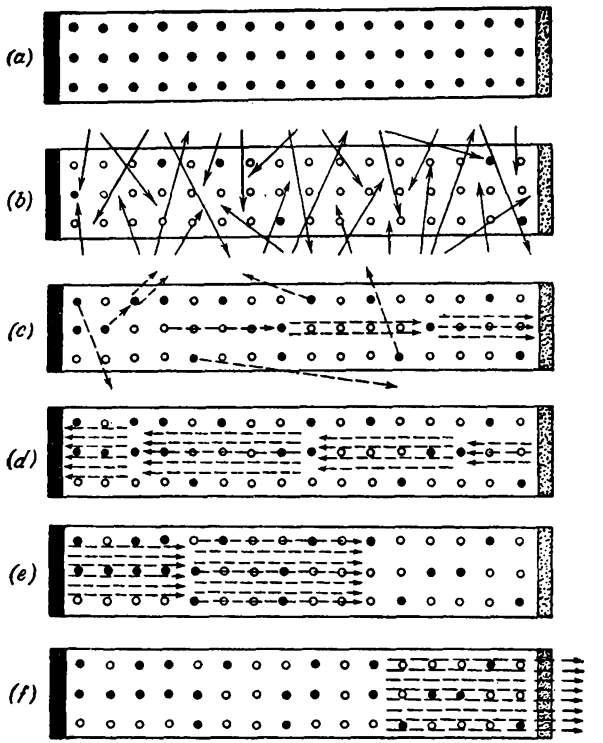


Fig. 5.35

were designed. At present, the list of laser materials includes many scores of solid and gaseous substances.

Laser radiation is distinguished by a number of remarkable features. It is characterized by (1) a strictly monochromatic nature ($\Delta\lambda \sim 0.1$ Å), (2) a high temporal and space coherence, (3) a high intensity, and (4) narrowness of the beam. The angular width of the light beam generated by a laser is so small that by using telescopic focussing, it is possible to obtain a spot of light with a diameter of only three kilometres on the Moon's surface. The high power and narrowness of the beam make it possible, when focussing with the

aid of a lens, to obtain a density of the energy flux that is 1000 times greater than the density of the energy flux that can be obtained by focussing sunlight. Light beams with such a high density of their power can be employed for mechanical processing and welding, for acting on the course of chemical reactions, etc.

The high coherence of the radiation opens up broad prospects for the use of lasers for radio communication, in particular for directed radio communication in outer space. If a method of modulation and demodulation of light is found, one laser will be able to replace with respect to the volume of transmitted information the entire system of communication between the east and west coasts of the USA.

The high coherence of a laser beam made it possible to bring to life such a remarkable phenomenon as holography.

What has been said above far from exhausts all the possibilities of the laser. It is an absolutely new type of light source, and it is meanwhile difficult to imagine all the possible fields of its application.

5.17. Non-Linear Optics

In a light wave produced with the aid of conventional (non-laser) light sources, the electric field strength E is negligibly small in comparison with the strength of the internal microscopic field acting on the electrons in a substance. For this reason, the optical properties of the medium (in particular, the refractive index) and the nature of the overwhelming majority of optical phenomena do not depend on the intensity of light. In this case, the propagation of light waves is described by linear differential equations. Therefore, prelaser optics can be called linear. We must note that the principle of light wave superposition (expressed in geometrical optics by the law of the independence of light rays) holds only in the region of linear optics. True, non-linear phenomena were also known in optics before the development of lasers. They include, for example, the combination scattering of light (the Raman effect). In combination scattering, the transformation of the frequency of a monochromatic light wave is observed, which is a feature of the non-linear nature of the process. In the predominating majority of cases, however, the optical processes were linear.

After the appearance of lasers, matters in optics changed quite appreciably. Quantum generators (lasers) make it possible to produce light waves with a field strength of almost the same magnitude as the strength of the microscopic field in atoms. For such fields, the refractive index depends on the strength E. In this case, the superposition principle is violated, the different waves propagating in a medium affect one another, and a number of non-linear optical phenomena appear. We shall briefly describe some of them.

Non-Linear Reflection of Light. At high intensities, radiation on the second harmonic of the incident radiation appears in the reflected light. In other words, in addition to the reflected ray having the frequency ω equal to the frequency of the incident light, a reflected ray of intensity 2ω is observed. This ray does not obey the conventional law of reflection, and as a consequence the direction of the reflected ray of frequency 2ω does not coincide with the direction of the reflected ray of frequency ω.

Self-Focussing of Light. At conventional intensities, an initially parallel narrow beam of light when propagating in a vacuum or in a medium undergoes so-called diffraction spreading, as a result of which diffraction divergence of the beam appears. It was found that when light beams propagate in liquids and in certain crystals, an increase in the power of a beam is attended by diminishing of its divergence. At a certain power called the critical one, the beam propagates without any divergence. Finally, at a power higher than the critical one, the beam contracts—self-focussing of the beam takes place in the medium. This phenomenon is due to the fact that the refractive index increases with a growth in the strength E. Therefore, the medium becomes optically denser in the region occupied by the beam, and this leads to bending of the rays toward the beam axis, i.e. to contraction of the beam.

Optical Harmonics. In the scattering of a laser beam in liquids and crystals, in addition to the light with the frequency of the incident radiation ω, scattered light is observed having frequencies that are multiples of the initial frequency (i.e. the frequencies 2ω, 3ω, etc.). These components of the scattered light are known as optical harmonics. The intensity of the optical harmonics may be quite considerable; in some crystals up to 50% of the power of the scattered radiation may transform into radiation of the harmonics.

Multiple-Photon Processes. At conventional intensities, only one photon is absorbed in an elementary event of interaction of light with a substance. The energy $\hbar\omega$ of the photon coincides with the difference between the energy levels $E_2 - E_1$ of the relevant atom or molecule. At high intensities, two or more photons may be absorbed in an elementary event of interaction. In this case, light not only of the frequency $\omega = (E_2 - E_1)/\hbar$, but also of the frequencies $\omega/2$, $\omega/3$, etc. may be absorbed. Such absorptionis called multiple-photon (in particular, two-photon, three-photon, etc.).

In one elementary event of interaction of light with a substance, two photons of different frequencies may be absorbed. This occurs when the light field is set up by two independent monochromatic sources. If the sum of the frequencies of these sources satisfies the condition $\omega_1 + \omega_2 = (E_2 - E_1)/\hbar$, an appreciable absorption of the radiation of both frequencies is observed. For this to occur, the

two radiations do not necessarily have to be of a high power. It is sufficient that their total intensity will be high. Therefore, multiple-photon absorption can be observed when light from a laser and a non-laser source with a continuous spectrum is superposed.

Multiple-photon processes also include the **multiple-photon photo-electric effect** (the multiple-photon ionization of atoms). Whereas the conventional (single-photon) photo-electric effect is observed at frequencies at which the energy of a photon is greater than the energy of ionization of an atom, the multiple-photon photo-electric effect can occur at frequencies that are $1/n$-th of these frequencies (n is the number of photons participating in an elementary event of interaction). The seven-photon ionization of inert gases has been reliably registered.

We have given a far from complete list of already discovered non-linear phenomena. It is sufficient, however, to form an idea of how rapidly the new branch of optics—non-linear optics—is developing.

PART III SOLID STATE PHYSICS*

CHAPTER 6 OSCILLATIONS OF A CRYSTAL LATTICE

6.1. Crystal Lattice. Miller Indices

An ideal crystal lattice is formed of identical unit cells. Each of them in the general case is an oblique parallelepiped constructed on the three vectors **a**, **b**, **c**. The latter can be taken as the unit vectors of the coordinate axes. The magnitudes of the vectors are the periods of identity in the directions of the relevant axes.

The choice of the coordinate axes, generally speaking, is not unique. The same crystal can be represented as consisting of different unit parallelepipeds. It is customary practice to choose the axes in the simplest way with account taken of the symmetry of the crystal.

Special symbols are used for an analytical description of the geometrical elements of a crystal, i.e. of its points, straight lines (directions) and planes.

Let us take a point with the coordinates x, y, z**. We shall adopt the combination of the quantities $\frac{x}{a}$, $\frac{y}{b}$, $\frac{z}{c}$, which are confined in double brackets, as the **indices of a point:** $\left[\left[\frac{x}{a}\,\frac{y}{b}\,\frac{z}{c}\right]\right]$. Usually, we have in mind points within the limits of a cell adjoining the origin of coordinates. In this case, the indices will be numbers not exceeding 1. For example, the indices $\left[\left[\frac{1}{2}\,\frac{1}{2}\,\frac{1}{2}\right]\right]$ correspond to the centre

---

* In the present part of the course, by solids we mean crystalline substances. Some information on crystals was given in Chapter 13 of Vol. I. There we set out information on the classification of crystals, the physical kinds of crystal lattices, and defects in crystals. In Vol. II we treated the classical theory of the electrical conductance of metals (Chapter 11), and also gave some information on the magnetic properties of bodies (Chapter 7). Naturally, we shall not repeat all this information here and shall limit ourselves to settling out the material that was not dealt with in the preceding volumes of the course.

** In the general case, these coordinates are oblique, and not Cartesian ones.

of the crystal cell, and the indices $\left[\left[0\,\frac{1}{2}\,\frac{1}{2}\right]\right]$ to the centre of the face in plane *yz*.

A direction in a crystal can be set with the aid of a straight line passing through the origin of coordinates. The direction of such a straight line is determined by the smallest integers m, n, p that are proportional to the indices of any point through which the line passes*.

$$m : n : p = \frac{x}{a} : \frac{y}{b} : \frac{z}{c}$$

The numbers m, n, p are called **direction indices** and are confined in single brackets: $[m, n, p]$. Thus, the direction of a straight line passing through the origin of coordinates and the point $\left[\left[\frac{1}{3}\,\frac{1}{2}\,1\right]\right]$ is designated by the symbol [2 3 6].

If one of the numbers m, n, p is negative, the minus sign is put not in front of, but on top of the relevant number. For example, the direction opposite to the y-axis is designated by the symbol $[0\ \bar{1}\ 0]$.

The position of a plane in a crystal can be determined by setting the intercepts u, v, w formed by the plane on the coordinate axes. For planes passing through the lattice points, however, it is more convenient to set the position of a plane with the aid of the smallest integers h, k, l that are the reciprocals of the intercepts u, v, w:

$$h : k : l = \frac{1}{u} : \frac{1}{v} : \frac{1}{w}$$

The numbers, h, k, l are known as the **Miller indices.** When writing the symbol of a plane, we confine the Miller indices in parentheses: $(h\ k\ l)$. Assume, for instance, that the intercepts formed by a plane on the coordinate axes are $\frac{1}{2}$, $\frac{2}{3}$, and 1. Their reciprocals will be 2, $\frac{3}{2}$, and 1. Multiplying these numbers by 2, we get the Miller indices: (4 3 2).

When an intercept formed by a plane on a coordinate axis is negative, the minus sign is put not in front, but on top of the corresponding Miller index. If a plane is parallel to a coordinate axis, the intercept on this axis is infinitely great, so that the relevant index is zero.

We must note that for cubic crystals the plane $(h\ k\ l)$ is perpendicular to the straight line $[h\ k\ l]$. This does not occur, generally speaking, for crystals of other systems.

---

* The directions passing through the crystal lattice points are usually of interest. This is why the indices form an integral proportion.

Figure 6.1 gives the Miller indices for the basic planes of a cubic crystal (the x-axis is directed toward us, the y-axis to the right, and the z-axis upward).

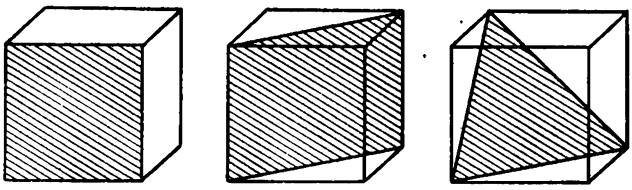


Fig. 6.1

6.2. Heat Capacity of Crystals. Einstein's Theory

According to classical notions, a crystal consisting of N atoms is a system with $3N$ vibrational degrees of freedom, to each of which there falls on an average the energy kT ($\frac{1}{2}kT$ in the form of kinetic and $\frac{1}{2}kT$ in the form of potential energy). These notions lead to the **Dulong and Petit law**, which states that the molar heat capacity of all chemically simple bodies in the crystalline state is the same and equals $3R$ (see Sec. 13.5 of Vol. I, p. 375). This law is obeyed sufficiently well only at comparatively high temperatures. At low temperatures, the heat capacity of crystals diminishes and tends to zero upon approaching 0 K.

The value of kT for the average energy of vibrational motion is obtained assuming that the energy of a harmonic oscillator can take on a continuous series of values. We established in Sec. 4.10 that vibrational energy is quantized. The result is that the average energy of vibration differs from kT. According to Eq. (4.60), the energy of a harmonic oscillator can have the values

$$\varepsilon_n = \left(n + \frac{1}{2}\right)\hbar\omega \qquad (n = 0, 1, 2, \ldots)$$

Assuming that the distribution of the oscillators by states with different energies obeys the Boltzmann law, we can find the average value of the energy of a harmonic oscillator $\langle\varepsilon\rangle$. Performing calculations similar to those that led us to formula (1.60), we shall get an expression for $\langle\varepsilon\rangle$ which differs from Eq. (1.60) only in having the additional addend $\frac{1}{2}\hbar\omega$. Thus,

$$\langle\varepsilon\rangle = \frac{1}{2}\hbar\omega + \frac{\hbar\omega}{\exp(\hbar\omega/kT) - 1} \qquad (6.1)$$

The theory of the heat capacity of crystalline bodies taking into account the quantization of the vibrational energy was created by Einstein (1907) and was later improved by Debye (1912).

Einstein considered that a crystal lattice consisting of N atoms is identical to a system of $3N$ independent harmonic oscillators with an identical natural frequency ω. The existence of the zero energy of oscillations was established much later, only after the appearance of quantum mechanics. Therefore, Einstein proceeded from Planck's value of the energy of a harmonic oscillator $\varepsilon_n = n\hbar\omega$. Accordingly, in the expression for $\langle\varepsilon\rangle$ used by Einstein, the addend $\frac{1}{2}\hbar\omega$ was absent.

Multiplying the second addend of Eq. (6.1) by $3N$, Einstein obtained the following formula for the internal energy of a crystal:

$$U = \frac{3N\hbar\omega}{\exp(\hbar\omega/kT) - 1} \tag{6.2}$$

Differentiating Eq. (6.2) with respect to the temperature, Einstein found the heat capacity of a crystal:

$$C = \frac{\partial U}{\partial T} = \frac{3N\hbar\omega}{[\exp(\hbar\omega/kT) - 1]^2} \exp(\hbar\omega/kT) \frac{\hbar\omega}{kT^2} \tag{6.3}$$

Let us consider two extreme cases.

1. High temperatures ($kT \gg \hbar\omega$). In this case, we may assume that $\exp(\hbar\omega/kT) \approx 1 + \hbar\omega/kT$ in the denominator and $\exp(\hbar\omega/kT) \approx \approx 1$ in the numerator of Eq. (6.3). As a result we get the following value of the heat capacity:

$$C = 3Nk$$

We have thus arrived at the Dulong and Petit law.

2. Low temperatures ($kT \ll \hbar\omega$). For this condition, we may disregard unity in the denominator of Eq. (6.3). The formula for the heat capacity thus becomes

$$C = \frac{3N(\hbar\omega)^2}{kT^2} \exp\left(-\frac{\hbar\omega}{kT}\right) \tag{6.4}$$

The exponential multiplier varies much more rapidly than T^2. Therefore when approaching absolute zero, Eq. (6.4) will tend to zero practically according to an exponential law. Experiments show that the heat capacity of crystals varies near absolute zero not exponentially, but according to the law T^3. Hence, Einstein's theory gives only a qualitatively correct course of the heat capacity at low temperatures. P. Debye succeeded in achieving quantitative agreement with experimental results.

6.3. Oscillations of Systems with a Large Number of Degrees of Freedom

To gain an understanding of Debye's theory, we must know the solution of the problem on small-amplitude oscillations of a system with a large number of degrees of freedom. In the present section, we shall consider the results of solving this problem without touching on the ways of solving it.

The position of a system with s degrees of freedom can be set with the aid of s quantities q_i called the **generalized coordinates** of the system. The part of the generalized coordinates can be played by lengths, angles, areas, etc. The generalized coordinates for the same system can be chosen in different ways. It can be shown that such a system has s natural frequencies ω_α (α is the number of the natural frequency running through the values 1, 2, ... , s). With an arbitrary choice of the generalized coordinates q_i, the general solution of the equations of motion has the form

$$q_i = \sum_{\alpha=1}^{s} A_{i\alpha} \cos(\omega_\alpha t + \delta_\alpha) \qquad (i = 1, 2, \ldots, s)$$

Hence, each of the functions q_i is, generally speaking, a superposition of s harmonic oscillations with the frequencies ω_α.

The energy of a system is determined by the expression

$$E = \frac{1}{2} \sum_{i,k=1}^{s} a_{ik} \dot{q}_i \dot{q}_k + \frac{1}{2} \sum_{l,m=1}^{s} b_{lm} q_l q_m$$

where the first sum gives the kinetic and the second the potential energy of the system; a_{ik} and b_{lm} are dimension coefficients. Thus, the expression for the energy includes, in general, not only the squares of the generalized coordinates q_i or the generalized velocities $\dot{q}_i$, but also the products of the coordinates or velocities corresponding to different degrees of freedom of the system.

It was found that we can choose the generalized coordinates of a system so that the change in each of them is a simple harmonic oscillation occurring with one of the natural frequencies ω_α. Denoting these coordinates by ξ_α, we can write that

$$\xi_\alpha = B_\alpha \cos(\omega_\alpha t + \delta_\alpha) \quad (\alpha = 1, 2, \ldots, s)$$

The generalized coordinates ξ_α perform a harmonic oscillation independently of one another, and each with its own frequency ω_α. The generalized coordinates chosen in this way are called **normal** (or **principal**), while the harmonic oscillations they perform are known as the **normal oscillations** of the system.

We must note that the changes in time of the arbitrarily chosen generalized coordinates q_i can be represented in the form of the superposition of the normal oscillations ξ_α

$$q_i = \sum_{\alpha=1}^{s} C_{i\alpha}\xi_\alpha \qquad (i = 1, 2, \ldots, s)$$

The expression for the energy in normal coordinates has the form

$$E = \frac{1}{2}\sum_{\alpha=1}^{s} a_\alpha \dot{\xi}_\alpha^2 + \frac{1}{2}\sum_{\alpha=1}^{s} b_\alpha \xi_\alpha^2 = \sum_{\alpha=1}^{s}\left(\frac{1}{2}a_\alpha\dot{\xi}_\alpha^2 + \frac{1}{2}b_\alpha\xi_\alpha^2\right)$$

Consequently, the energy of a system equals the sum of the energies falling to each of the normal oscillations separately.

Let us consider as an example a system consisting of two identical mathematical pendulums joined by a weightless spring (Fig. 6.2). Let us assume that the pendulums can oscillate only in the plane of the drawing, so that the system has two degrees of freedom. The position of the system can be set by the angles of deviation of both pendulums from the vertical position, or by the angle of deviation of one of the pendulums and the length of the spring, and so on. The solution of the equations of motion gives the following expressions for the angles φ_1 and φ_2 of deviation of the pendulums from their equilibrium position:

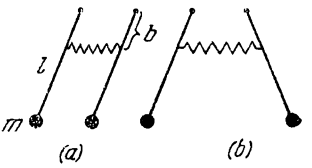


Fig. 6.2

$$\varphi_1 = A_1 \cos(\omega_1 t + \delta_1) + A_2 \cos(\omega_2 t + \delta_2)$$
$$\varphi_2 = A_1 \cos(\omega_1 t + \delta_1) - A_2 \cos(\omega_2 t + \delta_2)$$

Here A_1, A_2, δ_1, and δ_2 are constants determined from the initial conditions, and ω_1 and ω_2 are the natural frequencies of the system equal to

$$\omega_1 = \sqrt{\frac{g}{l}}, \qquad \omega_2 = \sqrt{\frac{g}{l} + 2\frac{kb^2}{ml^2}}$$

(m is the mass, l is the length of the pendulums, k is the spring constant, and b is the distance from the point of suspension to the point of fastening of the spring).

It is a simple matter to represent the oscillations φ_1 and φ_2 in the form

$$\varphi_1 = \xi_1 + \xi_2, \quad \varphi_2 = \xi_1 - \xi_2$$

where

$$\left.\begin{aligned} \xi_1 &= \frac{\varphi_1 + \varphi_2}{2} = A_1 \cos(\omega_1 t + \delta_1) \\ \xi_2 &= \frac{\varphi_1 - \varphi_2}{2} = A_2 \cos(\omega_2 t + \delta_2) \end{aligned}\right\} \qquad (6.5)$$

These two functions represent the normal oscillations of the given system. If the pendulums are deflected to the same side through the same angle $\varphi_{10} = \varphi_{20}$ and released without a push, then only the first normal oscillation will be completed ($A_1 \neq 0$, $A_2 = 0$), and $\varphi_1 = \varphi_2 = \xi_1$ (Fig. 6.2*a*). If we deflect the pendulums through the same angle in opposite directions ($\varphi_{10} = -\varphi_{20}$), then only the second normal oscillation will be completed ($A_1 = 0$, $A_2 \neq 0$), and $\varphi_1 = -\varphi_2 = \xi_2$ (Fig. 6.2*b*). In the first case, the pendulums oscillate with the frequency ω_1, in the second with the frequency ω_2 that is greater than ω_1. In other initial conditions, both normal oscillations will be performed simultaneously.

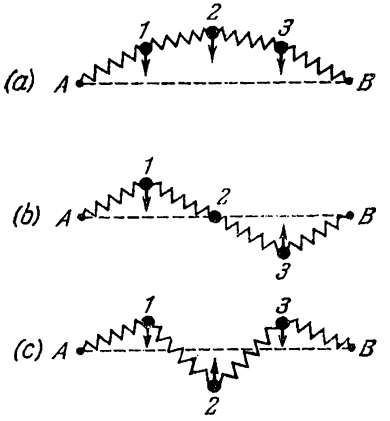


Fig. 6.3

Let us consider as a second example a system of three identical spheres joined by identical weightless springs (Fig. 6.3). Spring ends *A* and *B* are rigidly fixed. It is assumed that the spheres can move only in the plane of the drawing in directions at right angles to line *AB*. The system has three degrees of freedom in these conditions. The normal oscillations are shown in Fig. 6.3. In case *a*, all the spheres move in the same phase; in case *b*, spheres *1* and *3* oscillate in counterphase, and sphere *2* is stationary; in case *c*, spheres *1* and *3* oscillate in the same phase, and sphere *2* moves in counterphase with respect to them.

6.4. Debye's Theory

Debye took into account that the oscillations of the atoms in a crystal lattice are not independent. Displacement of one of the atoms from its equilibrium position results in the displacement of other atoms neighbouring with it. Thus, a crystal is a system of N elastically joined atoms with $s = 3N$ degrees of freedom.

In Sec. 14.8 of Vol. II, p. 293 et seq., we found out that the arbitrary oscillation of a string is the superposition of harmonic standing waves. Consequently, every normal oscillation of a string is a standing wave. Similarly, a standing wave that sets in within a crystalline body corresponds to every normal oscillation of the crystal lattice. Indeed, owing to the bond between the atoms, an oscillation appearing at one place in a crystal is transmitted from one atom to another, as a result of which an elastic wave is produced. Upon reaching the

boundary of the crystal, the wave is reflected. When the direct and reflected waves are superposed, a standing wave is formed. Standing waves can be produced only for frequencies (or wavelengths) satisfying definite conditions. If we take a crystalline body in the form of a parallelepiped with sides of a, b, and c, then these conditions are expressed by Eqs. (1.46).

The number of standing waves, i.e. normal oscillations, whose frequencies are confined within the interval from ω to $\omega + d\omega$ is determined by Eq. (1.49). The volume of the crystal V enters this expression in the form of a separate multiplier. We can therefore speak of the number of normal oscillations per unit volume of a crystal. In accordance with Eq. (1.50), this number is

$$dN_\omega = \frac{\omega^2 d\omega}{2\pi^2 v^3} \tag{6.6}$$

where v is the phase velocity of the wave in the crystal. We shall stress that now by dN_ω we understand the number of standing waves per unit volume; in Sec. 1.5 this number was designated by the symbol dn_ω. In connection with the fact that we shall need the letter n to denote the number of atoms in unit volume, however, it is expendient to write dN_ω instead of dn_ω.

Equation (6.6) takes no account of the possible kinds of polarization of a wave. Three different waves with the same value of ω can propagate in a solid medium along a certain direction. These waves differ in the direction of polarization: one is longitudinal and two are transverse with mutually perpendicular directions of oscillations. Accordingly, Eq. (6.6) must be modified as follows:

$$dN_\omega = \frac{\omega^2 d\omega}{2\pi^2}\left(\frac{1}{v_\parallel^3} + \frac{2}{v_\perp^3}\right)$$

Here $v_\parallel$ is the phase velocity of the longitudinal and $v_\perp$ of the transverse elastic waves. Let us assume for simplicity that $v_\parallel = v_\perp = v$. Hence,

$$dN_\omega = \frac{3\omega^2 d\omega}{2\pi^2 v^3} \tag{6.7}$$

We can find the maximum frequency ω_m of the normal oscillations of a lattice by equating the total number of oscillations to the number of degrees of freedom equal to $3n$ (n is the number of atoms in unit volume of the crystal; we remind our reader that the calculation is being performed for unit volume):

$$3n = \int dN_\omega = \int_0^{\omega_m} \frac{3\omega^2 d\omega}{2\pi^2 v^3} = \frac{\omega_m^3}{2\pi^2 v^3}$$

Hence

$$\omega_m = v\sqrt[3]{6\pi^2 n} \tag{6.8}$$

We shall note that in accordance with Eq. (6.8), the smallest length of a wave excited in a crystal is

$$\lambda_{min} = \frac{2\pi v}{\omega_m} \approx \frac{2}{\sqrt[3]{n}} \approx 2d$$

where d is the distance between neighbouring atoms in a lattice. This result agrees with the fact that waves whose length is less than the double interatomic distance have no physical meaning.

Deleting the velocity v from Eqs. (6.7) and (6.8), we obtain the following expression for the number of normal oscillations dN_ω in the frequency interval $d\omega$ per unit volume of a crystal:

$$dN_\omega = 9n \frac{\omega^2 d\omega}{\omega_m^3} \tag{6.9}$$

The internal energy of unit volume of a crystal can be represented in the form

$$U = \int \langle \varepsilon(\omega) \rangle \, dN_\omega$$

where $\langle \varepsilon(\omega) \rangle$ is the average value of the energy of normal oscillation at the frequency ω. Introducing Eq. (6.1) for $\langle \varepsilon(\omega) \rangle$ and (6.9) for dN_ω, we arrive at the formula

$$U = \frac{9n}{\omega_m^3} \int_0^{\omega_m} \left(\frac{1}{2} \hbar\omega + \frac{\hbar\omega}{\exp(\hbar\omega/kT) - 1} \right) \omega^2 d\omega =$$

$$= U_0 + \frac{9n\hbar}{\omega_m^3} \int_0^{\omega_m} \frac{\omega^3 d\omega}{\exp(\hbar\omega/kT) - 1} \tag{6.10}$$

Here $U_0 = 3n\left(\frac{3}{8}\hbar\omega_m\right)$ is the energy of zero oscillations of a crystal.

The derivative of U with respect to T gives the heat capacity of unit volume of a crystal

$$C = \frac{\partial U}{\partial T} = \frac{9n\hbar}{\omega_m^3} \int_0^{\omega_m} \frac{\exp(\hbar\omega/kT)\, \hbar\omega^4 d\omega}{[\exp(\hbar\omega/kT) - 1]^2 kT^2}$$

The quantity Θ determined by the condition $\hbar\omega_m = k\Theta$ is called the **Debye characteristic temperature.** By definition,

$$\Theta = \frac{\hbar\omega_m}{k}. \tag{6.11}$$

The Debye temperature indicates for each substance the region where the quantization of the energy of oscillations becomes appreciable.

Let us introduce the variable $x = \hbar\omega/kT$. Hence, the expression for the heat capacity becomes

$$C = 9nk \left(\frac{T}{\Theta}\right)^3 \int_0^{x_m} \frac{e^x x^4 dx}{(e^x - 1)^2} \tag{6.12}$$

where $x_m = \hbar\omega_m/kT = \Theta/T$. When $T \ll \Theta$, the upper limit of the integral will be very great, so that it can be assumed approximately equal to infinity ($x_m \approx \infty$). Therefore, the integral will be a certain number, and the heat capacity C will be proportional to the cube of the temperature: $C \propto T^3$. This approximation is known as the **Debye T^3 law.** At sufficiently low temperatures, this law is obeyed very well in many cases.

When $T \gg \Theta$, i.e. when $\hbar\omega_m/kT \ll 1$, Eq. (6.10) can be simplified by assuming that $\exp(\hbar\omega/kT) \approx 1 + \hbar\omega/kT$. Hence for the internal energy, we get the expression

$$U = U_0 + \frac{9n\hbar}{\omega_m^3} \int_0^{\omega_m} \frac{kT}{\hbar\omega} \omega^3 \, d\omega = \\ = U_0 + 3nkT$$

while for the heat capacity we get the value $C = 3nk$ figuring in the Dulong and Petit law.

That Debye's theory agrees with experimental data can be seen in Fig. 6.4 which gives data for the heat capacity of aluminium (Θ = 396 K) and copper (Θ = 309 K); C_∞ is the classical value of the heat capacity obtained from quantum formulas for $T \to \infty$. The curves have been plotted according to Eq. (6.12), the circles show the experimental points.

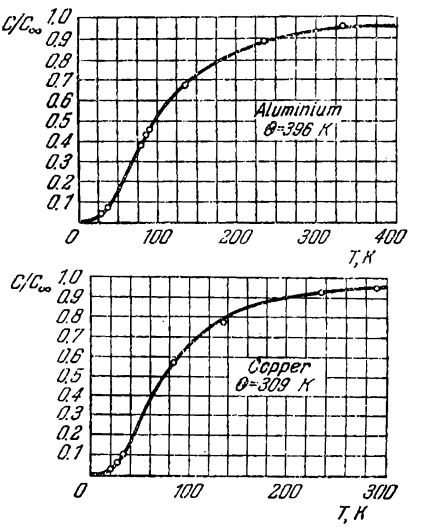


Fig. 6.4

Debye's formula gives the change in the heat capacity with the temperature quite well only for bodies with simple crystal lattices, i.e. for chemical elements and some simple compounds. The formula may not be applied for bodies with a more intricate structure. This is due to the spectrum of oscillations in such bodies being extremely complicated. For the simple crystal lattice considered above (in which a unit cell contains only one atom) three values of the natural frequency of oscillations of the lattice (one for the longitudinal wave and two values coinciding with each other* for the transverse waves) corresponded to each value of the wave vector k. If the number of atoms in a unit cell of a crystal is r, then in the general case $3r$ different values of ω correspond to each value of k. Consequently, the frequency is a multiple-valued function of the wave vector and has $3r$ branches. For example, for a unidimensional chain construct-

* In a greatly anisotropic crystal, all three frequencies will be different.

ed of alternating atoms of two species ($r = 2$), the dependence of ω on k has the form shown in Fig. 6.5. One of the branches is called the **acoustic**, and the other the **optical** one. These branches are distinguished by their dispersion, i.e. the nature of the dependence of ω on k. The acoustic branch tends to zero when k diminishes, while the optical branch has the terminal value of ω_{20} as its limit.

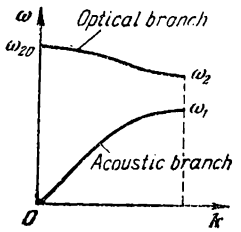


Fig. 6.5

In the three-dimensional case, of $3r$ branches, three are acoustic, and the remainder $(3r - 3)$ are optical. Sound frequencies correspond to the acoustic branches, and frequencies in the infrared region of the spectrum, to the optical ones. In normal oscillations of an acoustic frequency, similar atoms in different unit cells oscillate relative to one another. In normal oscillations of an optical frequency, different atoms within each of the unit cells oscillate relative to one another; similar atoms of different cells are at constant distances from one another.

6.5. Phonons

We established in the preceding section that the energy of a crystal can be represented as the sum of the energies of normal oscillations of a lattice:

$$U = \sum_{i=1}^{3Nr} \left(n_i + \frac{1}{2}\right) \hbar\omega_i$$

(N is the number of unit cells in a crystal, and r is the number of atoms in a cell).

With deduction of the energy of zero oscillations, the energy of a normal oscillation of frequency ω_i consists of portions of the magnitude

$$\varepsilon_i = \hbar\omega_i \tag{6.13}$$

This portion (quantum) of energy is called a **phonon.** Many processes in a crystal (for example, the scattering of X-rays or neutrons) proceed as if a phonon had the quasimomentum

$$\mathbf{p} = \hbar\mathbf{k} \tag{6.14}$$

where $\mathbf{k}$ is the wave vector of the corresponding normal oscillation.

A phonon in many respects behaves as if it were a particle having the energy determined by Eq. (6.13) and the momentum determined by Eq. (6.14). Unlike ordinary particles (electrons, protons, photons, etc.), however, a phonon cannot appear in a vacuum—it needs

a medium to appear and to exist. Particles like this are known as **quasiparticles**. Thus, a phonon is a quasiparticle. Accordingly, the quantity **p** [Eq. (6.14)] for a phonon is called the **quasimomentum**.

For conditions of thermal equilibrium, the average number of phonons $\langle n_i \rangle$ of frequency ω_i is determined by the expression

$$\left\langle \left(n_i + \frac{1}{2}\right) \hbar\omega_i \right\rangle = \frac{1}{2} \hbar\omega_i + \frac{\hbar\omega_i}{\exp(\hbar\omega_i/kT) - 1}$$

[see Eq. (6.1)]. Hence,

$$\langle n_i \rangle = \frac{1}{\exp(\hbar\omega_i/kT) - 1} \tag{6.15}$$

It can be seen from Eq. (6.15) that an unlimited number of identical phonons can be excited in a crystal simultaneously. Hence, the Pauli principle does not extend to phonons.

We must note that the quanta of an electromagnetic field—photons in a state of equilibrium with the walls of a cavity (see Chap. 1), also obey distribution (6.15).

Thus, the oscillations of a crystal lattice can be represented as a phonon gas confined within the limits of a crystal specimen like electromagnetic radiation can be represented as a photon gas filling a cavity. Formally, both notions are very similar—both photons and phonons obey the same statistics. But there is a significant distinction between photons and phonons: whereas photons are true particles, phonons are quasiparticles.

The combination (Raman) scattering of light by crystals (see Sec. 5.14) can be interpreted as a process of interaction of a photon with phonons. A photon flying through a crystal lattice may excite in it a phonon of one of the frequencies of the optical branch of the crystal. The photon uses part of its energy to do this. Consequently, its frequency decreases—a red satellite appears. If a phonon had already been excited in a crystal, a photon flying through the lattice may absorb it and increase its energy as a consequence—a violet satellite appears.

Distribution (6.15) is a particular case of the **Bose-Einstein distribution** that is obeyed by particles having an integral (in particular, a zero) spin. The general expression of this distribution has the form

$$\langle n_i \rangle = \frac{1}{\exp[(E_i - \mu)/kT] - 1} \tag{6.16}$$

where $\langle n_i \rangle$ = average number of particles in the state numbered i
E_i = energy of a particle in this state
μ = so-called **chemical potential** determined from the condition that the sum of all the $\langle n_i \rangle$'s equals the total number of particles N in the system: $\sum \langle n_i \rangle = N$.

The values of μ in distribution (6.16) cannot be positive because if it were, at $E_i < \mu$ the average number $\langle n_i \rangle$ would be negative, whichis deprived of a physical meaning. Thus, $\mu \leqslant 0$. For systems with a varying number of particles (among which are both a system of photons and a system of phonons), $\mu = 0$, and Eq. (6.16) transforms into Eq. (6.15).

Distribution (6.16) is the cornerstone of the **Bose-Einstein statistics.** Particles obeying this statistics are known as **bosons.** Thus, both photons and phonons are bosons. Bosons include all particles having a zero or integral spin.

Bosons are characterized by the fact that the probability P of the appearance ("birth") of a boson in a state in which there are already n particles is proportional to the square root of n:

$$P \propto \sqrt{n} \tag{6.17}$$

Thus, bosons "like" to accumulate in one state—they are "collectivists".

6.6. The Mössbauer Effect

Atoms absorb light of a frequency corresponding to the transition from the ground state to the nearest excited state especially intensively. This phenomenon is called **resonance absorption.** Returning later to the ground state, the atoms emit photons of the resonance frequency. The corresponding radiation is known as **resonance emission** or **resonance fluorescence.** The phenomenon of resonance fluorescence was discovered by the American physicist Robert Wood (1868-1955) in 1904. He found that sodium vapour when irradiated with light corresponding to the yellow line of sodium begins to glow, emitting radiation of the same wavelength. Later, similar glowing was observed in mercury vapour and in many other cases. Owing to resonance absorption, the light passing through the fluorescing substance weakens.

Like atoms, atomic nuclei have discrete energy levels. The lowest of them is called the ground (or normal) level, and the others are excited levels. Transitions between these levels lead to the production of short-wave electromagnetic radiation that has been called **gamma-rays** (see Sec. 10.5). The existence of the phenomenon of nuclear resonance fluorescence similar to the atomic resonance fluorescence observed in visible light could be expected for gamma-rays. For a long time, however, no investigators succeeded in observing resonance fluorescence with these rays. The explanation of this is as follows. It was shown in Sec. 5.3 that the emission line and the absorption line corresponding to a transition of a quantum system

between two states are shifted relative to each other by $2\Delta\omega_R = 2R/\hbar$, where R is the recoil energy determined by Eq. (5.23). For visible light, the shift $2\Delta\omega_R$ is many orders of magnitude smaller than the breadth of a spectral line $\delta\omega$ so that the emission and absorption lines are virtually superimposed. Matters are different with gamma-rays. The energy and momentum of a gamma-photon are many times greater than those of a photon of visible light. Therefore, the recoil energy R is considerably greater too, and in this case must be written as follows

$$R = \frac{(\hbar\omega)^2}{2m_{\text{nuc}}c^2} \tag{6.18}$$

where m_{nuc} is the mass of a nucleus.

It is customary practice in the spectroscopy of gamma-rays to use energies instead of frequencies. We shall therefore express the breadth of a spectral line, the shift of the lines, and the likein energy units, multiplying the relevant frequencies by Planck's constant $\hbar$ for this purpose. In these units, the natural breadth of a spectral line will be characterized by the quantity Γ [see Eq. (5.15)], the shift of the emission and absorption lines by the quantity $2R$, and the Doppler broadening of the line by the quantity

$$2D = 2\frac{v}{c}\Delta E_{nm} \approx 2\frac{v}{c}\hbar\omega \tag{6.19}$$

[see Eq. (5.27)].

The energy of gamma-quanta usually ranges from about 10 keV to about 5 MeV (which corresponds to frequencies ranging from 10^{19} to 10^{22} rad/s and wavelengths from about 1 Å to about 10^{-3} A). Let us calculate the recoil energy R for the case $\hbar\omega = 100$ keV and $m_{\text{nuc}} = 1.7 \times 10^{-22}$ g (an atomic mass of the order of 100). The value of $m_{\text{nuc}}c^2$ is $1.7 \times 10^{-22} \times 9 \times 10^{20} = 0.15$ erg, i.e. $0.15/1.6 \times \times 10^{-12} \approx 10^{11}$ eV. Consequently, in accordance with Eq. (6.18),

$$R = \frac{(10^5)^2}{2 \times 10^{11}} = 0.5 \times 10^{-1}\,\text{eV}$$

and the shift of the lines $2R$ is 10^{-1} eV.

The natural breadth of spectral lines Γ is determined by formula (6.14). The typical lifetime of the excited states of nuclei is 10^{-12} s. A value of

$$\Gamma = \frac{\hbar}{\tau} = \frac{1.05 \times 10^{-27}}{10^{-12}} = 1.05 \times 10^{-15}\,\text{erg} \approx 10^{-3}\,\text{eV}$$

corresponds to this lifetime.

For nuclei with a mass of $\sim 10^{-22}$ g, the average velocity of thermal motion at room temperature is about 300 m/s. At such a velocity,

the Doppler breadth of a line with $\hbar\omega = 100$ keV has the value

$$2D = 2 \times \frac{3 \times 10^2}{3 \times 10^8} \times 10^5 = 2 \times 10^{-1} \text{ eV}$$

[see Eq. (6.19)].

A comparison of the obtained values of Γ and $2D$ leads to the conclusion that the breadth of the spectral lines emitted by nuclei at room temperature is mainly determined by the Doppler breadth and is about 0.2 eV. For a shift of the emission and absorption lines of $2R$, we obtained a value of $\sim$0.1 eV. Thus, even for comparatively

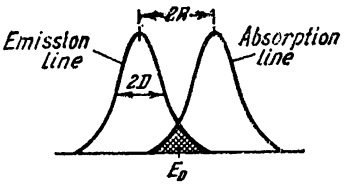


Fig. 6.6

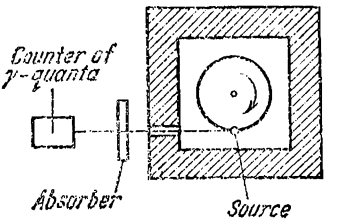


Fig. 6.7

soft gamma-rays with an energy of 100 keV, the shift of the emission and absorption lines is of the same order as the breadth of a spectral line. With an increase in the energy of a photon, R grows more rapidly [in proportion to ω^2, see expression (6.18)] than D [which is proportional to ω, see expression (6.19)]. Figure 6.6 contains a typical picture for gamma-photons showing the mutual arrangement of the emission and absorption lines. It is quite obvious that only a small part of the emitted photons (their relative number is determined by the corresponding ordinates of the emission line) can experience resonance absorption, the probability of their absorption being low (this probability is determined by the ordinates of the absorption line).

Prior to 1958, investigators succeeded in observing the resonance absorption of gamma-rays with the aid of devices in which a source of gamma-radiation travelled with the velocity v toward the absorbing substance. This was achieved by placing a radioactive substance on the rim of a rotating disk (Fig. 6.7). The disk was inside a massive lead shield absorbing gamma-rays. The irradiated beam emerged through a narrow channel and impinged on the absorbing substance. A counter of gamma-quanta installed after the absorber registered the intensity of radiation that had passed through the latter. Owing to the Doppler effect, the frequency of the gamma-rays emitted by the source increased by $\Delta\omega = \omega\,(v/c)$, where v is the velocity of the source relative to the absorber. By properly choosing the speed of

rotation of the disk, it was possible to observe resonance absorption. The latter was detected according to the reduction in the intensity of the gamma-rays measured by the counter.

In 1958, the German physicist Rudolf Mössbauer (born 1929) studied the nuclear resonance absorption of gamma-rays from Ir^{191} (the iridium isotope with a mass number of 191, see Sec. 10.1). The energy ΔE_{nm} of the relevant transition is 129 keV, the recoil energy is 0.05 eV, and the Doppler broadening at room temperature is about 0.1 eV. Thus, the emission and absorption lines partly overlap, and resonance absorption could be observed. To reduce the absorption, Mössbauer decided to cool the source and the absorber, thus expecting to reduce the Doppler breadth and, consequently, the overlapping of the lines. Instead of the expected reduction, however, he detected amplification of the resonance absorption.

Mössbauer devised an arrangement in which the source and the absorber were placed inside a vertical tube cooled by liquid helium. The source was fastened to the end of a long rod performing reciprocating motion. Working with this arrangement, Mössbauer observed that the resonance absorption vanished at linear velocities of the source of the order of several centimetres a second. The results of the experiment indicated that in cooled Ir^{191} the gamma-ray absorption and emission lines coincide and have a very small breadth equal to the natural breadth Γ. This phenomenon of the elastic (i.e. not attended by a change in the internal energy of a body) emission or absorption of gamma-quanta was called the **Mössbauer effect.**

The Mössbauer effect was soon discovered in Fe^{57} and for a number of other substances. The nucleus of Fe^{57} is remarkable in the respect that the effect is observed for it at temperatures up to 1000 °C, so that no cooling is needed. In addition, Fe^{57} is distinguished by the exceedingly small natural breadth of a line.

Let us now uncover the physical essence of the Mössbauer effect. When a nucleus at a crystal point emits a gamma-quantum, the transition energy ΔE_{nm} in principle may be distributed between the gamma-quantum, the nucleus that emitted the quantum, the solid as a whole, and, finally, between the oscillations of the lattice. In the latter case, phonons will be produced in addition to the gamma-quantum. Let us analyse these possibilities. The energy needed for a nucleus to leave its site in a lattice is at least about 10 eV, whereas the recoil energy R does not exceed several tenths of an electron-volt. Therefore, an atom whose nucleus has emitted a gamma-quantum cannot change its position in the lattice. The recoil energy which a solid body can receive as a whole is exceedingly small, so that it may be disregarded [this energy can be assessed by substituting the mass of a body for the mass of a nucleus in Eq. (6.18)]. Thus, the transition energy can be distributed only between the gamma-quantum and phonons. A Mössbauer transition occurs if the vibra-

tional state of the lattice does not change, and the gamma-quantum receives the entire energy of transition.

Thus, when a nucleus at a crystal lattice point emits or absorbs a gamma-quantum, two processes may occur: (1) a change in the vibrational state of the lattice, i.e. phonon excitation, and (2) the transition of the momentum of the gamma-quantum to the lattice as a whole without a change in its vibrational state, i.e. the elastic emission and absorption of a gamma-quantum. Each of these processes has a definite probability whose value depends on the particular properties of the crystal, the energy of the gamma-quantum, and the temperature. The relative probability of the elastic processes grows with lowering of the temperature.

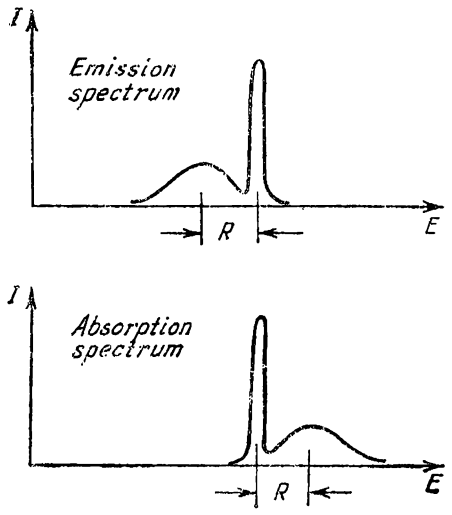


Fig. 6.8

It is easy to show that in inelastic processes, phonons with an energy of the order of $\hbar\omega_m = k\Theta$ should be mainly excited (ω_m is the maximum frequency of oscillations of the lattice, and Θ is the Debye temperature; see Sec. 6.4). The wavelength $\lambda_{min} \approx 2d$ corresponds to oscillation of the frequency ω_m [see the paragraph following Eq. (6.8)]. In this case, neighbouring atoms move in counterphase, which can occur when the atom emitting a gamma-quantum receives the entire recoil energy R and then collides with the neighbouring atom. To produce longer wavelengths (lower frequencies), several atoms must be brought into motion simultaneously, which has a low probability. Thus, the probability of producing oscillations of the lattice will be great provided that the recoil energy R received in radioactive decay by an individual atom is equal to or greater than the energy of a phonon of the maximum frequency: $R \geqslant \hbar\omega_m = k\Theta$.

For Ir^{191}, the recoil energy R is of the order of $k\Theta$. Consequently, to obtain a measurable resonance absorption, it is necessary to reduce the probability of exciting oscillations of the lattice. For Fe^{57}, the recoil energy $R \ll k\Theta$. Owing to this circumstance, already at room temperature, an appreciable fraction of the nuclear transitions occurs elastically.

Figure 6.8 shows typical emission and absorption spectra of gamma-quanta (E is the energy of a gamma-quantum, I is the intensity,

and R is the average recoil energy). The two spectra contain practically coinciding very narrow lines corresponding to elastic processes. These lines are on the background of broad shifted ones due to processes attended by a change in the vibrational state of the lattice. The background becomes weaker with lowering of the temperature, and the fraction of the elastic processes grows, but it never reaches unity.

The Mössbauer effect found numerous applications. In nuclear physics, it is used to find the lifetime of excited states of nuclei (through Γ), and also to determine the spin, magnetic moment, and electric quadrupole moment of nuclei. In solid state physics, the Mössbauer effect is used to study the dynamics of a crystal lattice and to study the intrinsic electric and magnetic fields in crystals.

Owing to the extremely small breadth of the Mössbauer lines, the method of a moving source makes it possible to measure the energy of gamma-quanta with an enormous relative accuracy (up to the 15-th significant digit). The U.S. physicists R. Pound and G. Rebka, Jr. took advantage of this circumstance to detect the **gravitational red shift** of the frequency of photons predicted by the general theory of relativity*. It follows from the general theory of relativity that the frequency of a photon should change with a change in the gravitational potential. According to the equivalence principle (see Sec. 6.3 of Vol. I, p. 181), a photon has a gravitational mass equal to its inert mass $m_r = \hbar\omega/c^2$ [see Eq. (8.54) of Vol. I, p. 254]. When a photon travels the path l in a direction opposite to that of the force $m_r g$ in a homogeneous gravitational field characterized by the strength g, the energy of the photon must diminish by $m_r gl = \hbar\omega gl/c^2$. Consequently, the energy of the photon will become equal to

$$\hbar\omega' = \hbar\omega - \frac{\hbar\omega gl}{c^2} = \hbar\omega\left(1 - \frac{gl}{c^2}\right)$$

Hence,

$$\frac{\Delta\omega}{\omega} = \frac{\omega - \omega'}{\omega} = \frac{gl}{c^2} = \frac{\Delta\varphi}{c^2}$$

where $\Delta\varphi$ is the change in the gravitational potential. The formula we have obtained also holds for a photon travelling in an inhomogeneous gravitational field (in this case $\Delta\varphi = \int g_l dl$).

The light reaching the Earth from stars overcomes their strong attracting field. Near the Earth, on the other hand, it experiences the action of only a very weak accelerating field. Consequently, all the spectral lines of stars must be slightly shifted toward the red end of the spectrum. Such a shift, called the gravitational red shift, was confirmed qualitatively by astronomical observations.

---

* We mentioned these experiments in Sec. 8.10 of Vol. I, p. 248.

Pound and Rebka undertook an attempt to detect this phenomenon in the conditions of the Earth. They put a source of gamma-radiation (Fe^{57}) and an absorber in a high tower at a distance of 21 metres apart (Fig. 6.9). The relative change in the energy of a gamma-photon when it covers this distance is only

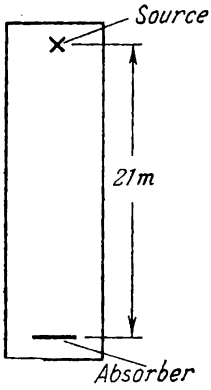


Fig. 6.9

$$\frac{\Delta\varepsilon}{\varepsilon} = \frac{\Delta\omega}{\omega} = \frac{gl}{c^2} = \frac{9.81 \times 21}{9 \times 10^{16}} \approx 2 \times 10^{-15}$$

This change* gives rise to a relative shift in the absorption and emission lines and should manifest itself in a slight weakening of resonance absorption. Notwithstanding the extreme smallness of the effect (the shift was about 10^{-2} of the breadth of a line), Pound and Rebka succeeded in detecting and measuring it with sufficient accuracy. The result they obtained was 0.99 ± 0.05 of what was predicted by theory. Thus, they gave a convincing proof of the existence of a gravitational shift in the frequency of photons in the conditions of a laboratory on the Earth.

* If the source is placed on top and the receiver below it, the energy of a photon grows, so that a violet frequency shift occurs.

CHAPTER 7 THE BAND THEORY OF SOLIDS

7.1. The Quantum Theory of Free Electrons in a Metal

In Sec. 11.2 of Vol. II, p. 230 et seq., we set out the elementary classical theory of free electrons in a metal. Now let us acquaint ourselves with the fundamentals of the quantum theory.

According to the free-electron model, the valence electrons of the atoms of a metal can almost freely travel within the confines of a specimen. It is exactly the valence electrons that give rise to the electrical conduction of metals, and this is why they are called conduction electrons.

Let us consider a specimen of a metal which for simplicity we shall consider to have the shape of a cube with the side L. Assume that the conduction electrons travel absolutely freely within the confines of the specimen. Assuming in Eq. (4.12) that $U = 0$, we get the Schrödinger equation for a free electron

$$-\frac{\hbar^2}{2m}\nabla^2\psi = E\psi \tag{7.1}$$

(m is the mass of an electron).

It is a simple matter to verify by substitution that the solution of Eq. (7.1) has the form

$$\psi = Ce^{i\,\mathbf{kr}} \tag{7.2}$$

where $\mathbf{k} = \mathbf{p}/\hbar$ is the wave vector of an electron associated with the energy by the relation

$$E = \frac{p^2}{2m} = \frac{\hbar^2 k^2}{2m} \tag{7.3}$$

The condition of normalization of the psi-function will be written as follows (integration is performed over the volume V of the specimen equal to L^3):

$$\int \psi^*\psi\, dV = C^*C \int dV = C^*CL^3 = 1$$

Assuming C to be real, we get the value $1/L^{3/2}$ for it. Substitution in Eq. (7.2) yields

$$\psi = \frac{1}{L^{3/2}} e^{i\,\mathbf{kr}} \tag{7.4}$$

The psi-function must satisfy the boundary conditions consisting in the requirement that it be periodic with respect to x, y, z with the period L. We can see that function (7.4) will satisfy these conditions at values of the wave vector components equal to

$$k_x = \frac{2\pi}{L} n_1, \quad k_y = \frac{2\pi}{L} n_2, \quad k_z = \frac{2\pi}{L} n_3 \tag{7.5}$$

where n_1, n_2, and n_3 are integers taking on the values 0, ± 1, ± 2, etc. independently of one another. Indeed, introduction of the values (7.5) into Eq. (7.4) yields

$$\psi = \frac{1}{L^{3/2}} \exp\left[i \frac{2\pi}{L} (n_1 x + n_2 y + n_3 z)\right]$$

Substitution of $x + L$ for x and $y + L$ for y, etc. leaves the function unchanged (only a multiplier equal to 1 appears).

Thus, the values of the wave vector are quantized. Accordingly, the energy of a conduction electron in a metal is quantized too. Introduction of the values (7.5) into Eq. (7.3) leads to the following expression for the energy:

$$E_{\mathbf{k}} = \frac{\hbar^2}{2m} \left(\frac{2\pi}{L}\right)^2 (n_1^2 + n_2^2 + n_3^2) \tag{7.6}$$

The state of a conduction electron is determined by the value of the wave vector $\mathbf{k}$ (i.e. by the values of k_x, k_y, k_z) and by the spin quantum number $m_s = \pm\frac{1}{2}$. Hence, the state can be set by the four quantum numbers n_1, n_2, n_3, m_s. The energy of an electron is determined by the sum of the squares of the quantum numbers n_i.

Table 7.1

n_1	n_2	n_3	m_s
1	0	0	$\pm\frac{1}{2}$
0	1	0	$\pm\frac{1}{2}$
0	0	1	$\pm\frac{1}{2}$
−1	0	0	$\pm\frac{1}{2}$
0	−1	0	$\pm\frac{1}{2}$
0	0	−1	$\pm\frac{1}{2}$

Several different combinations of the numbers n_i (except for the case $n_1 = n_2 = n_3 = 0$) correspond to the same sum of the squares. Consequently, the energy levels are degenerate. The level E_0 ($n_1 = n_2 = n_3 = 0$) has a degree of degeneracy equal to two $\left(m_s = \pm\frac{1}{2}\right)$. The next level E_1 is realized at 12 different combinations of the quantum numbers (see Table 7.1), the level E_2—at 24 combinations, etc. Thus, a growth in the energy is attended by an increase in the number of different states corresponding to a given value of E.

Let us introduce an imaginary space along whose axes we shall lay off the values of the quantum numbers n_1, n_2, n_3. In this space, a point corresponds to each pair of states (differing in the values of m_s). A surface of equal energy values has the shape of a sphere of radius $n^* = \sqrt{n_1^2 + n_2^2 + n_3^2}$. The number of states ν_E whose energy does not exceed the value $E = (\hbar^2/2m)\,(2\pi/L)^2 n^{*2}$ [see Eq. (7.6)] equals the double number of points contained within a sphere of radius n^*. Since the points are arranged with a density equal to unity, ν_E is determined by the double volume of the sphere:

$$\nu_E = 2 \times \frac{4}{3}\pi n^{*3} = \frac{8}{3}\pi\,(n_1^2 + n_2^2 + n_3^2)^{3/2} \tag{7.7}$$

Deleting the sum of the squares of the numbers n_i from Eqs. (7.6) and (7.7), we get

$$\nu_E = \frac{8}{3}\pi\left(\frac{2m}{\hbar^2}\right)^{3/2}\left(\frac{L}{2\pi}\right)^3 E^{3/2} = \frac{8}{3}\pi V\frac{(2m)^{3/2}}{(2\pi\hbar)^3}E^{3/2} \tag{7.8}$$

(V is the volume of the metal specimen). The formula we have obtained determines the number of states whose energy does not exceed the value E.

It follows from Eq. (7.8) that

$$d\nu_E = 4\pi V\frac{(2m)^{3/2}}{(2\pi\hbar)^3}E^{1/2}\,dE$$

Here $d\nu_E$ is the number of states with an energy within the interval from E to $E + dE$. Consequently, the density of the states $g(E) = d\nu/dE$, i.e. the number of states per unit interval of energy, is

$$g(E) = 4\pi V\frac{(2m)^{3/2}}{(2\pi\hbar)^3}E^{1/2} \tag{7.9}$$

Let the number of free electrons in unit volume of the metal be n. Hence, the metal specimen will contain nV free electrons. Owing to the Pauli principle, at absolute zero there will be one of these electrons in each state at the lowest energy levels. Therefore, all the states with an energy E less than a certain value $E_F(0)$ will be filled with electrons, whereas the states with $E > E_F(0)$ will be vacant. The energy $E_F(0)$ is known as the **Fermi level at absolute zero**. We shall show in the following section that the Fermi level plays the part of the parameter E_F in the distribution of the elec-

trons by states with different energies. This parameter depends slightly on the temperature. The quantity E_F (0) is the value of the parameter E_F at $T = 0$ K.

An isoenergetic surface* in k-space (or, which is the same, in p-space; $p = \hbar k$) corresponding to the value of the energy equal to E_F is called a **Fermi surface**. For free electrons, this surface is described by the equation

$$\frac{p^2}{2m} = \frac{\hbar^2 k^2}{2m} = E_F$$

[see Eq. (7.3)] and, consequently, has the form of a sphere. At absolute zero of temperatures, the Fermi surface separates states filled with electrons from the unfilled states.

The value of E_F (0) can be found by assuming in Eq. (7.8) that $\nu_E = nV$:

$$nV = \frac{8}{3}\pi V \frac{(2m)^{3/2}}{(2\pi\hbar)^3} [E_F(0)]^{3/2}$$

whence

$$E_F(0) = \frac{\hbar^2}{2m}(3\pi^2 n)^{2/3} \tag{7.10}$$

Let us assess the value of E_F (0). The concentration of conduction electrons in metals ranges from 10^{22} to 10^{23} cm^{-3}. Taking the average value of 5×10^{22} cm^{-3} for n, we get

$$E_F(0) = \frac{(1.05 \times 10^{-27})^2}{2 \times 0.91 \times 10^{-27}} (3 \times 3.14^2 \times 5 \times 10^{22})^{2/3} = 8 \times 10^{-12} \text{ erg} = 5 \text{ eV}$$

Let us find the average energy of the electrons at absolute zero. The total energy of the electrons filling states with energies from E to $E + dE$ is determined by the expression

$$E\, d\nu_E = Eg(E)\, dE$$

The total energy of all the conduction electrons is

$$\int E\, d\nu_E = \int_0^{E_F(0)} Eg(E)\, dE$$

Dividing this energy by the total number of electrons equal to $\int g(E)\, dE$, we get the average energy of one electron:

$$\langle E \rangle = \frac{\int_0^{E_F(0)} Eg(E)\, dE}{\int_0^{E_F(0)} g(E)\, dE}$$

* I.e., a surface of constant energy

Introduction of Eq. (7.9) for $g(E)$ yields

$$\langle E \rangle = \frac{\int_0^{E_F(0)} E^{3/2}\,dE}{\int_0^{E_F(0)} E^{1/2}\,dE} = \frac{3}{5} E_F(0) \qquad (7.11)$$

We obtained a value of the order of 5 eV for $E_F(0)$. Consequently, the average energy of the conduction electrons at absolute zero is about 3 eV. This is a tremendous value. To impart such an energy to a classical electron gas, it must be heated to a temperature of about 25 000 kelvins.

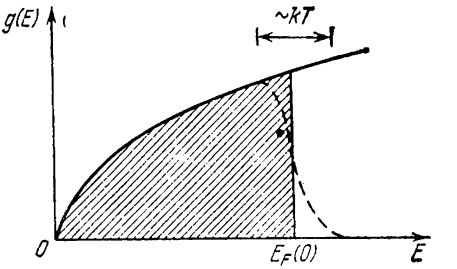


Fig. 7.1

Now we can explain why an electron gas contributes very little to the heat capacity of metals. The average thermal energy equal in its order of magnitude to kT is 0.025 eV at room temperature. Such an energy can excite only electrons at the highest levels adjoining the Fermi level. The main body of electrons at the deeper levels will retain their previous states and absorb no energy in heating. Thus, only an insignificant portion of the conduction electrons participate in the process of heating of a metal, and this is exactly what explains the low heat capacity of the electron gas in metals.

Figure 7.1 shows a graph of function (7.9). The hatched area gives the number of states filled with electrons at absolute zero. Heating of a metal is attended by the transition of electrons from levels adjoining the Fermi level to ones above $E_F(0)$. As a result, the sharp edge of the hatched figure will be blurred. The curve of filling of the levels by electrons will acquire the form in this region shown by the dash line. The area under this curve remains the same as it was at absolute zero (the area equals nV). The blurred region has a width of the order of kT. Hence, the fraction of the electrons participating in the process of heating of the metal is approximately T/T_F, where

$$T_F = \frac{E_F(0)}{k} \qquad (7.12)$$

is a quantity called the **Fermi temperature.** As a result, the heat capacity of the electrons will be

$$C_{el} = C_{class} \frac{T}{T_F}$$

At room temperature, C_{el} is about one-hundredth of the classical value ($T \approx 300$ K, $T_F \approx 25\,000$ K).

7.2. The Fermi-Dirac Distribution

At absolute zero, there is one electron in each of the states whose energy does not exceed E_F (0); there are no electrons in states with $E > E_F$ (0). Consequently, the function of the distribution of electrons by states with different energies has the form shown in Fig. 7.2 for absolute zero. Let us find the distribution function for a temperature other than absolute zero.

Following C. Kittel*, let us consider inelastic collisions of an equilibrium electron gas with an impurity atom implanted into the crystal lattice of a metal. Assume that the impurity atom can only be in two states whose energy we shall take equal to 0 and ε.

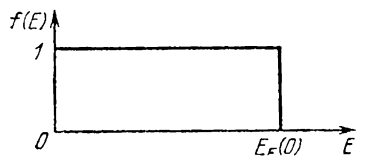


Fig. 7.2

Of the multitude of collision processes, we shall consider the one as a result of which an electron passes from state $\mathbf{k}$ with the energy E to state $\mathbf{k}'$ with the energy $E + \varepsilon$. The impurity atom here passes from the level with the energy ε to the level with an energy equal to zero. The probability $P_{\mathbf{kk}'}$ of the transition $\mathbf{k}\,(E) \to \mathbf{k}'\,(E + \varepsilon)$ is proportional to: (1) the probability $f\,(E)$ of the state $\mathbf{k}\,(E)$ being occupied by an electron, (2) the probability $[1 - f\,(E + \varepsilon)]$ of the state $\mathbf{k}'\,(E + \varepsilon)$ being free, and (3) the probability $p\,(\varepsilon)$ of the impurity atom being in the state with the energy ε. Thus,

$$P_{\mathbf{kk}'} \propto f\,(E)\,[1 - f\,(E + \varepsilon)]\,p\,(\varepsilon) \tag{7.13}$$

The probability $P_{\mathbf{k}'\mathbf{k}}$ of the reverse process is proportional to the expression

$$P_{\mathbf{k}'\mathbf{k}} \propto f\,(E + \varepsilon)\,[1 - f\,(E)]\,p\,(0) \tag{7.14}$$

where $p\,(0)$ is the probability of the impurity atom being in the state with the energy equal to zero.

The coefficient of proportionality in expressions (7.13) and (7.14) is the same owing to the detailed balancing principle**.

In the equilibrium state, the probabilities of the transitions $\mathbf{k} \to \mathbf{k}'$ and $\mathbf{k}' \to \mathbf{k}$ must be the same. Hence,

$$f\,(E)\,[1 - f\,(E + \varepsilon)]\,p\,(\varepsilon) = f\,(E + \varepsilon)\,[1 - f\,(E)]\,p\,(0)$$

Thus,

$$\frac{f\,(E+\varepsilon)}{1-f\,(E+\varepsilon)} \cdot \frac{1-f\,(E)}{f\,(E)} = \frac{p\,(\varepsilon)}{p\,(0)} = \exp\left(-\frac{\varepsilon}{kT}\right) \tag{7.15}$$

* See C. Kittel. *Elementary Solid State Physics.* New York, Wiley (1962).

** The principle of detailed balancing is the name given to the statement that in a state of statistical equilibrium the number of transitions of a system from state *1* to state *2* equals the number of reverse transitions from state *2* to state *1*.

(we have taken into account that the probabilities of the impurity atom being at the levels 0 and ε obey the Boltzmann distribution law).

Functional equation (7.15) must be obeyed at any temperature T. This will occur if we assume that

$$\frac{1-f(E)}{f(E)}=\exp\left(\frac{E-\mu}{kT}\right) \tag{7.16}$$

where μ is a quantity not depending on E. Accordingly,

$$\frac{f(E+\varepsilon)}{1-f(E+\varepsilon)}=\exp\left[-\frac{(E+\varepsilon)-\mu}{kT}\right] \tag{7.17}$$

The product of expressions (7.16) and (7.17) for any temperature is $\exp(-\varepsilon/kT)$.

Having solved Eq. (7.16) with respect to $f(E)$, we shall obtain the following expression for the function of the distribution of the electrons by states with different energies:

$$f(E)=\frac{1}{\exp[(E-\mu)/kT]+1} \tag{7.18}$$

This expression is called the **Fermi-Dirac distribution function.** The parameter μ is known as the **chemical potential.**

In accordance with the meaning of function (7.18), the quantity $f(E_i)$ is the average number $\langle n_i\rangle$ of electrons in the state with the energy E_i. Therefore, Eq. (7.18) can be written in the form

$$\langle n_i\rangle=\frac{1}{\exp[(E_i-\mu)/kT]+1} \tag{7.19}$$

[compare with Eq. (6.16)]. Unlike Eq. (6.16), the parameter μ in distribution (7.19) has positive values (in the given case this does not lead to negative values of the numbers $\langle n_i\rangle$).

Distribution (7.19) underlies the **Fermi-Dirac statistics.** Particles obeying this statistics are called **fermions.** They include all particles with a half-integral spin.

A characteristic of fermions is that they never occupy states in which there is a particle already. Thus, fermions are "individualists". We remind our reader that bosons, on the contrary, are "collectivists" (see the end of Sec. 6.5).

The parameter μ having the dimension of energy is frequently designated by the symbol E_F and is called the **Fermi level** or the **Fermi energy.** When this symbol is used, function (7.18) has the form

$$f(E)=\frac{1}{\exp[(E-E_F)/kT]+1} \tag{7.20}$$

Let us study the properties of function (7.20). For absolute zero, we have

$$f(E) = 1 \text{ if } E < E_F$$

and

$$f(E) = 0 \text{ if } E > E_F$$

Thus, at 0 K, the Fermi level E_F coincides with the upper level $E_F(0)$ filled by electrons (see the preceding section).

Regardless of the value of the temperature, when $E = E_F$, the function $f(E)$ equals 1/2. Consequently, the Fermi level coincides with the energy level whose probability of being filled is one-half.

The value of E_F can be found from the condition that the total number of electrons filling the levels must equal the number nV of free electrons in a crystal (n is the density of the electrons and V is the volume of the crystal). The number of states falling within the interval of energies dE is $g(E)\,dE$, where $g(E)$ is the density of the states. The average number of electrons in these states with thermal equilibrium prevailing is determined by the expression $f(E)\,g(E)\,dE$. The integral of this expression gives the total number of free electrons in a crystal

$$\int_0^\infty f(E)\,g(E)\,dE = nV \tag{7.21}$$

This expression is in essence the condition of normalization of the function $f(E)$.

Introduction of Eqs. (7.9) and (7.20) into Eq. (7.21) yields

$$4\pi V \frac{(2m)^{3/2}}{(2\pi\hbar)^3} \int_0^\infty \frac{E^{1/2}\,dE}{\exp[(E-E_F)/kT]+1} = nV \tag{7.22}$$

This relation makes it possible in principle to find E_F as a function of T and n. The integral in Eq. (7.22) cannot be taken. Provided that $kT \ll E_F$, an approximate value of the integral can be found. As a result, the following expression for the Fermi level is obtained:

$$E_F \approx E_F(0)\left[1 - \frac{\pi^2}{12}\left(\frac{kT}{E_F(0)}\right)^2\right] \tag{7.23}$$

[we remind our reader that $E_F(0)$ depends on n; see Eq. (7.10)].

It follows from expression (7.23) that at low temperatures (and these are the only ones for which this expression is true), the temperature dependence of the Fermi level, although it does exist, is very slight. We can therefore often assume that $E_F = E_F(0)$. To understand, for example, thermoelectric phenomena (see Sec. 9.4), however, the dependence of E_F on T is of a fundamental significance.

For temperatures other than absolute zero, the graph of function (7.20) has the form shown in Fig. 7.3. When the energies are high (i.e. when $E - E_F \gg kT$, which is observed in the region of the "tail" of the distribution curve), we may disregard unity in the denominator of the function. Hence, the distribution of the electrons by states with different energies acquires the form

$$f(E) = \exp\left(-\frac{E-E_F}{kT}\right) = \text{const} \cdot \exp\left(-\frac{E}{kT}\right) \qquad (7.24)$$

i.e. transforms into the Boltzmann distribution function.

We must note that an appreciable difference of the curve in Fig. 7.3 from the graph depicted in Fig. 7.2 is observed only in the region of the order of kT. The higher the temperature, the more gentle is the slope of the descending portion of the curve.

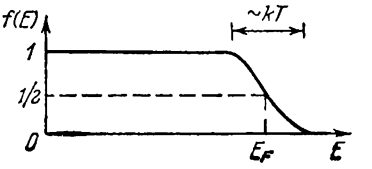


Fig. 7.3

The behaviour of an electron gas depends to a great extent on the relation between the temperature of the crystal and the Fermi temperature equal to E_F/k. Two extreme cases are distinguished:

1. $kT \ll E_F$. In this case, the electron gas is called **degenerate.**
2. $kT \gg E_F$. In this case, the electron gas is called **non-degenerate.**

We established in the preceding section that the Fermi temperature for metals is several tens of thousands of kelvins. Therefore even at a temperature close to the melting point of a metal (about 10^3 K), the electron gas in it is degenerate. In semiconductors, the density of the free electrons is much smaller than in metals. Accordingly, E_F is small [E_F is approximately proportional to $n^{2/3}$; see expressions (7.23) and (7.10)]. Hence already at room temperature, the electron gas in many semiconductors is non-degenerate and obeys classical statistics.

7.3. Energy Bands in Crystals

We established in Sec. 7.1 that in the approximation of free electrons, the energy of the valence electrons in a crystal changes **quasicontinuously.** This signifies that the spectrum of the allowed values of the energy consists of a multitude of closely arranged discrete levels. Actually, the valence electrons in a crystal do not have entirely free motion—the periodic field of the lattice acts on them. The result of this circumstance is that the spectrum of possible values of the energy of the valence electrons breaks up into a number

of allowed and forbidden bands (Fig. 7.4). The energy changes quasi-continuously within the limits of the allowed bands. The energy values belonging to the forbidden bands cannot be realized.

To understand the origin of the bands, let us consider an imaginary process of the combination of atoms into a crystal. Suppose we originally have N isolated atoms of a substance. As long as the atoms are isolated from one another, they have completely coinciding schemes of their energy levels. Electrons fill the levels in each atom independently of how similar levels are filled in the other atoms. As

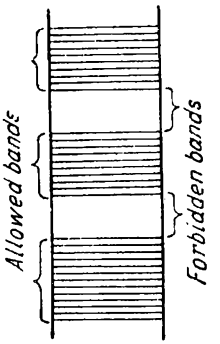


Fig. 7.4

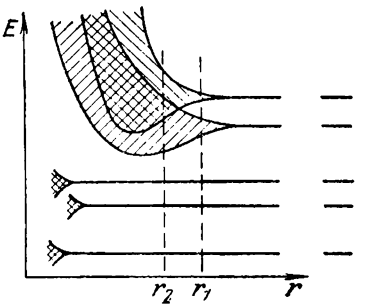


Fig. 7.5

the atoms approach one another, a constantly increasing interaction appears between them that results in a change in the position of the levels. Instead of a single level identical for all the N atoms, there appear N very close, but not coinciding, levels. Thus, each level of an isolated atom breaks up in a crystal into N densely arranged levels forming a **band**.

The amount of splitting is not the same for different levels. The levels filled by the outer or valence electrons in an atom are disturbed to a greater extent. The levels filled by the inner electrons are disturbed only slightly. Figure 7.5 shows the splitting of the levels as a function of the distance r between the atoms. Examination of the diagram reveals that the splitting of the levels occupied by the inner electrons is very small in a crystal. Only the levels occupied by the valence electrons split noticeably. The higher levels not occupied by electrons in the ground state of an atom are also subjected to similar splitting.

Depending on the particular properties of the atoms, the equilibrium state between neighbouring atoms in a crystal may be either of type r_1 or of type r_2 (see Fig. 7.5). With a distance of type r_1 between the allowed bands set up from adjacent levels of an atom, there is a forbidden band. With a distance of type r_2, the adjacent bands overlap. The number of levels in such a merged band equals

the sum of the numbers of levels into which both levels of the atom split up.

The band structure of the energy levels is obtained directly from the solution of the Schrödinger equation for an electron moving in a periodic force field. The latter is produced by the crystal lattice. The Schrödinger equation taking the lattice field into consideration has the form

$$-\frac{\hbar^2}{2m}\nabla^2\psi + U\psi = E\psi$$

where U is a function having the properties

$$U(x + a, y, z) = U(x, y, z)$$
$$U(x, y + b, z) = U(x, y, z)$$
$$U(x, y, z + c) = U(x, y, z)$$

(a, b, c are the lattice constants along the axes x, y, z).

The American physicist Felix Bloch (born 1905) proved that the solution of the Schrödinger equation with a periodic potential has the form

$$\psi_{\mathbf{k}} = u_{\mathbf{k}}(\mathbf{r})\, e^{i\mathbf{k}\mathbf{r}} \tag{7.25}$$

where $u_{\mathbf{k}}(\mathbf{r})$ is a function having the periodicity of the potential, i.e. the periodicity of the lattice. The solutions (7.25) are called **Bloch functions**. They differ from Eq. (7.2) in the presence of the periodic multiplier $u_{\mathbf{k}}(\mathbf{r})$.

In the approximation of free electrons, the dependence of the energy of an electron on the wave number (the magnitude of the wave vector) is described by the graph depicted in Fig. 7.6 [see Eq. (7.3)]. The values of the energy form a quasicontinuous sequence. Consequently, the graph $E(k)$ consists of discrete points. These points are so dense, however, that they visually merge into a continuous curve.

When the field is periodic, the dependence of E on k has the form shown in Fig. 7.7. A glance at the figure shows that the bands of quasicontinuously changing energy (allowed bands) depicted by solid lines alternate with the forbidden bands. Each allowed band consists of closely arranged discrete levels whose number equals the number of atoms in the crystals specimen.

The region of k-space in which the energy of an electron in a crystal changes quasicontinuously is called a **Brillouin zone**. There is an interruption in the energy at the boundaries of the zones. Figure 7.7 depicts the Brillouin zones for a one-dimensional crystal. For three-dimensional crystals, the boundaries of the Brillouin zones are closed polyhedral surfaces contained one within another.

We remind our reader that a Fermi surface is defined as an isoenergetic surface in k-space (or in p-space) corresponding to a value of

E equal to E_F (see Sec. 7.1). For free electrons, this surface is a sphere. The shape of the surface for the conduction electrons of a metal depends on the properties of the crystal lattice and is intricate, sometimes being quite odd. The shape of the Fermi surface for a number of metals has been established experimentally with a high accuracy

The Fermi surface is an important characteristic of a metal. The shape of this surface determines the nature of motion of the electrons

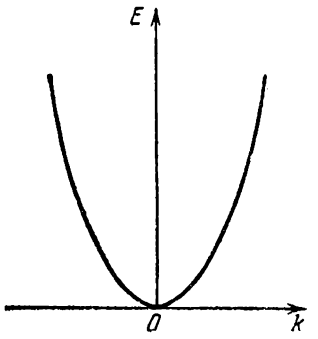


Fig. 7.6

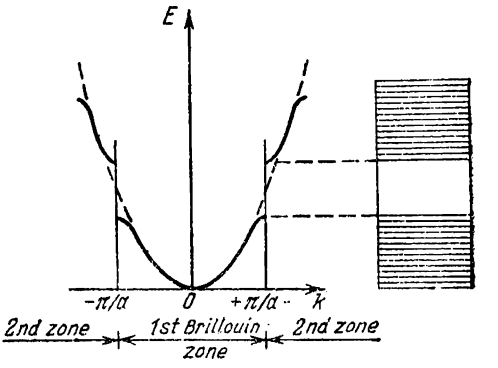


Fig. 7.7

with an energy close to E_F. The nature of motion of the electrons, in turn, determines the physics of the various phenomena observed when a magnetic field acts on a metal.

Thus, the spectrum of the possible values of the energy of the valence electrons in a crystal is divided into a number of allowed and forbidden bands. The width of the bands does not depend on the dimensions of the crystal. Hence, the greater the number of atoms in a crystal, the closer are the levels in a band. The width of the allowed bands has a value of the order of several electron-volts. Consequently, if a crystal contains 10^{23} atoms, then the distance between adjacent levels in a band is about 10^{-23} eV.

Each energy level corresponds to a definite value of **k**. Since the quantum number m_s can take on two values, at any allowed level there can be two electrons having opposite spins.

The existence of the energy bands makes it possible to explain the existence of metals, semiconductors, and dielectrics from a single viewpoint.

The allowed band appearing from the level at which the valence electrons are in the ground state of an atom is called the **valence band**. At absolute zero, the valence electrons fill the lower levels of the valence band in pairs. The higher allowed bands will be free of

electrons. The three cases shown in Fig. 7.8 are possible depending on the degree of filling of the valence band by electrons and on the width of the forbidden band. In case *a*, the electrons fill the valence band only partly. It is therefore sufficient to impart to the electrons at the upper levels a very small energy (of the order of 10^{-23} to 10^{-22} eV) to transfer them to higher levels. The energy of thermal motion (kT) is about 10^{-4} eV at 1 K. Hence, at temperatures other than absolute zero, some of the electrons are transferred to higher levels.

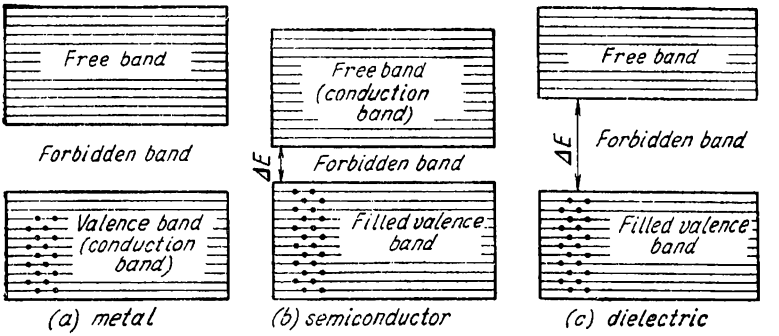


Fig. 7.8

The additional energy due to the action of an electric field on an electron is also sufficient to transfer the electron to higher levels. Consequently, the electrons can be accelerated by an electric field and acquire an additional velocity in a direction opposite to that of the field. Thus, a crystal with such an arrangement of its energy levels will be a metal.

Partial filling of the valence band (for a metal it is also called the **conduction band**) is observed when there is only one electron on the last occupied level in an atom, or when overlapping of the bands occurs (see Fig. 7.5, the distance r_2). In the first case, N conduction electrons fill in pairs only half of the valence band levels. In the second case, the number of levels in the conduction band will be greater than N so that even if the number of conduction electrons is $2N$, they will not be able to occupy all the levels of the band.

In cases *b* and *c* (see Fig. 7.8), the levels of the valence band are completely occupied by electrons—the band is filled. To increase the energy of an electron, it is necessary to impart to it an amount of energy not less than the width of the forbidden band ΔE. An electric field (at any rate of a strength such that no electric breakdown of the crystal occurs) is unable to impart such an energy to an electron. In these conditions, the electrical properties of a crystal are determined by the width of the forbidden band ΔE. If this width

is not great (of the order of several tenths of an electron-volt), the energy of the thermal motion will be sufficient to transfer part of the electrons to the upper free band. These electrons will be in conditions similar to those in which the valence electrons in a metal are. The free band will be a conduction band for them. Simultaneously, the transition of the valence band electrons to its freed upper levels will become possible. Such a substance is called an **electronic semiconductor**.

If the width of the forbidden band ΔE is great (of the order of several electron-volts), then thermal motion cannot feed an appreciable number of electrons into the free band. In this case, the crystal is a dielectric.

7.4. Dynamics of Electrons in a Crystal Lattice

The wave number k is associated with the momentum of an electron p by the equation $p = \hbar k$. Substituting the wave number for the momentum in the uncertainty relation $\Delta p \cdot \Delta x \sim \hbar$, we get an uncertainty relation for k and x:

$$\Delta k \cdot \Delta x \sim 1 \tag{7.26}$$

It follows from this relation that at a strictly definite k, the position of an electron in a crystal will be absolutely indeterminate. To be in a position to study the dynamics of an electron in a crystal, we must have expressions for its velocity and acceleration at our disposal. We can only speak about the velocity, however, if the electron will be at least approximately localized in space.

Let us assume that Δk is other than zero. Consequently, the electron will be localized within the region $\Delta x \sim 1/\Delta k$. According to the superposition principle (see Sec. 4.8), the psi-function of an electron can be represented in the form of the sum of plane waves of the kind e^{ikr}, the values of whose wave numbers are within the limits of Δk. If Δk is not great, the superposition of the plane waves forms a wave packet. The maximum of the resultant wave amplitude travels with the group velocity

$$v_{gr} = \frac{d\omega}{dk} \tag{7.27}$$

[see Eq. (20.15) of Vol. II. p. 461]. The most probable location of the electron coincides with the centre of the wave packet. Consequently, v_{gr} is the velocity of an electron in a crystal.

Taking advantage of the equation $E = \hbar\omega$, we shall substitute the energy for the frequency in Eq. (7.27). As a result, we find that

$$v_{gr} = \frac{1}{\hbar}\frac{dE}{dk} \tag{7.28}$$

Let us see how an electron will behave under the action of the external electric field $\mathscr{E}$ imposed on a crystal. In this case, apart from the forces F_{cryst} produced by the field of the lattice, the electron will experience the force F whose magnitude is $e\mathscr{E}$. During the time dt, this force does the work $dA = Fv_{\text{gr}}\,dt$ on the electron. The introduction of Eq. (7.28) for v_{gr} yields

$$dA = \frac{F}{\hbar}\frac{dE}{dk}\,dt \tag{7.29}$$

This work provides an increment of the energy of the electron in the crystal: $dA = dE$. Using dE instead of dA in Eq. (7.29) and taking into account that $dE = (dE/dk)\,dk$, we arrive at the expression

$$\frac{dE}{dk}\,dk = \frac{F}{\hbar}\frac{dE}{dk}\,dt$$

whence it follows that

$$\frac{dk}{dt} = \frac{F}{\hbar} \tag{7.30}$$

Time differentiation of Eq. (7.28) gives the acceleration of the electron in the crystal:

$$\frac{dv_{\text{gr}}}{dt} \frac{1}{\hbar}\frac{d}{dt}\left(\frac{dE}{dk}\right)\cdot = \frac{1}{\hbar}\frac{d^2E}{dk^2}\frac{dk}{dt}$$

Taking Eq. (7.30) into consideration, we obtain

$$\frac{dv_{\text{gr}}}{dt} = \frac{1}{\hbar}\frac{d^2E}{dk^2}\frac{F}{\hbar}$$

Let us write this formula as follows:

$$\left(\frac{\hbar^2}{d^2E/dk^2}\right)\frac{dv_{\text{gr}}}{dt} = F \tag{7.31}$$

Inspection of Eq. (7.31) shows that the acceleration of an electron in a crystal is proportional to the external force $e\mathscr{E}$. This result is non-trivial because the acceleration should be proportional to the sum of the forces $e\mathscr{E}$ and F_{cryst}, and only the peculiar nature of the force F_{cryst} leads to the fact that with proportionality of the acceleration to the sum of the forces $e\mathscr{E}$ and F_{cryst} it is also proportional to the addend $e\mathscr{E}$.

Comparing Eq. (7.31) with Newton's second law

$$m\frac{dv}{dt} = F$$

we arrive at the conclusion that the expression

$$m^* = \frac{\hbar^2}{d^2E/dk^2} \tag{7.32}$$

formally plays the part of the mass with respect to the external force $F=e\mathscr{E}$. In this connection, the quantity given by Eq. (7.32) is called the **effective mass** of an electron in a crystal.

The effective mass m^* may differ greatly from the actual mass of an electron m, in particular it may take on negative values. This is due to the fact that the equation of Newton's second law actually has the form

$$m\frac{d\mathbf{v}}{dt}=\mathbf{F}+\mathbf{F}_{\text{cryst}} \tag{7.33}$$

where $\mathbf{F}_{\text{cryst}}$ is the force due to the action of the lattice field on an electron. A comparison of Eq. (7.33) with the equation

$$m^*\frac{d\mathbf{v}}{dt}=\mathbf{F}$$

clearly shows that m^* may noticeably differ from m. Notwithstanding this circumstance, it is exactly the value of m^* that determines the nature of the motion of an electron in a lattice under the action of the force $e\mathscr{E}$. The introduction of the effective mass makes it possible to determine the nature of the motion of an electron under the action of the external field, disengaging ourselves from the interaction of the electrons with the lattice. Ascribing the mass m^* to an electron, we can study its behaviour under the action of the force $e\mathscr{E}$ considering the electron to be free. It follows from what has been said above that the relations obtained for the approximation of free electrons also hold for an electron travelling in a periodic field if we replace the true mass m in them with the effective mass m^*.

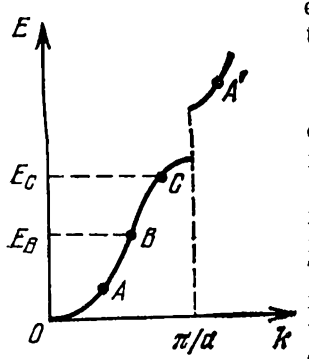


Fig. 7.9

In particular, Eq. (7.3) for a periodic field has the form

$$E=\frac{\hbar^2}{2m^*}k^2 \tag{7.34}$$

Indeed, double differentiation with respect to k yields

$$\frac{d^2E}{dk^2}=\frac{\hbar^2}{m^*}$$

that agrees with the definition of m^* [see Eq. (7.32)].

Thus, the action of the lattice on the motion of an electron can be taken into account by replacing in the equation of motion including only the external force $e\mathscr{E}$ the true mass m with the effective mass m^*.

Let us see how the effective mass m^* depends on the "location" of an electron inside the allowed energy band. Near the bottom of the band (see points A and A' in Fig. 7.9), the course of the curve

$E(k)$ differs only slightly from the course of the curve for free electrons (see Fig. 7.6). Accordingly, $m^* \approx m$.

At the point of inflection (point B in Fig. 7.9), the quantity d^2E/dk^2 is zero. Consequently, m^* becomes infinite. This signifies that an external field can exert no action on the motion of an electron in a state with the energy E_B.

Near the ceiling of the allowed band (point C in Fig. 7.9), the derivative $d^2E/dk^2 < 0$ (the quantity dE/dk diminishes with increasing k). Accordingly, the effective mass m^* of the electrons occupying levels near the ceiling of the band is negative. Actually, this signifies that under the joint action of the forces $e\mathscr{E}$ and F_{cryst} an electron in the state with the energy E_C receives an acceleration opposite in direction to the external force $e\mathscr{E}$.

CHAPTER 8 THE ELECTRICAL CONDUCTANCE OF METALS AND SEMICONDUCTORS

8.1. The Electrical Conductance of Metals

The relevant quantum mechanical calculations show that with a perfect crystal lattice, the conduction electrons would not experience any resistance in their motion, and the electrical conductance of metals would be infinitely great. A crystal lattice is never perfect, however. Violations of the strict periodicity of a lattice may be due to the presence of impurities or vacancies (i.e. the absence of atoms at a point), and also to thermal oscillations of the lattice. Scattering of the electrons on the impurity atoms and on phonons leads to the appearance of electrical resistance of metals. The purer the metal and the lower its temperature, the smaller is this resistance.

The resistivity of metals can be represented in the form

$$\rho = \rho_{osc} + \rho_{imp}$$

where ρ_{osc} is the resistivity due to thermal oscillations of the lattice, and ρ_{imp} is the resistivity due to scattering of the electrons on the impurity atoms. The addend ρ_{osc} diminishes with lowering of the temperature and vanishes at $T = 0$ K. The addend ρ_{imp} at a low concentration of the impurities does not depend on the temperature and forms the so-called **residual resistivity** of a metal* (i.e. the resistivity which a metal has at 0 K; see Fig. 5.5 of Vol. II, p. 105).

Assume that a unit volume of a metal contains n free electrons. We shall call the average velocity of these electrons the **drift velocity** $\mathbf{v}_{dr}$. By definition

$$\mathbf{v}_{dr} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{v}_i \tag{8.1}$$

In the absence of an external field, the drift velocity is zero, and there is no electric current in the metal. When an external electric field E is imposed on the metal, the drift velocity becomes other than zero—an electric current appears in the metal. According to Ohm's law, the drift velocity is finite and is proportional to the force $-e\mathbf{E}$.

It is known from mechanics that the velocity of steady motion is proportional to the external force F applied to a body when in

* We are speaking of metals that do not pass over into the superconducting tate (see the following section).

addition to the force F the body experiences the force of resistance of the medium proportional to the velocity of the body (an example is the falling of a small sphere in a viscous medium). Hence, we conclude that in addition to the force $-e\mathbf{E}$, the conduction electrons in a metal experience a force of "friction" whose average value is

$$\mathbf{F}_f = -r\mathbf{v}_{dr} \tag{8.2}$$

(r is a proportionality constant).

The equation of motion for an "average" electron has the form

$$m^* \frac{d\mathbf{v}_{dr}}{dt} = -e\mathbf{E} - r\mathbf{v}_{dr} \tag{8.3}$$

where m^* is the effective mass of an electron (see Sec. 7.4). This equation allows us to find the steady value of $\mathbf{v}_{dr}$.

If the external field $\mathbf{E}$ is switched off after a steady state sets in, the drift velocity begins to diminish, and completely vanishes when a state of equilibrium between the electrons and the lattice is achieved. Let us find the law of diminishing of the drift velocity after the external field is switched off. Assuming in Eq. (8.3) that $\mathbf{E} = 0$, we get

$$m^* \frac{d\mathbf{v}_{dr}}{dt} + r\mathbf{v}_{dr} = 0$$

We are well acquainted with an equation of this kind. Its solution has the form

$$\mathbf{v}_{dr}(t) = \mathbf{v}_{dr}(0) \exp\left(-\frac{r}{m^*} t\right) \tag{8.4}$$

where $\mathbf{v}_{dr}(0)$ is the value of the drift velocity at the moment when the field is switched off.

Equation (8.4) shows that during the time

$$\tau = \frac{m^*}{r} \tag{8.5}$$

the value of the drift velocity drops to $1/e$-th of its initial value. Thus, the quantity τ given by Eq. (8.5) is the **relaxation time** (see Sec. 10.3 of Vol. I, p. 270) characterizing the process of the establishment of equilibrium between the electrons and the lattice violated by the action of the external field $\mathbf{E}$.

With a view to Eq. (8.5), formula (8.2) can be written as follows:

$$\mathbf{F}_{fr} = -\frac{m^*}{\tau} \mathbf{v}_{dr} \tag{8.6}$$

The steady value of the drift velocity can be found by equating to zero the sum of the force $-e\mathbf{E}$ and the friction force given by Eq. (8.6):

$$-e\mathbf{E} - \frac{m^*}{\tau} \mathbf{v}_{dr} = 0$$

whence

$$\mathbf{v}_{dr} = -\frac{e\mathbf{E}\tau}{m^*}$$

We get the steady value of the current density by multiplying the value of $\mathbf{v}_{dr}$ by the charge of an electron $-e$ and the density of the electrons n:

$$\mathbf{j} = -\frac{e\mathbf{E}\tau}{m^*}(-e)n = \frac{ne^2\tau}{m^*}\mathbf{E}$$

The proportionality constant between E and j is the conductivity σ. Hence,

$$\sigma = \frac{ne^2\tau}{m^*} \tag{8.7}$$

In Sec. 11.2 of Vol. II, p. 232, we obtained the following classical expression for the conductivity of metals:

$$\sigma = \frac{ne^2\tau'}{2m} \tag{8.8}$$

where τ' is the average time of flight of an electron between collisions, and m is the conventional (not effective) mass of an electron [sel Eq. (11.9) of Vol. II, p. 232, we have substituted the average time τ' between collisions for l/v in this equation].

A comparison of Eqs. (8.7) and (8.8) reveals that the relaxation time coincides in the order of its magnitude with the average time of flight of electrons in a metal between collisions.

On the basis of physical considerations, it is possible to assess the quantities in Eq. (8.7) and thus calculate the conductivity σ with respect to its order of magnitude. The values obtained by this method are in good agreement with experimental data. The fact that σ varies with the temperature according to the law $1/T$ also agrees with experimental results. We remind our reader that according to the classical theory, σ is inversely proportional to $T^{1/2}$ (see Sec. 11.2 of Vol. II, p. 234).

We must note that the calculations which led us to Eq. (8.7) are equally suitable both in the classical interpretation of the motion of conduction electrons in a metal and in the quantum mechanical interpretation. The distinction between these two interpretations is as follows. In the classical treatment, it is assumed that all the electrons are disturbed by the external electric field, and in this connection each addend in Eq. (8.1) receives an addition in the direction opposite to E. In the quantum mechanical interpretation, it is necessary to take into account that only the electrons occupying states near the Fermi level are disturbed by the field and change their velocity. The electrons at lower levels are not disturbed by the field, and their contribution to the sum in Eq. (8.1) does not change. Apart from the above, in the classical interpretation, the denominator of

Eq. (8.7) must contain the conventional mass of an electron m, whereas in the quantum mechanical interpretation it must contain the effective mass of an electron m^* instead of the conventional one. This is a manifestation of the general rule noted in Sec. 7.4 according to which relations obtained for the approximation of free electrons are also correct for electrons travelling in the periodic field of a lattice if we substitute the effective mass of an electron m^* for its true mass m.

8.2. Superconductivity

At a temperature of the order of several kelvins, the electrical resistance of a number of metals and alloys vanishes in a jump—the substance passes into a **superconducting state** (see Sec. 5.4 of Vol. II, p. 106). The temperature at which this transition occurs is known as the **critical** or **transition temperature** and is designated by the symbol T_c. The highest observed value of T_c is of the order of 20 K.

Superconductivity can be observed experimentally in two ways:

1. By connecting a section of a superconductor to a general electric circuit. At the moment of transition to the superconducting state, the potential difference across the ends of this section vanishes.

2. By placing a loop of a superconductor into a magnetic field perpendicular to it. After the loop has been cooled to below T_c, the field is switched off. As a result, a non-attenuating electric current is induced in the loop. The current circulates in such a loop an unlimitedly long time. The Dutch scientist Heike Kamerlingh Onnes who discovered this phenomenon demonstrated this by taking a superconducting loop with a current flowing through it from Leyden to Cambridge. In a number of experiments, no current attenuation in a superconducting loop was observed for about a year. In 1959, G. Collins reported that he had observed no current attenuation during two and a half years.

In addition to the absence of electrical resistance, the superconducting state is characterized by no magnetic field penetrating into the body of a superconductor. This phenomenon was discovered by W. Meissner and R. Ochsenfeld in 1933 and is known as the **Meissner effect.** If a superconducting sample is cooled when placed in a magnetic field, at the moment of transition to the superconducting state the field is ejected from the sample, and the magnetic induction in the latter vanishes. We can say formally that a superconductor has a zero permeability ($\mu = 0$). Substances with $\mu < 1$ are called diamagnetics. Thus, a superconductor is a perfect diamagnetic.

A sufficiently strong external magnetic field destroys the superconducting state. The value of the magnetic induction at which this occurs is called the **critical** or **threshold field** and is designated by the symbol B_c. The value of B_c depends on the temperature of a sample.

At the transition temperature, $B_c = 0$. With lowering of the temperature, the value of B_c grows, tending to $B_{c,0}$—the value of the critical field at the zero temperature. An approximate plot of this relation is shown in Fig. 8.1.

If we increase the current flowing through a superconductor connected to a conventional circuit, then at a value of the current of I_c the superconducting state is destroyed. This value of the current is called the **critical current**. The value of I_c depends on the temperature. The form of this relation is similar to that of B_c against T (see Fig. 8.1).

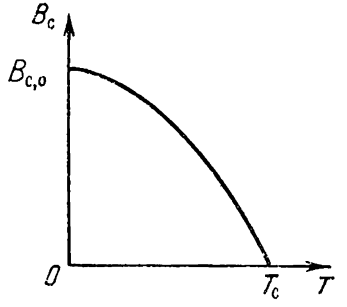


Fig. 8.1

Superconductivity is a phenomenon in which quantum mechanical effects are detected not on microscopic, but on large macroscopic scales*. The theory of superconductivity was developed in 1957 by the American physicists John Bardeen, Leon Cooper and J. Robert Schrieffer. It is called briefly the BCS theory (an acronym). This theory is very complicated. We are therefore forced to treat it at the level of popular science books which apparently will not completely satisfy our exacting readers.

The clue to superconductivity is that the electrons in a metal, apart from Coulomb repulsion, experience a special kind of mutual attraction which in the superconducting state predominates over the repulsion. As a result, the conduction electrons combine to form the so-called **Cooper pairs**. The electrons forming such a pair have oppositely directed spins. Consequently, the spin of a pair is zero, and it is a boson. Bosons are inclined to accumulate in the ground energy state, from which it is comparatively difficult to transfer them to an excited state. Hence, the Cooper pairs, after coming into coordinated motion, stay in this state for an unlimitedly long time. It is exactly this coordinated motion that is the superconduction current.

Let us explain the above in greater detail. An electron moving in a metal deforms (polarizes) the crystal lattice consisting of positive ions. As a result of this deformation, the electron is surrounded by a "cloud" of a positive charge moving along the lattice together with the electron. The electron and the cloud surrounding it are a positively charged system which another electron will be attracted to. The ionic lattice thus plays the part of an intermediate medium whose presence results in attraction between the electrons.

---

* Another phenomenon of this kind is the superfluidity of liquid helium II.

In the language of quantum mechanics, the attraction between electrons is explained as a result of their exchanging quanta of lattice excitation—phonons. An electron moving in a metal violates the conditions of lattice vibrations—it produces phonons. The excitation energy is transmitted to another electron that absorbs a phonon. Owing to this exchange of phonons, additional interaction appears between the electrons that has the nature of attraction. At low temperatures, this attraction exceeds the Coulomb repulsion in substances that are superconductors.

The interaction due to the exchange of phonons manifests itself to the greatest extent in electrons having opposite momenta and spins. As a result, two such electrons combine into a Cooper pair. This pair must not be imagined as two electrons adhering to each other. On the contrary, the distance between the electrons in a pair is quite great and is about 10^{-4} cm, i.e. it is greater by four orders of magnitude than the interatomic distances in a crystal. Approximately 10^6 Cooper pairs appreciably overlap, i.e. occupy a common volume.

Not all conduction electrons combine into Cooper pairs. At a temperature T other than absolute zero, there is a probability of a pair being destroyed. Consequently, in addition to the pairs, there are always "normal" electrons moving in a usual way through a crystal. The closer T is to T_c, the greater is the fraction of normal electrons, which becomes equal to unity when $T = T_c$.

The formation of Cooper pairs leads to reconstruction of the energy spectrum of a metal. To excite an electron system in the superconducting state, it is necessary to dissociate at least one pair. This needs an energy equal to the binding energy E_b of the pair. This energy is the minimum amount of energy that the system of electrons in a superconductor can pick up. Hence, the energy spectrum of the electrons in the superconducting state has a gap whose width is E_b in the region of the Fermi level. The values of the energy belonging to this gap are forbidden. The existence of the gap was proved experimentally.

Thus, the excited state of an electron system in the superconducting state is separated from the ground state by an energy gap of width E_b. Therefore, quantum transitions of this system will not always be possible. At low velocities of its motion (corresponding to a current less than I_c), the electron system will not be excited, and this is exactly what signifies motion without friction, i.e. without electrical resistance.

The width of the energy gap E_b diminishes with increasing temperature and vanishes at the transition temperature T_c. Accordingly, all the Cooper pairs dissociate, and the substance passes over to its normal (non-superconducting) state.

It follows from the theory of superconductivity that the magnetic flux Φ associated with a superconducting loop (or cylinder) in which a current is circulating must be an integral multiple of the quantity $2\pi\hbar/q$, where q is the charge of a current carrier:

$$\Phi = n\frac{2\pi\hbar}{q}$$

The quantity

$$\Phi_0 = \frac{2\pi\hbar}{q} \tag{8.9}$$

is a **quantum of magnetic flux.**

The quantization of the magnetic flux was detected experimentally in 1961 by B. Deaver and W. Fairbank, and independently of them by R. Doll and M. Näbauer. In the experiments run by Deaver and Fairbank, the sample was a belt of tin applied onto a copper wire about 10^{-3} cm in diameter. The wire played the part of a framework and did not pass over to the superconducting state. The measured values of the magnetic flux in these experiments, as in the ones conducted by Doll and Näbauer, were found to be integral multiples of quantity (8.9) in which the double charge of an electron must be substituted for q ($q = -2e$). This is an additional confirmation of the correctness of the BCS theory according to which the Cooper pairs having a charge of $-2e$ are the current carriers in a superconductor.

8.3. Semiconductors

Semiconductors are crystalline substances in which the valence band is completely filled with electrons (see Fig. 7.8*b*), and the width of the forbidden band is not great (not over 1 eV in intrinsic semiconductors). Semiconductors owe their name to the circumstance that with respect to the value of their electrical conductance they occupy an intermediate position between metals and dielectrics. Their feature, however, is not the magnitude of the conductance, but the fact that their conductance grows with increasing temperature (we remind our reader that in metals it diminishes).

Intrinsic and **impurity** (or **extrinsic**) semiconductors are distinguished. The intrinsic semiconductors include chemically pure ones. The electrical properties of the impurity semiconductors are determined by the impurities they have artificially been doped with.

In considering the electrical properties of semiconductors, a great part is played by the concept of "holes". Let us stop to deal with the physical meaning of this concept.

In an intrinsic semiconductor at absolute zero, all the levels of the valence band are completely filled with electrons, while the

latter are absent in the conduction band (Fig. 8.2*a*). An electric field cannot transfer electrons from the valence band to the conduction one. Therefore, intrinsic semiconductors behave at absolute zero like dielectrics. At temperatures other than 0 K, a part of the electrons from the upper levels of the valence band transfer to the lower levels of the conduction band as a result of thermal excitation (Fig. 8.2*b*). In these conditions, an electric field obtains the possibility of changing the state of the electrons in the conduction band. In addition, owing to the formation of vacant levels in the valence

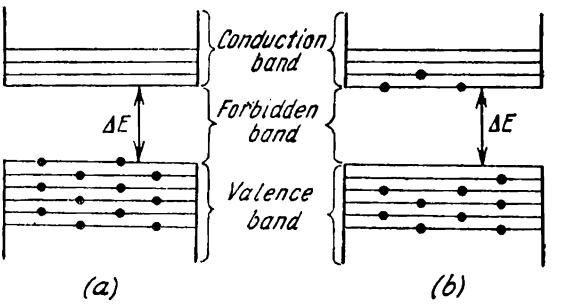


Fig. 8.2

band, the electrons of this band can also change their velocity under the action of the external field. As a result, the electrical conductance of the semiconductor becomes other than zero.

It was found that when vacant levels are present, the behaviour of the valence band electrons can be represented as the motion of positively charged quasiparticles that have been named "holes". It follows from the equality to zero of the conductance of a completely filled valence band that the sum of the velocities of all the electrons of such a band equals zero:

$$\sum_i \mathbf{v}_i = 0$$

Let us separate from this sum the velocity of the k-th electron:

$$\sum_{i \neq k} \mathbf{v}_i + \mathbf{v}_k = 0$$

whence

$$\sum_{i \neq k} \mathbf{v}_i = -\mathbf{v}_k$$

This relation shows that if the k-th electron is absent in the valence band, then the sum of the velocities of the remaining electrons is $-\mathbf{v}_k$. Hence, all these electrons will set up a current equal to $(-e) \cdot (-\mathbf{v}_k) = e\mathbf{v}_k$. Thus, the produced current is found to be equivalent

to the current that would be set up by a particle with the charge $+e$ having the velocity of the absent electron. It is exactly this imaginary particle that is a hole.

We can also arrive at the concept of holes as follows. Vacant levels are formed at the top of the valence band. In Sec. 7.4, we established that the effective mass of an electron at the ceiling of an energy band is negative. The absence of a particle having a negative charge $-e$ and a negative mass m^* is equivalent to the presence of a particle having a positive charge $+e$ and a positive mass $|m^*|$, i.e. of a hole.

Thus, a valence band having a small number of vacant states is equivalent in its electrical properties to a vacant band containing a small number of positively charged quasiparticles called holes.

We must stress the fact that the motion of a hole is not the motion of a real positively charged particle.The notion of holes reflects the nature of motion of the entire multiple-electron system in a semiconductor.

8.4. Intrinsic Conductance of Semiconductors

Intrinsic conductance appears as a result of the transition of electrons from the upper levels of the valence band to the conduction band. A certain number of current carriers—electrons appear in the conduction band that occupy levels near the bottom of the band; simultaneously, the same number of sites are freed in the valence band at its upper levels, the result being the appearance of holes (see the preceding section).

The distribution of the electrons among the levels of the valence band and the conduction band is described by the Fermi-Dirac function [see Eq. (7.20)]. This distribution can be made very illustrative by depicting, as has been done in Fig. 8.3, a graph of the distribution function combined with a diagram of the energy bands.

The relevant calculations indicate that for intrinsic semiconductors the value of the Fermi level measured from the ceiling of the valence band is

$$E_F = \frac{1}{2}\Delta E + \frac{3}{4}kT \ln \frac{m_h^*}{m_e^*}$$

where ΔE is the width of the forbidden band, and m_h^* and m_e^* are the effective masses of a hole and an electron in the conduction band. The second addend is usually negligibly small, and we may assume that $E_F = \frac{1}{2}\Delta E$. This signifies that the Fermi level is at the middle of the forbidden band (see Fig. 8.3). Consequently, for the electrons that have passed into the conduction band, the quantity $E - E_F$ differs only slightly from half the width of the forbidden band. The

levels of the conduction band are on the tail of the distribution curve. Therefore, the probability of their being filled can be found by Eq. (7.24). Assuming in this equation that $E - E_F \approx \Delta E/2$, we find that

$$f(E) \propto \exp\left(-\frac{\Delta E}{2kT}\right) \tag{8.10}$$

The number of electrons that have passed into the conduction band and, consequently, the number of holes formed too, will be proportional to the probability given by expression (8.10). These electrons

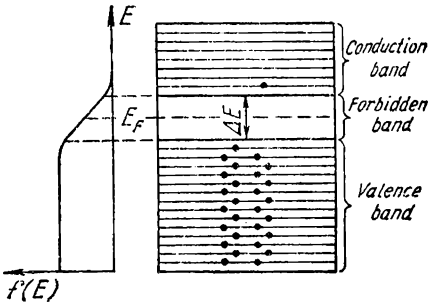


Fig. 8.3

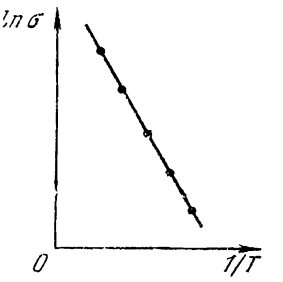


Fig. 8.4

and holes are current carriers. Since the conductance is proportional to the number of carriers, it must also be proportional to expression (8.10). Hence, the electrical conductance of intrinsic semiconductors rapidly grows with the temperature, varying according to the law

$$\sigma = \sigma_0 \exp\left(-\frac{\Delta E}{2kT}\right) \tag{8.11}$$

where ΔE is the width of the forbidden band, and σ is a constant.

If we plot $\ln \sigma$ against $1/T$ on a graph, then for intrinsic semiconductors we get the straight line shown in Fig. 8.4. We can determine the width of the forbidden band ΔE according to the slope of this line.

The elements of group IV of Mendeleev's periodic table germanium and silicon are typical semiconductors. They form a diamond-type lattice in which each atom is connected by covalent (electron-pair) bonds to four neighbouring atoms at equal distances from it (see Fig. 13.6*a* of Vol. I, p. 371). Such a mutual arrangement of the atoms can conditionally be represented in the form of the plane structure shown in Fig. 8.5. The circles with the sign "+" designate positively charged atom cores (i.e. the part of an atom that remains after de-

pature of the valence electrons), the circles with the sign "—" designate the valence electrons, the double lines show the covalent bonds.

At a sufficiently high temperature, thermal motion may dissociate separate pairs, releasing one electron. The site left by the electron stops being neutral, and a surplus positive charge $+e$ appears in its vicinity—a hole is formed (it is depicted by a dash circle in Fig. 8.5). An electron from one of the adjacent pairs may jump over to this site. As a result, the hole also begins to wander along the crystal like the released electron.

When a free electron meets a hole, they **recombine.** This signifies that the electron neutralizes the surplus positive charge in the vicinity of the hole and loses its freedom of motion until it again receives the energy sufficient for its release from the crystal lattice. Recombination leads to the simultaneous vanishing of the free electron and the hole. In a level diagram (Fig. 8.3), the transition of an electron from the conduction band to one of the free levels of the valence band corresponds to the recombination process.

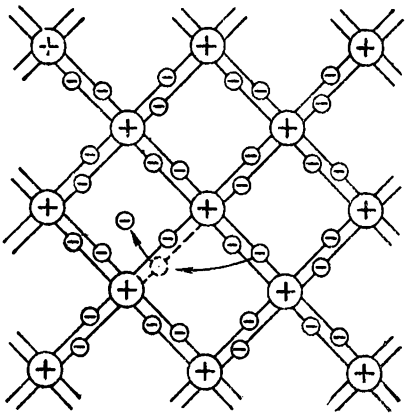


Fig. 8.5

Thus, two processes occur simultaneously in an intrinsic semiconductor: the birth of free electron-hole pairs and recombination leading to the vanishing of these pairs. The probability of the first process grows rapidly with the temperature. The probability of recombination is proportional both to the number of free electrons and to the number of holes. Hence, a definite equilibrium concentration of electrons and holes corresponds to each temperature, and it varies with the temperature in proportion to expression (8.10).

When an external electric field is absent, the conduction electrons and holes move chaotically. When a field is switched on, ordered motion is imposed onto the chaotic motion: the electrons move against the field, and the holes in the direction of the field. Both motions—of the holes and the electrons—result in transferring the charge along the crystal. Hence, intrinsic conductance is due, as it were, to charge carriers of two signs—negative electrons and positive holes.

We must note that at a sufficiently high temperature, intrinsic conductance is observed in all semiconductors without any exception. In semiconductors containing an impurity, however, the conductance consists of the intrinsic and impurity conductances.

8.5. Impurity Conductance of Semiconductors

Impurity conductance appears if some atoms of a given semiconductor are replaced at the points of a crystal lattice by atoms whose valence differs by unity from that of the normal atoms. Figure 8.6 shows schematically a germanium lattice doped with pentavalent phosphorus atoms. A phosphorus atom needs only four electrons to form covalent bonds with its neighbours. Consequently, the fifth valence electron is surplus, as it were, and is easily detached from the atom at the expense of the energy of thermal motion, forming

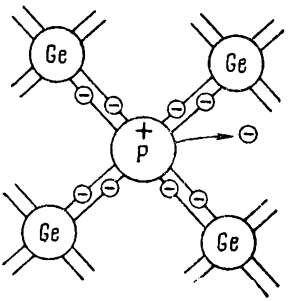


Fig. 8.6

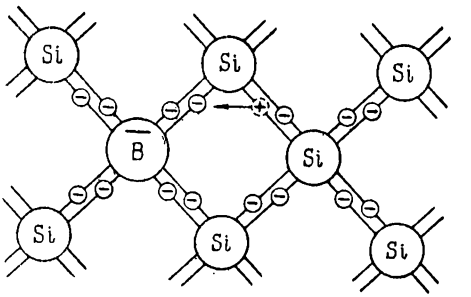


Fig. 8.7

a wandering free electron. Unlike the case treated in the preceding section, the formation of a free electron is not attended by the violation of the covalent bonds, i.e. by the formation of a hole. Although an excess positive charge does appear in the vicinity of the impurity atom, it is bound to this atom and cannot travel along the lattice. Owing to this charge, the impurity atom can capture an electron approaching it, but the bond of the captured electron with the atom will be weak and can easily be broken again at the expense of the thermal oscillations of the lattice.

Thus, a semiconductor with an impurity whose valence is greater by unity than that of the normal atoms has only one kind of current carriers—electrons. Accordingly, such a semiconductor is said to have electronic conductance or to be a semiconductor of the *n*-type (negative). The impurity atoms supplying the conduction electrons are called **donors**.

Now let us consider the case when the valence of the impurity is less by unity than that of the normal atoms. Figure 8.7 shows schematically a silicon lattice doped with trivalent boron atoms. The three valence electrons of a boron atom are not enough to form bonds with all four neighbours. One of the bonds is therefore not completed and will be a place capable of capturing an electron. When an electron from one of the neighbouring pairs passes to this place, a hole

will be formed that will travel along the crystal. An excess negative charge will appear near the impurity atom, but it will be associated with the given atom and cannot become a current carrier. Thus, current carriers of only one kind—holes—are produced in a semiconductor with an impurity whose valence is less by unity than that of the normal atoms. Such a semiconductor is said to have hole conductance, and the semiconductor is said to be of the *p*-type (positive). The impurity atoms causing holes to appear are called **acceptors.**

The electronic nature of the conductance of *n*-type semiconductors and the hole nature of the conductance of *p*-type semiconductors are

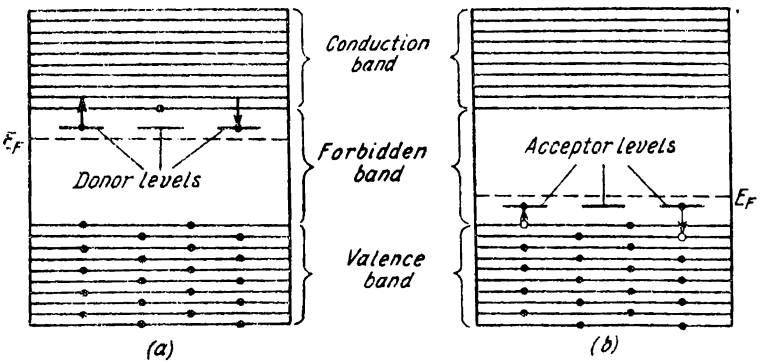


Fig. 8.8

confirmed experimentally when studying the Hall effect (see Sec. 11.3 of Vol. II, p. 234 et seq.). The observed sign of the Hall potential difference corresponds to negative current carriers in *n*-type semiconductors, and to positive carriers in *p*-type ones.

Impurities distort the field of a lattice. This results in the appearance of **impurity levels** in the energy scheme. They are in the forbidden band of the crystal. The impurity levels of *n*-type semiconductors are called **donor levels** (Fig. 8.8*a*), and of *p*-type semiconductors **acceptor levels** (Fig. 8.8*b*).

The Fermi level in *n*-type semiconductors is in the upper half of the forbidden band, and in *p*-type semiconductors, in the lower half of this band. With elevation of the temperature, the Fermi level in semiconductors of both types is displaced toward the middle of the forbidden band.

If the donor levels are not far from the ceiling of the valence band*, they cannot appreciably affect the electrical properties of the crystal. Matters are different when the distance from such levels

* This signifies that the fifth valence electron is firmly bound to its atom.

to the bottom of the conduction band is much smaller than the width of the forbidden band. In this case, the energy of thermal motion even at ordinary temperatures is sufficient to transfer an electron from a donor level to the conduction band (see Fig. 8.8*a*). Detachment of the fifth valence electron from an impurity atom corresponds to this process. The transition of an electron from the conduction band to one of the donor levels corresponds in Fig. 8.8*a* to the capture of a free electron by an impurity atom.

The acceptor levels noticeably affect the electrical properties of a crystal when they are not far from the ceiling of the valence zone (see Fig. 8.8*b*). The transition of an electron from the valence zone to an acceptor level corresponds to the formation of a hole. The reverse transition corresponds to breaking of one of the four covalent bonds of an impurity atom with its neighbours and recombination of the electron and hole formed.

Upon elevation of the temperature, the concentration of the impurity current carriers rapidly reaches saturation. This signifies that practically all the donor levels are freed of electrons or all the acceptor levels are filled with them. At the same time with elevation of the temperature, the intrinsic conductance of the semiconductor, owing to the transition of electrons directly from the valence band to the conduction band, begins to tell more and more. Thus, at high temperatures, the conductance of the semiconductor will consist of impurity and intrinsic conductances. At low temperatures, impurity conductance prevails, and at high ones intrinsic conductance.

CHAPTER 9 CONTACT AND THERMOELECTRIC PHENOMENA

9.1. Work Function

Conduction electrons do not leave a metal arbitrarily in noticeable numbers. The explanation is that the metal is a potential well for them. Only those electrons succeed in escaping from the metal whose energy is sufficient to surmount the potential barrier on its surface. The forces giving rise to this barrier have the following origin. The chance removal of an electron from the external layer of positive

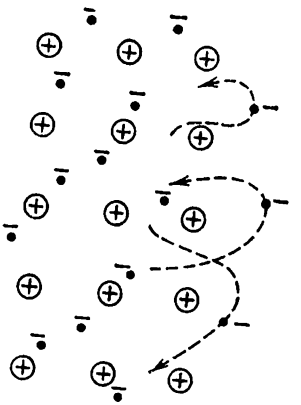


Fig. 9.1

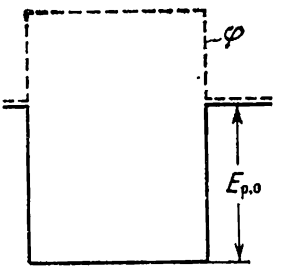


Fig. 9.2

ions of the lattice results in the appearance of an excess positive charge at the site left by the electron. The Coulomb interaction with this charge causes an electron whose velocity is not very high to return. Thus, individual electrons are constantly leaving the surface of the metal, travelling several interatomic distances away from it and then returning. As a result, the metal is surrounded by a thin cloud of electrons. This cloud together with the external layer of ions forms an electrical double layer (Fig. 9.1; the circles depict ions, and the black dots, electrons). The forces exerted on an electron in such a layer are directed into the metal. The work done against these forces in transferring an electron out of the metal goes to increase the potential energy E_p of the electron.

Thus, the potential energy of the valence electrons* inside a metal is greater than that outside the metal by an amount equal to the depth of the potential well $E_{p,\,0}$ (Fig. 9.2). The energy changes over a length of the order of several interatomic distances, therefore we can consider that the walls of the well are vertical.

The potential energy of an electron $E_p = -e\varphi$ and the potential φ of the point where the electron is have opposite signs. It thus follows

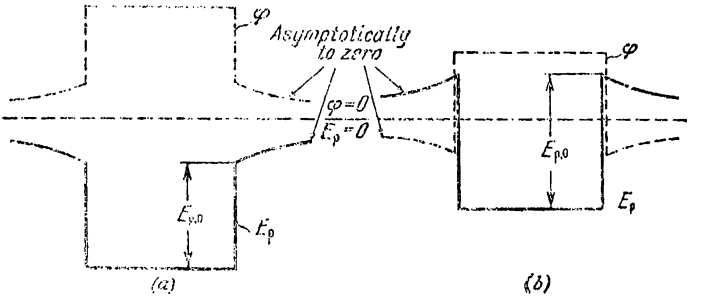


Fig. 9.3

that the potential inside a metal is higher than that in direct proximity to its surface (we shall simply say "on its surface" for brevity) by the amount $E_{p,\,0}/e$.

The imparting of an excess positive charge to a metal increases the potential both on the surface and inside the metal. The potential energy of an electron diminishes accordingly (Fig. 9.3*a*). We remind our reader that the values of the potential and the potential energy at infinity have been taken as the reference point. The imparting of a negative charge lowers the potential inside and outside the metal. The potential energy of an electron grows accordingly (Fig. 9.3*b*).

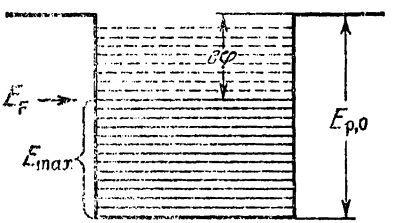


Fig. 9.4

The total energy of an electron in a metal consists of its potential and kinetic energies. We established in Sec. 7.1 that at absolute zero the values of the kinetic energy of conduction electrons range from zero to the energy E_{max} coinciding with the Fermi level. In Fig. 9.4, the energy levels of the conduction band are inscribed in the potential well (the dash lines depict the levels that are not occupied at 0 K).

* The potential well for the electrons filling the levels of the lower bands (i.e. firmly bound to their atoms) has a greater depth. All the reasoning in this section relates to valence electrons.

Different electrons need different energies to escape from a metal. For example, an electron on the lowest level of the conduction band needs the energy $E_{p,0}$; an electron on the Fermi level needs only the energy $E_{p,0} - E_{max} = E_{p,0} - E_F$.

The smallest energy that must be imparted to an electron in order to remove it from a solid or liquid in a vacuum is called the **work function.** The work function is customarily designated by $e\varphi$, where φ is a quantity known as the **emission potential.**

In accordance with what has been said above, the work function* in the emission of an electron from a metal is determined by the expression

$$e\varphi = E_{p,0} - E_F \tag{9.1}$$

We have obtained this expression assuming that the temperature of the metal is 0 K. For other temperatures, the work function is also determined as the difference between the depth of the potential well and the Fermi level, i.e. definition (9.1) is extended to other temperatures. It is also employed for semiconductors.

The Fermi level depends on the temperature [see formula (7.23)]. Moreover, owing to the change in the average interatomic distances due to thermal expansion, the depth of the potential well $E_{p,0}$ changes slightly. The result is that the work function has a slight temperature dependence.

The work function is very sensitive to the state of a metal surface, particularly to its purity. By appropriately choosing the coating of a surface, we can greatly diminish the work function. For example, the application of a layer of an alkaline-earth metal oxide (CaO, SrO, BaO) to the surface of tungsten reduces the work function from 4.5 eV (for pure tungsten) to 1.5-2 eV.

9.2. Thermionic Emission. Electronic Tubes

At temperatures other than absolute zero, there is a certain number of electrons whose energy is sufficient to surmount the potential barrier on the boundary of a metal. The number of such electrons sharply grows with elevation of the temperature and becomes quite noticeable. The emission of electrons by a heated metal is known as **thermionic emission.**

Thermionic emission is studied with the aid of the circuit shown in Fig. 9.5. The main element of the circuit is a **two-electrode tube,** also called a **vacuum-tube diode.** It is a well evacuated metal or glass bulb containing two electrodes—cathode C and anode A. The electrodes can be designed in different ways. In the simplest design,

* The quantity determined by Eq. (9.1) is sometimes called the effective work function, while E_{p0}, is called the total work function.

the cathode has the shape of a thin straight filament, and the anode, of a cylinder coaxial with it (Fig. 9.6).

The cathode is heated by the current provided by the filament battery B_f. The temperature of the cathode can be varied by regulating the filament current with the aid of rheostat R. Anode battery B_a feeds a voltage to the electrodes. The anode voltage U_a can be varied with the aid of potentiometer P and measured with voltmeter V

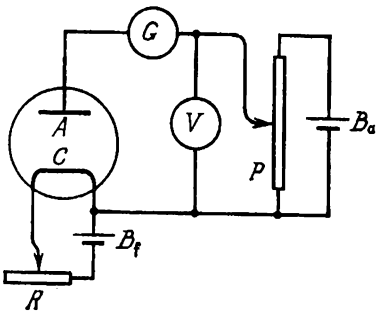


Fig. 9.5

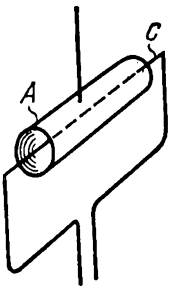


Fig. 9.6

(the voltage U_a is considered to be positive if the anode potential is higher than that of the cathode). Galvanometer G is intended for measuring the anode current I_a.

At a constant cathode filament current, the curve showing how the anode current I_a depends on the anode voltage U_a has the form shown in Fig. 9.7. This curve is called the **volt-ampere characteristic of a diode.** The different curves in Fig. 9.7 correspond to different cathode temperatures. At low values of U_a, these curves coincide.

Let us consider the features of the curves $I_a = f(U_a)$. When $U_a = 0$, the electrons flying out of the cathode form a negative space charge around it, i.e. an electron cloud. The latter repels the electrons flying out of the cathode and returns the greater part of them. A small number of electrons, nevertheless, succeed in flying up to the anode, and as a result a weak current is set up in the anode circuit. To completely stop the electrons from getting onto the anode, i.e. to make I_a equal to zero, a certain negative voltage must be applied between the cathode and the anode. This is why the volt-ampere characteristic of a diode begins not from zero, but somewhat to the left of the origin of coordinates.

A glance at Fig. 9.7 shows that Ohm's law is not obeyed for a vacuum diode. The initial portion of the curve follows quite well the **three-halves power law** obtained theoretically by I. Langmuir and S. Boguslavsky. According to this law, the anode current changes in proportion to $U_a^{3/2}$.

With a growth in U_a, a greater and greater number of electrons are drawn off by the electric field to the anode, and, finally, at a definite value of U_a, the electron cloud is completely dispersed and all the electrons flying out of the cathode can reach the anode. A further growth in U_a cannot increase the anode current—it reaches saturation.

It is obviously exactly the saturation current that characterizes thermionic emission. If N electrons fly out of unit surface area of the cathode in unit time, then the density of the saturation current (the saturation current related to unit surface area of the cathode) will be $j_{sat} = Ne$. Hence, by measuring the saturation current density at various filament currents, we can find the number of electrons flying out from unit surface area at different temperatures.

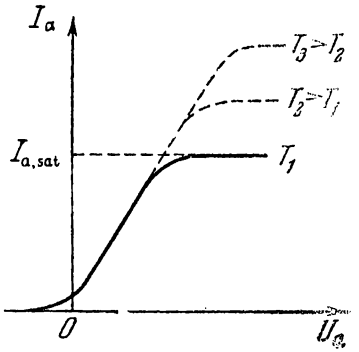


Fig. 9.7

Proceeding from quantum notions, S. Dushman in 1923 obtained the following formula for the saturation current:

$$j_{sat} = AT^2 \exp\left(-\frac{e\varphi}{kT}\right) \qquad (9.2)$$

Here $e\varphi$ is the work function, and A is a constant not depending on the kind of the metal. The theoretical value of A is 120 $A/(cm^2 \cdot K^2)$. The experimental values of the constant A are considerably lower than the theoretical one and differ greatly for various metals. Formula (9.2) shows the temperature dependence of the saturation current quite satisfactorily. A graph of function (9.2) is given in Fig. 9.8.

In 1901, J. Richardson derived a classical formula for thermionic emission. It differs from Eq. (9.2) only in including $T^{1/2}$ instead of T^2. Formula (9.2) is known as the Richardson-Dushman formula.

Inspection of Eq. (9.2) reveals that a decrease in $e\varphi$ sharply increases the emission (it is easy to see that at 1160 K, i.e. at $kT = 0.10$ eV, a decrease in $e\varphi$ from 3 to 1 eV leads to a growth in j_{sat} of almost 5×10^8 times). Therefore, in manufacturing electronic tubes, special coatings and ways of treating the cathodes are used that result in lowering of the work function. The modern so-called oxide cathodes made of nickel coated with barium or strontium oxide have a work function of the order of 1.0-1.2 eV.

A diode passes a current only when the potential of the anode is higher than that of the cathode. At a negative voltage, there is no current in the anode circuit. This property of a diode makes its use possible for rectifying an alternating current. A diode intended for this purpose is also called a kenotron. The solid line in Fig. 9.9 is a

graph of the current flowing through a kenotron if an alternating voltage that changes with time according to a harmonic law is applied to it. In this case, the current will flow in the circuit only during half a period, which is why this way of rectifying current is called half-wave.

By using simultaneously two kenotrons or a double diode assembled in one bulb, we can obtain full-wave rectification. The corres-

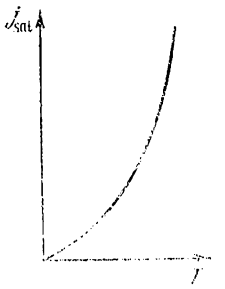


Fig. 9.8

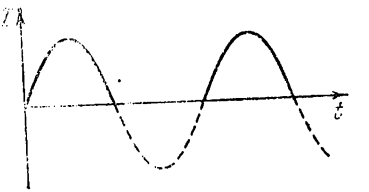


Fig. 9.9

ponding circuit is depicted in Fig. 9.10. The primary winding of the transformer is supplied with alternating current. There are two secondary windings. The smaller one is used to heat the cathode. The larger winding has a middle terminal that is connected to the cathode via load R. The ends of the winding are connected to the anodes. During one half of a period, one anode is under a higher potential than the cathode, and during the second half—the other anode. As a result, the current shown graphically in Fig. 9.11 flows through the load. Such a pulsating current can be smoothed.

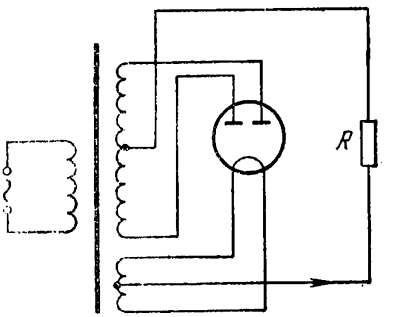


Fig. 9.10

If we place a third electrode in the form of a grid between the cathode and the anode, we get a three-electrode tube—a **triode** (Fig. 9.12; the filament circuit has been omitted in the diagram). The grid may be designed, for example, as a spiral winding around the cathode. When a slight positive potential relative to the cathode is imparted to the grid (in this case we shall consider the voltage U_g between the grid and the cathode to be positive), the electrons will be withdrawn from the cathode at a higher rate. Some of them will get onto the grid (as a result, a small grid current I_g is produced), but the major part of the electrons fly through the grid and reach

the anode. Owing to the grid being close to the cathode, slight changes in the voltage between the grid and the cathode greatly affect the anode current.

A negative grid voltage U_g reduces the anode current, and at a sufficiently high negative voltage U_g the current stops completely—the tube is wiped out.

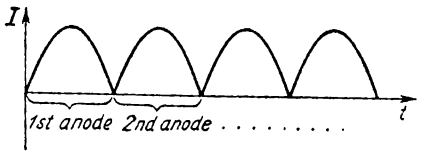


Fig. 9.11

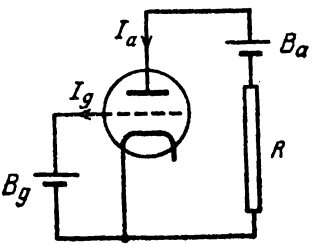


Fig. 9.12

If we plot the anode current I_a against the grid voltage U_g for a constant anode voltage U_a, we get the curve shown in Fig. 9.13.

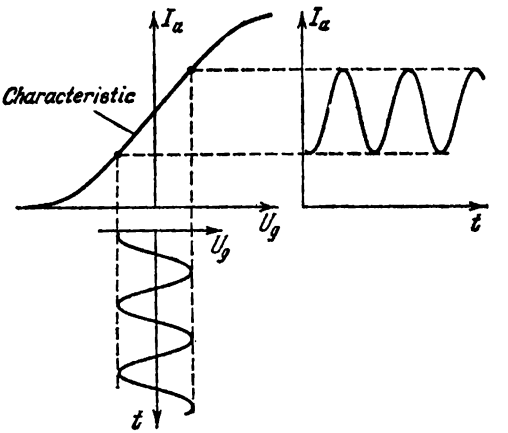


Fig. 9.13

A collection of such curves plotted for different values of U_a forms a family of the **grid characteristics** of a triode. The quantity

$$S = \frac{dI_a}{dU_g}$$

is known as the **slope of the characteristic.**

A considerable portion of the characteristic is straight. This makes it possible, by supplying a small sinusoidal voltage U_g to the grid,

to obtain a large sinusoidal change in the anode current. An alternating voltage with a much greater amplitude than that of U_g can be taken off resistor R. This underlies the operation of a triode as an amplifier. A triode can also be used for the conversion (changing the shape) and generation of varying currents and voltages.

Additional electrodes—grids—are introduced into electronic tubes to improve their characteristics. A tube with two grids, i.e. a four-electrode tube, is called a **tetrode**, a five-electrode one is called a **pentode**, and so on. Tubes in which one bulb accommodates two systems of electrodes have also come into great favour. Such a tube performs the functions of two conventional ones.

9.3. Contact Potential Difference

If we bring two different metals into contact, a potential difference is produced between them that is called a **contact** potential. The result is the appearance of an electric field in the space surrounding the metals. Figure 9.14 shows the equipotential surfaces (solid lines) and the strength lines (dash) of this field; the surface of each of the metals is an equipotential one.

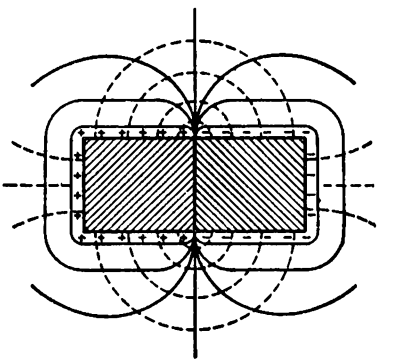


Fig. 9.14

The contact potential difference is due to the fact that when metals come into contact, part of the electrons pass from one metal into the other. The upper part of Fig. 9.15 shows two metals—at the left before they are brought into contact, and at the right, after contact. The lower part of the figure gives graphs of the potential energy of an electron. It is assumed that the Fermi level in the first metal is higher than in the second one. It is natural that when the metals come into contact the electrons from the highest levels in the first metal will begin to pass over to the lower free levels of the second metal. As a result, the potential of the first metal grows, and of the second one diminishes. Accordingly, the potential energy of an electron in the first metal diminishes, and in the second one grows (we remind our reader that the potential of a metal and the potential energy of an electron in it have different signs; see Fig. 9.3).

It is proved in statistical physics that the condition for equilibrium between metals in contact (and also between semiconductors or a metal and a semiconductor) is the equality of the total energies corresponding to the Fermi levels. In this condition, the Fermi

levels of both metals are at the same height in the diagram. Inspection of Fig. 9.15 reveals that in this case the potential energy of an electron in direct proximity to the surface of the first metal will be lower by $e\varphi_2 - e\varphi_1$ than near the second metal. Hence, the potential on the surface of the first metal will be higher by

$$U_{12} = \frac{e\varphi_2 - e\varphi_1}{e} = \varphi_2 - \varphi_1 \tag{9.3}$$

than on the surface of the second one. It is exactly the quantity U_{12} that is the contact potential difference between the first and the second metal.

According to Eq. (9.3), the contact potential difference between the first and the second metal equals the difference between the work

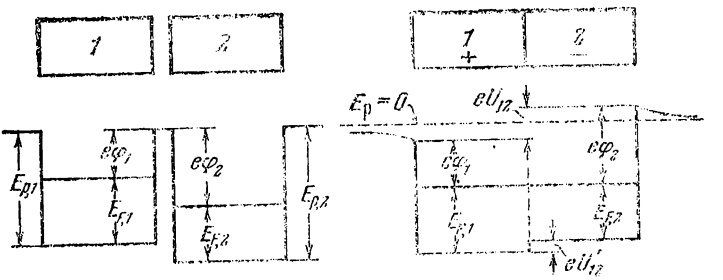


Fig. 9.15

functions for the second and the first metal divided by the elementary charge, or simply the difference of the emission (work function) potentials for the second and the first metal.

The potential difference given by Eq. (9.3) is established between points outside the metals in direct proximity to their surfaces. This is why it is known as the external contact potential difference. We most frequently speak simply of the contact potential difference, having in mind the external one. There is also a potential difference between the internal points of the metals called the internal one. Examination of Fig. 9.15 reveals that the potential energy of an electron in the first metal is lower than that in the second one by $E_{F,1} - E_{F,2}$. Accordingly, the potential inside the first metal is higher than that inside the second one by the amount

$$U'_{12} = \frac{E_{F,1} - E_{F,2}}{e} \tag{9.4}$$

This expression gives the internal contact potential difference. This is the amount by which the potential diminishes when passing from the first metal to the second one.

If we give two different metals the shape shown in Fig. 9.16 and bring them into contact, then an electric field is set up in gap *B-C*

whose strength lines are shown by dash lines. The change in the potential along the contour designated by the dot-and-dash line is shown at the right of the figure.

Figure 9.17 shows the change in the potential energy of an electron along three different metals *1*, *2*, *3* in contact with one another. A glance at the figure shows that the potential difference which

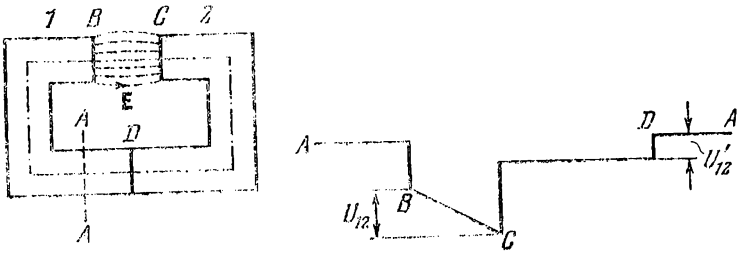


Fig. 9.16

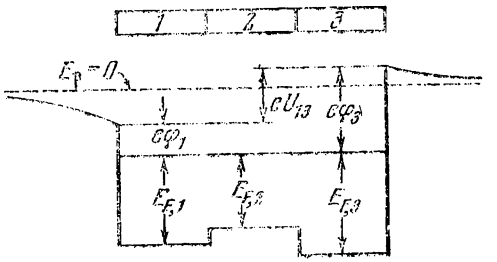


Fig. 9.17

sets in between metals *1* and *3* is exactly the same in this case as when they are in direct contact*. The same is true for any number of intermediate metals: the potential difference across the ends of a circuit is determined by the difference between the work functions for the metals forming the extreme links of the circuit.

The values of the external contact potential difference vary for different pairs of metals from several tenths of a volt to several volts.

We have considered the contact between two metals. A contact potential difference also appears, however, on the interface between a metal and a semiconductor, and also on the interface between two semiconductors.

In conclusion, we shall consider a closed circuit consisting of an arbitrary number of different metals or semiconductors (Fig. 9.18).

* The values of the potentials may change here. In particular, both extreme metals may have a potential of the same sign.

If all the junctions are maintained at the same temperature, the sum of the potential jumps will be zero. Therefore, no e.m.f. can appear in the circuit. The appearance of a current in such a circuit would contradict the second law of thermodynamics. Indeed, since the flow of a current in metals and semiconductors is not attended by chemical changes, the current would do work at the expense of the heat received from the medium surrounding the circuit. No auxiliary processes (for example, the transmission of part of the heat received to other bodies) would occur. Consequently, a perpetual motion machine of the second kind would be achieved here.

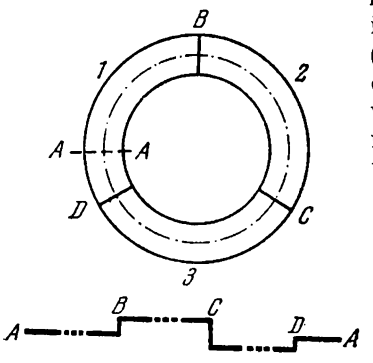


Fig. 9.18

9.4. Thermoelectric Effects

There is a relation between thermal and electrical processes in metals and semiconductors that underlies **effects** known as **thermoelectric** ones. Among them are the Seebeck effect, the Peltier effect, and the Thomson effect.

The Seebeck Effect. The German physicist Thomas Seebeck (1770-1831) discovered in 1821 that when junctions *1* and *2* of two different metals forming a closed circuit (Fig. 9.19) have different temperatures,

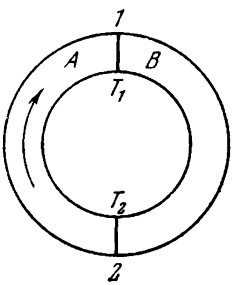


Fig. 9.19

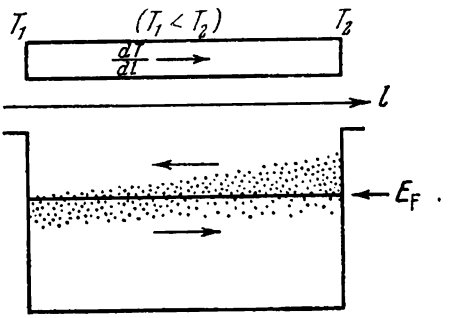


Fig. 9.20

an electric current flows in the circuit. A change in the sign of the difference between the junction temperatures is attended by a change in the direction of the current.

The thermal electromotive force (thermal e.m.f.) is due to three reasons: (1) the dependence of the Fermi level on the temperature, (2) the diffusion of electrons (or holes), and (3) the carrying along of electrons by phonons.

The Fermi level depends on the temperature [see formula (7.23)]. Therefore, the jump in the potential when passing from one metal to another [i.e. the internal contact potential difference, see Eq. (9.4)] is not the same for junctions at different temperatures, and the sum of the potential jumps differs from zero. This alone would be enough for the appearance of an e.m.f. acting in the direction shown in Fig. 9.19 and equal to

$$\mathscr{E}_{\text{cont}} = U'_{\text{AB}}(T_1) + U'_{\text{BA}}(T_2) =$$

$$= \frac{1}{e}\{[E_{\text{F, A}}(T_1) - E_{\text{F, B}}(T_1)] + [E_{\text{F, B}}(T_2) - E_{\text{F, A}}(T_2)]\} =$$

$$= \frac{1}{e}\{[E_{\text{F, B}}(T_2) - E_{\text{F, B}}(T_1)] - [E_{\text{F, A}}(T_2) - E_{\text{F, A}}(T_1)]\}$$

The last expression can be written as follows:

$$\mathscr{E}_{\text{cont}} = \int_{T_1}^{T_2} \left(\frac{1}{e}\frac{dE_{\text{F, B}}}{dT}\right) dT - \int_{T_1}^{T_2} \left(\frac{1}{e}\frac{dE_{\text{F, A}}}{dT}\right) dT \qquad (9.5)$$

To understand the second reason for the appearance of a thermal e.m.f., let us consider a homogeneous metal conductor along which there is a temperature gradient (Fig. 9.20). In this case, the concentration of electrons with $E > E_{\text{F}}$ at the warm end will be higher than at the cold one; conversely, the concentration of electrons with $E < E_{\text{F}}$ will be lower at the warm end. A gradient of concentration of electrons with a given value of the energy is set up along the conductor; this results in diffusion of the faster electrons to the cold end and of the slower electrons to the warm one. The diffusion flux of the fast electrons will be greater than the flux of the slow electrons. Therefore, a surplus of electrons will be formed near the cold end, and a shortage of them near the warm end. This leads to the setting up of a diffusion component of the thermal e.m.f.

The third reason for the appearance of a thermal e.m.f. is the carrying along of electrons by phonons. When there is a temperature gradient along a conductor, a drift of phonons is set up. Upon colliding with the electrons, the phonons impart to them directed motion from the warmer end of the conductor to the colder one. The result is accumulation of electrons on the cold end and a shortage of electrons on the warm end; this leads to the appearance of a "phonon" component of the thermal e.m.f.

Both processes—the diffusion of electrons and the carrying along of electrons by phonons—lead to the formation of a surplus of electrons near the cold end of the conductor and a shortage of them near the warm end. The result is the setting up of an electric field inside the conductor directed toward the temperature gradient. At a definite value of the field, which, generally speaking, is different for each

cross section of the conductor, the sum of the diffusion and phonon fluxes of the electrons becomes equal to zero and, consequently, a steady state sets in. The strength of this field can be represented in the form

$$E^* = -\frac{d\varphi}{dl} = -\frac{d\varphi}{dT}\frac{dT}{dl} = -\beta\frac{dT}{dl} \tag{9.6}$$

where

$$\beta = \frac{d\varphi}{dT} \tag{9.7}$$

Expression (9.6) relates the field strength E^* and the temperature gradient dT/dl. The field set up and the temperature gradient have opposite directions. Hence, E^* and dT/dl have opposite signs. Consequently, for metals $\beta > 0$*.

The process of the appearance of the field E^* in a non-uniformly heated conductor described above also occurs in semiconductors. For n-type semiconductors, we have $\beta > 0$. With hole conduction, the holes diffusing in a great number toward the cold end set up an excess positive charge near it. Carrying along of the holes by phonons leads to the same result. Therefore in p-type semiconductors, the potential of the cold end will be higher than that of the warm one, and, consequently, $\beta < 0$.

The field determined by Eq. (9.6) is one of extraneous forces. By integrating the strength of this field over the section of circuit A from junction *2* to junction *1*, we shall get the thermal e.m.f. acting on this section** in the direction indicated by the arrow in Fig. 9.19:

$$\mathscr{E}_{2A1} = -\int_2^1 \beta_A \frac{dT}{dl}\,dl = \int_{T_1}^{T_2} \beta_A\,dT \tag{9.8}$$

(we have exchanged the places of the integration limits). Similarly, the thermal e.m.f. acting on section B from junction *1* to junction *2* is

$$\mathscr{E}_{1B2} = -\int_1^2 \beta_B \frac{dT}{dl}\,dl = -\int_{T_1}^{T_2} \beta_B\,dT \tag{9.9}$$

The thermal e.m.f. $\mathscr{E}_{\text{therm}}$ consists of the e.m.f.s set up in the contacts, and the e.m.f.s acting on sections A and B:

$$\mathscr{E}_{\text{therm}} = \mathscr{E}_{\text{cont}} + \mathscr{E}_{2A1} + \mathscr{E}_{1B2}$$

---

* This holds for the overwhelming majority of metals. In some metals, however (beryllium, zinc, etc.), the conduction is of a hole nature. The sign of the Hall potential difference for these metals corresponds to positive current carriers. For such metals $\beta < 0$.

** See formula (5.15) of Vol. II, p. 103.

Introducing Eqs. (9.5), (9.8), and (9.9) and performing simple transformations, we get

$$\mathscr{E}_{\text{therm}} = \int_{T_1}^{T_2} \left(\beta_A - \frac{1}{e}\frac{dE_{F,A}}{dT}\right) dT - \int_{T_1}^{T_2} \left(\beta_B - \frac{1}{e}\frac{dE_{F,B}}{dT}\right) dT$$

The quantity

$$\alpha = \beta - \frac{1}{e}\frac{dE_F}{dT} \tag{9.10}$$

is known as the **thermoelectric coefficient**. Since both β and dE_F/dT depend on the temperature, the coefficient α is a function of T.

With a view to Eq. (9.10), the expression for the thermal e.m.f. can be written in the form

$$\mathscr{E}_{\text{therm}} = \int_{T_1}^{T_2} \alpha_A \, dT - \int_{T_1}^{T_2} \alpha_B \, dT \tag{9.11}$$

or

$$\mathscr{E}_{\text{therm}} = \int_{T_1}^{T_2} \alpha_{AB} \, dT \tag{9.12}$$

where

$$\alpha_{AB} = \alpha_A - \alpha_B \tag{9.13}$$

The quantity given by Eq. (9.13) is called the **differential** or **specific thermal electromotive force** of a given pair of metals or semiconductors. For most pairs of metals, α_{AB} is of the order of 10^{-5} to 10^{-4} V/K; for semiconductors it may be much higher (up to 1.5×10^{-3} V/K). The explanation is that for semiconductors with different kinds of conduction α has different signs, owing to which $|\alpha_{AB}| = |\alpha_A| + |\alpha_B|$.

In some cases, the specific thermal e.m.f. only slightly depends on the temperature. Therefore, formula (9.12) can be written approximately in the form

$$\mathscr{E}_{\text{therm}} = \alpha_{AB}(T_2 - T_1) \tag{9.14}$$

As a rule, however, with an increase in the difference between the junction temperatures, $\mathscr{E}_{\text{therm}}$ varies not according to a linear law, but in a quite complicated way, and may even change its sign. For example, if one junction of the pair iron-copper is kept at 0 °C, then at a temperature of the other junction of about 540 °C the thermal e.m.f. vanishes; at a lower temperature of the second junction $\mathscr{E}_{\text{therm}}$ has one sign, and at a higher temperature the opposite sign.

The Seebeck effect is taken advantage of to measure temperatures. The relevant device is called a **thermocouple**. One junction of a thermocouple is kept at a constant temperature (for example, at

0 °C), and the other is placed in the medium whose temperature is to be measured. The value of the temperature can be assessed according to the thermal current produced, which is measured with a galvanometer. A more accurate result is obtained if the produced thermal e.m.f. is measured according to the compensation method. A thermocouple is first graduated. Thermocouples can be employed to measure both low and high temperatures with an accuracy of the order of hundredths of a kelvin.

Thermocouples made from metals and their alloys are not used as current sources owing to their very low efficiency (not over 0.5%). Thermocouples made from semiconductor materials have a much higher efficiency (of the order of 10%). They have already found use as small generators for powering radio apparatus. Generators developing a power of hundreds of kilowatts are on the drawing board at present.

The Peltier Effect. This effect, discovered in 1834 by the French physicist Jean Peltier (1785-1845), consists in that when a current flows through a circuit formed of different metals or semiconductors, heat is liberated in some junctions and absorbed in others. Thus, the Peltier effect is the reverse of the Seebeck effect.

It was established experimentally that the amount of heat liberated or absorbed in a junction is proportional to the charge q passing through the junction:

$$Q_{AB} = \Pi_{AB} q = \Pi_{AB} It \tag{9.15}$$

(the subscripts indicate that the current flows from side A to side B). The proportionality constant Π_{AB} is known as the **Peltier coefficient.**

Equation (9.15) shows that unlike the Joule-Lenz heat, the Peltier heat is proportional to the first power of the current, and not to its square.

When the direction of the current changes, Q changes its sign, i.e. instead of liberation (absorption) of heat, the absorption (liberation) of the same amount of heat is observed (at the same q). Hence,

$$\Pi_{AB} = -\Pi_{BA}$$

It follows from the laws of thermodynamics that the Peltier coefficient and the specific thermal e.m.f. are related by the expression

$$\Pi_{AB} = \alpha_{AB} T \tag{9.16}$$

When two substances with the same kind of current carrier are in contact (metal-metal, metal-n-type semiconductor, two n-type semiconductors, two p-type semiconductors), the Peltier effect is explained as follows. The current carriers (electrons or holes) at different sides of a junction have a different average energy (we have in mind the total energy—kinetic plus potential). If the carriers when passing through a junction get into a region with a lower energy, they give

up their excess energy to the crystal lattice, and the resultis heating of the junction. At another junction, the carriers pass into a region with a higher energy; they borrow the lacking energy from the lattice, which results in cooling of the junction.

When two semiconductors with different kinds of conduction are in contact, the Peltier effect has a different explanation. In this case, the electrons and holes at one junction move toward one another. Upon meeting, they recombine: an electron in the conduction band of the *n*-semiconductor after getting into the *p*-semiconductor occupies the place of a hole in the valence band. When this occurs, the energy is released that is needed for the formation of a free electron in the *n*-semiconductor and of a hole in the *p*-semiconductor; the kinetic energy of the electron and the hole is also released. All this energy is transferred to the crystal lattice and heats the junction. At the other junction, the current flowing through it draws off the electrons and the holes from the boundary between the semiconductors. The decrease in the number of current carriers in the boundary region is replenished as a result of the birth of electron and hole pairs (here an electron from the valence band of the *p*-semiconductor passes over into the conduction band of the *n*-semiconductor). The energy used for the formation of a pair is borrowed from the lattice—the junction cools.

The Soviet physicist Abram Joffe (1880-1960) presented the idea of using the Peltier effect for designing refrigerating installations. The working element of such arrangements is a battery of alternating *n*- and *p*-type semiconductors. The junctions of one kind (corresponding, for example, to a transition from *n* to *p*) are introduced into the space to be cooled, and of the other kind (corresponding to a transition from *p* to *n*) are led out. With an appropriate direction of the current, the internal junctions absorb heat and lower the temperature of the space surrounding them, while the external junctions give up heat to the surroundings.

The Thomson Effect. In 1856, the British physicist William Thomson (Lord Kelvin) predicted on the basis of thermodynamic considerations that a heat similar to the Peltier heat should be liberated (or absorbed) when a current flows through a homogeneous conductor along which there is a temperature gradient. This effect was later discovered experimentally and was called the **Thomson effect.**

The amount of heat liberated as a result of the Thomson effect in unit time in a conductor element of length dl is

$$dQ = \tau I \frac{dT}{dl} dl \tag{9.17}$$

Here I = current
dT/dl = temperature gradient
τ = proportionality constant known as the **Thomson coefficient.**

The Thomson effect is explained by analogy with the Peltier effect. Assume that a current is flowing in the direction of increasing temperature. If the current carriers are electrons, they will pass during their motion from places having a higher temperature (and, consequently, a higher average energy) to places having a lower temperature (and a lower average energy). The electrons will give up their excess energy to the lattice, which will result in the liberation of heat. If holes are the current carriers, then the effect will have the opposite sign.

9.5. Semiconductor Diodes and Triodes

Currents can be rectified and voltages and powers amplified with the aid of semiconductor devices known as semiconductor (or crystal) diodes and triodes. Semiconductor triodes are also called **transistors.**

The main element of semiconductor devices is the so-called *p-n* **junction.** It is a thin layer on the boundary between two regions of the same crystal differing in the kind of impurity conduction. Such a junction is manufactured by taking, for example, a monocrystal of very pure germanium with an electronic conduction mechanism (due to negligible residues of impurities). A thin plate is cut out of the crystal and a small piece of indium is fused into it at one side. During this operation, which is conducted in a vacuum or in an atmosphere of an inert gas, the indium atoms diffuse into the germanium to a certain depth. In the region into which the indium atoms penetrate, the conductance of the germanium becomes of the hole type. A *p-n* junction appears at the boundary of this region. There are also other ways of obtaining *p-n* junctions.

Figure 9.21 shows how the impurity concentration changes in a direction at right angles to the boundary layer. The majority current carriers in the *p*-region are the holes formed as a result of the capture of electrons by impurity atoms; the acceptors become negative ions (Fig. 9.22; the circles are ions, the black dots are electrons, and the white dots are holes). In addition, the *p*-region contains a small number of minority carriers—electrons appearing owing to the transfer of electrons from the valence band directly into the conduction band by thermal motion (this process also increases the number of holes somewhat). In the *n*-region, the majority current carriers are the electrons given up by the donors into the conduction band (the donors themselves transform into positive ions); the transition of electrons from the valence band to the conduction band occurring at the expense of thermal motion leads to the formation of a small number of holes, minority carriers for this region.

Diffusing in opposite directions through the boundary layer, the holes and electrons recombine with one another. Therefore, the *p-n*

junction is greatly depleted of current carriers and acquires a high resistance. At the same time, an electrical double layer appears on the boundary between the regions. It is formed by the negative ions of the acceptor impurity whose charge is now no longer compensated

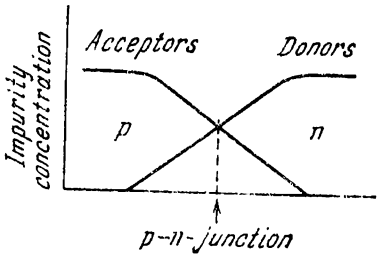


Fig. 9.21

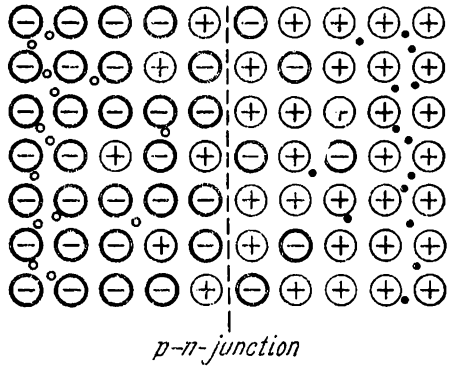


Fig. 9.22

by the holes, and by the positive ions of the donor impurity whose charge is now no longer compensated by the electrons (see Fig. 9.22). The electric field in this layer is directed so that it counteracts the further transition of the majority carriers through the layer. Equilibrium sets in at such a height of the potential barrier at which the Fermi levels of the two regions are at the same height (Fig. 9.23).

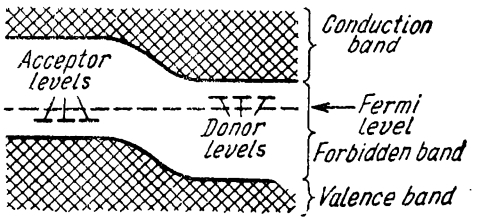


Fig. 9.23

The bending of the energy bands in the region of the junction is due to the fact that the potential of the p-region in the state of equilibrium is lower than that of the n-region; accordingly, the potential energy of an electron in the p-region is higher than in the n-region. The bottom boundary of the valence band gives the change in the potential energy of an electron $E_{p,e}$ in a direction at right angles to the junction (see the solid curve in Fig. 9.24*a*). The charge of a hole is opposite to that of an electron, therefore its potential energy $E_{p,h}$ is higher where $E_{p,e}$ is lower, and vice versa (see the dash curve in Fig. 9.24*a*).

In the state of equilibrium, a certain number of majority carriers succeed in surmounting the potential barrier, and as a result the small current I_{maj} flows through the junction (Fig. 9.24*a*). This

current is compensated by the counter current I_{min} set up by the minority carriers. There are very few minority carriers, but they easily penetrate through the boundary of the regions, "rolling down" from the potential barrier. The value of I_{min} is determined by the number of minority carriers given birth to every second and does not virtually depend on the height of the potential barrier. The value of I_{maj}, on the contrary, depends greatly on the height of the barrier. Equilibrium sets in exactly at such a height of the potential barrier at which both currents I_{maj} and I_{min} compensate each other.

Let us supply to a crystal an external voltage* directed so that the plus is connected to the p-region and the minus to the n-region

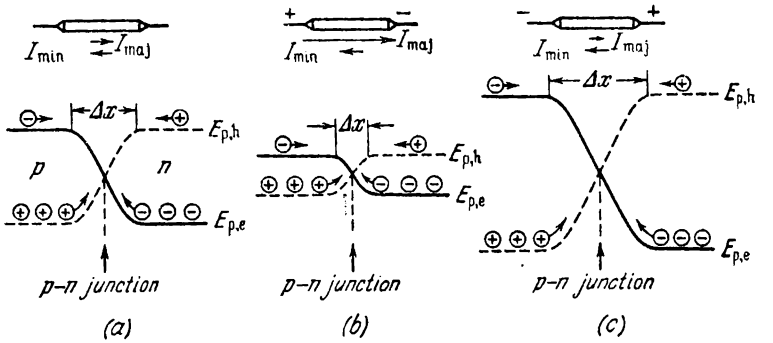


Fig. 9.24

(such a voltage is called **forward**). The result is a growth in the potential (i.e. an increase in $E_{p,h}$ and a decrease in $E_{p,e}$) of the p-region and a lowering of the potential (i.e. a decrease in $E_{p,h}$ and an increase in $E_{p,e}$) of the n-region (Fig. 9.24*b*). As a result, the height of the potential barrier will diminish, and the current I_{maj} will grow. The current I_{min}, however, will remain virtually unchanged (we have already noted that it is almost independent of the barrier height). Hence, the resultant current will become different from zero. The lowering of the potential barrier is proportional to the applied voltage (it equals eU). With lowering of the barrier height, the current of the majority carriers and, consequently, the resultant current, rapidly grow. Thus, in the direction from the p-region to the n-region, a p-n junction passes a current that rapidly grows with an increase in the applied voltage. This direction is called the forward one.

Figure 9.25 gives a volt-ampere characteristic of a p-n junction. The electric field set up in a crystal when a forward voltage is applied

* An external voltage violates equilibrium so that the Fermi levels of both regions become displaced relative to each other. With a forward voltage, the Fermi level in the p-region is lower than in the n-region.

"forces" the majority carriers toward the boundary between the regions, owing to which the width of the transition layer depleted of carriers (this layer is also known as the depletion layer—see Δx in Fig. 9.24) diminishes. The resistance of the junction falls off accordingly, the greater, the higher is the voltage. Hence, the volt-ampere characteristic in the transition region is not a straight line (see the right-hand branch of the curve in Fig. 9.25).

Now let us apply to a crystal a voltage of a direction such that the plus is connected to the *n*-region and the minus to the *p*-region (this voltage is called a **reverse** one). This will result in elevation

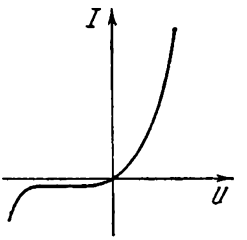


Fig. 9.25

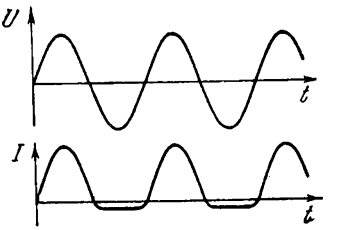


Fig. 9.26

of the potential barrier and a corresponding reduction in the current of the majority carriers I_{maj} (Fig. 9.24*c*). The resultant current set up (called the reverse one) rapidly reaches saturation (i.e. stops depending on U) and becomes equal to I_{min}. Thus, in the direction from the *n*-region to the *p*-region (called the reverse or cut-off direction), a *p-n* junction passes a weak current completely due to the minority carriers. Only at a very high reverse voltage does the current begin to grow very sharply, which is the result of electrical breakdown of the junction (see the left-hand branch of the curve in Fig. 9.25). Every *p-n* junction is characterized by its extreme value of the reverse voltage that it can withstand without breakdown.

Examination of Fig. 9.25 shows that a *p-n* junction has a much higher resistance in the reverse direction than in the forward one. The explanation is that the field produced in a crystal when a reverse voltage is imposed on it "pulls back" the majority carriers from the boundary between the regions, which leads to a growth in the width of the depletion layer. The resistance of the junction grows accordingly.

The different resistances in the forward and in the reverse direction make it possible to use *p-n* junctions for rectifying an alternating current. Figure 9.26 shows a graph of the current flowing through a junction when the applied voltage varies harmonically. In this case, the width of the depletion layer and the resistance of the junction pulsate, changing in step with the changes in the voltage.

A semiconductor triode or **transistor** is a crystal with two *p-n* junctions. Depending on the sequence in which the regions with different kinds of conduction alternate, *n-p-n* and *p-n-p* transistors are distinguished. The middle part of a transistor is called its **base.** The regions adjoining the base at both sides and having a different kind of conduction than it form the **emitter** and the **collector.**

Let us consider briefly the principle of operation of an *n-p-n* type transistor. Figure 9.27 shows how such a transistor is connected to an amplifier circuit. A constant forward bias voltage U_{em} is fed to the emitter-base junction, and a constant reverse bias voltage U_{col} is fed to the base-collector junction. The alternating voltage U_{in} being

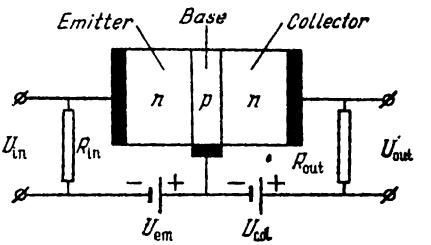


Fig. 9.27

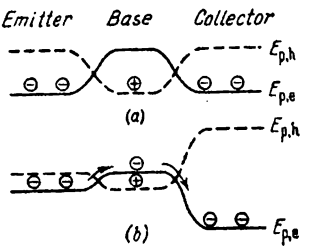


Fig. 9.28

amplified is fed to the small input resistor R_{in}. The amplified voltage U_{out} is taken off the output resistor R_{out}. With the signs of the bias voltages indicated in the diagram, the resistance of the emitter-base junction is not high, whereas the resistance of the base-collector junction, on the contrary, is very high. This makes it possible to use a resistor R_{out} having a high resistance.

Figure 9.28*a* shows the change in the potential energy of the electrons $E_{p,e}$ (the solid curve) and of the holes $E_{p,h}$ (the dash curve) when the bias voltages and input signal are absent. The connection of a forward voltage U_{em} lowers the potential barrier at the first junction, while the connection of a reverse voltage U_{col} elevates the potential barrier at the second junction (Fig. 9.28*b*). Flowing of a current in the emitter circuit is attended by the penetration of electrons into the region of the base. The electrons that have penetrated into the base diffuse toward the collector. If the thickness of the base is small, almost all the electrons, without managing to recombine, "roll down" the potential hill at the boundary between the base and the collector, and enter the collector circuit.

The change in the current I_{em} in the emitter circuit produced by the input voltage leads to a change in the number of electrons penetrating into the collector and, consequently, to almost the same change in the current I_{col} in the collector circuit. Assume that $I_{col} \approx \approx I_{em}$. Expressing these currents through the relevant voltages and

resistances, we find that $U_{in}/R_{in} \approx U_{out}/R_{out}$. Hence,

$$\frac{U_{out}}{U_{in}} \approx \frac{R_{out}}{R_{in}}$$

Since $R_{out} \gg R_{in}$, the voltage U_{out} considerably exceeds the input voltage U_{in}. Thus, a transistor amplifies the voltage and power. The increased power taken from the device appears at the expense of the current source connected to the collector circuit.

The operating principle of a *p-n-p* type transistor is similar to that described above for a type *n-p-n* transistor. The only difference is that the part of the electrons is played by the holes.

9.6. The Barrier-Layer Photoelectric Effect

Apart from the **extrinsic photoelectric effect** (usually called simply the photoelectric effect) treated in Sec. 2.2, there is also an **intrinsic photoelectric effect** observed in dielectrics and semiconductors. It consists in the redistribution of the electrons among the energy levels due to the action of light. If the energy of a quantum $\hbar\omega$ exceeds the width of the forbidden band, an electron that has absorbed a quantum passes from the valence band to the conduction band. The result is the appearance of an additional pair of current carriers—an electron and a hole, which manifests itself in an increase in the electrical conductance of the substance. If the latter contains impurities, the action of light may cause electrons to pass from the valence band onto levels of the impurity or from the impurity levels to the conduction band. In the first case, hole, and in the second, electron **photoconduction** appears.

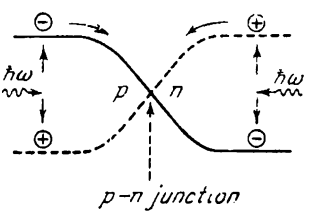


Fig. 9.29

The intrinsic photoelectric effect underlies the functioning of **photoresistors.** The number of current carriers formed is proportional to the incident light flux. This is why photoresistors are used for photometric purposes. Photoresistors made from cadmium sulphide (CdS) are used in the visible part of the spectrum. Photoresistors made from the semiconductors PbS, PbSe, PbTe, and InSb are used as detectors of infrared radiation.

The barrier-layer photoelectric effect is observed in the region of a *p-n* junction or on the boundary of a metal with a semiconductor. It consists in the setting up by light of an electromotive force (photo-e.m.f.). Figure 9.29 shows the change in the potential energy of the electrons (solid curve) and holes (dash curve) in the region of a *p-n* junction. The carriers that are the minority ones for the given

region (electrons in the *p*-region and holes in the *n*-region) produced under the action of light penetrate through the junction without hindrance. The result is the accumulation of an excess positive charge in the *p*-region and of an excess negative charge in the *n*-region. This leads to the appearance of a voltage applied to the junction that is exactly the **photoelectromotive force.**

If we connect a crystal with a *p-n* junction to an external load, a photocurrent will flow in it. When the illumination is not very great, this current is proportional to the light flux falling on the crystal. This underlies the operation of photoelectric photometers, in particular of the exposure meters used in photography. Several scores of series-connected silicon *p-n* junctions form a **solar battery.** Such batteries are used for supplying power to radio equipment on spaceships and on satellites of the Earth.

PART IV PHYSICS OF THE ATOMIC NUCLEUS AND ELEMENTARY PARTICLES

CHAPTER 10 THE ATOMIC NUCLEUS

10.1. Composition and Characteristic of the Atomic Nucleus

The nucleus of the simplest atom—the hydrogen atom—consists of a single elementary particle called a **proton.** The nuclei of all the other atoms consist of two kinds of elementary particles—**protons** and **neutrons.** These particles are known as **nucleons.**

The Proton. The proton (p) has a charge of $+e$ and a mass of

$$m_p = 938.26 \text{ MeV}^* \tag{10.1}$$

We shall indicate for comparison that the mass of an electron is

$$m_e = 0.511 \text{ MeV} \tag{10.2}$$

A comparison of Eqs. (10.1) and (10.2) shows that $m_p = 1836 m_e$.

A proton has a spin equal to one-half ($s = 1/2$), and an intrinsic magnetic moment

$$\mu_p = +2.79 \mu_{nuc} \tag{10.3}$$

where

$$\mu_{nuc} = \frac{e\hbar}{2m_p c} = 5.05 \times 10^{-24} \text{ erg/Gs} \tag{10.4}$$

is a unit of magnetic moment called the **nuclear magneton.** It follows from a comparison with Eq. (5.42) that μ_{nuc} is 1/1836-th of the Bohr magneton μ_B. Hence, the intrinsic magnetic moment of a proton is about 1/660-th of the magnetic moment of an electron.

The Neutron. The neutron (n) was discovered in 1932 by the British physicist James Chadwick (born 1891). The electric charge of this

---

* In nuclear physics, it is customary practice to express masses in units of energy, multiplying them by c^2 for this purpose. A unit of mass called the **atomic mass unit** is also employed (see Sec. 10.2 of Vol. I, p. 267); 1 amu = = 931.44 MeV.

particle is zero, and its mass

$$m_n = 939.55 \text{ MeV} \tag{10.5}$$

is very close to that of a proton. The difference between the masses of a neutron and a proton $m_n - m_p$ is 1.3 MeV, i.e. $2.5m_e$.

A neutron has a spin equal to one-half ($s = 1/2$) and (notwithstanding the absence of an electric charge) an intrinsic magnetic moment equal to

$$\mu_n = -1.91\mu_{nuc} \tag{10.6}$$

(the minus sign indicates that the directions of the intrinsic mechanical angular momentum and of the magnetic moment are opposite). This astonishing fact will be explained in Sec. 10.4.

We must note that the ratio of the experimental values of μ_p and μ_n with a high degree of accuracy equals $-\frac{3}{2}$. This was noted only after such a value had been obtained theoretically.

In the free state, a neutron is unstable (radioactive)—it spontaneously decays transforming into a proton and emitting an electron (e^-) and another particle called an **antineutrino** ($\tilde{\nu}$) (see Sec. 11.8). The half-life (i.e. the time during which half of the original number of neutrons decays) is about 12 minutes. The decay scheme can be written as follows:

$$n \to p + e^- + \tilde{\nu} \tag{10.7}$$

The mass of an antineutrino is zero*. The mass of a neutron is greater than that of a proton by $2.5m_e$. Hence, the mass of a neutron exceeds the total mass of the particles in the right-hand side of Eq. (10.7) by $1.5m_e$, i.e. by 0.77 MeV. This energy is liberated when a neutron decays as the kinetic energy of the particles formed.

Characteristics of an Atomic Nucleus. One of the most important characteristics of an atomic nucleus is its **charge number** (or **proton number**) Z. It equals the number of protons in the nucleus and determines its charge, which equals $+Ze$. The number Z determines the serial number of a chemical element in Mendeleev's periodic table. It is therefore also known as the **atomic number** of a nucleus.

The number of nucleons (i.e. the total number of protons and neutrons) in a nucleus is designated by the letter A and is called the **mass number** of the nucleus. The number of neutrons in a nucleus is $N = A - Z$.

Nuclei are designated by the symbol

$$_ZX^A$$

where X stands for the chemical symbol of a given element. The right-hand superscript is the mass number, and the left-hand sub-

* Here and in the following, by mass we understand the invariant quantity, i.e. the rest mass.

script is the atomic number (the latter symbol is often omitted). The mass number is sometimes written at the left of the symbol of a chemical element ($^{A}_{Z}X$) instead of at its right.

Nuclei having identical values of Z but different ones of A are called **isotopes.** Most chemical elements have several stable isotopes. For example, oxygen has three stable isotopes $_8O^{16}$, $_8O^{17}$, $_8O^{18}$, and tin has ten.

Hydrogen has three isotopes:

$_1H^1$—ordinary hydrogen, or protium ($Z = 1$, $N = 0$),

$_1H^2$—heavy hydrogen, or deuterium ($Z = 1$, $N = 1$), and

$_1H^3$—tritium ($Z = 1$, $N = 2$)*.

Protium and deuterium are stable, and tritium is radioactive.

Nuclei having the same mass number A are called **isobars.** We can cite $_{18}Ar^{40}$ and $_{20}Ca^{40}$ as an example. Nuclei having the same number of neutrons $N = A - Z$ are known as **isotones** ($_6C^{13}$, $_7N^{14}$). Finally, there are radioactive nuclei having identical Z's and A's, but differing in their half-lives. They are called **isomers.** For example, we know of two isomers of the nucleus $_{35}Br^{80}$, the half-life of one of which is 18 minutes, and of the other, 4.4 hours.

About 1500 nuclei are known differing either in Z, or in A, or in both together. About one-fifth of them are stable, the remaining ones are radioactive. Many nuclei were obtained artificially with the aid of nuclear reactions.

In nature, elements with an atomic number Z from 1 to 92, excluding technetium (Tc, $Z = 43$) and promethium (Pm, $Z = 61$) are encountered. Plutonium (Pu, $Z = 94$), after being obtained artificially, was detected in minute amounts in the natural mineral pitchblende. The other **transuranium** elements (i.e. elements after uranium), with Z from 93 to 107, were obtained artificially by means of various nuclear reactions.

The transuranium elements curium (96 Cm), einsteinium (99 Es), fermium (100 Fm) and mendelevium (101 Md) were named in honour of the outstanding scientists Pierre and Marie Curie, Albert Einstein, Enrico Fermi and Dmitri Mendeleev. Lawrencium (103 Lw) was named in honour of the inventor of the cyclotron Ernest Lawrence. Kurchatovium (104 Ku) was named in honour of the outstanding Soviet physicist Igor Kurchatov.

Some of the transuranium elements, including kurchatovium and the elements 106 and 107, were obtained in the Laboratory of Nuclear Reactions at the Joint Institute for Nuclear Research at Dubna by the Soviet scientist Georgi Flerov and his collaborators.

Dimensions of Nuclei. In a first approximation, a nucleus may be considered as a sphere whose radius is determined quite accurately

---

* Deuterium is also denoted by the symbol D, and tritium by the symbol T.

by the formula

$$r = 1.3 \times 10^{-13} A^{1/3} \text{ cm} = 1.3 A^{1/3} \text{ Fm} \tag{10.8}$$

(Fm—fermi—is the name of a length unit employed in nuclear physics and equal to 10^{-13} cm). It follows from Eq. (10.8) that the volume of a nucleus is proportional to the number of nucleons in it. Thus, the density of matter in all nuclei is approximately the same.

Spin of a Nucleus. The spins of the nucleons are summated into the resultant spin of the nucleus. The spin of a nucleon is $\frac{1}{2}$. Therefore, the nuclear spin quantum number I will be a half-integer with an odd number of nucleons A and an integer or zero with an even number of nucleons A. The spins I of nuclei do not exceed several units. This points to the fact that the spins of most of the nucleons in a nucleus mutually compensate one another, being antiparallel. In all even-even nuclei (i.e. nuclei with an even number of protons and an even number of neutrons), the spin is zero.

The mechanical angular momentum of a nucleus $\mathbf{M}_I$ is added to the momentum of the electron shell $\mathbf{M}_J$ to form the total angular momentum of an atom $\mathbf{M}_F$ that is determined by the quantum number F.

The interaction of the magnetic moments of the electrons and the nucleus leads to the fact that the states of an atom corresponding to different mutual orientations of $\mathbf{M}_I$ and $\mathbf{M}_J$ (i.e. to different F's) have a slightly differing energy. The interaction of the moments μ_L and μ_S is responsible for the fine structure of spectra (see Sec. 5.4). The interaction of μ_I and μ_J determines the **hyperfine structure** of atomic spectra. The splitting of the spectral lines corresponding to the hyperfine structure is so small (of the order of several hundredths of an angstrom) that it can be observed only with the aid of instruments having the highest possible resolving power.

10.2. Mass and Binding Energy of a Nucleus

The mass of a nucleus m_{nuc} is always smaller than the sum of the masses of the particles it consists of. The reason is that when nucleons combine to form a nucleus, the binding energy of the nucleons is liberated.

The rest energy of a particle is associated with its mass by the relation $E_0 = mc^2$ [see Eq. (8.40) of Vol. I, p. 241, and the beginning of the first paragraph on p. 242]. Hence, the energy of a nucleus at rest is less than the total energy of the non-interacting nucleons at rest by the amount

$$E_b = c^2 \{[Zm_p + (A - Z) m_n] - m_{\text{nuc}}\} \tag{10.9}$$

It is exactly this quantity that is the **binding energy** of the nucleons in a nucleus. It equals the work that must be done to separate the nucleons forming the nucleus and to remove them from one another to distances virtually excluding their interaction.

Equation (10.9) is practically not violated if we substitute the mass of a hydrogen atom m_H for the mass of a proton, and the mass of an atom m_a for that of its nucleus m_{nuc}. Indeed, if we disregard the comparatively negligible binding energy of the electrons to the nuclei, this substitution will signify the addition to the minuend and the subtrahend in braces of an identical quantity equal to Zm_e. Thus, Eq. (10.9) can be written in the form

$$E_b = c^2\{[Zm_H + (A-Z)\,m_n] - m_a\} \qquad (10.10)$$

The latter equation is more convenient than Eq. (10.9) because usually the masses of atoms, and not of nuclei, are tabulated.

The binding energy per nucleon, i.e. E_b/A, is sometimes called the **binding fraction.**

The quantity

$$\Delta = [Zm_p + (A-Z)\,m_n] - m_{nuc} \qquad (10.11)$$

is known as the **mass defect** of a nucleus*. The mass defect is associated with the binding energy by the relation $\Delta = E_b/c^2$.

Let us calculate the binding energy of the nucleons in the nucleus ${}_2He^4$ that includes two protons ($Z = 2$) and two neutrons ($A - Z = 2$). The mass of the atom ${}_2He^4$ is 4.002 60 amu, which 3728.0 MeV correspond to. The mass of a hydrogen atom ${}_1H^1$ is 1.008 15 amu [938.7 MeV; compare with Eq. (10.1)]. The mass of a neutron is 939.55 MeV [see Eq. (10.5)]. Using these values in Eq. (10.10), we get

$$E_b = (2 \times 938.7 + 2 \times 939.55) - 3728.0 = 28.5 \text{ MeV}$$

The binding energy of a helium nucleus per nucleon is 7.1 MeV. We shall indicate for comparison that the binding energy of the valence electrons in atoms has a value that is $1/10^6$ of this one (of the order of 10 eV). The binding energy per nucleon (E_b/A) for other nuclei has approximately the same value as for helium.

Figure 10.1 depicts a graph showing how the binding energy per nucleon (the binding fraction) E_b/A depends on the mass number A. The nucleons are bound most stongly in nuclei having mass numbers of the order of 50-60 (i.e. for the elements from Cr to Zn). The binding energy for these nuclei reaches 8.7 MeV/nucleon. A growth in A is

* The mass defect was originally defined as the difference between the numerical value of the mass of an atom m_a expressed in atomic mass units and the mass number A:

$$\Delta = m_a - A$$

This quantity has a less clear physical meaning than that determined by Eq. (10.11).

attended by gradual diminishing of the binding energy per nucleon; for the heaviest natural element—uranium—it is 7.5 MeV/nucleon. Such a dependence of the binding energy per nucleon on the mass number makes two processes possible from the energy viewpoint: (1) the fission of heavy nuclei into several lighter ones, and (2) the fusion of light nuclei into a single nucleus. Both processes should be attended by the liberation of a great amount of energy. For example, the fission of one nucleus with the mass number $A = 240$ (the binding

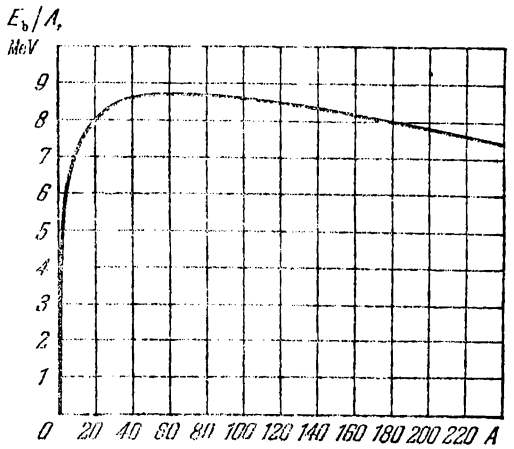


Fig. 10.1

energy per nucleon is 7.5 MeV) into two nuclei with mass numbers of $A = 120$ (the binding energy per nucleon is 8.5 MeV) would result in the liberation of energy amounting to 240 MeV. The fusion of two nuclei of heavy hydrogen ${}_1H^2$ into a helium nucleus ${}_2He^4$ would result in the liberation of energy equal to 24 MeV. We shall indicate for comparison that when one atom of carbon combines with two atoms of oxygen (the combustion of coal to CO_2), energy of the order of 5 eV is liberated.

Nuclei with values of the mass number A ranging from 50 to 60 are the most profitable from an energy viewpoint. In this connection, the question appears: why are nuclei with other values of A stable? The answer is as follows. To divide into several parts, a heavy nucleus must pass through a number of intermediate states whose energy exceeds that of the ground state of the nucleus. Hence, the nucleus needs additional energy (the **activation energy**) for the fission process. This energy is then returned, being added to the energy liberated upon fission as a result of a change in the binding energy. In ordinary conditions, a nucleus does not have where to take the activation energy from, and as a result heavy nuclei do not undergo

spontaneous fission. The activation energy can be communicated to a heavy nucleus by an additional neutron that it captures. The process of fission of uranium or plutonium nuclei under the action of the neutrons captured by the nuclei underlies the operation of nuclear reactors and the conventional atomic bomb.

As regards light nuclei, for fusion into a single nucleus they must approach one another to a very close distance ($\sim 10^{-13}$ cm). Such approaching of the nuclei is prevented by the Coulomb repulsion between them. To overcome this repulsion, the nuclei must travel with enormous speeds corresponding to temperatures of the order of several hundred millions of kelvins. For this reason, the process of fusion of light nuclei is called a **thermonuclear reaction.** Thermonuclear reactions proceed in the interior of the Sun and stars. In the conditions of the Earth, uncontrolled thermonuclear reactionns were meanwhile accomplished in the explosions of hydrogen bombs. Scientists of a number of countries are persistently working on the finding of ways of carrying out controllable thermonuclear fusion. Soviet physicists occupy one of the leading places in this field.

10.3. Models of the Atomic Nucleus

Attempts to construct a theory of the nucleus are confronted by two serious difficulties: (1) the inadequacy of ou knowledge of the forces acting between nucleons, and (2) the exceedingly great cumbersomeness of the quantum problem of many bodies (a nucleus with the mass number A is a system of A bodies). These difficulties make it necessary to follow the path of creating nuclear models. The latter permit us to describe a definite collection of properties of a nucleus with the aid of comparatively simple mathematical means. None of such models can give an exhaustive description of a nucleus. Therefore, several models have to be used, each of which describes its own collection of the properties of a nucleus and its own circle of phenomena. Each model contains arbitrary parameters whose values are chosen so as to obtain agreement with experimental results.

It is impossible to describe all the models of a nucleus that exist within the scope of a general course of physics. We are forced to restrict ourselves to a brief narration only about two of them—the liquid-drop and the shell models.

The Liquid-Drop Model. This model was proposed by the Soviet physicist Yakov Frenkel in 1939 and was then developed by the Danish physicist Niels Bohr and other scientists. Frenkel gave attention to the similarity between an atomic nucleus and a liquid drop, consisting in that in both cases the forces acting between the constituent particles—molecules in the liquid and nucleons in the nucleus—are short-range ones. In addition, the virtually identical density of

the matter in various nuclei points to the extremely low compressibility of the nuclear matter. Liquids have a compressibility that is just as low. This circumstance gives us grounds to consider a nucleus to be similar to a charged drop of a liquid.

The liquid-drop model made it possible to derive a semi-empirical formula for the binding energy of the particles in a nucleus. This model also assisted in explaining many other phenomena, in particular the process of fission of heavy nuclei.

The Shell Model. The nuclear shell model was developed by the German physicist Maria Goeppert-Mayer and other scientists. In this model, the nucleons are considered to move independently of one another in an averaged centrally symmetrical field. Accordingly, there are discrete energy levels (like the levels of an atom) filled with nucleons with account taken of the Pauli principle (we remind our reader that the spin of nucleons is 1/2). These levels are grouped into **shells**, each of which can contain a definite number of nucleons. A completely filled shell is an especially stable formation.

In accordance with experimental data, those nuclei are especially stable in which the number of protons, or the number of neutrons (or both these numbers) is

$$2, \quad 8, \quad 20, \quad 28, \quad 50, \quad 82, \quad 126$$

These numbers were named **magic.** Nuclei in which the number of protons Z or the number of neutrons N is magic (i.e. especially stable nuclei) are also called **magic.** Nuclei in which both Z and N are magic are called **doubly magic.**

Altogether five doubly magic nuclei are known:

${}_2He^4$ ($Z = 2$, $N = 2$), ${}_8O^{16}$ ($Z = 8$, $N = 8$),

${}_{20}Ca^{40}$ ($Z = 20$, $N = 20$), ${}_{20}Ca^{48}$ ($Z = 20$, $N = 28$),

${}_{82}Pb^{208}$ ($Z = 82$, $N = 126$)

These nuclei are especially stable. In particular, the especial stability of the helium nucleus ${}_2He^4$ manifests itself in that it is the only composite particle emitted by heavy nuclei in radioactive decay (it is called an alpha-particle).

10.4. Nuclear Forces

The tremendous binding energy of the nucleons in a nucleus indicates that there is very intensive interaction between nucleons. This interaction has the nature of attraction. It keeps the nucleons at distances of $\sim 10^{-13}$ cm from one another notwithstanding the strong Coulomb repulsion between protons. The nuclear interaction between nucleons has been named **strong interaction.** It can be described with the aid of a field of nuclear forces. Let us list the distinguishing features of these forces.

1. Nuclear forces are **short-range** ones. Their radius of action is of the order of 10^{-13} cm. At distances appreciably smaller than 10^{-13} cm, the attraction of nucleons is replaced by repulsion.

2. Strong interaction does not depend on the charge of nucleons. The nuclear forces acting between two protons, between a proton and a neutron, or between two neutrons, have the same magnitude. This property is called the **charge independence** of nuclear forces.

3. Nuclear forces depend on the mutual orientation of the spins of the nucleons. For example, a neutron and a proton are kept together, forming a nucleus of heavy hydrogen—a **deuteron** (or **deuton**) only if their spins are parallel to each other.

4. Nuclear forces are not central ones. They cannot be represented as directed along the straight line connecting the centres of the interacting nucleons. The non-central nature of nuclear forces follows, in particular, from the fact that they depend on the orientation of the nucleon spins.

5. Nuclear forces have the property of **saturation** (this signifies that each nucleon in a nucleus interacts with a limited number of nucleons). Saturation manifests itself in that the binding energy per nucleon does not grow with an increase in the number of nucleons, but remains approximately constant. In addition, the saturation of the nuclear forces is also indicated by the volume of a nucleus being proportional to the number of nucleons forming it [see Eq. (10.8)].

According to modern notions, strong interaction is due to the fact that nucleons virtually exchange particles that have been called **mesons.** To understand the essence of this process, let us first consider what electromagnetic interaction looks like from the point of view of quantum electrodynamics.

Charged particles interact via an electromagnetic field. We know that this field can be represented as a collection of photons. As quantum electrodynamics indicates, a process of interaction between two charged particles, for example, electrons, consists in an exchange of photons. Each particle sets up a field around itself that continuously emits and absorbs photons. The action of the field on the other particle manifests itself in its absorbing one of the photons emitted by the first particle. Such a description of the interaction cannot be understood literally. The photons by means of which the interaction is carried out are not ordinary real photons, but **virtual** ones. In quantum mechanics, the name virtual is applied to particles that cannot be detected during their lifetime. In this sense, virtual particles can be called imaginary ones.

To better understand the meaning of the term "virtual", let us consider an electron at rest. The process of its setting up a field in the surrounding space can be represented by the equation

$$e^- \rightleftarrows e^- + \hbar\omega \qquad (10.12)$$

The total energy of a photon and electron is greater than that of an electron at rest. Consequently, the transformation described by expression (10.12) is attended by violation of the law of energy conservation. For a virtual photon, however, this violation is seeming. According to quantum mechanics, the energy of a state existing during the time Δt is determinate only with the accuracy ΔE satisfying the uncertainty relation

$$\Delta E \cdot \Delta t \sim \hbar \tag{10.13}$$

[see formula (4.5)]. It follows from this expression that the energy of a system can experience the deviations ΔE whose duration Δt must not exceed the value determined by condition (10.13). Consequently, if a virtual photon emitted by an electron is absorbed by this or another electron before the time $\Delta t = \hbar/\varepsilon$ elapses ($\varepsilon = \hbar\omega$), then no violation of the law of energy conservation can be detected.

When additional energy is imparted to an electron (this may happen, for instance, if it collides with another electron), a real photon may be emitted instead of a virtual one, and it can exist for an unlimitedly long time.

During the time $\Delta t = \hbar/\varepsilon$ determined by condition (10.13), a virtual photon can transmit interaction between points separated by the distance

$$l = c\Delta t = c\,\frac{\hbar}{\varepsilon}$$

The energy of a photon $\varepsilon = \hbar\omega$ can be as small as desired (the frequency ω varies from 0 to ∞). Tnerefore, the radius of action of electromagnetic forces is unlimited. If the particles exchanged by interacting electrons had a mass m other than zero, then the radius of action of the corresponding forces would be limited by the quantity

$$r = c\Delta t_{\max} = c\,\frac{\hbar}{\varepsilon_{\min}} = c\,\frac{\hbar}{mc^2} = \frac{\hbar}{mc} = \lambda\!\!\!^{-}_{C}$$

where $\lambda\!\!\!^{-}_{C}$ is the Compton wavelength of a given particle [see Eq. (2.24)]. We have assumed that the particle which is the carrier of the interaction is moving with the speed c.

In 1934, the Soviet physicist Igor Tamm (1895-1971) advanced the assumption that the interaction between nucleons is also transmitted by means of virtual particles. At that time, only the photon, electron, positron, and neutrino were known in addition to nucleons. The heaviest of these particles—the electron—has a Compton wavelength $\lambda\!\!\!^{-}_{C} = 3.86 \times 10^{-11}$ cm [see Eq. (2.25)] that exceeds the radius of action of nuclear forces by two orders of magnitude. Moreover, the magnitude of the forces that could be due to virtual electrons, as

calculations have shown, is exceedingly low. Thus, the first attempt to explain nuclear forces with the aid of an exchange of virtual particles was unsuccessful.

In 1935, the Japanese physicist Hideki Yukawa (born 1907) advanced the bold hypothesis that particles having a mass from 200 to 300 times that of an electron exist in nature, although not yet detected, and that it is exactly these particles that play the part of carriers of nuclear interaction, in the same way as photons are carriers of electromagnetic interaction. Yukawa called these hypothetic particles heavy photons. In connection with the fact that as regards the magnitude of their mass these particles occupy an intermediate position between electrons and nucleons, they were later called mesons (the Greek word "mesos" means middle).

In 1936 the American physicists C. Anderson and S. Neddermeyer detected particles with a mass of $207m_e$ in cosmic rays. It was initially assumed that these particles, called **μ-mesons** or **muons**, are the carriers of interaction predicted by Yukawa. It was later established, however, that muons interact very weakly with nucleons, so that they cannot be responsible for nuclear interactions. Only in 1947 did C. Lattes, G. Occhialini, and C. Powell discover another kind of meson in cosmic radiation—the so-called **π-mesons**, or **pions**, which were found to be the carriers of nuclear forces predicted 12 years earlier by Yukawa.

There are positive (π^+), negative (π^-), and neutral (π^0) mesons. The charge of π^+- and π^--mesons equals the elementary charge e. The mass of charged pions is the same and equals $273m_e$ (140 MeV), the mass of a π^0-meson is $264m_e$ (135 MeV). The spin of both charged and of the neutral π-meson is zero ($s = 0$). All three particles are unstable. The lifetime of π^+- and π^--mesons is 2.60×10^{-8} s, and of a π^0-meson is 0.8×10^{-16} s.

The overwhelming part of charged π-mesons decay according to the scheme

$$\pi^+ \rightarrow \mu^+ + \nu, \qquad \pi^- \rightarrow \mu^- + \tilde{\nu} \tag{10.14}$$

(μ^+ and μ^- are a positive and a negative muon, respectively, ν is a neutrino, and $\tilde{\nu}$ an antineutrino). On an average, 2.5 decays in a million proceed according to other schemes (for example, $\pi \rightarrow e + \nu$; $\pi \rightarrow \pi^0 + e + \nu$, etc.; when π^+ decays, e^+, i.e. a positron, is formed, and when π^- decays, e^-, i.e. an electron, is formed).

On an average, 98.8 per cent of π^0-mesons decay into two gamma quanta:

$$\pi^0 \rightarrow \gamma + \gamma \tag{10.15}$$

The remaining 1.2 per cent of the decays follow the schemes

$$\pi^0 \rightarrow e^+ + e^- + \gamma; \quad \pi^0 \rightarrow e^+ + e^- + e^+ + e^-;$$
$$\pi^0 \rightarrow \gamma + \gamma + \gamma$$

The particles called μ-mesons or muons belong to the class of **leptons** (see Sec. 11.1), and not mesons. We shall therefore call them muons in the following. Muons have a positive (μ^+) or a negative (π^μ) charge equal to the elementary charge e (no neutral muons exist). The mass of a muon is $207m_e$ (106 MeV), its spin is one-half ($s = 1/2$). Muons, like π-mesons, are not stable. They decay according to the scheme

$$\mu^+ \to e^+ + \nu + \hat{\nu}, \qquad \mu^- \to e^- + \nu + \tilde{\nu} \tag{10.16}$$

The lifetime of both muons is the same and equals 2.2×10^{-6} s.

Let us now consider the exchange interaction between nucleons. As a result of the virtual processes

$$p \rightleftarrows n + \pi^+ \tag{10.17}$$

$$n \rightleftarrows p + \pi^- \tag{10.18}$$

$$p \rightleftarrows p + \pi^0; \quad n \rightleftarrows n + \pi^0 \tag{10.19}$$

a nucleon is surrounded by a cloud of virtual π-mesons forming a field of nuclear forces. The absorption of these mesons by another

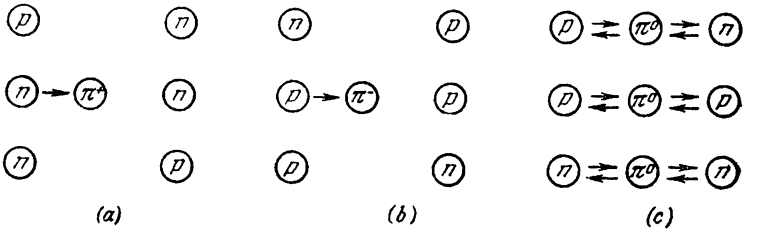


Fig. 10.2

nucleon leads to strong interaction between nucleons according to one of the following schemes:

$$(1) \quad p + n \rightleftarrows n + \pi^+ + n \rightleftarrows n + p$$

A proton emits a virtual π^+-meson and transforms into a neutron. The meson is absorbed by a neutron that, as a result, transforms into a proton. Next, the same process occurs in the reverse direction (Fig. 10.2*a*). Each of the interacting nucleons spends part of its time in the charged state and part in the neutral one.

$$(2) \quad n + p \rightleftarrows p + \pi^- + p \rightleftarrows p + n$$

A neutron and a proton exchange π^--mesons (Fig. 10.2*b*).

$$(3) \quad p + n \rightleftarrows p + \pi^0 + n \rightleftarrows p + n$$

$$p + p \rightleftarrows p + \pi^0 + p \rightleftarrows p + p$$

$$n + n \rightleftarrows n + \pi^0 + n \rightleftarrows n + n$$

The nucleons exchange π^0-mesons (Fig. 10.2*c*).

The first of the three processes described above is confirmed experimentally in the scattering of neutrons on protons. When a beam of neutrons passes through hydrogen, protons appear in the beam. Many of them have the same energy and direction of motion as the incident neutrons. A corresponding number of neutrons practically at rest is detected in the target. It is absolutely improbable that such a large number of neutrons completely transmitted their momentum to the protons previously at rest as a result of head-on collisions. It therefore becomes necessary to acknowledge that part of the neutrons flying near protons capture one of the virtual π^+ mesons. The result is the conversion of a neutron into a proton, while the proton that has lost its charge transforms into a neutron (Fig. 10.3).

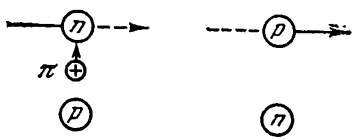


Fig. 10.3

If energy equivalent to the mass of a π-meson is communicated to a nucleon, then the virtual π-meson can become real. The required energy can be communicated upon the collision of sufficiently accelerated nucleons (or nuclei), or when a nucleon absorbs a gamma-quantum. At very high energies of the colliding particles, several real π-mesons may appear.

Now we are in a position to explain the existence of a magnetic moment of a neutron and the anomalous value of the magnetic moment of a proton (see Sec. 10.1). In accordance with process (10.18), a neutron spends part of its time in the virtual state $(p + \pi^-)$. The orbital motion of the π^--meson leads to the setting up of the negative magnetic moment observed in the neutron. The anomalous magnetic moment of a proton ($2.79\mu_{nuc}$ instead of one nuclear magneton) can also be explained by the orbital motion of a π^+-meson during the time interval when the proton is in the virtual state $(n + \pi^+)$.

10.5. Radioactivity

Radioactivity is defined as the spontaneous transformation of atomic nuclei into other ones attended by the emission of elementary particles. Only unstable nuclei undergo such transformations. Radioactive processes include (1) alpha decay, (2) beta decay (including the capture of an electron), (3) gamma radiation of nuclei, (4) spontaneous fission of heavy nuclei, and (5) proton radioactivity.

Radioactivity observed in nuclei existing in natural conditions is called **natural**. The radioactivity of nuclei obtained as a result of nuclear reactions is called **artificial**. There is no difference of principle between artificial and natural radioactivity. The process of radioactive transformation in both cases obeys the same laws.

Law of Radioactive Transformation. Individual radioactive nuclei transform independently of one another. We may therefore consider that the number of nuclei dN decaying during the small time interval dt is proportional both to the number of available nuclei N and to the time interval dt:

$$dN = -\lambda N\, dt \tag{10.20}$$

Here λ is a constant characteristic of a given radioactive substance and known as the **decay constant**. The minus sign has been taken to allow us to consider dN as an increment of the number of undecayed nuclei N.

Integration of Eq. (10.20) leads to the expression

$$N = N_0 e^{-\lambda t} \tag{10.21}$$

where N_0 is the number of nuclei at the initial moment, and N is the number of undecayed atoms at the moment t. Equation (10.21) expresses the law of radioactive transformation. This law is very simple: *the number of undecayed nuclei diminishes with time exponentially.*

The number of nuclei decaying during the time t is determined by the expression

$$N_0 - N = N_0 (1 - e^{-\lambda t}) \tag{10.22}$$

The time during which a half of the initial number of nuclei decays is called the **half-life** T. It is determined by the condition

$$\frac{1}{2} N_0 = N_0 e^{-\lambda T}$$

whence

$$T = \frac{\ln 2}{\lambda} = \frac{0.693}{\lambda} \tag{10.23}$$

The half-life for the radioactive nuclei known at present ranges from 3×10^{-7} s to 5×10^{15} years.

Let us find the average lifetime of a radioactive nucleus. The number of nuclei $dN(t)$ transforming during the time from t to $t + dt$ is determined by the magnitude of Eq. (10.20): $dN(t) = \lambda N(t)\, dt$. The lifetime of each of these nuclei is t. Hence, the sum of the lifetimes of all the N_0 initial nuclei is obtained by integration of the expression $t\, dN(t)$. Dividing this sum by the number of nuclei N_0, we get the **average lifetime** τ of a radioactive nucleus:

$$\tau = \frac{1}{N_0} \int_0^\infty t\, dN(t) = \frac{1}{N_0} \cdot \int_0^\infty t\lambda N(t)\, dt$$

Let us introduce into this equation expression (10.21) for $N(t)$:

$$\tau = \frac{1}{N_0} \int_0^\infty t\lambda N_0 e^{-\lambda' t}\, dt = \int_0^\infty t\lambda e^{-\lambda t}\, dt = \frac{1}{\lambda}$$

(it is necessary to pass over to the variable $x = \lambda t$ and integrate by parts). Thus, the average lifetime is a quantity that is the reciprocal of the decay constant λ:

$$\tau = \frac{1}{\lambda} \tag{10.24}$$

A comparison with Eq. (10.23) shows that the half-life T differs from τ by a numerical factor of ln 2.

It often happens that the nuclei appearing as a result of radioactive transformation are also radioactive and decay at a rate characterized by the decay constant λ'. The new decay products may again be radioactive, and so on. The result is a whole series of radioactive transformations. Three radioactive series (or families) exist in nature, whose parents (first members) are U^{238} (the uranium series), Th^{232} (the thorium series), and U^{235} (the actinouranium series). The final products in all three series are lead isotopes—in the first one Pb^{206}, in the second Pb^{208}, and, finally, in the third Pb^{207}.

Natural radioactivity was discovered in 1896 by the French scientist Antoine Henri Becquerel (1852-1908). A great contribution to the studying of radioactive substances was made by the French scientists Pierre Curie and Marie Sklodowska-Curie. Three kinds of radiation were found to exist. The first kind, called **alpha rays**, deflects under the action of a magnetic field in the same direction in which a stream of positively charged particles would. The second kind, called **beta rays**, deflects under the action of a magnetic field in the opposite direction, i.e. like a stream of negatively charged particles. Finally, the third kind of radiation, that shows no reaction at all to a magnetic field, was called **gamma rays**. It was later found that gamma rays are electromagnetic radiation of a very short wavelength (from 10^{-3} to 1 Å).

Alpha Decay. Alpha rays are a flux of helium ${}_2He^4$ nuclei. Decay proceeds according to the following scheme:

$${}_zX^A \rightarrow {}_{z-2}Y^{A-4} + {}_2He^4 \tag{10.25}$$

The letter X stands for the chemical symbol of the decaying (parent) nucleus, and the letter Y for the chemical symbol of the new (daughter) nucleus formed. Alpha decay is usually attended by the daughter nucleus emitting gamma rays. It can be seen from the decay scheme that the atomic number of the daughter substance is less by 2 and the mass number less by 4 than the relevant quantities of the parent substance. An example is the decay of the uranium isotope U^{238}

proceeding with the formation of thorium:

$$_{92}U^{238} \rightarrow {}_{90}Th^{234} + {}_{2}He^{4}$$

The velocities with which alpha particles (i.e. nuclei of $_2He^4$) fly out from decaying nuclei are very high ($\sim 10^9$ cm/s; the kinetic energy is of the order of several MeV). When flying through a substance, an alpha particle gradually loses its energy, using it to ionize the molecules of a substance, and finally stops. An average of 35 eV is needed to form one pair of ions in air. Thus, an alpha particle forms about 10^5 pairs of ions along its path. It is natural that the greater the density of a substance, the shorter is the path of an alpha particle before it stops. Thus, its path is several centimetres in air at normal pressure, and is of the order of 10^{-3} cm in a solid (alpha particles are completely retained by an ordinary sheet of paper).

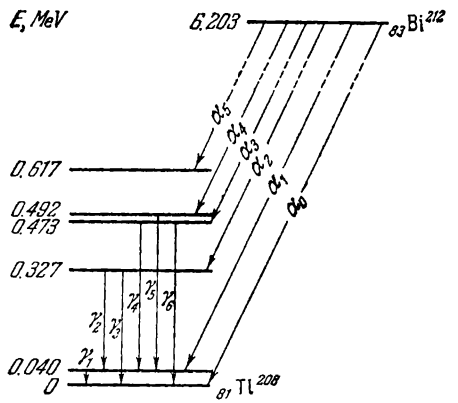


Fig. 10.4

The kinetic energy of an alpha particle is produced at the expense of the excess rest energy of the parent nucleus in comparison with the total rest energy of the daughter nucleus and the alpha particle. This excess energy is distributed between the alpha particle and the daughter nucleus in a ratio that is inversely proportional to their masses*. The energies (velocities) of the alpha particles emitted by a given radioactive substance are strictly definite. In the majority of cases, a radioactive substance emits several groups of alpha particles having close but different energies. This is due to the fact that the daughter nucleus may appear not only in the normal, but also in the excited state. Figure 10.4 explains schematically the appearance of different groups of alpha particles (the appearance of the fine structure of an alpha spectrum) emitted upon the decay of the nuclei $_{83}Bi^{212}$ (bismuth-212). The energy levels of the daughter nucleus $_{81}Tl^{208}$ (thallium-208) are shown at the left of the diagram. The energy of the ground state has been taken as zero. The excess rest energy of the parent nucleus above the rest energy of an alpha particle and the daughter nucleus in the normal state is 6.203 MeV. If the daughter nucleus appears in the unexcited state, all this energy

* The velocities which alpha particles fly out with are of the order of 0.1*c*. We may therefore use the classical expressions for the momentum and kinetic energy of a particle.

is liberated in the form of kinetic energy, the share of the alpha particle being

$$E_k = 6.203 \times \frac{208}{212} = 6.086 \text{ MeV}$$

(this group of particles is denoted by α_0 in the diagram). If the daughter nucleus appears in the fifth excited state whose energy exceeds that of the normal state by 0.617 MeV, then the liberated kinetic energy will be 6.203 — 0.617 = 5.586 MeV, and 5.481 MeV will fall to the share of the alpha particle (the group of particles α_5). The relative number of particles is about 27% for α_0, about 70% for α_1, and only about 0.01% for α_5. The relative numbers of α_2, α_3, and α_4 are also very small (of the order of 0.1 to 1%).

The average lifetime τ of the excited states for most nuclei ranges from 10^{-8} to 10^{-15} s*. During an average time of τ, the daughter nucleus passes over to the normal or to a lower excited state, emitting a gamma photon. Figure 10.4 shows the appearance of gamma photons of six different energies.

The excitation energy of the daughter nucleus can also be separated in other ways. An excited nucleus may emit a particle: a proton, neutron, electron, or an alpha particle. Finally, the excited nucleus formed as a result of alpha decay can give up its excess energy directly (without first emitting a gamma quantum) to one of the electrons of the *K*-, *L*-, or even the *M*-shell of the atom, as a result of which an electron flies out of the latter. This process is called **internal conversion**. The vacancy appearing as a result of the ejection of an electron will be filled by electrons from the higher energy levels. Consequently, internal conversion is always attended by the emission of characteristic X-rays.

In the same way as a photon "ready for use" does not exist inside an atom and appears only at the moment of its emission, an alpha particle also appears at the moment of radioactive decay of a nucleus. In leaving a nucleus, an alpha particle has to surmount a potential barrier whose height exceeds the total energy of an alpha particle equal on an average to 6 MeV (Fig. 10.5). The outer side of the barrier falling asymptotically to zero is due to Coulomb repulsion of the alpha particle and the daughter nucleus. The inner side of the barrier is due to nuclear forces. Experiments involving the scattering of alpha particles by heavy alpha-radioactive nuclei have shown that the height of the barrier noticeably exceeds the energy of the alpha particles flying out in decay. Classical notions say that a particle cannot surmount the potential barrier in these conditions. According to quantum mechanics, however, there is a probability other than zero that the particle will penetrate through

* It may sometimes be very great (up to several years). Levels with such a lifetime are called **isomeric**, and the excited nucleus, an **isomer**.

the barrier, passing, as it were, through a tunnel in the latter. We treated this phenomenon, called the tunnel effect, in Sec. 4.9. The theory of alpha decay based on the notion of the tunnel effect leads to results that well agree with experimental data.

Beta Decay. There are three varieties of beta decay. In one of them, the parent nucleus emits an electron, in another, a positron, and in the third variety, called **electron capture** (*e*-**capture**), the

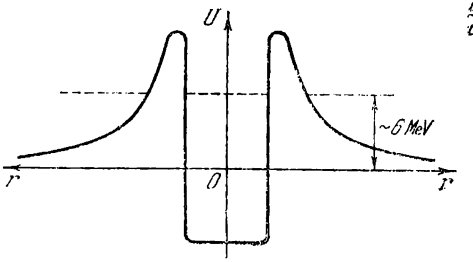


Fig. 10.5

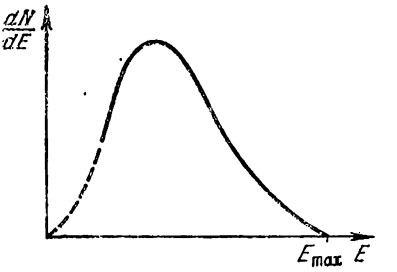


Fig. 10.6

nucleus absorbs one of the electrons of the *K*-shell, considerably more rarely of the *L*- or *M*-shell (accordingly, we speak of *K*-capture, *L*-capture, or *M*-capture instead of *e*-capture).

The first kind of beta decay (β^- **decay or electron decay**) follows the scheme

$$ {}_{Z}X^{A} \rightarrow {}_{Z+1}Y^{A} + {}_{-1}e^{0} + \tilde{\nu} \qquad (10.26) $$

To underline the conservation of the charge and number of nucleons in beta decay, we have assigned a charge number of $Z = -1$ and a mass number of $A = 0$ to a beta electron.

Inspection of scheme (10.26) reveals that the daughter nucleus has an atomic number that is greater by unity than that of the parent nucleus, while the mass numbers of both nuclei are the same. An antineutrino $\tilde{\nu}$ is emitted in addition to an electron. The entire process occurs as if one of the neutrons of the nucleus ${}_{Z}X^{A}$ transferred into a proton, changing according to scheme (10.7). In general, process (10.7) is a particular case of process (10.26). A free neutron is therefore said to be beta-radioactive.

Beta decay may be attended by the emission of gamma rays. The mechanism of their appearance is the same as in alpha decay—the daughter nucleus appears not only in the normal state, but also in excited ones. Next passing over to a state with a lower energy, the nucleus emits a gamma photon.

An example of β^- decay is the transformation of thorium Th^{234} into protactinium Pa^{234} with the emission of an electron and an

antineutrino:

$$_{90}Th^{234} \rightarrow {}_{91}Pa^{234} + {}_{-1}e^{0} + \tilde{\nu}$$

Unlike alpha particles, which have a strictly definite energy within the limits of each group, beta electrons have the most diverse kinetic energy from 0 to E_{max}. Figure 10.6 shows the energy spectrum of electrons emitted by nuclei in beta decay. The area covered by the curve gives the total number of electrons emitted in unit time, dN is the number of electrons whose energy is within the interval dE. The energy E_{max} corresponds to the difference between the mass of the parent nucleus and the masses of an electron and the daughter nucleus. Consequently, decays in which the energy of an electron E is lower than E_{max} occur with an apparent violation of the law of energy conservation.

To explain the vanishing of the energy $E_{max} - E$, the Swiss physicist Wolfgang Pauli (1900-1958) advanced the assumption in 1932 that in beta decay another particle is emitted apart from an electron, and this particle carries with it the energy $E_{max} - E$. Since this particle does not reveal itself in any way, it should be acknowledged that it is neutral and has a very small mass (it has been established at present that the rest mass of this particle is zero). According to the proposal of Enrico Fermi, this hypothetic particle was called a **neutrino*** (which means "tiny neutron").

There is another reason for the assumption on the existence of the neutrino (or antineutrino). The spin of a neutron, proton, and electron is the same and equals 1/2. If we write scheme (10.26) without an antineutrino, then the total spin of the appearing particles (which for two particles with $s = 1/2$ can be either zero or unity) will differ from that of the initial particle. Thus, the participation of another particle in beta decay is dictated by the law of angular momentum conservation, and a spin of 1/2 (or 3/2) must be ascribed to this particle. It has been established that the spin of a neutrino (and antineutrino) is 1/2. A direct experimental proof of the existence of neutrinos was obtained only in 1956.

Thus, the energy liberated in β^- decay is distributed between an electron and an antineutrino (or between a positron and a neutrino, see below) in the most diverse proportions.

The second kind of beta decay (β^+ **decay** or **positron decay**) proceeds according to the scheme

$$_{Z}X^{A} \rightarrow {}_{Z-1}Y^{A} + {}_{+1}e^{0} + \nu \qquad (10.27)$$

An example is the transformation of nitrogen N^{13} into carbon C^{13}:

$$_{7}N^{13} \rightarrow {}_{6}C^{13} + {}_{+1}e^{0} + \nu$$

* In accordance with the classification adopted at present, in β^- decay an antineutrino is emitted, and not a neutrino.

It can be seen from scheme (10.27) that the atomic number of the daughter nucleus is less by unity than that of the parent one. The process is attended by the emission of a positron e^+ [in formula (10.27) it is designated by the symbol ${}_{+1}e^0$] and a neutrino ν. The appearance of gamma rays is also possible. A positron is an electron's antiparticle. Consequently, both particles emitted in decay (10.27) are antiparticles with respect to the particles emitted in decay (10.26).

The process of β^+ decay occurs as if one of the protons of the parent nucleus transformed into a neutron, having emitted a positron and a neutrino:

$$p \rightarrow n + e^+ + \nu \qquad (10.28)$$

Such a process is impossible for a free proton from energy considerations because the mass of a proton is less than that of a neutron. A proton in a nucleus, however, can borrow the required energy from other nucleons in the nucleus.

The third kind of beta decay (**electron capture**) consists in that a nucleus absorbs one of the K-electrons (less often one of the L- or M-electrons) of its atom, and as a result one of the protons transforms into a neutron emitting a neutrino:

$$p + e^- \rightarrow n + \nu$$

The appearing nucleus may be in the excited state. Passing later into lower energy states, it emits gamma photons. The scheme of the process is as follows:

$${}_Z\mathrm{X}^A + {}_{-1}e^0 \rightarrow {}_{Z-1}\mathrm{Y}^A + \nu \qquad (10.29)$$

The site in the electron shell freed by the captured electron is filled by electrons from the higher layers. The result is the appearance of X-rays. Electron capture is easily detected by the X-radiation attending it. It is exactly in this way that the American physicist Luis Alvarez (born 1911) discovered K-capture in 1937.

An example of electron capture is the transformation of potassium K^{40} into argon Ar^{40}:

$${}_{19}\mathrm{K}^{40} + {}_{-1}e^0 \rightarrow {}_{18}\mathrm{Ar}^{40} + \nu$$

Spontaneous Fission of Heavy Nuclei. In 1940, the Soviet physicists Georgi Flerov and Konstantin Petrzhak discovered a process of the spontaneous fission of uranium nuclei into two approximately equal parts. Later, this phenomenon was also observed for many other heavy nuclei. Spontaneous fission in its characteristic features is close to stimulated fission, which will be treated in Sec. 10.7.

Proton Radioactivity. As follows from its name, in proton radioactivity, a nucleus transforms by emitting one or two protons (in the latter case we speak of double proton radioactivity). This kind

of radioactivity was first observed in 1963 by a group of Soviet physicists headed by Flerov.

Activity of a Radioactive Substance. The activity of a radioactive preparation is defined as the number of disintegrations occurring in it in unit time. If dN_{dis} nuclei decay during the time dt, then the activity is dN_{dis}/dt. According to Eq. (10.20),

$$dN_{dis} = |dN| = \lambda N\, dt$$

Hence it follows that the activity of a radioactive preparation equals λN, i.e. the product of the decay constant and the number of undecayed nuclei in the preparation.

The SI unit of activity is the disintegration per second (d/s). The use of non-system units such as the disintegration per minute (d/min) and the curie (Ci) is permitted. The unit of activity called the curie is defined as the activity of any substance in which 3.700×10^{10} atoms disintegrate per second. Submultiple (millicurie, microcurie, etc.) and multiple (kilocurie, megacurie) units are also used.

10.6. Nuclear Reactions

A nuclear reaction is defined as a process of strong interaction of an atomic nucleus with an elementary particle or with another nucleus resulting in the transformation of a nucleus (or nuclei). The reacting particles interact when they approach each other up to distances of the order of 10^{-13} cm owing to the action of nuclear forces.

The most widespread kind of nuclear reaction is the interaction of a light particle a with a nucleus X, the result being the formation of a light particle b and a nucleus Y:

$$X + a \rightarrow Y + b$$

The equation of such reactions is customarily written in the following abbreviated form:

$$X\ (a,\ b)\ Y \qquad (10.30)$$

The light particles participating in the reaction are indicated in parentheses, first the initial particle, then the final one.

The light particles a and b may be a neutron (n), proton (p), deuteron (d), alpha particle (α), and a gamma photon (γ).

Nuclear reactions may be attended either by the liberation or by the absorption of energy. The amount of liberated energy is called **the reaction energy**. It is determined by the difference between the masses (expressed in energy units) of the initial and final nuclei. If the sum of the masses of the nuclei formed exceeds the sum of the

masses of the initial nuclei, the reaction proceeds with the absorption of energy, and the reaction energy will be negative.

In 1936, Niels Bohr established that reactions initiated by not very fast particles proceed in two stages. The first one consists in the capture of the particle a approaching the nucleus X and in the formation of the compound nucleus C. The energy supplied by the particle a (it consists of the kinetic energy of the particle and the energy binding it to the nucleus) is redistributed in a very short time between all the nucleons of the compound nucleus, as a result of which this nucleus is in the excited state. In the second stage, the compound nucleus emits the particle b. Such a two-stage reaction is written as follows:

$$X + a \rightarrow C \rightarrow Y + b \qquad (10.31)$$

If the emitted particle is identical to the captured one ($b = a$), process (10.31) is called scattering. When the energy of the particle b equals that of the particle a (i.e. $E_b = E_a$), the scattering is **elastic**, otherwise (i.e. when $E_b \neq E_a$) it is **inelastic**. A nuclear reaction occurs if the particle b is not identical to a.

The time τ_{nuc} needed for a nucleon having an energy of the order of 1 MeV (which corresponds to a nucleon velocity $\sim 10^9$ cm/s) to cover a distance equal to the diameter of the nucleus ($\sim 10^{-13}$ cm) is called the **nuclear time** (or the **nuclear transit time**). The order of magnitude of this time is

$$\tau_{nuc} \sim \frac{10^{-13}\ \text{cm}}{10^9\ \text{cm/s}} = 10^{-21}\ \text{s} \qquad (10.32)$$

The average lifetime of a compound nucleus (equal to from 10^{-18} to 10^{-12} s) is many orders greater than the nuclear time τ_{nuc}. Hence, the decay of the compound nucleus (i.e. the emission of the particle b by it) is a process that does not depend on the first stage of the reaction consisting in the capture of the particle a (the compound nucleus "forgets", as it were, the way it was formed). The same compound nucleus may decay in different ways, the nature of these ways and their relative probability not depending on how the compound nucleus was formed.

Reactions caused by fast nucleons and deuterons proceed without the formation of a compound nucleus. Such reactions are known as **direct nuclear interactions**. A typical direct interaction reaction is **the stripping reaction** observed in non-central collisions of a deuteron with a nucleus. In such collisions, one of the deuteron nucleons may get into the zone of action of the nuclear forces and will be captured by the nucleus, whereas the other nucleon will remain outside the zone of action of the nuclear forces and will fly past the nucleus. This reaction can be represented symbolically in the form (d, p) or (d, n).

The reverse of a stripping reaction is a **pickup reaction**—a bombarding nucleon (n or p) splits off a nucleon (p or n) from a nucleus, transforming into a deuteron: (n, d) or (p, d).

It is customary practice in nuclear physics to characterize interaction with the aid of the effective cross section σ. The meaning of this quantity is as follows. Assume that a flux of particles, for example, neutrons, falls on a target so thin that the nuclei of the target do not overlap (Fig. 10.7). If the nuclei were rigid spheres with a cross section of σ and the falling particles were rigid spheres with a vanishingly small cross section, then the probability of a falling particle grazing one of the target nuclei would be

$$P = \sigma n \delta$$

where n is the concentration of the nuclei, i.e. their number in unit volume of the target, and δ is the thickness of the target ($\sigma n \delta$ determines the relative fraction of the target area covered by the nuclei-spheres).

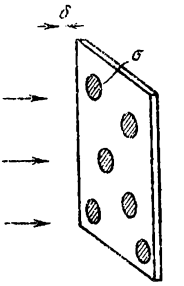


Fig. 10.7

Assume that a particle flux N* falls on a target at right angles to its surface. Hence, the number of particles colliding with the target nuclei in unit time ΔN is determined by the formula

$$\Delta N = NP = N\sigma n \delta \tag{10.33}$$

Consequently, having determined the relative number of particles experiencing collisions $\Delta N/N$, it would be possible to calculate the cross section $\sigma = \pi r^2$ of a nucleus by the formula

$$\sigma = \frac{\Delta N}{N n \delta} \tag{10.34}$$

Actually, neither the target nuclei nor the particles falling on the target are rigid spheres. By analogy with the model of colliding spheres, however, the probability of interaction is characterized by the quantity σ determined by Eq. (10.34) in which by ΔN is meant not the number of colliding particles, but the number of particles that have interacted with the target nuclei. It is exactly this quantity that is called the effective cross section for a given reaction (or process).

When the target is thick, the flux of particles will gradually weaken as it passes through it. After dividing the target into thin layers, let us write Eq. (10.33) for a layer of thickness dx at the depth x from the surface:

$$dN = -N(x)\, \sigma n\, dx$$

* We remind our reader that a flux of particles is defined as the number of particles flying through a surface in unit time.

where $N(x)$ is the particle flux at the depth x. We have written the minus sign to permit dN to be considered as an increment (and not as a decrement) of the flux along the path dx. Integration of this equation gives the expression

$$N(\delta) = N_0 e^{-\sigma n \delta}$$

in which N_0 is the initial flux, and $N(\delta)$ is the flux at the depth δ. Thus, by measuring the weakening of the particle flux when it passes through a target of thickness δ, we can find the interaction cross section by the formula

$$\sigma = \frac{1}{n\delta} \ln \frac{N_0}{N(\delta)} \tag{10.35}$$

It is usual practice to express the effective cross sections of nuclear processes in units called **barns:**

$$1 \text{ barn} = 10^{-24} \text{ cm}^2 \tag{10.36}$$

A nuclear reaction was carried out for the first time by Ernest Rutherford in 1919. In irradiating nitrogen with the alpha particles emitted by a radioactive source, some of the nitrogen nuclei transformed into oxygen nuclei, emitting a proton. The equation of this reaction has the form

$$_7N^{14}(\alpha, p)\,_8O^{17}$$

Rutherford used natural projectiles—alpha particles—for disintegration of an atomic nucleus. The first nuclear reaction induced by artificially accelerated particles was carried out by the British physicist John Cockcroft (1897-1967) and the Irish physicist Ernest Walton (born 1903) in 1932. Using the so-called voltage multiplier, they accelerated protons to an energy of the order of 0.8 MeV and observed the reaction

$$_3Li^7(p, \alpha)\,_2He^4$$

The development of the equipment and techniques used for accelerating charged particles was attended by multiplication of the number of nuclear transformations accomplished artificially.

Of the greatest significance are reactions induced by neutrons. Unlike charged particles (p, d, α), neutrons do not experience Coulomb repulsion, owing to which they can penetrate into nuclei while having a very low energy. The effective cross sections of reactions usually grow with a decrease in the energy of neutrons. This can be explained by the fact that the lower the velocity of a neutron, the greater is the time it spends in the sphere of action of the nuclear forces when flying near a nucleus and, consequently, the greater is the probability of its capture. Therefore, many effective cross sections vary like $1/v \propto E^{-1/2}$. Cases are often observed, however, when

the cross section of neutron capture has a sharply expressed maximum for neutrons of a definite energy E_r. Figure 10.8 shows as an example a curve giving the dependence of the cross section of neutron capture by a U^{238} nucleus on the energy of a neutron E. A logarithmic scale is used on both axes. In this case, the dependence $\sigma \propto E^{-1/2}$ is depicted by a straight line described by the equation $\ln \sigma = \text{const} - \frac{1}{2} \ln E$. A glance at the figure reveals that apart from the region of energies near 7 eV, the dependence of $\ln \sigma$ on $\ln E$ is indeed close to a linear one. When $E = E_r = 7$ eV, the capture cross section sharply grows and reaches 23 000 barns. The shape of the curve indicates that the phenomenon has a resonance nature. Such resonance absorption occurs when the energy brought by a neutron into a compound nucleus exactly equals the energy needed for transferring the compound nucleus to an excited energy level. Similarly, the probability of the absorption of photons whose energy equals the difference between the energies of the first excited level and the ground level of an atom is especially great (the resonance absorption of light).

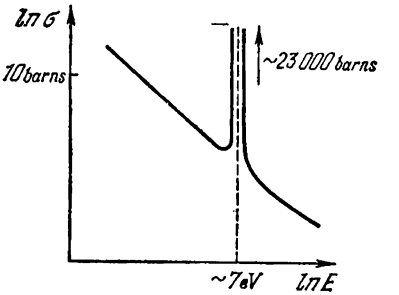


Fig. 10.8

Of interest is the reaction

$$_7N^{14}(n, p)\,_6C^{14}$$

that constantly proceeds in the atmosphere under the action of the neutrons formed by cosmic rays. The carbon isotope $_6C^{14}$ produced in this reaction is called radiocarbon because it is β^--radioactive; its half-life is 5730 years. Radiocarbon is assimilated by plants in photosynthesis and participates in the cycle of substances in nature.

The number of radiocarbon nuclei ΔN_+ produced in the atmosphere in unit time remains constant on an average. The number of decaying nuclei ΔN_- is proportional to the number of nuclei N present:

$$\Delta N_- = kN$$

Since the half-life is very great, an equilibrium concentration of the nuclei of C^{14} in ordinary carbon sets in that meets the condition

$$\Delta N_- = \Delta N_- \quad \text{or} \quad \Delta N_+ = kN$$

Special research has shown that owing to the action of winds and ocean currents, the equilibrium concentration of C^{14} at different places on the globe is the same and corresponds to about 14 disintegrations a minute per gramme of carbon.

In a living organism, the diminishing of C^{14} in it owing to radioactivity is replenished as a result of its participation in the cycle of substances in nature. When an organism dies, the process of assimilation immediately stops, and the concentration of C^{14} in ordinary carbon begins to diminish according to the law of radioactive decay. Consequently, by measuring the concentration of C^{14} in the remains of organisms (in wood, bones, etc.), we can determine the date when they died or, as is said, their age. Verification of this method using ancient specimens whose age was determined exactly by historical methods gave quite satisfactory results.

10.7. The Fission of Nuclei

In 1938, the German scientists Otto Hahn (1879-1968) and Fritz Strassmann (born 1902) discovered that the bombardment of uranium with neutrons results in the formation of elements from the middle of the periodic table—barium and lanthanum. This phenomenon was explained by the German scientists Otto Frisch (born 1904) and Lise Meitner (1878-1968). They put forward the assumption that an uranium nucleus which has captured a neutron decays into two approximately equal parts called **fission fragments**.

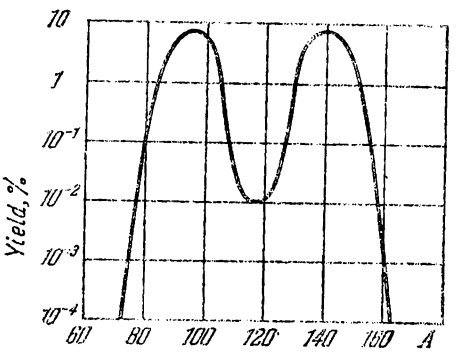


Fig. 10.9

Further investigations showed that fission may occur in different ways. A total of about 80 different fragments are formed, the most probable being fission into fragments the ratio of whose masses is 2 : 3. The curve shown in Fig. 10.9 gives the relative yield (in percent) of fragments of different mass obtained in the fission of U^{235} produced by slow (**thermal***) neutrons (a logarithmic scale is used along the axis of ordinates). Inspection of the curve reveals that the relative number of fission events when two fragments of equal mass are formed ($A \approx 117$) is 10^{-2} per cent, whereas the formation of fragments with mass numbers of the order of 95 and 140 ($95 : 140 \approx \approx 2 : 3$) is observed in 7 per cent of all cases.

The binding energy per nucleon for nuclei of a medium mass is greater than that of heavy nuclei by about 1 MeV (see Fig. 10.1).

* The name thermal is applied to neutrons that are in thermal equilibrium with the atoms of a substance. Their energy is about 0.03 eV.

Hence, it follows that the fission of nuclei must be attended by the liberation of a great amount of energy. But of special importance is the fact that several neutrons are freed upon the fission of every nucleus. The relative number of neutrons in heavy nuclei is appreciably greater than in medium-size ones. Consequently, the fragments formed are greatly overloaded with neutrons, and as a result they liberate several neutrons each. Most of the neutrons are emitted instantaneously (during a time less than $\sim 10^{-14}$ s). Part (about 0.75%) of the neutrons called **delayed neutrons** are emitted not instantaneously, but with a delay ranging from 0.05 s to 1 min. *On an average, 2.5 neutrons are released per fission event.*

The liberation of instantaneous and delayed neutrons does not completely eliminate the overloading of the fission fragments with neutrons. Therefore, the fragments are mainly radioactive and undergo a chain of β^- transformations attended by the emission of gamma rays. Let us explain what has been said above with an example. One way in which fission may proceed is as follows:

$$_{92}U^{235} + n \rightarrow {}_{55}Cs^{140} + {}_{37}Rb^{94} + 2n$$

The fission fragments—cesium and rubidium—undergo transformations:

$$_{55}Cs^{140} \rightarrow {}_{56}Ba^{140} \rightarrow {}_{57}La^{140} \rightarrow {}_{58}Ce^{140}$$

$$_{37}Rb^{94} \rightarrow {}_{38}Sr^{94} \rightarrow {}_{39}Y^{94} \rightarrow {}_{40}Zr^{94}$$

The products—cerium Ce^{140} and zirconium Zr^{94}—are stable.

In addition to uranium, bombardment by neutrons* causes the fission of thorium ($_{90}Th^{232}$) and protactinium ($_{91}Pa^{231}$), and also of the transuranium element plutonium ($_{94}Pu^{239}$). Neutrons of superhigh energies (of the order of several hundred MeV) produce the fission of lighter nuclei too. The fission of the nuclei of U^{235} and Pu^{239} is produced by neutrons of any energy, but especially well by slow neutrons. The fission of U^{233} and Th^{230} is also produced by thermal neutrons, but these isotopes are not encountered in nature, they are prepared artificially.

Only fast neutrons (with energies not lower than $\sim$1 MeV) can cause the fission of nuclei of U^{238}. At lower energies, the neutrons are absorbed by the nuclei of U^{238} without their following fission. The result is the formation of a nucleus of U^{239} whose excitation energy is liberated in the form of a gamma photon. This is why such a process is called **radioactive capture** [the reaction (n, γ)]. The effective cross section of this process sharply grows at an energy of the neutrons equal to $\sim$7 eV, reaching 23 000 barns (see Fig. 10.8).

* The fission of heavy nuclei may be produced not only by neutrons, but also by other particles—protons, deuterons, alpha particles, and also by gamma photons. In the latter case, we have to do with the photofission of nuclei.

The cross section of the capture by a nucleus of U^{238} of thermal neutrons is less than three barns.

The nucleus of U^{239} formed as a result of the capture of a neutron is not stable (its half-life T is 23 min). Emitting an electron, an antineutrino, and a gamma photon, it transforms into a nucleus of the transuranium element neptunium Np^{239}. Neptunium also experiences β^- decay ($T = 2.3$ days) and transforms into plutonium Pu^{239}. This chain of transformations can be written as follows:

$$_{92}U^{239} \xrightarrow{(25\ \text{min})} {}_{93}Np^{239} \xrightarrow{(2.3\ \text{days})} {}_{94}Pu^{239} \qquad (10.37)$$

Plutonium is alpha-radioactive, but its half-life is so great (24 400 years) that it may be considered virtually stable.

The radiative capture of neutrons by a nucleus of Th^{232} leads to the formation of the fissioning uranium isotope U^{233} that is absent in natural uranium:

$$_{90}Th^{232} + n \rightarrow {}_{10}Th^{233} \xrightarrow{(22\ \text{min})} {}_{91}Ac^{233} \xrightarrow{(27\ \text{days})} {}_{92}U^{233}$$

Uranium-233 is alpha-radioactive ($T = 162\,000$ years).

The emission of several neutrons in the fission of nuclei of U^{235}, Pu^{239}, and U^{233} makes it possible to achieve a **chain nuclear reaction.** Indeed, the z neutrons emitted in the fission of one nucleus may cause the fission of z nuclei; as a result z^2 new neutrons will be emitted, which will cause the fission of z^2 nuclei, and so on. Thus, the number of neutrons born in each generation grows in a geometrical progression. The neutrons emitted in the fission of nuclei of U^{235} have an average energy of ~ 2 MeV, which corresponds to a velocity of $\sim 2 \times 10^9$ cm/s. Therefore, the time that elapses between the emission of a neutron and its capture by a new fissioning nucleus is very small. Hence, the process of multiplication of neutrons in a fissioning substance goes on quite rapidly.

The picture we have drawn above is ideal. Neutrons would multiply as described above provided that all the liberated neutrons are absorbed by fissioning nuclei. Actually, matters are far from being so. First of all, owing to the finite dimensions of the fissioning body and the high penetrating ability of the neutrons, many of them will leave the reaction zone before they are captured by a nucleus and cause its fission. Moreover, part of the neutrons will be absorbed by non-fissioning impurities owing to which they drop out of the game without causing fission and, consequently, without giving birth to new neutrons.

The volume of a body grows as the cube and its surface as the square of its linear dimensions. Therefore, the relative part of the neutrons flying out of a body diminishes with an increasing mass of the fissioning substance.

Natural uranium contains 99.27% of the isotope U^{238}, 0.72% of U^{235} and about 0.01% of U^{234}. Hence, for each nucleus of U^{235} that fissions under the action of slow neutrons there are 140 nuclei of U^{238} that capture not too fast neutrons without fission. This is the reason why no chain fission reaction occurs in natural uranium.

A chain nuclear reaction can be achieved in uranium in one of two ways. The first consists in recovering the fissioning isotope U^{235} from natural uranium. Owing to the chemical indistinguishability of the isotopes, it is a very difficult task to separate them. It was solved, however, in several ways.

In a piece of pure U^{235} (or Pu^{239}) each neutron captured by a nucleus produces fission with the emission of $\sim$2.5 new neutrons. If the mass of such a piece, however, is less than a definite critical value, most of the emitted neutrons will fly out without producing fission, so that a chain reaction does not occur. At a mass greater than the critical one, the neutrons rapidly multiply, and the reaction acquires an explosive nature. The functioning of an **atomic bomb** is based on this principle. The nuclear charge of such a bomb is two or more pieces of almost pure U^{235} or Pu^{239} (they are denoted by the reference number 1 in Fig. 10.10). The mass of each piece is less than critical, as a result of which no chain reaction occurs.

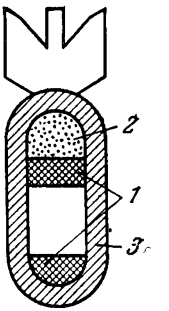


Fig. 10.10

The Earth's atmosphere always contains a certain number of neutrons produced by cosmic rays. Therefore, to call forth an explosion, it is sufficient to connect the parts of the nuclear charge into one piece with a mass greater than the critical one. This must be done very rapidly, and the pieces must be connected very tightly. Otherwise, the nuclear charge will fly apart into fragments before an appreciable portion of the fissioning substance has time to react. Ordinary explosive substance *2* (an igniter), by means of which one part of the nuclear charge is shot into the other, is used for connecting them. The entire device is confined in massive shell *3* of a high-density metal. The shell reflects neutrons and, in addition, prevents scattering of the nuclear charge until the maximum possible number of its nuclei have liberated their energy in fission. The chain reaction in an atomic bomb proceeds on fast neutrons. Only a part of the nuclear charge has time to react upon an explosion.

A different way of carrying out a chain reaction is used in **nuclear reactors.** The fissioning substance employed in reactors is natural uranium (or uranium enriched somewhat with the isotope U^{235}). To prevent radiative capture of neutrons by the nuclei of U^{238} (which becomes especially intensive at an energy of the neutrons of $\sim$7 eV), comparatively small rods of the fissioning substance are spaced at

a certain distance apart, and the spaces between these rods are filled with a moderator, i.e. a substance in which the neutrons are retarded to thermal velocities. The cross section of thermal neutron capture by a nucleus of U^{238} is only three barns, whereas the fission cross section of U^{235} by thermal neutrons is almost 200 times greater (580 barns). Therefore, although the neutrons collide with U^{238} nuclei 140 times more often than with U^{235} nuclei, radiative capture occurs less frequently than fission, and at large critical dimensions of the entire device, the neutron multiplication factor (i.e. the ratio of the number of neutrons born in each of two consecutive generations) may reach values greater than unity.

The neutrons are retarded at the expense of elastic scattering. In this case, the energy lost by a particle being retarded depends on the ratio of the masses of the colliding particles. The maximum amount of energy is lost when both particles have the same mass (see Sec. 3.11 of Vol. I, p. 104 et seq). From this viewpoint, a substance containing ordinary hydrogen, for example, water (the masses of a proton and a neutron are approximately the same) ought to be an ideal moderator. Such substances were found to be unsuitable as a moderator, however, because protons absorb neutrons, entering with them into the reaction

$$p\ (n,\ \gamma)\ d$$

The moderator nuclei must have a small cross section of neutron capture and a large cross section of elastic scattering. This condition is satisfied by a deuteron (a nucleus of heavy hydrogen—deuterium D), and also by nuclei of graphite (C) and beryllium (Be). About 25 collisions are sufficient to reduce the energy of a neutron from 2 MeV to thermal energies in heavy water (D_2O), and about 100 collisions in C or Be.

The first uranium-graphite reactor was started in December, 1942, at the University of Chicago under the supervision of the Italian physicist Enrico Fermi. In the Soviet Union, a reactor of the same kind was placed into operation under the supervision of Igor Kurchatov in December, 1946, in Moscow.

A schematic view of an uranium-graphite reactor is shown in Fig. 10.11. Reference number *1* designates the moderator—graphite; *2*—uranium elements; and *3*—rods containing cadmium or boron. These rods are used to control the process in the reactor. Cadmium and boron are intensive absorbers of neutrons. Therefore, when the rods are pushed into the reactor, the neutron multiplication factor diminishes, and when they are pulled out, this factor grows. A special automatic device controlling the rods makes it possible to maintain the power developed in the reactor at the preset level. Control is considerably facilitated by the circumstance that part of the neutrons, as we have already noted, are emitted upon the fission of nuclei not instantaneously, but with a delay of up to 1 min.

The first industrial reactors were intended to produce a fissioning material for atomic bombs—plutonium. In such reactors, part of the neutrons emitted in the fission of U^{235} nuclei maintains the chain reaction, while the other part experiences radiative capture by nuclei of U^{238}, which, as we have seen, leads in the long run to the formation of Pu^{239} [see scheme (10.37)]. After a sufficient amount of

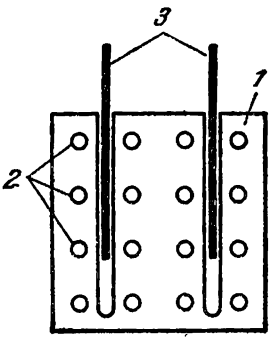


Fig. 10.11

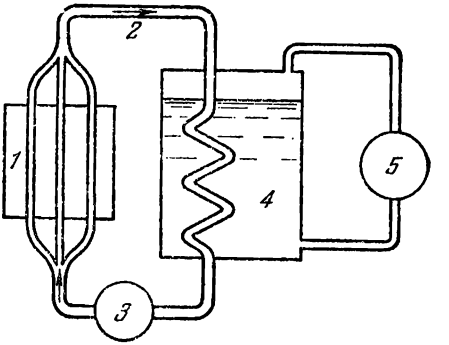


Fig. 10.12

Pu has accumulated in the uranium elements, the latter are extracted from the reactor and delivered for chemical treatment to recover the Pu from them.

The use of nuclear energy for peaceful purposes was first achieved in the USSR under the supervision of Igor Kurchatov. In 1954, the first atomic power plant with a capacity of 5000 kW was placed into service in the USSR. A schematic view of an atomic power plant is shown in Fig. 10.12. The energy liberated in the active zone of reactor *1* is picked up by a heat carrying agent flowing in circuit *2*. Circulation is ensured by pump *3*. Water or alkali metals having a low melting point, for example, sodium ($T_{melt} = 98$ °C) are used as the heat carrying agent. In heat exchanger *4*, the heat carrying agent gives up its heat to water and transforms it into steam that rotates turbine *5*.

Reactors with a moderator operate on slow (thermal) neutrons. By using fuel enriched with fissioning isotopes (U^{235} or Pu^{239}), it is possible to construct a reactor operating on fast neutrons. Part of the neutrons in such reactors are used to transform U^{238} into Pu^{239} or Th^{232} into U^{233}. Here the number of nuclei formed that are capable of fission under the action of thermal neutrons may exceed the number of fissioning nuclei used for maintaining operation of the reactor. Hence a greater amount of nuclear fuel is reproduced than burns in the reactor. Such nuclear reactors are therefore called **breeder reactors** or **breeders.**

We shall note in conclusion that by-products of the processes occurring in nuclear reactors are radioactive isotopes of many chemical elements that find widespread use in biology, medicine, and engineering.

10.8. Thermonuclear Reactions

Nuclear fusion, i.e. the fusion of light nuclei into a single nucleus, is attended, like the fission of heavy nuclei, by the liberation of enormous amounts of energy. Since very high temperatures are needed for the synthesis of nuclei, this process is called a **thermonuclear reaction.**

To surmount the potential barrier due to Coulomb repulsion, nuclei with the atomic numbers Z_1 and Z_2 must have the energy

$$E = \frac{Z_1 Z_2 e^2}{r_{\text{nuc}}}$$

where r_{nuc} is the radius of action of nuclear forces equal to $\sim 2 \times \times 10^{-13}$ cm. Even for nuclei with $Z_1 = Z_2 = 1$, this energy is

$$E = \frac{e^2}{r_{\text{nuc}}} = \frac{(4.8 \times 10^{-10})^2}{2 \times 10^{-13}} = 1.15 \times 10^{-6} \text{ erg} \approx 0.7 \text{ MeV}$$

An energy of 0.35 MeV falls to the part of each colliding nucleus. A temperature of the order of 2×10^9 K corresponds to an average energy of thermal motion equal to 0.35 MeV. The fusion of light nuclei, however, can proceed at considerably lower temperatures. The matter is that owing to the random distribution of particles by velocities, there is always a certain number of nuclei whose energy considerably exceeds the average value. Moreover, which is especially important, the fusion of nuclei may occur owing to the tunnel effect. Therefore, some thermonuclear reactions proceed with an appreciable intensity already at temperatures of the order of 10^7 K.

The conditions for the fusion of nuclei of deuterium and tritium are especially favourable because the reaction between them has a resonance nature. It is exactly these substances that form the charge of a **hydrogen** (or **thermonuclear**) **bomb***. The igniter in such a bomb is a conventional atomic bomb upon whose explosion a temperature of the order of 10^7 K is produced. The reaction of fusion of a deuteron (d) and a tritium nucleus ($_1H^3$)

$$_1d^2 + {}_1H^3 \rightarrow {}_2He^4 + {}_0n^1$$

is attended by the liberation of energy equal to 17.6 MeV, which is ~ 3.5 MeV per nucleon. We shall indicate for comparison that the

* The first thermonuclear explosion was conducted in the Soviet Union in 1953.

fission of a uranium nucleus leads to the liberation of about 0.85 MeV per nucleon.

There was no doubt up to recent times that the fusion of hydrogen nuclei into helium nuclei is the source of energy of the Sun* and stars, the temperature inside of which reaches 10^7 to 10^8 K. This fusion may occur in two ways. At lower temperatures, the **proton-proton cycle** occurs, which proceeds as follows. First the fusion of two protons occurs with the formation of a deuteron, a positron, and a neutrino:

$$p + p \rightarrow d + e^+ + \nu$$

The deuteron formed collides with a proton and combines with it to form a He^3 nucleus:

$$d + p \rightarrow {}_2He^3 + \gamma$$

The last link in the cycle is formed by the reaction

$${}_2He^3 + {}_2He^3 \rightarrow {}_2He^4 + p + p$$

At higher temperatures, the **carbon** (or **carbon-nitrogen**) **cycle** proposed by the German physicist Hans Bethe (born 1906) has a greater probability. It consists of the following reactions:

$$\begin{aligned}
{}_6C^{12} + {}_1p^1 &\rightarrow {}_7N^{13} + \gamma \\
{}_7N^{13} &\rightarrow {}_6C^{13} + e^+ + \nu \\
{}_6C^{13} + {}_1p^1 &\rightarrow {}_7N^{14} + \gamma \\
{}_7N^{14} + {}_1p^1 &\rightarrow {}_8O^{15} + \gamma \\
{}_8O^{15} &\rightarrow {}_7N^{15} + e^+ + \nu \\
{}_7N^{15} + {}_1p^1 &\rightarrow {}_6C^{12} + {}_2He^4
\end{aligned}$$

The result of the carbon cycle is the vanishing of four protons and the formation of one alpha particle. The number of carbon nuclei remains constant; these nuclei participate in the reaction as a catalyst.

In a hydrogen bomb, the thermonuclear reaction has an uncontrolled nature. To carry out **controlled thermonuclear reactions**, it is necessary to set up and maintain a temperature of the order of 10^8 K in a certain volume. At such a high temperature, a substance is a completely ionized **plasma** (see Sec. 12.5 of Vol. II, p. 249 et seq). There are tremendous difficulties in the path of carrying out a controlled thermonuclear reaction. In addition to the need of obtaining

* In the middle of the seventies, grounds appeared to doubt the correctness of this statement. It can be seen from the equations of the reactions given below that the fusion of protons is attended by the appearance of neutrinos whose number can be assessed. Measurements have shown, however, that the number of neutrinos liberated on the Sun is extremely small. In this connection, the question of the nature of solar energy remains unclear.

extremely high temperatures, the problem appears of keeping the plasma within the preset volume. Contact of the plasma with the walls of the vessel will result in its cooling. In addition, a wall of any substance will immediately evaporate. In this connection, a magnetic field has to be used to retain the plasma within the preset volume. The forces acting in this field on moving charged particles make them travel along trajectories arranged in a limited part of space.

The achievement of controlled thermonuclear fusion will provide mankind with a virtually inexhaustible source of energy. This is why work on mastering controlled thermonuclear reactions is going on in many countries. The scope of this work is the greatest in the USSR and the USA.

CHAPTER 11 ELEMENTARY PARTICLES

11.1. Kinds of Interactions and Classes of Elementary Particles

It is rather difficult to give a strict definition of the concept of an elementary particle. As a first approximation, we can define elementary particles to be such microparticles whose internal structure at the present stage of development of physics cannot be presented as a combination of other particles. In all phenomena observed to date, each such particle behaves like a single whole. Elementary particles can transform into one another. We encountered such transformations in the preceding chapter [see (10.7), (10.14), (10.15), and (10.16)].

To explain the properties and the behaviour of elementary particles, we have to supply them, in addition to mass, electric charge and spin, with a number of additional quantities characterizing them (quantum numbers), which will be treated on a later page.

Four kinds of interactions between elementary particles are known: strong, electromagnetic, weak, and gravitational (we have listed them in the order of diminishing of their intensity).

The intensity or strength of an interaction is customarily characterized with the aid of the so-called **coupling constant.** The latter is a dimensionless parameter determining the probability of the processes due to the given kind of interaction. The ratio of the values of the constants gives the relative strength of the corresponding interactions.

Strong Interaction. This kind of interaction ensures the binding of nucleons in a nucleus. The coupling constant in strong interaction has a value of the order of 10. The greatest distance over which strong interaction manifests itself (the radius of action r) is about 10^{-13} cm.

Electromagnetic Interaction. The coupling constant is $1/137 \approx \approx 10^{-2}$ [see formula (5.37) and the text following it]. The radius of action is unlimited ($r = \infty$).

Weak Interaction. This interaction is responsible for all kinds of beta decay of nuclei (including e-capture), for many disintegrations of elementary particles, and also for all the processes of interaction of neutrinos with a substance. The coupling constant is 10^{-14} in its order of magnitude. Weak interaction, like its strong counterpart, is short-range.

Gravitational Interaction. The coupling constant has a value of the order of 10^{-39}. The radius of action is unlimited ($r = \infty$). Gravitational interaction is universal. All elementary particles without

any exception are subjected to it. In the processes of the microworld, however, gravitational interaction does not play an appreciable part.

Table 11.1 gives the values (the order of magnitude) of the coupling constants for the different kinds of interaction. The last column

Table 11.1

Kind of interaction	Coupling constant	Lifetime, s
Strong	10	10^{-23}
Electromagnetic	10^{-2}	10^{-16}
Weak	10^{-14}	10^{-8}
Gravitational	10^{-39}	—

of the table indicates the average lifetime of particles that decay as a result of the given kind of interaction (this time is also known as the decay time).

Elementary particles are usually divided into four classes*. The first of them includes only a single particle—the **photon**. The second class includes **leptons**, the third—**mesons**, and, finally, the fourth class includes **baryons**. Mesons and baryons are often combined into a single class of strongly interacting particles called **hadrons** (the Greek word "hadros" means large, massive).

Let us briefly characterize the listed classes of particles.

1. Photons, γ (quanta of an electromagnetic field), participate in electromagnetic interactions, but do not have strong and weak interactions.

2. Leptons derive their name from the Greek word "leptos", which means "light-weight". They include particles having no strong interaction: muons (μ^-, μ^+), electrons (e^-, e^+), electron neutrinos (ν_e, $\tilde{\nu}_e$) and muon neutrinos (ν_μ, $\tilde{\nu}_\mu$) (see Sec. 11.8). All leptons have a spin equal to $\frac{1}{2}$, and are therefore fermions. All leptons have weak interaction. Those of them that carry an electric charge (i.e. muons and electrons) also have electromagnetic interaction.

3. Mesons are strongly interacting unstable particles not carrying a so-called baryon charge (see below). They include π-mesons or pions (π^+, π^-, π^0), K-mesons or kaons (K^+, K^-, K^0, $\tilde{K}^0$), and the eta-meson (η). Pions were treated in Sec. 10.4. The mass of K-mesons is about 970 m_e (494 MeV for charged and 498 MeV for neutral K-mesons). The lifetime of K-mesons is of the order of 10^{-8} s. They decay with

* There presumably exists another class of particles—**gravitons** (quanta of a gravitational field). These particles have not yet been discovered experimentally.

the formation of π-mesons and leptons or only leptons. The mass of an eta-meson is 549 MeV (1074 m_e), its lifetime is of the order of 10^{-19} s. Eta-mesons decay with the formation of π-mesons and γ-photons.

Unlike leptons, mesons have not only weak (and, if charged, electromagnetic), but also strong interaction. The latter manifests itself when they interact with one another, and also in interaction between mesons and baryons. The spin of all mesons is zero, so that they are bosons.

4. The class of baryons combines nucleons (p, n) and unstable particles having a mass greater than that of nucleons and called **hyperons** (Λ, Σ^+, Σ^0, Σ^-, Ξ^0, Ξ^-, Ω^-). All baryons have strong interaction and, consequently, readily interact with atomic nuclei. The spin of all baryons is 1/2, so that baryons are fermions. Except for the proton, all baryons are unstable. When a baryon decays, a baryon is formed without fail in addition to other particles. This is one of the manifestations of the **law of baryon charge conservation**, which we shall deal with in Sec. 11.4.

In addition to the particles listed above, a great number of strongly interacting short-lived particles called **resonances** have been discovered. These particles are resonance states formed by two or more elementary particles. The lifetime of resonances is only about 10^{-23} to 10^{-22} s. Some of the resonances are bosons and must be included in the class of mesons. Others are fermions and must be included in the class of hyperons. We shall not consider resonances in the following.

11.2. Methods for Detecting Elementary Particles

We succeed in observing elementary particles, and also complex microparticles (α, d, etc.) owing to the traces they leave in passing through a substance. The nature of the traces makes it possible to judge about the charge sign of a particle, its energy, momentum, etc. Charged particles ionize the molecules along their path. Neutral particles leave no traces, but they may reveal their presence at the moment of decaying into charged particles or at the moment of colliding with a nucleus. Hence, in the long run, neutral particles are also detected according to the ionization produced by the charged particles they give birth to.

The instruments used to detect ionizing particles are divided into two groups. The first group includes devices that register the fact of a particle flying through them and, in addition, sometimes make it possible to judge about its energy. The second group includes **track-detecting instruments**, i.e. instruments that make it possible to observe the tracks of particles in a substance.

The registering instruments include **ionization chambers** and **gas-discharge counters** (see Sec. 12.3 of Vol. II, p. 240 et seq), and also **Cerenkov counters** (see Sec. 20.6 of Vol. II, p. 471), **scintillation counters**, and **semiconductor counters.**

The operation of scintillation counters is based on the fact that a charged particle when flying through a substance produces not only ionization, but also excitation of the atoms. Upon returning to their normal state, the atoms emit visible light. Substances in which charged particles produce a noticeable flash of light (scintillation) are called **phosphors** or **scintillators.** A scintillation counter consists of a phosphor from which light is supplied via a special light guide to a photomultiplier. The pulses obtained at the output of the photomultiplier are counted. The amplitude of the pulses (which is proportional to the intensity of the scintillations) is also determined. This provides additional information on the particles being registered.

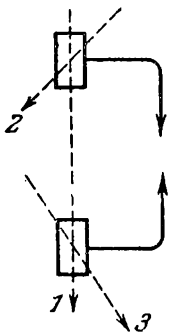


Fig. 11.1

A semiconductor counter is a semiconductor diode to which a voltage is fed of a sign such that the main current carriers are pulled back from the junction layer. Consequently, in the normal state, the diode is cut off. A fast charged particle when passing through the junction layer gives birth to electrons and holes that are drained off to the electrodes. The result is the appearance of an electric pulse proportional to the number of current carriers produced by the particle.

Counters are often combined into groups and are connected so as to register only those events that are recorded by several instruments simultaneously, or, conversely, that are recorded only by one of the instruments. In the first case, the counters are said to form a **coincidence system** or **circuit**, and in the second one an **anticoincidence system** or **circuit.** By using different systems of connection, one can single out the phenomenon of interest from a multitude of phenomena. For example, two coincidence counters (Fig. 11.1) installed one after the other register particle *1* flying along their common axis and do not register particles *2* and *3*.

Track-detecting instruments include Wilson chambers, diffusion chambers, bubble chambers, spark chambers, and emulsion chambers.

The Wilson Cloud Chamber. This instrument was developed by the British physicist Charles Wilson (1869-1959) in 1912. A track of the ions formed by a charged particle flying through the cloud chamber becomes visible in it because the supersaturated vapour of a liquid condenses on the ions. The instrument operates in cycles, and not continuously. The comparatively short (about 0.1 to 1 s) time of sensitivity of the cloud chamber alternates with the dead time (from

100 to 1000 times greater) during which the chamber is prepared for the next working cycle. Supersaturation is achieved by sudden cooling produced by the sharp (adiabatic) expansion of the working mixture consisting of a non-condensing gas (helium, nitrogen, argon) and the vapour of water, ethyl alcohol, etc. At the same moment, the working volume of the cloud chamber is photographed stereoscopically (i.e. from several points). Stereophotography makes it possible to reproduce a three-dimensional picture of the recorded phenomenon. Since the ratio of the sensitive time to the dead time is very small, sometimes tens of thousands of photographs have to be made before an event having a low probability will be recorded. To increase the probability of observing rare phenomena, controlled Wilson chambers are used in which the operation of the expansion mechanism is controlled by particle counters included in the electronic circuit singling out the required event.

In 1927, the Soviet scientist Dmitri Skobeltsyn (born 1892) for the first time placed a Wilson cloud chamber between the poles of an electromagnet, which greatly extended its possibilities. The curvature of the trajectory produced by the action of the magnetic field makes it possible to determine the sign of the charge of a particle and its momentum. An example of a photograph obtained with the aid of a Wilson cloud chamber placed in a magnetic field is Fig. 11.7 (see p. 275), in which the tracks of an electron and a positron can be seen.

The Diffusion Chamber. Like the Wilson cloud chamber, the working substance in a diffusion chamber is a supersaturated vapour. The state of supersaturation, however, is produced not by adiabatic expansion, but as a result of the diffusion of alcohol vapour from the lid of the chamber, which is at a temperature of about 10 °C, to its bottom, which is cooled by solid carbon dioxide (to about —70 °C). A layer of supersaturated vapour a few centimeters thick appears not far from the bottom. It is exactly in this layer that tracks are formed. Unlike the Wilson cloud chamber, a diffusion chamber functions continuously.

The Bubble Chamber. In the bubble chamber invented by the American physicist Donald Glaser (born 1926) in 1952, the supersaturated vapour is replaced with a transparent superheated liquid (i.e. a liquid under an external pressure that is less than its saturated vapour pressure; see Sec. 15.5 of Vol. I, p. 394 et seq). An ionizing particle flying through the chamber causes violent boiling of the liquid, owing to which the track of a particle is indicated by a chain of vapour bubbles. A bubble chamber, like a Wilson cloud chamber, operates in cycles. The chamber is started by a sharp lowering of the pressure, owing to which the working liquid transforms into the metastable superheated state. The capacity of the working liquid, which simultaneously is a target for the particles flying through it, is filled

by hydrogen, xenon, propane (C_3H_8) and certain other substances. The working volume of a chamber reaches 1000 litres.

The Spark Chamber. In 1957, T. Cranshaw and J. De Beer developed an instrument for registering the trajectories of charged particles called a spark chamber. The instrument consists of an array of plane parallel metal electrodes (Fig. 11.2). The even-numbered electrodes are grounded, and the odd-numbered ones are periodically supplied with a short (lasting 10^{-7} s) high-voltage pulse (10-15 kV). If an ionizing particle flies through the chamber at the moment when a pulse is supplied to it, its track will be marked by a chain of sparks jumping between the electrodes. The chamber is switched on automatically with the aid of additional counters connected in a coincidence circuit that register the passage of the particles being investigated through the working volume of the chamber.

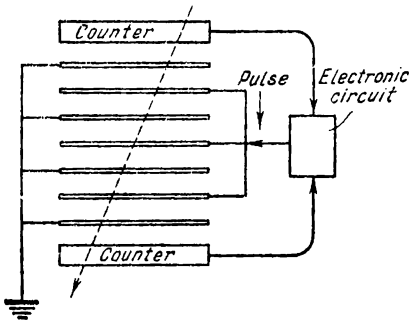


Fig. 11.2

An improved variant of the spark chamber is the **streamer chamber.** In this chamber the high voltage is removed before

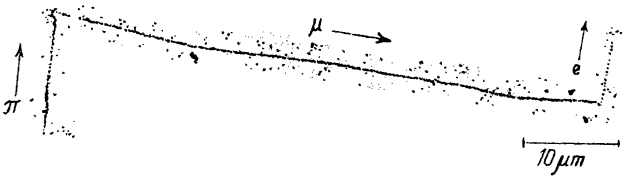


Fig. 11.3

a spark manages to develop completely. Therefore, only embryonic sparks develop that form a clearly visible track.

The Emulsion Chamber. The Soviet physicists Lev Mysovsky and Aleksandr Zhdanov were the first to use photographic plates for recording microparticles. Charged particles act on a photographic emulsion in the same way as photons. Therefore, after the development of a plate, a visible track of a particle that has flown past it is formed. A shortcoming of the photographic plate method was the

small thickness of the emulsion layer, as a result of which only the tracks of particles flying parallel to the plane of the layer were obtained completely. In **emulsion chambers,** thick bundles (with a mass of up to several scores of kilogrammes and a thickness of several hundred millimetres) made up from separate layers of photographic emulsions (without a substrate) are subjected to radiation. After irradiation, the bundle is taken apart into layers, each of which is developed and examined under a microscope. To be able to trace the path of a particle when passing from one layer to another, before the bundle is taken apart an identical coordinate network is inscribed on all the layers by means of X-rays. The tracks of particles obtained in this way are shown in Fig. 11.3, in which the consecutive transformation of a π-meson into a muon and then into a positron has been recorded.

11.3. Cosmic Rays

Before the development of powerful accelerators of charged particles, cosmic radiation was the only source of particles having an energy sufficient for the formation of mesons and hyperons. The positron, muons, π-mesons, and many strange particles (see Sec. 11.6) were discovered in the composition of cosmic rays.

Primary and **secondary cosmic rays** are distinguished. The primary rays are a flux of atomic nuclei (mainly protons) of a high energy (on an average of about 10 GeV, the energy of individual particles reaching 10^{10} GeV*) continuously falling on the Earth. The particles of primary cosmic rays collide inelastically with atomic nuclei in the upper layers of the atmosphere, the result being secondary radiation. At altitudes below 20 km, cosmic rays are virtually completely of a secondary nature. All the elementary particles known at present are encountered in the secondary rays.

The intensity of the primary cosmic rays at the atmosphere's boundary (i.e. at an altitude of about 50 km) is approximately 1 particle/(cm^2·s). The flux of charged particles at sea level averages about 2×10^{-2} particle/(cm^2·s). The existence of the Earth's magnetic field leads to the fact that the intensity of cosmic rays varies with the latitude. This phenomenon is known as the **latitude effect.**

Instruments installed on artificial satellites of the Earth and on spaceships helped scientists discover **radiation belts** near the Earth. These are two zones with a sharply increased intensity of ionizing radiation surrounding the Earth. Their existence is due to the capture and retaining of charged cosmic particles by the Earth's magnetic field. In the plane of the equator, the internal radiation belt ex-

* We remind our reader that 1 GeV (gigaelectron-volt) equals 10^9 eV.

tends from 600 to 6000 km, and the external belt from 20 000 to 60 000 km. At latitudes of 60 to 70 degrees, both belts approach the Earth to a distance of several hundred kilometres.

Secondary cosmic rays contain two components. One of them is greatly absorbed by lead and was therefore called **soft**; the second one penetrates through thick layers of lead and was called **hard**.

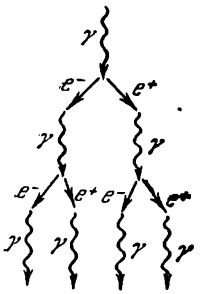


Fig. 11.4

The soft component consists of **cascades** or **showers** of electron-positron pairs. A gamma photon produced as a result of the decay of a π^0-meson [see (10.15)] or of the sharp retardation of a fast electron when flying near an atomic nucleus gives birth to an electron-positron pair (Fig. 11.4). The retardation of these particles again results in the formation of gamma photons, and so on. The processes of the birth of pairs and of the production of gamma photons alternate until the energy of the gamma photons becomes inadequate for the formation of pairs. Since the energy of the initial photon is very high, many generations of secondary particles have time to appear before the development of a shower stops.

The hard penetrating component of cosmic rays consists mainly of muons. It is formed predominatingly in the upper and middle layers of the atmosphere as a result of the decay of charged π-mesons [see (10.14)].

With the appearance of accelerators making it possible to accelerate particles up to energies of hundreds of GeV (see Sec. 10.5 of Vol. II, p. 223 et seq.), cosmic rays have lost their exclusive significance in studying elementary particles. As previously, however, they remain the only source of particles having superhigh energies.

11.4. Particles and Antiparticles

The Schrödinger equation does not satisfy the requirements of the theory of relativity—it is not invariant with respect to the Lorentz transformations. In 1928, the British physicist Paul Dirac succeeded in finding a **relativistic quantum-mechanical equation** for an electron from which a number of remarkable corollaries follow. First of all, this equation naturally, without any additional assumptions, gives us the spin and the numerical value of the intrinsic magnetic moment of an electron. The spin was thus found to be a quantity that is simultaneously a quantum and a relativistic one.

But this does not exhaust the significance of Dirac's equation. It also made it possible to predict the existence of an electron's antiparticle—the **positron**. Dirac's equation gives not only positive,

but also negative values for the total energy of a free electron. Investigation of the equation shows that at a given momentum of a particle p it has solutions corresponding to the energies

$$E = \pm \sqrt{c^2p^2 + m_e^2c^4} \tag{11.1}$$

There is an interval of energy values between the maximum negative energy ($-m_ec^2$) and the minimum positive energy ($+m_ec^2$) that cannot be realized. The width of this interval is $2m_ec^2$ (Fig. 11.5). Hence, two regions of energy eigenvalues are obtained, one begins from $+m_ec^2$ and extends to $+\infty$ and the other begins from $-m_ec^2$ and extends to $-\infty$.

In non-quantum relativistic mechanics, the energy is expressed through the momentum with the aid of an expression coinciding with Eq. (11.1) [see Eq. (8.42) of Vol. I, p. 242], so that it formally can also have negative values. In the non-quantum theory, however, energy changes continuously and therefore cannot intersect the forbidden band and pass from positive values to negative ones. In the quantum theory, the energy can change not only continuously, but also in a jump, so that the existence of a forbidden band cannot prevent the transition of a particle to states with a negative energy (compare with the transition of an electron in a semiconductor from the valence band to the conduction band, Fig. 8.3).

A particle with a negative energy must have very strange properties. Upon passing over to states with a decreasing energy (i.e. with a negative energy increasing in magnitude), it could liberate energy, say, in the form of radiation, and since $|E|$ is restricted by nothing, a particle having a negative energy could emit an infinitely great amount of energy. A similar conclusion can be arrived at as follows. A glance at the equation $E = mc^2$ reveals that the mass of a particle having a negative energy will also be negative. Under the action of a retarding force, a particle with a negative mass should accelerate instead of retarding, doing an infinitely great amount of work on the source of the retarding force.

These difficulties should seem to make us acknowledge that states with a negative energy must be excluded from consideration as leading to absurd results. This, however, would contradict some of the general principles of quantum mechanics. Therefore, Dirac chose a different way. He assumed that transitions of electrons to states with a negative energy are usually not observed because all the available levels with a negative energy are already occupied by electrons. We remind our reader that electrons obey the Pauli principle which prohibits more than one particle from being in the same state.

According to Dirac, a vacuum is a state in which all the levels of negative energy are populated by electrons, while the levels with a positive energy are vacant (Fig. 11.6*a*). Since all the levels below the forbidden band are occupied without any exception, the electrons at

these levels do not reveal their presence in any way*. If the energy

$$E \geqslant 2m_e c^2 \qquad (11.2)$$

is imparted to one of the electrons at negative levels, then this electron will transfer to a state with a positive energy and will behave in the usual way like a particle with a positive mass and a negative charge. The vacancy ("hole") formed in the collection of negative

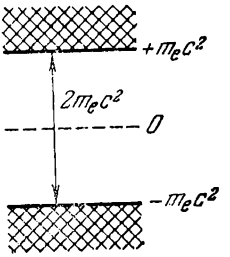


Fig. 11.5

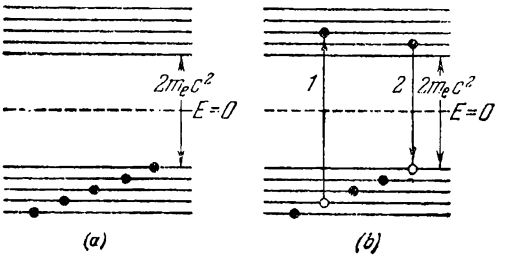


Fig. 11.6

levels must behave like an electron having a positive charge. Indeed, the absence of a particle having a negative mass and charge will be perceived as the presence of a particle having a positive mass and a positive charge. This first particle of those predicted theoretically was called a **positron**.

When a positron and an electron meet, they **annihilate** (vanish)—the electron transfers from a positive level to a vacant negative one**. The energy corresponding to the difference between these levels is liberated in the form of radiation. In Fig. 11.6*b*, arrow *1* depicts the process of the birth of an electron-positron pair, and arrow *2*—their annihilation. The term "annihilation" must not be understood literally. In essence, the particles (electron and positron) do not vanish, but transform into other particles (gamma photons).

Dirac's theory was so "crazy" that most physicists were very distrustful of it. It won recognition only after the American physicist Carl Anderson in 1932 detected a positron in the composition of cosmic rays. In a Wilson cloud chamber placed between the poles of an electromagnet, a positron left the same track as an electron born simultaneously with it, except that this track was curled in the opposite direction (Fig. 11.7).

Electron-positron pairs are born when gamma photons pass through a substance. This is one of the main processes resulting in a substance

* Similarly, in a dielectric, the electrons completely filling the valence band do not react in any way to the action of an electric field.

** This process is similar to the recombination of an electron and a hole in a semiconductor.

absorbing gamma rays. In complete accordance with Dirac's theory, the minimum energy of a gamma photon at which the birth of a pair is observed is $2m_ec^2 = 1.02$ MeV [see expression (11.2)]. To observe

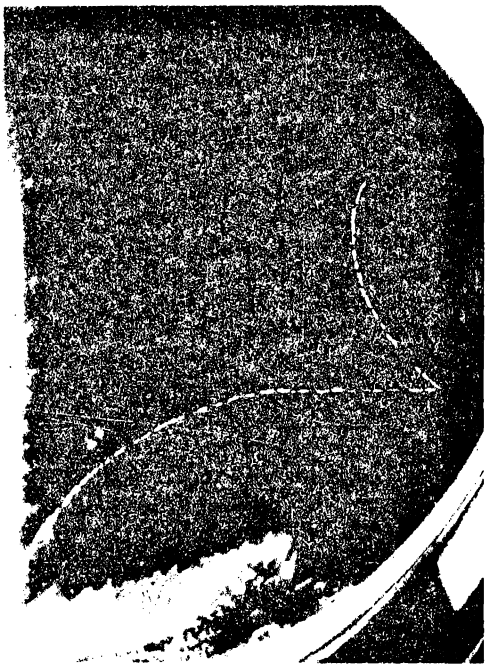


Fig. 11.7

the law of momentum conservation in the process of the birth of a pair, another particle (electron or nucleus) must participate in it that receives the excess momentum of a gamma photon over the total momentum of an electron and a positron. Hence, a pair is born as follows:

$$\gamma + e^- \rightarrow e^- + e^- + e^+ \qquad (11.3)$$

or

$$\gamma + X \rightarrow X + e^- + e^+ \qquad (11.4)$$

where X is the nucleus in whose force field the pair is born.

Electron-positron pairs may also be produced when two charged particles, for example, electrons, collide:

$$e^- + e^- \rightarrow e^- + e^- + e^- + e^+ \qquad (11.5)$$

In annihilation, the requirements of the law of momentum conservation are observed in that two (more rarely three) gamma photons are produced that fly away in different directions:

$$e^- + e^+ \rightarrow \gamma + \gamma\ (+\gamma) \qquad (11.6)$$

The fraction of the energy received by the nucleus X in process (11.4) is so small that the threshold of the reaction of pair formation (i.e. the minimum energy of a gamma photon needed for it) virtually equals $2m_ec^2$. The threshold of reaction (11.3) is $4m_ec^2$, and of reaction (11.5) is $7m_ec^2$ (in the latter case by the threshold of the reaction is meant the minimum total energy of the colliding electrons). Thus, the requirements of the simultaneous conservation of energy and momentum result in the fact that the threshold of a reaction (the minimum energy of the initial particles) may be appreciably greater than the total rest energy of the born particles.

Dirac's equation in a somewhat modified form may be applied not only to electrons, but also to other particles having a spin of 1/2. Consequently, for each such particle (for example, a proton or a neutron) there must exist an antiparticle*. By analogy with process (11.5), the birth of a proton-antiproton pair (p-$\tilde{p}$) or of a neutron-antineutron pair (n-$\tilde{n}$) can be expected when nucleons having a sufficiently high energy collide. The total rest energy of a proton and antiproton, like that of a neutron and antineutron, is almost 2 GeV [see Eqs. (10.1) and (10.5)]. The threshold of the reaction determined by the requirements of energy and momentum conservation is 5.6 GeV. In 1955, an accelerator (synchrophasotron; see Sec. 10.5 of Vol. II, p. 227) was put into service at Berkeley, California (USA) that made it possible to accelerate particles up to an energy of 6.3 GeV. By irradiating a copper target with a beam of accelerated protons, the American physicists O. Chamberlain, E. Segré, C. Wiegand, and T. Ypsilantis observed the formation of a p-$\tilde{p}$ pair. The reaction proceeded according to one of the following schemes:

$$p + p \rightarrow p + p + p + \tilde{p} \quad \text{or} \quad p + n \rightarrow p + n + p + \tilde{p} \qquad (11.7)$$

The second nucleon in the left-hand side is in a Cu nucleus. Since the

* An antiparticle is designated by the same symbol as the particle corresponding to it, with the addition of a tilde ($\sim$) or a bar over the symbol. For example, an antiproton is designated by the symbol $\tilde{p}$ or $\bar{p}$.

nucleons in a nucleus are in motion, the threshold energy of the impinging particle in this case is about 4.3 GeV.

An antiproton differs from a proton in the sign of its electric charge and in its intrinsic magnetic moment (in an antiproton the magnetic moment is negative, i.e. directed oppositely to the mechanical angular momentum). The main feature distinguishing an antiproton from a proton (and in general a particle from an antiparticle) is their ability of mutual annihilation, as a result of which other particles are produced. An antiproton may annihilate when it encounters not only a proton, but also a neutron. The collection of particles produced in separate events of annihilation is different. For example, the following processes are possible:

$$\left.\begin{aligned} &\tilde{p}+p\to\pi^{+}+\pi^{-}+\pi^{+}+\pi^{-}+\pi^{0}\\ &\tilde{p}+p\to\pi^{+}+\pi^{-}+\pi^{0}+\pi^{0}+\pi^{0}\\ &\tilde{p}+n\to\pi^{+}+\pi^{-}+\pi^{-}+\pi^{0}+\pi^{0} \end{aligned}\right\} \qquad (11.8)$$

In 1956, using the same accelerator at Berkeley, B. Cork, G. Lambertson, O. Piccioni, and W. Wenzel observed antineutrons obtained by the recharging of antiprotons, i.e. as a result of the processes

$$\begin{aligned} &\tilde{p}+p\to\tilde{n}+n\\ &\tilde{p}+n\to\tilde{n}+n+\pi^{-} \end{aligned} \qquad (11.9)$$

An antineutron differs from a neutron in the sign of its intrinsic magnetic moment (in an antineutron the direction of the magnetic moment coincides with that of the mechanical angular momentum) and in its ability to annihilate when it encounters a nucleon (neutron or proton). Annihilation results in the birth of new particles (mainly π-mesons).

Not only fermions, but also bosons have antiparticles. For example, a π^{-}-meson is the antiparticle with respect to a π^{+}-meson.

There are particles that are identical with their antiparticles (i.e. that have no antiparticles). Such particles are called **absolutely neutral**. They include the photon, π^{0}-meson, and the η-meson. Particles identical with their antiparticles are not capable of annihilation. This, however, does not signify that they in general cannot transform into other particles.

If a **baryon charge*** (**or baryon number**) $B = +1$ is ascribed to baryons (i.e. nucleons and hyperons), a baryon charge $B = -1$ is ascribed to antibaryons, and a baryon charge $B = 0$ to all other particles, then all the processes occurring with the participation of baryons and antibaryons [for example, processes (11.7), (11.8) and

* The baryon charge is one of the quantum numbers mentioned in the second paragraph of Sec. 11.1.

(11.9)] will be characterized by **conservation of the baryon charge**, like processes (11.3)-(11.6) are characterized by conservation of the electric charge.

The law of baryon charge conservation results in the stability of the lightest of the baryons—the proton. The other conservation laws (of energy, momentum, angular momentum, electric charge, etc.) do not forbid, for example, the process

$$p \to e^+ + \nu + \tilde{\nu} \tag{11.10}$$

that in the long run would lead to annihilation of atoms. Such a process, however, would be attended by a reduction in the baryon charge by unity, and is therefore not observed. Similarly, the law of electric charge conservation results in stability of the lightest charged particle—the electron, forbidding, for example, the process

$$e^- \to \gamma + \gamma + \nu \tag{11.11}$$

To explain the features of processes with the participation of leptons and antileptons, it is necessary to introduce the quantum number L called the **lepton charge** (or **lepton number**). For leptons, $L = +1$, for antileptons $L = -1$, and for all other particles $L = 0$. When this condition is observed, conservation of the total lepton charge of the physical system being considered is observed in all processes without any exception.

The transformation of all the quantities describing a physical system in which all the particles are replaced with antiparticles (for example, electrons with positrons, positrons with electrons, etc.) is called **charge conjugation**. Which of two charge-conjugated particles is to be considered a particle and which an antiparticle is, generally speaking, a purely conditional matter. Having made a choice for one pair of charge-conjugated particles, however, the choice for the other pairs must be made so as to conserve the baryon and lepton charges in the observed interactions. The electron and the proton are conventionally considered to be particles, and the positron and the antiproton to be antiparticles. When this condition is observed, the choice for the remaining baryons and leptons is unambiguous. For example, to conserve the baryon charge in the course of process (10.7), we must consider the neutron to be a particle. The results obtained when account of the requirements of conserving B and L for other particles is taken are given in Table 11.2.

Table 11.2 indicates all the particles discovered up to 1977 except for resonances. The first column gives the names of the particles. When the antiparticle is designated with the aid of a tilde ($\sim$) or a bar ($-$), the name of the antiparticle is obtained by adding the prefix "anti" to the name of the relevant particle. For example, the antiparticle of the lambda hyperon is called the antilambda hyperon. The antiparticle of the electron is the positron. In the remaining cases,

Table 11.2

Name of particle	Particle	Antiparticle	m, MeV	τ, s	Decay scheme
Photon	γ		0	stable	
		Leptons			
Electron	e^-	e^+	0.511	stable	
Muon	μ^-	μ^+	106	2.2×10^{-6}	$e^-+\tilde{\nu}_e+\nu_\mu$
Electron neutrino	ν_e	$\tilde{\nu}_e$	0	stable	
Muon neutrino	ν_μ	$\tilde{\nu}_\mu$	0	stable	
		Mesons			
Positive pi-meson	π^+	π^-	140	2.6×10^{-8}	$\mu^+ + \nu_\mu$
Neutral pi-meson	π^0		135	0.8×10^{-16}	$\gamma+\gamma$ $e^+ + e^- + \gamma$
Positive K-meson	K^+	K^-	494	1.2×10^{-8}	$\mu^+ + \nu_\mu$ $\pi^+ + \pi^0$ $\pi^+ + \pi^+ + \pi^-$
Neutral K-meson	K^0	$\tilde{K}^0$	498	10^{-10}-10^{-8}	$\pi^+ + \pi^-$ $\pi^0 + \pi^0$ $\pi^+ + e^- + \tilde{\nu}_e$ $\pi^- + e^+ + \nu_e$
Eta-meson		η	549	2.4×10^{-19}	$\gamma+\gamma$ $\pi^+ + \pi^- + \pi^0$ $\pi^0 + \pi^0 + \pi^0$
		Baryons			
Proton	p	$\tilde{p}$	938.2	stable	
Neutron	n	$\tilde{n}$	939.6	0.9×10^{3}	$p + e^- + \tilde{\nu}_e$
Lambda hyperon	Λ	$\tilde{\Lambda}$	1116	2.5×10^{-10}	$p+\pi^-$ $n+\pi^0$
Positive sigma hyperon	Σ^+	$\tilde{\Sigma}^+$	1189	0.8×10^{-10}	$p+\pi^0$ $n+\pi^+$
Neutral sigma hyperon	Σ^0	$\tilde{\Sigma}^0$	1192	$<10^{-14}$	$\Lambda+\gamma$
Negative sigma hyperon	Σ^-	$\tilde{\Sigma}^-$	1197	1.5×10^{-10}	$n+\pi^-$
Neutral xi hyperon	Ξ^0	$\tilde{\Xi}^0$	1315	3×10^{-10}	$\Lambda+\pi^0$
Negative xi hyperon	Ξ^-	$\tilde{\Xi}^-$	1321	1.7×10^{-10}	$\Lambda+\pi^-$
Negative omega hyperon	Ω^-	$\tilde{\Omega}^-$	1672	1.3×10^{-10}	$\Xi^0+\pi^-$ $\Xi^-+\pi^0$ $\Lambda+K^-$

the names of particles and antiparticles are distinguished by adding the words "positive" and "negative". For example, the negative pi-meson is the antiparticle of the positive pi-meson*. The second and third columns give the symbols of the particle and its antiparticle. The symbols of absolutely neutral particles are inserted in interruptions of the vertical line dividing the particle and antiparticle columns. The fourth and fifth columns give the mass of a particle m and its average lifetime τ. Finally, the last column indicates the main schemes of decay of the particles. To obtain the scheme of decay of an antiparticle, the particles must be replaced with antiparticles and the antiparticles with particles For example, the scheme of the decay of a positive muon has the form $e^+ + \nu_e + \tilde{\nu}_\mu$.

Now we can explain why the particle produced in decays (10.7) and (10.26) should be called an antineutrino, and that produced in decay (10.27) a neutrino. This follows from the requirement of conservation of the lepton charge. For an electron and a neutrino, $L = +1$, and for a positron and an antineutrino, $L = -1$. Hence, the total lepton charge does not change if an electron is produced together with an antineutrino, and a positron together with a neutrino.

Ascribing $L = +1$ to an electron, we must also ascribe $L = +1$ to a negative muon in accordance with decay reaction (10.16), i.e. consider μ^- to be a particle, and a positive muon to be an antiparticle with a value of $L = -1$. It is easy to see that the lepton charge is also conserved in π-meson decay processes [see (10.14)].

11.5. Isotopic Spin

It follows from the charge independence of nuclear forces (see Sec. 10.4) that a proton and a neutron display much more similarities than distinctions. They participate equally in strong interaction, the spin of both particles is the same, their masses are very close. This gives us grounds to consider a proton and a neutron as two different states of the same particle—a nucleon. If electromagnetic interaction is "switched off", then both these states coincide completely (the slight difference between the masses of a proton and a neutron is due to electromagnetic interaction).

Let us turn to the diagram of the sodium atom levels (see Fig. 5.6). We remind our reader that the multiplet structure of the levels is due to the interaction between the spin and the orbital moments of the electrons. "Switching off" of the spin-orbital interaction would result in vanishing of the difference between, for example, the levels

* The terms pion and kaon are sometimes used instead of π-meson and K-meson, respectively.

$3^2P_{1/2}$ and $3^2P_{3/2}$ and in their merging into the single level $3P$. Switching on of the spin-orbital interaction, on the contrary, results in the formation of multiplets like the switching on of electromagnetic interaction results in the appearance of differences between a proton and a neutron. This analogy served as the grounds for calling a proton and a neutron a **charge multiplet** (doublet). Other particles can also be combined into charge multiplets. For example, a lambda hyperon forms a singlet (see Table 11.2), π-mesons a triplet (when electromagnetic interaction is switched off, all three π-mesons become indistinguishable).

A definite value of the spin S corresponds to each spectral multiplet (the number of components in a multiplet is $2S + 1$). The separate components of a multiplet are distinguished by the values of the projection of the spin onto the z-axis. By analogy with conventional spin, to each charge multiplet there is ascribed a definite value of the **isotopic spin** (or **isospin**)* T selected so that $2T + 1$ equals the number of particles in a multiplet. Different values of T_z—the projection of the isotopic spin onto the z-axis in imaginary **isotopic space**—are ascribed to different particles. For example, for nucleons $T = 1/2$, $T_z = +1/2$ corresponds to a proton, and $T_z = -1/2$ to a neutron. For π-mesons, $T = 1$, the projections T_z equal $+1$, 0, and -1 for π^+-, π^0-, and π^--mesons, respectively.

To avoid misunderstanding, we must note that the quantum number T called the isotopic spin has no relation to isotopes or to conventional spin. The word "isotopic" appeared in the name of the quantum number T because a proton and a neutron form different "varieties" of a nucleon, like real isotopes form varieties of a given chemical element. The word "spin" appeared in the name because the mathematical apparatus describing the quantum number T was found to be exactly the same as the mathematical apparatus describing conventional spin. Otherwise, there is nothing in common between isotopic and conventional spins.

Table 11.3 gives the values of T and T_z for different particles. Each line in this table gives a charge multiplet. Hence, if there are, for example, two lines for a nucleon, this signifies that nucleons form two charge multiplets.

Let us consider two charge multiplets differing in that the particles forming one multiplet are antiparticles with respect to the particles forming the other multiplet. The isotopic spins of both multiplets are obviously the same ($2T + 1$ gives the number of particles in a multiplet). As regards the projections of the isotopic spin T_z, they differ in sign for a particle and antiparticle. Thus, for a proton, $T_z =$

* Isotopic spin was first introduced into consideration by Werner Heisenberg in 1932 in order to describe a proton and a neutron as different states of a nucleon.

Table 11.3

Particle	Isotopic spin T	Projection of isotopic spin T_z				
		-1	$-1/2$	0	$+1/2$	$+1$
π-Meson	1	π^-		π^0		π^+
K-Meson	1/2 1/2		K^0 K^-		K^+ $\widetilde{K}^0$	
η-Meson	0			η		
Nucleon	1/2 1/2		n $\widetilde{p}$		p $\widetilde{n}$	
Λ-Hyperon	0 0			Λ $\widetilde{\Lambda}$		
Σ-Hyperon	1 1	Σ^- $\widetilde{\Sigma}^+$		Σ^0 $\widetilde{\Sigma}^0$		Σ^+ $\widetilde{\Sigma}^-$
Ξ-Hyperon	1/2 1/2		Ξ^- $\widetilde{\Xi}^0$		Ξ^0 Ξ^+	
Ω-Hyperon	0 0			Ω^- $\widetilde{\Omega}^-$		

$= +1/2$, for an antiproton, $T_z = -1/2$, for a neutron $T_z = -1/2$, for an antineutron, $T_z = +1/2$.

At first sight, it may appear strange that for π-mesons, both a particle (π^+) and its antiparticle (π^-) combine to form a single charge multiplet, whereas, for example, a Λ-hyperon and an anti-Λ-hyperon form two different charge multiplets. The explanation is that a charge multiplet unites particles differing only in the magnitude or the sign of their electric charge; all the other quantities characterizing the particles must be the same*. The hyperons Λ and $\widetilde{\Lambda}$ differ in the value of their baryon number, and therefore cannot be included in one multiplet. The baryon number of all π-mesons is zero, and the other quantum numbers are also the same; hence, there are no obstacles preventing their combination into one multiplet.

A conservation law is associated with the isotopic spin. In strong interactions, both the isotopic spin T and its projection T_z are con-

* The difference between charged and neutral particles due to electromagnetic interaction, for example, their slight difference in mass, is not taken into consideration.

served. In electromagnetic interactions, only T_z is conserved, while the isotopic spin T itself is not conserved. Weak interactions proceed, as a rule, with a change in the isotopic spin.

The concept of isotopic spin was a very fruitful one. It played a major part in systematizing elementary particles. In particular, it suggested the idea to the American physicist Murray Gell-Mann (born 1929) and independently of him to the Japanese physicist Kazuhiko Nishijima (born 1926) of combining particles into charge multiplets and then led them to the concept of **strangeness** (see the following section).

11.6. Strange Particles

K-mesons and hyperons (Λ, Σ, Ξ) were discovered in cosmic rays in the early 1950's . Beginning from 1953, scientists have been using accelerators to produce them. The behaviour of these particles was found to be so unusual that they were called **strange.**

The singularity of the behaviour of strange particles consists in that they are obviously born as a result of strong interactions with a characteristic time of the order of 10^{-23} s, whereas their lifetimes were found to be of the order of 10^{-8} to 10^{-10} s. The latter circumstance pointed to the fact that the particles decay as a result of weak interactions. It was absolutely incomprehensible why strange particles live so long, and what hinders them from decaying at the expense of the strong interaction giving birth to them. For example, one of the processes producing strange particles has the form

$$\pi^- + p \rightarrow K^0 + \Lambda \tag{11.12}$$

while the lambda hyperon decays according to the scheme

$$\Lambda \rightarrow \pi^- + p \tag{11.13}$$

(see the photograph of tracks of particles obtained in a liquid hydrogen bubble chamber given in Fig. 11.8). Since the same particles (a π^--meson and a proton) participate both in the birth and in the decay of a lambda hyperon, it is surprising that the speed (i.e. probability) of the two processes is so different.

Further investigations showed that strange particles are born in pairs [see (11.12)]. This suggested the idea that strong interactions cannot play a part in the decay of these particles because the presence of two strange particles is needed for their manifestation. For the same reason, the single birth of strange particles is forbidden.

A law of conservation always underlies the forbiddenness of a process. Thus, the decay of a free proton according to the scheme $p \rightarrow n + e^+ + \nu$ is forbidden by the law of energy conservation, ac-

cording to the scheme $p \rightarrow e^+ + \nu$ by the law of baryon charge conservation, etc.

To explain the forbiddenness of the single birth of strange particles, Gell-Mann and Nishijima introduced a new quantum number S

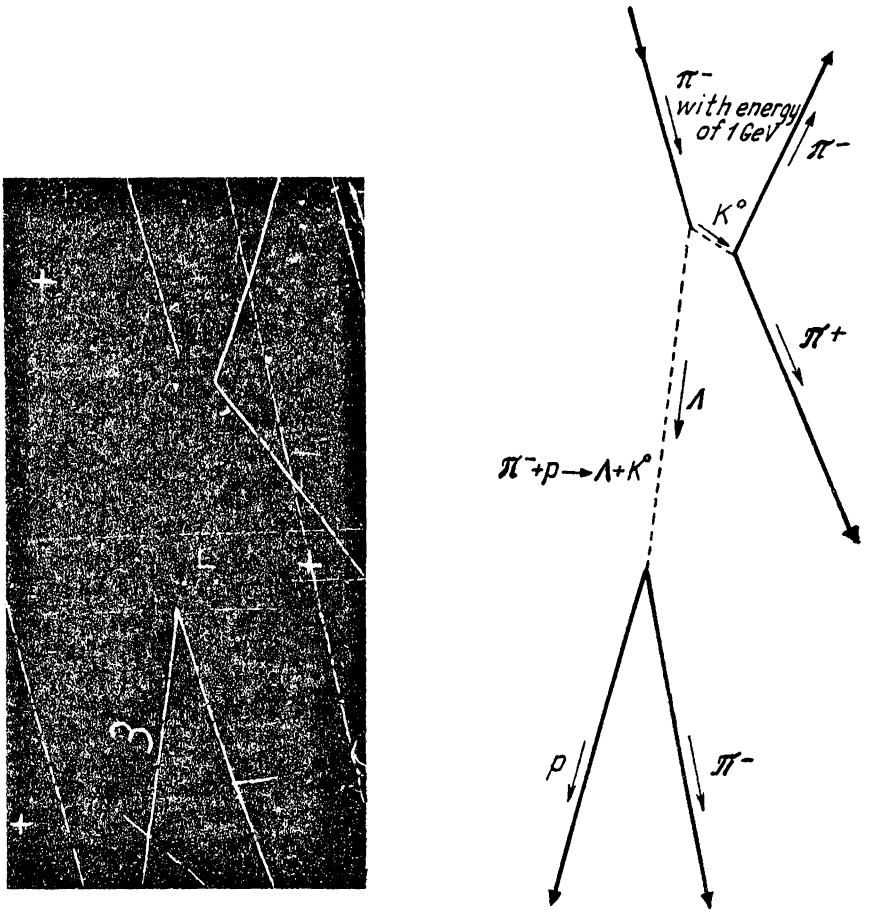


Fig. 11.8

whose total value, according to their assumption, must be conserved in strong interactions. This quantum number was called the **strangeness*** of a particle. In weak interactions, the strangeness may not

* Names of the quantum numbers such as "strangeness" and "charm" (such a quantum number also exists) prove that physicists, as a rule, have a sense of humour. Naturally, terms such as "quantum number No. 1", "quantum number No. 2", etc. could have been introduced instead of such exotic names. But this would be awfully boring.

be conserved. It is therefore ascribed only to strongly interacting particles—mesons and baryons. Thus, for K-mesons, $S = +1$, and for Λ-hyperons, $S = -1$. Consequently, process (11.12) occurs with conservation of the strangeness (the total strangeness of both the initial and the formed particles is zero), whereas in the course of process (11.13) the strangeness changes by unity. Therefore, process (11.13) cannot proceed with the participation of strong interactions.

Gell-Mann and Nishijima related the strangeness to the average electric charge $\langle Q \rangle$ of the particles forming a charge multiplet and to the baryon charge B of a particle:

$$S = 2 \langle Q \rangle - B \tag{11.14}$$

The data of Table 11.3 can be used to find $\langle Q \rangle$ for each of the multiplets, and from Table 11.2 the value of B can be determined for different particles (we remind our reader that for particles $B = +1$, and for antiparticles $B = -1$). It is easy to see that for nucleons, antinucleons, π-mesons, and the η-meson, we get $S = 0$. For example, for nucleons, $\langle Q \rangle = 1/2$, $B = +1$, for antinucleons $\langle Q \rangle = -1/2$, $B = -1$. The substitution of these values in Eq. (11.14) gives $S = 0$ for both cases. Particles with $S = 0$ are conventional, non-strange ones.

At that time, not all K-mesons and hyperons were known. Gell-Mann and Nishijima ascribed such values of the quantum number S to the known strange particles which with the aid of the law of strangeness conservation could explain the features of their birth and decay. This made it possible to establish the possible number of particles in charge multiplets and to predict the existence and properties of new particles. In this way, the Σ^0 and Ξ^0-hyperons and the $\tilde{K}^0$-meson were predicted. They were later discovered experimentally.

The average charge $\langle Q \rangle$ for many multiplets is a half-integer. To avoid dealing with fractions, the quantum number

$$Y = 2 \langle Q \rangle \tag{11.15}$$

was introduced, and it was called the **hypercharge**. According to Eq. (11.14),

$$Y = B + S \tag{11.16}$$

Since the baryon charge is conserved in all interactions, the hypercharge behaves in the same way as the strangeness: it is conserved in strong and electromagnetic interactions and may not be conserved in weak interactions.

It is a simple matter to see that the three quantum numbers $\langle Q \rangle$ Y and S are in essence absolutely equivalent—the value of one of them determines the values of the other two (B is assumed to be known). The most convenient of these three quantum numbers is the hypercharge Y; this is why it is ordinarily used instead of S.

Table 11.4 gives the values of the hypercharge Y, the baryon charge B, and the strangeness S for various charge multiplets.

Table 11.4

Charge multiplet	Composition of multiplet	Y	B	S
π-Mesons	$\pi^+ \pi^0 \pi^-$	0	0	0
K-Mesons	$K^+ K^0$	+1	0	+1
Anti-K-mesons	$K^- \widetilde{K}^0$	−1	0	−1
Eta-meson	η	0	0	0
Nucleons	$p\ n$	+1	+1	0
Antinucleons	$\widetilde{p}\ \widetilde{n}$	−1	−1	0
Λ-Hyperon	Λ	0	+1	−1
Anti-Λ-hyperon	$\widetilde{\Lambda}$	0	−1	+1
Σ-Hyperons	$\Sigma^+ \Sigma^0 \Sigma^-$	0	+1	−1
Anti-Σ-hyperons	$\widetilde{\Sigma}^+ \widetilde{\Sigma}^0 \widetilde{\Sigma}^-$	0	−1	+1
Ξ-Hyperons	$\Xi^- \Xi^0$	−1	+1	−2
Anti-Ξ-hyperons	$\Xi^+ \widetilde{\Xi}^0$	+1	−1	+2
Ω-Hyperon	Ω^-	−2	−1	+3
Anti-Ω-hyperon	$\widetilde{\Omega}^-$	+2	−1	+3

It must be noted that the electric charge Q of a particle can be expressed through the projection of the isotopic spin T_z and the hypercharge Y (or the baryon charge B and the strangeness S):

$$Q = T_z + \frac{Y}{2} = T_z + \frac{B+S}{2} \tag{11.17}$$

We can convince ourselves that this relation is true by using the data of Tables 11.3 and 11.4.

11.7. Non-Conservation of Parity in Weak Interactions

Among the quantities characterizing microparticles, there is a purely quantum-mechanical quantity called the **parity** (P). We know that the state of a particle is described in quantum mechanics by

the function $\psi(x, y, z)$. Let us see how the function ψ can behave in the so-called **inversion of space**, i.e. upon transition to the coordinates x', y', z' associated with x, y, z by the relations

$$x' = -x, \quad y' = -y, \quad z' = -z$$

Examination of Fig. 11.9 shows that such a transformation signifies a transition from a right-handed coordinate system to a left-handed one. The same transition occurs upon reflection in a mirror (Fig. 11.10). Hence, the inversion transformation results in a right-handed reference frame being replaced with a left-handed one. The

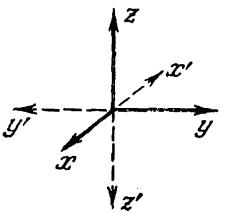


Fig. 11.9

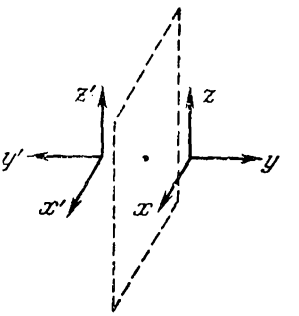


Fig. 11.10

two reference frames x, y, z and x', y', z' differ from each other in the same way as the right-hand and left-hand gloves of a pair do. If we turn right-hand glove, for example, inside out (i.e. subject it to inversion), it will coincide with the left-hand one.

The operation of inversion done twice obviously returns a system of coordinates to its initial form. Assume that the operation of inversion leads to multiplication of the function ψ by a certain number a:

$$\psi(x', y', z') = a\psi(x, y, z)$$

Applying the operation of inversion once more to the expression obtained, we arrive at the function

$$a\psi(x', y', z') = a^2\psi(x, y, z)$$

that must coincide with the initial function $\psi(x, y, z)$. Hence, a^2 must be 1, and a itself may be $+1$ or -1.

It follows from the above that the operation of inversion either leaves the function ψ unchanged or reverses the sign of ψ. In the first case, the state described by the function ψ is called **even**, in the second one **odd**. The behaviour of the function ψ upon inversion depends on the internal properties of the particles described by this function. Particles described by even functions are said to have

positive (even) intrinsic parity ($P = +1$); particles described by odd functions have **negative (odd) intrinsic parity** ($P = -1$). The parity of a system of particles equals the product of the parities of the separate particles in the system.

The **law of parity conservation** follows from quantum mechanics. According to it, in all the transformations experienced by a system of particles, the parity of the state remains unchanged. Conservation of parity signifies invariance of the laws of nature with respect to replacement of right with left (and vice versa).

There was no doubt up to 1956 that the law of parity conservation is observed in all interactions. In 1956, the American physicists Tsung Dao Lee (born 1926) and Chen Ning Yang (born 1922) advanced the assumption that parity may not be conserved in weak interactions. This assumption was based on the following. At that time two mesons designated τ and θ were known. Both mesons were absolutely identical in all respects except one: the τ-meson decayed into three π-mesons, and the θ-meson only into two π-mesons. It could naturally have been assumed that both mesons are the same particle capable of decaying in two different ways. This assumption, however, contradicted the law of parity conservation. The parity of a π-meson $P = -1$. Therefore, the parity of a system of two π-mesons is $(-1)^2 = +1$, and of a system of three π-mesons is $(-1)^3 = -1$. It followed from the law of parity conservation that τ- and θ-mesons differ in their intrinsic parity (for a τ-meson decaying into three π-mesons, $P = -1$, and for a θ-meson decaying into two π-mesons, $P = +1$), and are therefore two different particles.

It was authetically established with time that the τ- and θ-mesons are the same particle now called a K^0-meson, for which $P = -1$. Consequently, the process

$$K^0 \rightarrow \pi^+ + \pi^-$$

occurs with violation of parity.

Lee and Yang proposed the idea of an experiment for verifying the non-conservation of parity that was carried out at the Columbia University (USA) by Chien-Shiung Wu (born 1913) and her collaborators. The idea of the experiment was as follows. If right and left are indistinguishable in nature, then in beta decay the flying out of an electron in the direction of spin of the nucleus and in the opposite direction must be equally probable. Indeed, upon reflection of a nucleus in a mirror, the direction of its "rotation", i.e. the direction of its spin, is reversed (Fig. 11.11). If a nucleus emits beta electrons with equal probability in both directions (Fig. 11.11*a*), then the mirror image of the system nucleus-electrons will be indistinguishable from the system itself (they are only turned relative to each other through 180 degrees). If the beta electrons are emitted mainly in one direction (Fig. 11.11*b*), then "left" and "right" become distinguishable.

In Wu's experiment, nuclei of radioactive cobalt Co^{60} were oriented with their spins in one direction with the aid of a magnetic field. To keep thermal motion from hindering such orientation, the radioactive preparation was cooled to superlow temperatures (~0.1 K). A considerable difference in the numbers of beta electrons emitted in both directions was detected. The beta electrons were found to be emitted mainly in the direction opposite to that of the nuclear spins. It was thus proved experimentally that right and left are not equal in weak interactions (we remind our reader that beta decay is due to weak interaction).

After it had been established that space parity (P) is not conserved in weak interactions, the Soviet physicist Lev Landau and independently of him Lee and Yang advanced the hypothesis that any interactions are invariant relative to a complex transformation consisting in the simultaneous inversion of space and in the replacement of particles with antiparticles. Such a transformation was called **combined inversion**. According to this hypothesis, symmetry between right and left is conserved if in the mirror image of space, particles are replaced with antiparticles. Indeed, if we replace the mirror image of the nucleus in Fig. 11.11*b* with an antinucleus, then the direction of the spin will be reversed and the mirror image of the system will not differ from the system itself.

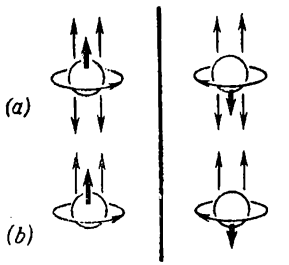


Fig. 11.11

Let us designate the operation of space inversion by the symbol P, and the operation of charge conjugation (i.e. the replacement of particles with antiparticles) by the symbol C. The symbol of combined inversion will therefore be CP. This is why the invariance relative to combined inversion is called **CP-invariance**. The parity of the state of a particle relative to combined inversion is called **combined parity**. Therefore, the two previously existing laws—the law of invariance relative to charge conjugation* and the law of space parity conservation—form a single law of **combined parity conservation** for weak interactions.

The combined parity is indeed conserved in a number of processes in which space parity is violated. In 1964, however, data were obtained in studying the decays of K^0-mesons that point to the violation in these decays of the law of combined parity conservation. We do not have the possibility of delving into the details of this matter.

---

* I.e., the inalterability of the laws of nature when particles are replaced with ant particles.

11.8. The Neutrino

The neutrino is the only particle that does not participate in strong or in electromagnetic interactions. Excluding gravitational interaction, which all particles participate in, a neutrino can take part only in weak interactions.

It was not clear for a long time in what a neutrino differs from an antineutrino. The discovery of the law of combined parity conservation (see the preceding section) made it possible to answer this question: they differ in their **helicity**.

By helicity is meant a definite relation between the directions of the momentum **p** and spin **s** of a particle. The helicity is considered positive if the spin and the momentum have the same direction. In this case, the direction of motion of a particle (**p**) and the direction of "rotation" corresponding to the spin form a right-handed screw (Fig. 11.12*a*). With oppositely directed spin and momentum (Fig. 11.12*b*), the helicity will be negative (the forward motion and "rotation" form a left-handed screw). It is obvious that the helicity can be determined as the sign of the scalar product **sp**.

Helicity can have an absolute value, i.e. be an intrinsic property, only for a particle with a zero rest mass (such a particle exists only when travelling with the speed c). A particle whose rest mass differs from zero will travel with a speed v less than c. The helicity of such a particle in reference frames travelling with speeds less than v and with speeds exceeding v (but less than c) will be different (the momentum of a particle in such reference frames has opposite directions). Thus, of all particles, only a neutrino can have helicity as an intrinsic property*.

According to the **theory of the longitudinal neutrino** developed by Yang and Lee, Landau, and also by the Indian physicist Abdus Salam (born 1926), all neutrinos existing in nature, regardless of how they were produced, are always completely longitudinally polarized (i.e. their spin is either parallel or antiparallel to the momentum **p**). A neutrino has a negative (left-handed) helicity (the directions of **s** and **p** shown in Fig. 11.12*b* correspond to it) and an antineutrino has a positive (right-handed) helicity (Fig. 11.12*a*). Thus, the helicity is what distinguishes a neutrino from an antineutrino.

Upon reflection in a mirror, a right-handed helix transforms into a left-handed one. Hence, the existence of helicity in neutrinos contradicts the law of space parity conservation (the particle does not coincide with its image). But if simultaneously with reflection in a

* A photon also has a zero rest mass, but unlike a neutrino, the two helicity values obtained for a photon (positive and negative) correspond not to a particle and antiparticle, but to two different states of polarization of the same particle.

mirror we replace a neutrino (having left-handed helicity) with an antineutrino (having right-handed helicity), then the requirements of the law of combined parity conservation will be observed.

That a neutrino has helicity is observed in the chain of transformations $\pi \to \mu \to e$. At the end of its path, a π^+-meson decays into a muon and a neutrino:

$$\pi^+ \to \mu^+ + \nu$$

The spin of a π^+-meson is zero, the momentum at the end of its path also vanishes. Therefore, the muon and the neutrino must fly away

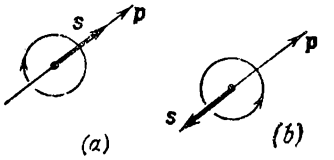


Fig. 11.12

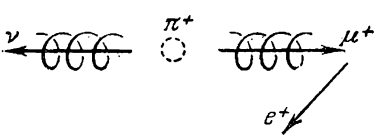


Fig. 11.13

in opposite directions, the neutrino "imposing" its helicity on the muon* (Fig. 11.13); otherwise the spin of the system will not remain equal to zero.

The muon at the end of its path decays as follows:

$$\mu^+ \to e^+ + \nu + \tilde{\nu}$$

Since here we have to do with the decay of polarized muons, the same phenomenon should be observed in their decay as in the beta decay of polarized nuclei (in Wu's experiment)—the angular distribution of the positrons should be anisotropic relative to the direction of polarization of the muon, i.e. relative to the direction of its motion before stopping. Indeed, studying of the photographs registering the processes of $\pi \to \mu \to e$ decay in a bubble chamber shows that positrons are most often emitted in a direction opposite to that of motion of the muons (see Fig. 11.13).

The hypothesis on the existence of neutrinos was advanced in 1932. During the following quarter of a century, numerous indirect proofs of the correctness of this hypothesis were obtained, but no-one succeeded in directly observing neutrinos. The reason is that neutrinos do not have an electric charge and mass and therefore interact extremely weakly with a substance. For example, a neutrino with an energy of about 1 MeV has a path of about 10^{20} cm or 100 light years in lead. Only after the development of nuclear reactors, which

* The spin of a muon is usually not fixed relative to the direction of its motion.

are sources of powerful streams of neutrinos [about 10^{13} part/(cm^2·s)], did the possibility appear of observing reactions with the participation of these elusive particles.

Antineutrinos were directly observed in a series of experiments run in 1953-1956 by the American physicists F. Reines and C. Cowan, Jr. The reaction

$$\tilde{\nu} + p \rightarrow n + e^+ \qquad (11.18)$$

was observed, which is in essence an inversion of neutron decay reaction (10.7)*. That an antineutrino entered into a reaction with a proton is indicated by the simultaneous appearance of a neutron

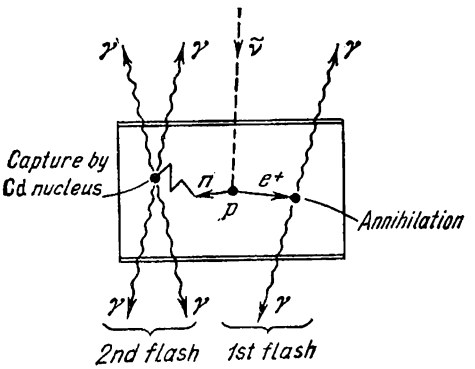


Fig. 11.14

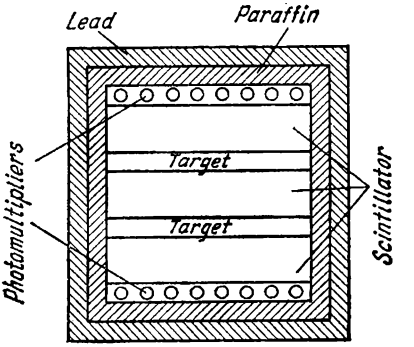


Fig. 11.15

and a positron (Fig. 11.14). The positron virtually immediately annihilated with an electron, which resulted in the production of two gamma quanta each having an energy of 0.51 MeV. The neutron after retardation was captured by a cadmium nucleus. The excited nucleus formed as a result emitted luminescently several gamma photons with a total energy of 9.1 MeV.

The installation is shown schematically in Fig. 11.15. Two tanks (190 × 130 × 7 cm) filled with an aqueous solution of cadmium chloride were the target. Three tanks (190 × 130 × 60 cm) were filled with a liquid capable of scintillation under the action of gamma photons. The scintillation flashes were registered by 110 photomultipliers. The tanks were confined in a paraffin and then in a lead shield for protection against cosmic radiation and against neutrons

* The inversion of the reaction of neutron decay in the literal meaning of the word would be the reaction $\tilde{\nu} + p + e^- \rightarrow n$, but such a reaction requires the meeting of three particles and is therefore practically impossible. The "subtract on" of a particle is equivalent to the addition of an antiparticle; by subtracting e^- from the left-hand side and adding e^+ to the right-hand one, we get reaction (11.18).

leaving the reactor. The entire installation was embedded deep in the earth near a large reactor. The scintillation flash produced by the captured gamma photons lagged behind the flash caused by annihilation gamma photons by several scores of microseconds. Both flashes were registered according to a delayed coincidence circuit. In addition, the energy of the gamma photons producing each flash (1.02 MeV and 9.1 MeV) was also assessed. This made it possible to reliably separate the effect being studied from the background due to other processes. The experiment lasted 1371 hours (57 days). Every hour, an average of about three double flashes of the expected intensity were registered. These results are a direct proof of the existence of antineutrinos.

In some processes, a neutrino (or antineutrino) appears together with an electron (positron), in others together with a muon (examples can be found in Table 11.2). It was assumed for a long time that the former (electron) neutrinos ν_e are identical with the latter (muon) neutrinos ν_μ. In 1962, it was proved experimentally that this is not correct. The idea of the experiment belongs to Bruno Pontecorvo (born 1913). The inversion of reaction (10.28) will be the process

$$\nu_e + n \rightarrow p + e^- \qquad (11.19)$$

(see the footnote on page 292). A similar process is possible in which a muon appears instead of an electron

$$\nu_\mu + n \rightarrow p + \mu^- \qquad (11.20)$$

(the particle participating in this reaction must obviously be a muon neutrino, and not an electron one). Pontecorvo proposed to irradiate the substance with the muon neutrinos formed in the decay $\pi^+ \rightarrow \mu^+ + \nu_\mu$ and observe the particles produced. The presence of both e^- and μ^- among them would indicate that ν_e and ν_μ are identical. The presence of only μ^- would indicate that electron and muon neutrinos differ.

The experiment was carried out by the American physicists L. Lederman, M. Schwartz and others at Brookhaven (USA). The accelerator produced π^+-mesons with an energy of 15 GeV. The process of π-μ-decay [see (10.14)] resulted in the formation of muon neutrinos with an energy of about 500 MeV. The stream of these neutrinos was directed into a spark chamber with massive iron plates (with a total mass of 10 tonnes). During 800 hours, 51 cases of the birth of muons were registered, and not a single case of the birth of electrons. This result proves the existence of four different neutrinos: ν_e, $\tilde{\nu}_e$, ν_μ, $\tilde{\nu}_\mu$.

In connection with the need to distinguish electron and muon neutrinos, the symbol of the neutrino in formulas (10.7), (10.26), (10.27) and (10.28) must be supplemented with the subscript "e", and in

formulas (10.14) with the subscript "μ". Formulas (10.16) must be written as follows:

$$\mu^- \to e^- + \tilde{\nu}_e + \nu_\mu; \qquad \mu^+ \to e^+ + \nu_e + \tilde{\nu}_\mu$$

11.9. Systematization of Elementary Particles

The regularities observed in the world of elementary particles can be formulated as conservation laws. Quite a lot of such laws have accumulated (see Table 11.5). Some of them are not accurate, but

Table 11.5

Law of conservation of	Kind of interaction		
	strong	electro-magnetic	weak
energy E	+	+	+
momentum **p**	+	+	+
angular momentum **M**	+	+	+
electric charge Q	+	+	+
baryon charge B	+	+	+
lepton charge L	+	+	+
isotopic spin T	+	−	−
hypercharge Y (or strangeness S)	+	+	−
charge conjugation C	+	+	−
parity P	+	+	−
combined parity CP	+	+	−

only approximate. For example, the law of hypercharge Y (or strangeness S) conservation is obeyed for strong and electromagnetic interactions and is violated for weak interactions (observance of a law in a given kind of interaction is indicated in Table 11.5 by a plus sign, and violation by a minus sign).

Every conservation law expresses a definite symmetry of a system. The laws of conservation of momentum **p**, angular momentum **M**, and energy E reflect the properties of symmetry of space and time: the conservation of E is a result of the uniformity of time, the conservation of **p** is due to the uniformity of space, and the conservation of **M** to its isotropy. The law of parity conservation is associated with symmetry between right and left (P is the invariance). Symmetry relative to the charge conjugation (symmetry of particles and antiparticles) leads to conservation of the charge parity (C-invariance). The laws of conservation of the electric, baryon, and lepton charges express the special symmetry of the ψ-function. Finally, the law of

isotopic spin conservation reflects the isotropy of isotopic space. The failure to observe one of the conservation laws signifies violation of the corresponding kind of symmetry in the given interaction. For example, electromagnetic interaction violates the symmetry of isotopic space, owing to which the isotopic spin T is not conserved in electromagnetic interactions.

The introduction of the isotopic spin made it possible to combine particles into charge multiplets (see Sec. 11.5). Extension of the scheme of isotopic spin led Gell-Mann and independently of him Yu. Ne'eman to the creation in 1961 of the **theory of unitary symmetry**. It is assumed in this theory that strong interaction is invariant relative to special transformations* in a certain three-dimensional complex vector space (the space of **unitary spin**) that keep the isotopic spin T and the hypercharge Y unchanged. In this way, it becomes possible to group charge multiplets into **supermultiplets** (or **unitary multiplets**). The system of symmetry of particles established by the unitary theory is also called the **eightfold way**.

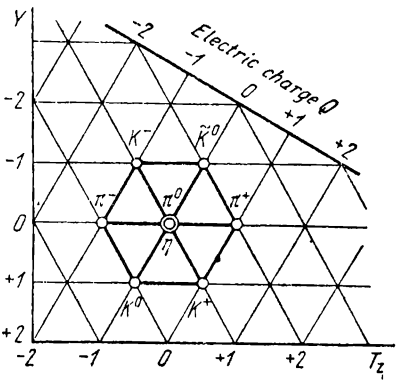


Fig. 11.16

The particles** forming a supermultiplet must have the same spin and the same parity P. They may differ in mass, electric charge, hypercharge, and isotopic spin, but these quantities must be related to one another by definite rules.

Figure 11.16 depicts an octet (a supermultiplet including eight particles) combining mesons (except for resonances). All of them have a spin equal to zero and a negative parity. The hypercharge Y is laid off along the vertical axis, the projection of the isotopic spin T_z (the third component of the isotopic spin) along the horizontal axis, and the electric charge Q along the inclined axis. The particles of the meson octet are arranged at the apices and at the centre of a regular hexagon. The centre accommodates two particles: π^0 and η. At the time when the theory of unitary symmetry was developed, only seven mesons (except for η) were known. In accordance with the conclusions of this theory, the existence of an eighth meson and its properties were predicted. In 1961, the predicted meson (η) was discovered, and its properties agreed quite well with the theoretical predictions.

---

* Belonging to the so-called $SU(3)$ group.

** We have in mind only strongly interacting particles.

Figure 11.17 shows an octet of long-lived baryons. All the particles have a spin of 1/2 and a positive parity. The centre of the hexagon accommodates the hyperons Σ^0 and Λ.

Finally, Fig. 11.18 shows a baryon decuplet (a supermultiplet combining 10 particles). It includes nine resonances and a long-lived "genuine" particle—an Ω^--hyperon. The spin of all the particles is 3/2, the parity is positive. The particles are arranged in the diagram in the form of an equilateral triangle. The mass m of the particles and the mass difference Δm expressed in MeV are indicated at the right

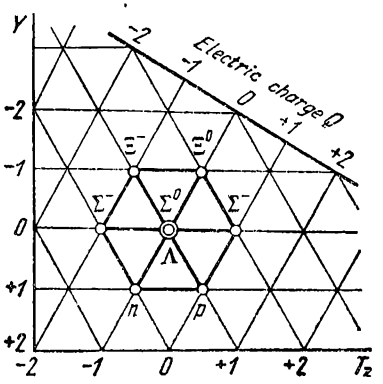


Fig. 11.17

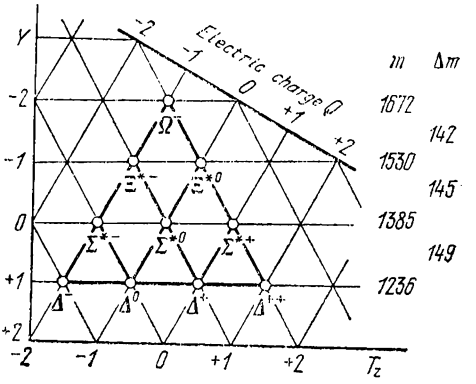


Fig. 11.18

of the figure. It is worthy of note that when passing from one group of particles to another, the mass changes by almost the same amount (about 145 MeV).

At the time when the theory was created, Ξ^*-hyperons and the Ω^--particle were not yet known. The resonances Ξ^{*-} and Ξ^{*0} were discovered in 1962. The apex of the pyramid remained unfilled. Gell-Mann predicted that the particle corresponding to it should have a spin of 3/2, a hypercharge of $Y = -2$, and a mass of about 1675 MeV (greater by 145 MeV than the mass of a Ξ^*-particle). Almost immediately systematic searches of this particle were begun, the particle being called the Ω^--hyperon. At the Brookhaven laboratory, an accelerator for 33 GeV and a two-metre bubble chamber containing 900 litres of liquid hydrogen were used for this purpose. About 300 000 photographs were made before the process of the birth and decay of an Ω^--particle was recorded on one of them in January, 1964. Its properties, in particular its mass, exactly coincided with those predicted by theory. Thus, the discovery of the Ω^--hyperon was a triumph of the theory of unitary symmetry.

The question arises as to why a "genuine" particle, an Ω^--hyperon, which lives approximately 10^{-10} s, got into one decuplet with resonances whose lifetime is of the order of 10^{-23} s. The reason for such

a "vitality" of the Ω^--hyperon is that its hypercharge is -2 (its strangeness $S = -3$). As a result, the Ω^--hyperon cannot decay at the expense of strong interactions, with whose participation the remaining particles of the decuplet decay.

The resonances in the decuplet decay as follows:

$$\Delta \to N + \pi, \quad \Delta \to N + \pi^+ + \pi^-, \quad \Delta \to N + \gamma$$
$$\Sigma^* \to \Lambda + \pi, \quad \Sigma^* \to \Sigma + \pi$$
$$\Xi^* \to \Xi + \pi$$

where N is a nucleon, and Λ, Σ, and Ξ are the corresponding hyperons. In the course of all these processes, the hypercharge (and, therefore, the strangeness) is conserved (this can easily be seen by comparing the diagrams shown in Figs. 11.17 and 11.18). Consequently, the decays occur at the expense of strong interactions with a characteristic time of about 10^{-23} s.

Conservation of the hypercharge (strangeness) of an Ω^--hyperon could occur upon its decay into two or more strange particles. Such processes, in which the electric and the baryon charges are also conserved in addition to the hypercharge, include:

$$\Omega^- \to \Xi^- + \tilde{K}^0$$
$$\Omega^- \to \Sigma^- + \tilde{K}^0 + \tilde{K}^0$$
$$\Omega^- \to n + K^- + \tilde{K}^0 + \tilde{K}^0$$
$$\Omega^- \to p + K^- + K^- + \tilde{K}^0$$

These processes, however, are forbidden by the law of energy conservation. Thus, an Ω^--hyperon can decay only by violating the law of hypercharge conservation, i.e. at the expense of weak interactions. Accordingly, its lifetime is 10^{-10} s. Decay occurs in one of the following ways:

$$\Omega^- \to \Xi^0 + \pi^-, \quad \Omega^- \to \Xi^- + \pi^0, \quad \Omega^- \to \Lambda + K^-$$

The spin of an Ω^--hyperon is 3/2, that of a Ξ- and Λ-hyperons is 1/2, and the spin of π- and K-mesons is zero. The law of angular momentum conservation is not violated, however, because the formed pair of particles has an orbital angular momentum equal to 1. Consequently, the total angular momentum of these particles is 3/2.

11.10. Quarks

The number of particles called elementary has become so great that serious doubts concerning their actually being elementary have appeared. Each of the strongly interacting particles is characterized

by three independent additive quantum numbers: the charge Q, hypercharge Y, and baryon charge B. In this connection, a hypothesis was advanced that all particles are built up of three fundamental particles—the carriers of these charges. The first model of such a kind was proposed by the Japanese physicist Shoichi Sakata (born 1911). He considered the proton p, neutron n, and Λ-hyperon to be the fundamental particles*. Sakata's model, however, was found to be inapplicable to the field of strong interactions.

In 1964, Gell-Mann advanced the hypothesis that all elementary particles are built up of three particles which he named **quarks****. Fractional quantum numbers are ascribed to these particles, in particular an electric charge equal to $+2/3$, $-1/3$, $-1/3$, respectively, for each of the three quarks. Quarks are usually designated by the letters P, N, and Λ (other symbols are also used). In addition to quarks, antiquarks ($\widetilde{P}$, $\widetilde{N}$, $\widetilde{\Lambda}$) are also considered. The properties assigned to quarks are indicated in Table 11.6.

Mesons are formed from a quark-antiquark pair, and baryons from three quarks. Table 11.7 gives some of these formations. The letter Λ in the first column of this table signifies a Λ-hyperon, and the same letter in the second column signifies a Λ-quark.

An identical magnetic moment μ_{qk} is ascribed to each quark. Its magnitude is not determined from theory. Calculations performed on the basis of such an assumption give the value of the magnetic moment of $\mu_p = \mu_{qk}$ for a proton, and $\mu_n = -\frac{2}{3}\mu_{qk}$ for a neutron.

Table 11.6

Quark	Electric charge, Q	Baryon charge, B	Spin	Isotopic spin, T	Strangeness, S
P	$+2/3$	$+1/3$	1/2	1/2	0
N	$-1/3$	$+1/3$	1/2	1/2	0
Λ	$-1/3$	$+1/3$	1/2	0	-1
$\widetilde{P}$	$-2/3$	$-1/3$	1/2	1/2	0
$\widetilde{N}$	$+1/3$	$-1/3$	1/2	1/2	0
$\widetilde{\Lambda}$	$+1/3$	$-1/3$	1/2	0	$+1$

* The fact that many particles have a mass considerably smaller than the sum of the masses of p, n, and Λ should not confuse us because the mass of a system of bound particles may be much smaller than the sum of the masses of the particles in the system (compare with the binding energy of particles in a nucleus, Sec. 10.2).

** Gell-Mann took the name "quark" from James Joyce's science-fiction novel "Finnegan s Wake", from the line "three quarks for Mr. Marks".

Table 11.7

Particle	Composition	Mutual orientation of quark spins	Mutual "orientation" of quark isotopic spins
π^+	$P\tilde{N}$	↑↓	↑↑
π^-	$\tilde{P}N$	↑↓	↑↑
K^+	$P\tilde{\Lambda}$	↑↓	↑
p	PPN	↑↓↑	↑↓↑
n	PNN	↑↓↑	↑↓↑
Σ^+	$PP\Lambda$	↑↓↑	↑↑
Λ	$PN\Lambda$	↑↓↑	↑↓
Δ^{++}	PPP	↑↑↑	↑↑↑
Δ^-	NNN	↑↑↑	↑↑↑
Ω^-	$\Lambda\Lambda\Lambda$	↑↑↑	—

Thus, for the ratio of the magnetic moments, we get the value

$$\frac{\mu_p}{\mu_n} = -\frac{3}{2}$$

which excellently agrees with the experimental value (see Sec. 10.1).

It later became necessary to extend the system of quarks. The reason for this, in particular, was the fact that bound states of three quarks such as PPP (Δ^{++}), NNN (Δ^-), and $\Lambda\Lambda\Lambda$ (Ω^-) contradict the Pauli principle. Indeed, inspection of Table 11.8 reveals that all

Table 11.8

Quark	Electric charge, Q	Baryon charge, B	Strangeness, S	Charm, C	Colour
P	+2/3	1/3	0	0	Red, yellow, blue
N	−1/3	1/3	0	0	Ditto
Λ	−1/3	1/3	−1	0	Ditto
C	+2/3	1/3	0	1	Ditto

the quantum numbers of quarks in these formations are the same. But since the spin of quarks is 1/2, one system cannot contain not only three, but even two quarks with the same set of quantum numbers.

For a number of considerations, in particular to eliminate the contradiction with the Pauli principle, the concept of the "colour" of a quark was introduced. Physicists began to say that each quark can

exist in three "coloured" forms: red, yellow, and light blue (we must note that a mixture of these colours gives the "neutral" white colour). Therefore, let us say, the Λ-quarks forming the Ω^--hyperon have different colours, and the Pauli principle is not violated.

The combination of the colours of quarks in hadrons must be such that the average colour of a hadron is neutral. For example, the composition of a proton includes the quarks P (red), P (yellow), and N (light blue). The sum gives the neutral (white) colour.

Antiquarks are considered to have the anticolour that together with the colour gives neutrality. Accordingly, mesons, consisting of a quark and an antiquark, also have a neutral colour.

Basically, however, the colour of a quark (like the sign of an electric charge) began to express the difference in its property determining the mutual attraction and repulsion of quarks. By analogy with the quanta of fields of different interactions (photons in electromagnetic interactions, π-mesons in strong interactions, etc.), particles were introduced that are carriers of interaction between quarks. These particles were named **gluons** (from the word "glue"). They transfer colour from one quark to another, as a result of which the quarks are kept together.

In 1974-1975, particles (resonances) with enormous masses of 3.1, 3.7 and 4.1 GeV (from three to four nucleon masses) were discovered in powerful accelerators in various laboratories of the world. Thus, a new family of strongly interacting ψ-particles was discovered. This discovery confirmed the earlier proposed model of particles consisting of four quarks. In addition to the P-, N-, and Λ-quarks mentioned above, a fourth "charmed" C-quark figures in this model. It differs from the other quarks in that the quantum number C*, called "charm", equals 1 for it, whereas it is zero for the other quarks. The properties of all four quarks are given in Table 11.8 (only the quarks are indicated, but there are also four antiquarks).

The C-quark does not enter the composition of ordinary "uncharmed" particles (mesons and baryons). The structure $C\bar{C}$ is ascribed to the recently discovered ψ-particles. The correspondingly named "charm" quantum number C is zero for these particles. The ψ-particles are said to have a concealed charm.

Theory predicts the existence of charmed particles, i.e. of particles having a quantum number C other than zero. The following structures are examples: $\Lambda\bar{C}$ (the charm $C = +1$), and ΛCC ($C = +2$). Such charmed particles have meanwhile not been observed.

The hypothesis of quarks was quite fruitful. It made it possible to predict new particles in addition to systematizing the already

---

* The symbol C is used to designate both a charmed quark and the quantum number called the charm.

known ones. In particular, the existence of Ω^--hyperons was predicted with the aid of the quark model. The quark hypothesis also made it possible to explain many properties of particles and relate different processes to one another. It is quite natural that attempts were made to discover quarks. Several sensational reports appeared on the experimental discovery of quarks. But these reports were not confirmed subsequently. To date, quarks have not been observed, and their existence is problematic.

11.11. Conclusion

The situation in the physics of elementary particles reminds one of the situation in the physics of the atom after Dmitri Mendeleev discovered the periodic law in 1869. Although the essence of this law was determined only when about 60 years had passed, after the advent of quantum mechanics, the law made it possible to systematize the chemical elements known at that time and, in addition, led to the prediction of the existence of new elements and their properties. In exactly the same way, physicists have learned how to systematize elementary particles, and in a number of cases this has made it possible to predict the existence of new particles and anticipate their properties.

The establishment of a classification of elementary particles, however, "... will not at all solve the fundamental problem of understanding all the laws of the microworld. This understanding will evidently arrive only when a new physical theory will be developed... At present, we are approaching a new state in the cognition of the fundamental laws of structure of nature, from which the quantum theory, the theory of relativity, and Newton's theory should follow as a particular case of the general one... We are not able to predict when and how a new comprehensive physical theory will be created... But the fact that an enormous army of experimenters and theoreticians all over the world are working on this front line for physics allows us to hope that this time is not far distant."

The quotation belongs to Academician Igor Tamm. It will complete our tale of the physics of elementary particles.

APPENDIX

List of Symbols

A amplitude; mass number of nucleus; work; work function
a absorptivity; amplitude
$\mathbf{a}$ acceleration
B baryon charge
$\mathbf{B}$ magnetic induction
b constant in Wien's displacement law; impact parameter
C charge conjugation; charm; charmed quark; constant; heat capacity
c speed of light
d dimension; distance
E energy
E_F Fermi level
$\mathbf{E}$ electric field strength
$\mathscr{E}$ electric field strength;electromotive force
e base of natural logarithms; electron; elementary charge
F quantum number of angular momentum of atom
$\mathbf{F}$ force
f function
g density of states; Landé g factor
$\mathbf{g}$ acceleration of free fall
$\hat{H}$ Hamiltonian operator
h Planck's constant
$\hbar$ Planck's constant h divided by 2π
I current; intensity of light; moment of inertia; nuclear spin quantum number
i imaginary unity
J quantum number of angular momentum of electron shell
j density of energy flux; quantum number of angular momentum of electron
$\mathbf{j}$ current density
K K-meson (kaon)
k wave number
$\mathbf{k}$ wave vector
L azimuthal (orbital) quantum number of electron shell; lepton charge
l azimuthal (orbital) quantum number of electron

M angular momentum (in Vols. I and II the symbol **L** was used for the angular momentum)
m magnetic quantum number; mass
m_s spin quantum number
m^* effective mass
N number; quark
n integer; neutron; number; principal quantum number
P parity; probability; quark; radiant power; space inversion
$\mathcal{P}$ pressure of light
p proton
p momentum
Q amount of heat; charge of a particle
q charge
R average recoil energy of atom; distance; radiant emittance; radius; reflection coefficient; Rydberg constant
r distance; emissivity
r_0 Bohr radius
S area; spin quantum number of electron shell; slope of characteristic; strangeness
s spin quantum number of electron
T absolute temperature; half-life; isotopic spin; term; transmission coefficient
t time
U internal energy; potential energy; voltage
u radiant energy density
V volume
v vibrational quantum number
v velocity
w energy density
x Cartesian coordinate
Y hypercharge
y Cartesian coordinate
Z charge number
z Cartesian coordinate
α alpha particle; fine structure constant; initial phase of oscillations; Rydberg correction; thermoelectric coefficient
β beta particle
Γ breadth of energy level
γ gamma radiation; photon
Δ increment; mass defect
δ increment
ε energy
η eta-meson
Θ Debye characteristic temperature
θ angle
$\varkappa$ wave absorption coefficient

Λ lambda hyperon; quark
λ decay constant; wavelength
μ chemical potential; mu-meson (muon); permeability
μ_B Bohr magneton
μ magnetic moment
ν neutrino
ν' wave number
Ξ ksi hyperon
ξ displacement of medium particles in wave
Π Peltier coefficient
π pi-meson (pion)
ρ resistivity
Σ sigma hyperon
σ conductivity; shielding factor; Stefan-Boltzmann constant
τ Thomson coefficient; time
Φ flux; magnetic flux
φ angle; function; polar angle; potential
Ψ psi-function
ψ angle; psi-function
Ω omega hyperon; solid angle
ω angular velocity; cyclic frequency
$\boldsymbol{\omega}$ angular velocity

NAME INDEX

Alvarez, L. W., 250
Anderson, C. D., 241, 274

Balmer, J. J., 48
Bardeen, J., 198
Barnett, S. J., 118
Basov, N. G., 150
Becquerel, A. H., 245
Bethe, H. A., 263
Biberman, L. M., 68
Bloch, F., 187
Boguslavsky, S. A., 211
Bohr, N., 57, 60, 61, 62, 63, 95, 237, 252
Boltzmann, L., 19
Born, M., 78
Bothe, W. W. G. F., 40

Chadwick, J., 231
Chamberlain, O., 276
Cockcroft, J. D., 254
Collins, G. B., 197
Compton, A. H., 44, 47
Cooper, L. N., 198
Cork, B., 277
Cowan, C. L., Jr., 292
Cranshaw, T. E., 270
Curie, M., 233, *see also* Sklodowska-Curie, M.
Curie, P., 233, 245

Davisson, C. J., 65
Deaver, B. S., 200
De Beer, J. F., 270
De Broglie, L. V., 65, 74, 76
Debye, P. J. W., 161, 162, 164
De Haas, W. A., 118
Dirac, P. A. M., 110, 272, 273, 274, 275
Doll, R., 200

Einstein, A., 38, 40, 118, 148, 149, 150, 161, 233

Fabrikant, V. A., 68, 150
Fairbank, W. M., 200
Fermi, E., 233, 249, 260
Feynman, R. P., 68
Flerov, G. N., 233, 250, 251
Franck, J., 57
Frenkel, Ya. I., 237
Frisch, O. R., 256

Gell-Mann, M., 283, 284, 285, 295, 296, 298
Gerlach, W., 85
Germer, L. H., 65
Glaser, D. A., 269
Goeppert-Mayer, M., 238
Goudsmit, S. A., 109

Hahn, O., 256
Heisenberg, W., 71, 281
Heitler, W. H., 139
Hertz, G., 58
Hertz, H., 36

Irodov, I. E., 55

Javan, A., 154
Jeans, J. H., 28, 29
Joffe, A. F., 223
Joyce, J., 298

Kamerlingh Onnes, H., 197
Kelvin, Lord (W. Thomson), 223
Kirchhoff, G. R., 14
Kittel, C., 182
Kurchatov, I. V., 233, 260, 261

Lambertson, G., 277
Landau, L. D., 286
Landsberg, G. S., 146
Langmuir, I., 211

Lattes, C. M. G., 241
Lawrence, E. O., 233
Lederman, L. M., 293
Lee, T. D., 288, 289, 290
Leighton, R. B., 68
Lenard, P. E. A., 37
London, F., 139
Lorentz, H. A., 123
Lukirsky, P. I., 39

Mandelshtam, L. I., 146
Meiman, T., 150, 152
Meissner, W., 197
Meitner, L., 256
Mendeleev, D. I., 131, 233, 301
Millikan, R. A., 38, 39
Moseley, H. G. J., 137, 138
Mössbauer, R. L., 173
Mysovsky, L. V., 270

Näbauer, M., 200
Neddermeyer, S. H., 241
Ne'eman, Y., 295
Newton, I., 301
Nishijima, K., 283, 284, 285

Occhialini, G., 241
Ochsenfeld, R., 197

Pauli, W., 130, 249
Peltier, J. C. A., 222
Petrzhak, K. A., 250
Piccioni, O., 277
Planck, M., 30, 38, 40, 60, 94
Pontecorvo, B. M., 293
Pound, R., 175, 176
Powell, C. F., 241
Prilezhaev, S. S., 39
Prokhorov, A. M., 150

Raman, C. V., 146
Rayleigh, Lord, 28, 29
Rebka, G., Jr., 175, 176
Reines, F., 292
Rutherford, E., 52, 53, 56, 57, 254
Rydberg, J. R., 48, 104

Sakata, S., 298
Salam, A., 290
Sands, M., 68
Schrieffer, J. R., 198
Schrödinger, E., 74, 75
Schwartz, M., 293
Seebeck, T. J., 218
Segré, E., 276
Sklodowska-Curie, M., 245
Skobeltsyn, D. V., 269
Stefan, J., 19
Stern, O., 67, 85
Stoletov, A. G., 36, 37
Strassmann, F., 256
Strutt, J. W., *see* Rayleigh, Lord
Sushkin, N. G., 68

Tamm, I. E., 240, 301
Tartakovsky, P. S., 67
Thomson, G. P., 67
Thomson, J. J., 37, 51
Thomson, W. (Lord Kelvin), 223
Townes, C. H., 150

Uhlenbeck, G. E., 109

Vavilov, S. I., 44

Walton, E. T. S., 254
Weber, J., 150
Wenzel, W. A., 277
Wiegand, C. E., 276
Wien, W., 19, 29
Wilson, C. T. R., 268
Wood, R. W., 170
Wu, C. S., 288

Yang, C. N., 288, 289, 290
Ypsilantis, T., 276
Yukawa, H., 241

Zavoisky, Y. K., 127
Zeeman, P., 122
Zhdanov, A. P., 270

SUBJECT INDEX

Absorption,
  multiple-photon, 156f
  resonance, 170, 174
    gamma rays, 172f
Absorptivity, 14
  blackbody, 15
  and emissivity, 14f
Acceptors, 206, 224
Activity, radioactive substance, 251
  units, 251
Alpha particle(s), 52, 238
  impact parameter, 53
Angular momentum,
  atom, 103, 115f, 234
  atomic residue, 103
  direction, 85
  electron, 110, 117
  mechanical, and magnetic moment, 109, 110
  microparticle system, 86
  operators, 84
  orbital, 117
  quantization, 83ff
  summation, 111
Annihilation,
  particle-antiparticle, 277
  positron-electron pair, 274, 276
Anode, 211
Antineutrino, 232, 241, 249, 280, 290ff
Antineutrons, 277
Antiparticles, 277ff
  decay scheme, 280
Antiproton, 277
Antiquarks, 300
Atom(s),
  angular momentum, 103
    alkali metal, 103
    resultant, 115f
    total, 234
  electron configuration, 131ff
  energy, 116
  energy levels, *see* Energy level(s)
  excitation, 137
  excited states, 105
  ground state, 105
Atom(s),
  hydrogen, 95ff
    Bohr's theory, 62ff
    electron distance from nucleus, 100f
    energy, 96
    energy levels, 63
  internal energy, 62f
    allowed values, 63
  light absorption, 170
  light emission, 170
  magnetic moment, 118
  metastable states, 105
  model,
    nuclear, 52ff
    Rutherford's, 52ff
    Thomson's, 51
    vector, 119f
  nucleus, *see* Nucleus(i)
  recoil energy, 106
    average, 107
  size assessment, 51
  term, 116f

Bands,
  allowed, 186, 187
    width, 188
  conduction, 189
    in semiconductor, 201
  energy level, 186f
  forbidden, 186
    width, 203
  spectrum,
    edge, 142f
    electron-vibrational, 142f
    rotational, 143f
    vibrational-rotational, 143, 144f
  valence, 188f
    in semiconductor, 200f
Barn, 254
Baryons, 266, 267, 279, 285
  charge, 277, 285f
  number, 277
Battery, solar, 230
Belts, Earth's radiation, 271f

Blackbody, 14, 15
absorptivity, 15
emissivity, 15f, 29
and equilibrium radiation density, 42f
radiant emittance, 16, 19, 33
Body, gray, 14
Bomb,
atomic, 237, 259, 262
hydrogen, 237, 262
thermonuclear, 262
Bosons, 170, 198, 267, 277
Breadth,
energy level, 105
spectral line, 105, 108
Doppler, 107f
natural, 106
Breeders, 261
Bremsstrahlung, 35, 136
spectrum, short wavelength limit, 34, 35f
Broadening, Doppler, spectral line, 106, 107

Capture,
e^-, 248
electron, 248, 250
K, 248, 250
L, 248
M, 248
radioactive, 257f
Cascades, electron-positron pairs, 272
Catastrophe, ultraviolet, 29
Cathode(s), 211
oxide, 212
Cathodoluminescence, 11
Chamber(s),
bubble, 269f, 296
diffusion, 269
emulsion, 270f
ionization, 268
spark, 270, 293
streamer, 270
Wilson, cloud, 268f
Charge,
baryon, 277, 285f, 298
conjugation, 278, 289
lepton, 278
Charm, 300
Chemiluminescence, 11, 12
Coefficient,
absorption, negative, 152
Einstein's, 149
Peltier, 22
and specific thermal e.m.f., 222
reflection, from potential barrier, 90
Coefficient,
thermoelectric, 221
Thomson, 223f
transmission, through potential barrier, 90, 92
Conductance, electrical, 194ff
Conductivity, electrical, 196
Conjugation, charge, 278, 289
Constant,
coupling, 265f
decay, 244
electron coupling with electromagnetic field, 115
fine structure, 114f
Planck's, 30, 86
determination, 36, 39
Rydberg, 48, 49, 63
Stefan-Boltzmann, 19, 33
Wien's displacement law, 33
Conversion, internal, 247
Coordinate(s),
generalized, 61
normal, 162
principal, 162
Correction, Rydberg, 104
Counters,
anticoincidence circuit, 268
Cerenkov, 268
coincidence circuit, 268, 270
gas-discharge, 268
scintillation, 268
semiconductor, 268
Coupling,
jj, 116
LS, 116
Russel-Saunders, 116
Cross section,
effective, 253, 254f
neutron capture and optical frequency, 168 energy, 255
Crystal(s),
combination scattering of light, 169
direction indices, 159
electrons in, 184, 190ff
energy, 168
heat capacity, 160f, 166f
high temperatures, 161
low temperatures, 161
internal energy, 161, 166, 167
lattice, *see* Crystal lattice
Miller indices, 159f
point indices, 158f
standing wave in, 164f
Crystal lattice,
normal oscillations, 164f
acoustic frequency, 168
maximum frequency, 165f

Crystal lattice,
  oscillation frequency,
    acoustic branch, 168
    optical branch, 168
  periods of identity, 158
  unit cells, 158
Curie (Ci), 251
Curium, 233
Current,
  carriers,
    majority, 224, 225
    minority, 224, 226
  critical, 198
  density, 196
  rectification,
    full-wave, 213f
    half-wave, 213
  saturation, 212
  superconduction, 198
Cycle,
  carbon, 263
  carbon-nitrogen, 263
  proton-proton, 263

Decay,
  alpha, 245ff
  antiparticles, 280
  beta, 248ff, 288
  electron, 248
  muons, 242
  omega hyperon, 297
  particles, 279
  positron, 249f
Decuplet, baryon, 296
Degree, degeneracy, 96f
Density,
  probability, 78f, 82f, 100
  states, 179
Deuterium, 133, 262
Deuteron, 239, 260
Deuton, 239
Diagram, energy level, 81f, 97
  cesium, 113f
  diatomic molecule, 142
  hydrogen, 98f, 101
  sodium, 101, 102, 112f
Dielectric, 190
Diode(s),
  semiconductor, 224, 268
  vacuum-tube, 210f
  volt-ampere characteristic, 211f
Distribution, Bose-Einstein, 169
Donors, 205, 224
Doublet, 109
  complex, 114
  sodium, splitting, 124

Edge, spectrum band, 142f
Effect(s),
  Compton, 44ff
  latitude, 271
  Meissner, 197
  Mössbauer, 173ff
  non-linear, in optics, 148
  Paschen-Back, 126
  Peltier, 222f
    use for refrigeration, 223
  photoelectric, 36ff
    barrier-layer, 229f
    external, 40
    internal, 40
    intrinsic, 229
    multiple-photon, 39f, 157
  Raman, *see* Light, combination scattering
  Seebeck, 218ff
  Stark, 122
  thermoelectric, 218ff
  Thomson, 223f
  tunnel, 92, 248
  Zeeman, 122ff
    anomalous, 124
    complicated, 124
    normal, 123
    simple, 123f
Eigenfunctions, 79f, 82, 96, 99
  complete set, 87
  graphs, 82f
  normalized, 100
  for particle energy eigenvalues, 80
Eigenvalues, 79f, 83
  energy, 80f, 139f
  particle energy, 80f
Einsteinium, 233
Electroluminescence, 11
Electron(s), 266, 278, 279
  acceleration in crystal, 191
  angular momentum,
    intrinsic, 110
    orbital, 117
  beta, energy, 249
  carrying by phonons, 219
  in circular orbit, 61f
  Compton wavelength, 47, 115, 240
  conduction, 177, 180
    Cooper pairs, 198f
    escape from metal, 208
  configuration, atom, 131ff
  diffraction, 67f, 70
  diffusion in metal conductor, 219
  distance from nucleus, 100f
  distribution by energy states, 185
  drift velocity, 194ff
  dynamics in crystal, 190ff

Electron(s),
effective mass, 192f
in crystal, 202
energy, 96
in diatomic molecule, 140
in hydrogen atom, 74
and wave number, 187f
equation of motion, 195
equivalent, 134
free, 177
number in crystal, 184
gas,
degenerate, 185
non-degenerate, 185
heat capacity, in metals, 181
magnetic moment, intrinsic, 111, 272
motion in cathode-ray tube, 73
optical, 132
outer, 101
paramagnetic resonance, 127ff
photoconduction, 229
potential energy, 95, 209
at *p-n* junction, 229f
psi-function, 190
quantum numbers in atom, 129, *see also* Quantum number(s)
shells, 130ff
symbols, 131
spin, 111, 272
state(s),
in atom, 129
energy, 129
stationary, 57
superconductivity, 198f
symbols, 97
subshells, 130ff
total energy, 209
valence, 101, 132
energy, 101, 185
Elements, transuranium, 233
Emission,
induced, 148
resonance, 170
stimulated, 148ff
and stimulating radiation, 148
thermionic, 210ff
saturation current, 212
Emissivity, 13
and absorptivity, 14f
blackbody, 15f, 29
and equilibrium density of radiation, 42f
Emittance, radiant, 12
blackbody, 16, 19, 33
and emissivity, 13
Energy,
activation, nucleus, 236
alpha particles, 246f
atom, 116
beta electrons, 249
binding,
average per nucleon, 235, 239, 256
and mass number, 235f
nucleons, 1234f
nucleus, 234f
conduction electron, 178
crystal, 168
eigenvalues, 80f, 139f
electron(s), 96
at absolute zero, 180f
diatomic molecules, 140
and electron configuration, 140
in hydrogen atom, 74
state, 129
total, 209
and wave number, 187f
Fermi, 183
gamma quanta, 171
harmonic oscillator, 60, 141, 160f
hydrogen atom, 96
internal,
atom, 62f
crystal, 161, 166, 167
levels, *see* Energy level(s), Level(s)
molecule, 138ff
rotational, 141
total, 142
vibrational, 141
in normal coordinates, 163
nuclear reaction, 251f
nucleus at rest, 234
operator, 78
phonons, 174
photon, 41, 65, 143, 145, 175, 176
potential,
electron, 95, 209, 229f
hydrogen molecule, 139
molecule, 141
quantization, 79ff
quantum, 30
radiant,
equilibrium density, 16ff
and radiant emittance of blackbody, 17
recoil, 106, 171, 173, 174
average, 107
visible light, 106
rest, particle, 234
spectrum, 81f
valence electron, 185ff
system, 162f
total, in quantum mechanics, 92

Energy,
valence electron, 101
in crystal, 185
zero, 93, 94
Energy level(s), *see also* **Level(s)**
breadth, 105
degeneracy degree, 96f, 179
diagram, 81f, 97, 98f, 101f, 112ff, 142
discrete, 60
harmonic oscillator, 93f
hydrogen atom, 63
inverse population, 151
population, 151
splitting, 122
transitions between, 148f
Equation(s), *see also* **Formula**
Dirac's, 272f, 276
dynamic variables, 78
electron motion, 195
motion, 162
relativistic quantum-mechanical, 272f
Schrödinger, 74ff, 95
with consideration of lattice field, 187
free electron, 177
harmonic oscillator, 93
hydrogen molecule, 139
solution with periodic potential, 187
stationary states, 76
and theory of relativity, 272
Equilibrium, substance and radiation, 149
Experiment(s),
Bothe's, 40f
Compton's, 44ff
Davisson's and Germer's, 65f
Franck's and Hertz's, 57ff
gedanken, *see* **Experiment(s), mental**
Lederman's and Schwartz's, 293
mental (thought), 69
Millikan's, 38f
Pound's and Rebka's, 176
Reines's and Cowan's, 292f
Rutherford's, 52ff
Stern's, 67
Stoletov's, 36f
Tartakovsky's, 67
Thomson's, 67
Wu's, 288f, 291

Factor,
Lande *g*, 118, 124
neutron multiplication, 260
shielding, 138
Families, radioactive, 245
Fermi (Fm), 234
Fermions, 183, 266, 267, 277
Fermium, 233
Field,
critical, 197f
and temperature, 198
threshold, 197
Fission, nuclei, 236, 256ff
fragments, 256
spontaneous, 250
Fluorescence,
resonance, 170
X-ray, 40
Force(s),
nuclear, 238
charge independence, 239, 280
radius of action, 239, 262
saturation, 239
photo-electromotive, 229f
short-range, 239
thermal electromotive, 218f, 220, 221
differential, 221
specific, 221
Formula, *see also* **Equation(s)**
Balmer, 49, 50
generalized, 50, 63
Debye's, 166f
Einstein's, 38
for multiple-photon photoelectric effect, 40
Planck's, 32
Rayleigh-Jeans, 29
Richardson-Dashman, 212
Rutherford, 56
Rydberg's, 104
Fraction, binding, 235
Fragments, fission, 256
Function(s),
Block, 183
Boltzmann distribution, 185
Fermi-Dirac distribution, 183, 202
normalized, 78
psi-, *see* **Psi-function**
spectral distribution, 19
wave, 74, *see also* **Psi-function**
work, 38, 210, 212
Fusion, nuclei, 236, 237, 262f

Gamma quantum(a), 173
absorption spectrum, 171f
emission spectrum, 174f
energy, 171
Gas, electron,
degenerate, 185
non-degenerate, 185

Generators,
molecular, 150
optical quantum, 150
Gluons, 300
Gravitons, 266
Grids, 215

Hadrons, 266, 300
Half-breadth, spectral line, 105
Half-life,
neutrons, 232
radioactive nuclei, 244, 245
Hamiltonian, 77f
Harmonic oscillator, 93f
energy, 141, 160f
average value, 160
energy levels, 93f
phase trajectory, 60
possible states, 60
Schrödinger equation, 93
selection rule, 94
total energy, 60
Harmonics, optical, 156
Heat, Peltier, 222
Heat capacity,
crystals, 160f, 166f
electrons in metals, 181
Holes,
photoconduction, 229
in semiconductors, 200ff, 204
Hydrogen,
atom, 62ff, 95ff, 100f, *see also* Atom, hydrogen
heavy, 233
molecule, potential energy, 139
Hypercharge, 285f, 297, 298
Hyperons, 267, 279, 283
lambda, 279
omega, 279, 296, 301
decay, 297
sigma, 279
xi, 279
Hypothesis,
de Broglie's, 65ff
Planck's, 30, 94
Yukawa's, 241

Impurity,
acceptor, 225
donor, 225
Indices, crystal, 158ff
Instruments,
registering, 268
track-detecting, 267
Interaction(s),
electromagnetic, 265f, 280, 283
from viewpoint of quantum electrodynamics, 239ff
exchange, between nucleons, 242f
gravitational, 265f
nuclear, direct, 252
spin-orbital, 111, 280f
strong, 238f, 251, 265f, 282
between nucleons, 242f
weak, 265f, 283, 288f, 290
parity conservation law, 288
Invariance, *CP*-, 289
Inversion,
combined, 289
space, and psi-function, 287f
Ion, hydrogen-like, 95
Isobars, 233
Isomers, 233
Isospin, 281, *see also* Spin, isotopic
Isotones, 233
Isotopes, 233

Junction, *p-n*, 224ff
cut-off direction, 227
depletion layer, 227
forward direction, 226
forward voltage, 226
resistance, 227
reverse direction, 227
reverse voltage, 227
transition layer, 227
use for current rectification, 227
volt-ampere characteristic, 226f

Kaons, 266, 280
Kenotron, 212f
Kurchatovium, 233

Lasers, 148, 150ff, 155
gas, 154
pumping, 153
radiation, 154f
ruby, 152ff
Law(s),
Boltzmann's, 30, 150, 151
conservation,
angular momentum, 249, 294
baryon charge, 267, 278, 294
combined parity, 289
electric charge, 278, 294

and electromagnetic interaction, 294f
Law(s),
energy, 294
hypercharge, 285, 294
isotopic spin, 282f, 295
lepton charge, 280, 294
momentum, 275, 276, 294
parity, 288, 294
space parity, 289, 290
strangeness, 285, 294
and strong interaction, 294
and weak interaction, 294
Debye T^3, 167
Dulong and Petit, 160, 161
invariance relative to charge conjugation, 289
Kirchhoff's, 14
Moseley's, 137f
Newton's second, 192
Ohm's, 194
periodic, 301
radioactive transformation, 244f
Stefan-Boltzmann, 19, 33
three-halves power, 211
Wien's displacement, 20
Lawrencium, 233
Layer,
depletion, 227
transition, 227
Leptons, 242, 266, 278, 279
charge, 278
number, 278
Level(s), *see also* Energy level(s)
acceptor, 206f
donor, 206f
Fermi, 183f, 196, 199, 202, 206 210, 215, 225, 226
at absolute zero, 179
temperature dependence, 184
impurity, 206
Lifetime,
compound nucleus, 252
elementary particles, 266, 279
excited states,
atoms, 105
nuclei, 171, 175, 247
metastable states, atoms, 105
radioactive nuclei, 244f
Light,
beam, critical power, 156
combination scattering, 146ff, 155
by crystals, 169
corpuscular-wave duality, 43, 65
non-linear reflection, 156
pressure, 42
recoil energy, 106
Light,
resonance absorption, 170, 255
resonance emission, 170
self-focussing, 156
Line,
anti-Stokes, 147
spectral,
breadth, 105, 108
natural, 106, 171
Doppler breadth, 107f, 171, 172
Doppler broadening, 106, 171, 172
relative, 107
half-breadth, 105
Stokes, 147
Luminescence, 11

Magneton,
Bohr, 117
nuclear, 231
Masers, 150
Mass,
atomic nucleus, 234
critical, radioactive substance, 259
defect, nucleus, 235
elementary particles, 279
neutron, 232
proton, 231
Mendelevium, 233
Mesons, 239, 241, 266f, 279, 285, 288, 300
eta-, 266f, 277, 279, 295
K-, 266f, 279, 283, 288
mu-, 241
octet, 295
pi-, 241, 266, 279, 300
decay, 241
Metal(s), 189
electrical conductance, 194ff
electrical conductivity, 196
resistivity, 194
residual, 194
Microparticles, 65, 68ff
diffraction by slit, 71f
trajectory, 69f, 73
Model,
atom, 51ff, 119f
nucleus, 237f
Sakata's, 298
Moderator, 260
Molecule(s),
diatomic, 138f
covalent bond, 139
electron energy, 140
heteropolar bond, 139
homopolar bond, 139
ionic bond, 139
energy, 138ff

Molecule(s),
energy,
potential, 141
rotational, 141
total, 142
vibrational, 141
hydrogen, potential energy, 139
moment of inertia, 144, 145
Moment, magnetic,
atom, 118
electron, 111, 272
neutron, 232, 243, 298f
proton, 231, 243, 298f
Momentum,
angular, *see* Angular momentum
photon, 41f
Multiplet(s), 109, 114
charge, 281ff
baryon charge, 286
electric charge, 285
hypercharge, 286
strangeness, 286
inverted, 135
normal, 135
unitary, 295
Muons, 241, 242, 266, 272, 279
decay, 242

Neptunium, 258
Neutrino(s), 241, 249, 280, 290ff
electron, 266, 279, 293
longitudinal, 290
muon, 266, 279, 293
Neutron(s), 231, 254, 280
magnetic moment, 232, 243, 298f
mass, 232
radioactivity, 248
resonance absorption, 255
spin, 232
thermal, 256, 261
Normalization condition, 78, 100
psi-function, 177f
Nucleons, 231, 267, 280, *see also* Neutron(s), Protons
binding energy in nucleus, 234f
Nucleus(i),
activation energy, 236
atomic, *see* Nucleus(i)
atomic number, 232
binding energy, 234f
characteristics, 232f
charge number, 232
composition, 231
compound, 252
cross section, 253
daughter, 245

Nucleus(i),
doubly magic, 238
energy, 234
even-even, 234
excited states, lifetime, 171, 175
average, 247
fission, 236, 256ff
fragments, 256
spontaneous, 250
fusion, 236, 237, 262f
internal conversion, 247
light, fusion, 262f
magic, 238
mass, 234
mass defect, 235
mass number, 232
models, 237f
liquid-drop, 237f
shell, 238
parent, 245
photofission, 257
proton number, 232
radioactive,
average lifetime, 244f
half-life, 244, 245
radius, 233f
spin, 175, 234
theory, 237
velocity of thermal motion, 171
Number(s),
atomic, determination, 138
baryon, 277
lepton, 278
magic, 238
quantum, *see* Quantum number(s)
wave, 49

Octet,
baryons, 296
mesons, 295
Operator, 78
definition, 77
energy, 78
Laplacian, 75
Optics,
linear, 155
non-linear, 155ff
Orbits, stationary, 60
Oscillations, normal, 162ff
crystal lattice, 164f
acoustic frequency, 168
maximum frequency, 165f
optical frequency, 168
Oscillator, harmonic, *see* Harmonic oscillator

Packet, wave, 190
Pairs,
  Cooper, 198f
  electron-positron, 272, 274f
    annihilation, 274, 276
    birth, 274f
    cascades, 272
    showers, 272
Parameter, impact, 53
Parity, 286, 288
  combined, 289
  non-conservation, 288
Particles,
  absolutely neutral, 277, 279, 280
  alpha, 246f
    energies, 246f
    scattering, 247
  and antiparticles, 277ff
  charmed, 300
  elementary, 265ff, 297
    average lifetime, 266, 279
    classes, 266
    decay scheme, 279
    detecting methods, 267ff
    helicity, 290f
    mass, 279
    with even intrinsic parity, 288
    with negative intrinsic parity, 288
    with odd intrinsic parity, 288
    with positive intrinsic parity, 288
  penetration through potential barrier, 88ff
  psi, 300
  rest energy, 234
  scattering, 252
    elastic, 252
    inelastic, 252
  strange, 283ff
  strangeness, 283, 284ff
  virtual, 240
Phonon(s), 168, 173, 199
  average number, 169
  carrying of electrons by, 219
  energy, 174
  and photons, 169
  quasimomentum, 168, 169
Photoconduction,
  electron, 229
  hole, 229
Photocurrent, 37, 230
  and voltage, 37
Photoelectrons, maximum velocity, 38
Photoluminescence, 11
Photon(s), 40, 41, 99, 247, 266, 277, 279, 290, 300
  cascade formation in stimulated emission, 153f
Photon(s),
  collisions with molecules, 147
  distribution, 169
  energy, 41, 65, 143, 145, 175
  flux, relative fluctuation, 44
  frequencies, 145
    gravitational red shift, 175
  gamma, 247, 248
    absorption lines, 172, 176
    emission lines, 172, 176
    energy, 176
  gravitational mass, 175
  momentum, 41f, 65
  and phonons, 169
  rest mass, 42
  scattering, 147
  speed, 42
  spin, 97f
  virtual, 239f
Photoresistors, 229
Pions, 241, 266, 280
Plane, phase, 60
Plant, atomic power, 261
Plasma, 263
Plutonium, 233, 258, 261
Population, energy level, 151
  inverse, 151
Positron, 249f, 272, 274
Postulates, Bohr's, 57
Potential,
  chemical, 169f
  contact, 215
    difference, 215ff
      external, 216f
      internal, 216
  emission, 210
Power, critical, light beam, 156
Pressure, light, 42
Principle,
  exclusion, *see* Principle, Pauli
  Heisenberg uncertainty, 71
  Pauli, 129ff, 134, 169, 179, 273, 299f
  superposition, 190
    of states, 87
Process,
  forbiddenness and conservation laws, 283f
  multiple-photon, 39
  single-photon, 39
Protium, 233
Proton, 231, 280
  charge, 231
  magnetic moment, 231, 243, 298f
  mass, 231
  spin, 231

Psi-function, 74f
  meaning, 78f
  normalization condition, 177f
  in space inversion, 287f
  standard conditions, 79, 96
Pumping, laser, 153

Quantities, canonically conjugate, 71
Quantization, angular momentum, 83ff
Quantum(a),
  action, 30
  energy, 30
    maximum value, 36
  light, 40
  light energy, 57
  magnetic flux, 200
Quantum number(s),
  angular momentum,
    resultant, 116
    resultant spin, 116
    total, 111, 113, 234
    total orbital, 116
  azimuthal, 84, 86, 96, 97, 103, 111, 129
  charm, 300
  electron in atom, 129
  hypercharge, 285
  isotopic spin, 281f
  magnetic, 85, 96, 123, 129
  nuclear spin, 234
  orbital, 84, 86
  principal, 62, 96, 129
  rotational, 141
  spin, 110, 111, 129, 178
  and state of conduction electron, 178f
  strangeness, 284f
  vibrational, 141
Quarks, 297ff
  baryon charge, 298
  charmed, 300
  colour, 299f
  electric charge, 298
  spin, 298f
  strangeness, 298
Quasiparticles, 169, 202

Radiation,
  braking, *see* Bremsstrahlung
  characteristic, 35, 136
  equilibrium, in cavity, 28
  mean energy, 31
  minimum wavelength, 36
  thermal, 11ff
Radiation,
  thermal,
    equilibrium with emitting bodies, 11f
    equilibrium energy density and radiant emittance of blackbody, 18
  wavelength, 13
  X-ray, 136ff
Radioactivity, 243ff
  artificial, 243
  natural, 243, 245
  proton, 250f
Radiocarbon, 255f
Radius,
  action, 265
  Bohr, 62
Ratio, gyromagnetic, 117f
Rays,
  alpha, 245
  beta, 245
  cosmic, 241, 271f
    primary, 271
    secondary, 271
      hard component, 272
      soft component, 272
  gamma, 170, 245, 248
    resonance absorption, 172f
Reaction(s),
  nuclear, 251ff
    chain, 258f
  effective cross sections, 254f
    energy, 251f
    pickup, 253
    stripping, 252
  thermonuclear, 237, 262ff
    controlled, 263f
  threshold, 276
Reactors,
  breeder, 261
  nuclear, 237, 259ff
Recombination, electron-hole, 204, 207
  probability, 204
Relation, uncertainty, 71ff, 190, 240
Resistivity, metals, 194
Resonance,
  electron paramagnetic, 127ff
  nuclear magnetic, 127
Resonances, 267, 300
Rule(s),
  Hund's, 135f
  selection, 94, 152
    for J, 141
    for j, 113
    for l, 97, 103
    for m_J, 123, 127

Rule(s),
selection,
for m_L, 126
for v, 141

Satellite,
red, 147, 169
violet, 147, 169
Scintillation, 52
Semiconductor(s), 200ff
conductance,
electronic, 205
hole, 206
impurity, 205ff
intrinsic, 202ff, 207
and temperature, 200f
electronic, 190
extrinsic, 200
holes, 200ff, 204
impurity, 200
intrinsic, 200f
n-type, 205, 206
non-uniformly heated, 220
p-type, 206
non-uniformly heated, 220
typical, 203f
Series,
Balmer, 49, 99
Bergmann, 101
Brackett, 49
diffuse, 101, 103, 104
fine structure, 114
fundamental, 101, 103, 104
limit, 50
Lyman, 49, 99
Paschen, 49
Pfund, 49, 50
principal, 101, 103, 104
radioactive, 245
sharp, 101, 103, 104
X-ray spectra, 136
Shells,
electron, 130ff
symbols, 131
nuclear, 238
Shift,
gravitational, red, 175
Lorentz, 123
normal, 123
Showers, electron-positron pairs, 272
Singlets, 109, 122
Space, isotopic, 281
Spectroscope, microwave, 128
Spectrum(a),
absorption, gamma quanta, 174f
alpha, fine structure, 246
Spectrum(a),
atomic, 48
fine structure, 234
hyperfine structure, 234
band, 142ff
bremsstrahlung X-ray, short wavelength limit, 34, 35f
combination scattering, 146f
continuous, 80
discrete, 80, 81, 85
emission,
alkali metals, 101
gamma quanta, 174f
energy, 81f
valence electron, 185ff
fine structure, 109
hydrogen atom, 48ff
frequencies of lines, 49f
line, 48
molecular, 142ff
of quantity, 80
series of lines, 48, *see also* Series
X-ray, 34, 136ff
Spin, 110, 272
atomic nucleus, 234
double magnetism, 118
electron, 111, 272
isotopic, 281ff, 298f
neutron, 232
nuclei, 175
photon, 97f, 110
proton, 231
quantum number, 110, 111, 129, 178
quarks, 298f
State(s),
degenerate, 96
electron, symbols, 97
even, 287
excited, 99, 105
ground, 99, 105
metastable, 105
with negative temperatures, 151
odd, 287
stationary, 79
electron, 57
Schrödinger equation, 76
superconducting, 197ff
Statistics,
Bose-Einstein, 170
Fermi-Dirac, 183
Strangeness, 297, 298
quarks, 298
Strength, electric field, and temperature gradient, 220
Subshells, electron, 130ff
Superconductivity, 197ff
Superconductor, 197

Superfluidity, 198
Supermultiplets, 295
Surface, Fermi, 180, 187f
Symbols,
  electron shells, 131
  electron states, 97
  terms, 112f
System,
  generalized coordinates, 162
  periodic, of elements, 131ff

Temperature,
  characteristic Debye, 166
  critical, 197
  Fermi, 181, 185
  transition, 197
Term(s), 50, 60
  atom, 116f
  multiplicity, 117
  spectral, 50, *see also* Term(s)
  symbols, 112f, 116f
Theory,
  BCS, 198ff
  Bohr's, hydrogen atom, 62ff
  Debye's, 164ff
  Dirac's, 273ff
  longitudinal neutrino, 290
  superconductivity, 198ff
  unitary symmetry, 295
Thermocouple, 221f
Threshold,
  photoelectric, 39, 40
  reaction, 276
Time,
  decay, 266
  nuclear, 252
    transit, 252
  relaxation, 195, 196
Trajectory,
  microparticle, 69f, 73
  particle, 79
  and particle mass, 72f
  phase, 60
Transistor(s), 224, 228
  base, 228
  collector, 228
Transistor(s),
  emitter, 228
  n-p-n, 228f
  p-n-p, 229
Triode(s), 213ff
  grid characteristics, 214
    slope, 214
  semiconductor, 224, 228
Tritium, 233, 262
Tube,
  electronic, 210ff
  three-electrode, 213f
  two-electrode, 210f
  X-ray, 34

Unit(s),
  activity, radioactive substances, 251
  atomic mass (amu), 231
  natural system, 115
Uranium, natural, 259

Variables, dynamic, 70
  equation, 78
Velocity,
  drift, 194ff
  group, 190
Voltage, retarding, 37f

Water, heavy, 260
Wave(s),
  number, 49
  packet, 190
  standing,
    in crystal, 164f
    in three-dimensional space, 20ff
    two-dimensional, 23
    wave vector, 25
Wavelength, Compton, 47
  electron, 47, 115, 240
Way, eightfold, 295
Work function, 38, 210, 212

Zone, Brillouin, 187

УДК 539.18(075.8) = 20

Учебное издание

Игорь Владимирович Савельев

КУРС ОБЩЕЙ ФИЗИКИ

Том III

Квантовая оптика. Атомная физика. Физика твердого тела.
Физика атомного ядра и элементарных частиц

Заведующий редакцией А. П. Ястребов. Контрольный редактор Г. Г. Анохина. Редакторы Н. М. Привато, Н. Ф. Шлифер. Художник Л. М. Муратова. Художественный редактор И. И. Шиф. Технический редактор Л. Х. Абдулла. Корректоры Е. Л. Трошнева, И. В. Горбунова

ИБ № 7886

Подписано к печати 19.09.88. Фотоофсет. Формат $60 \times 90^1/_{16}$. Бумага офсетная № 1. Гарнитура обыкновенная. Объем 10,0 бум. л. Усл. печ. л. 20,0. Усл. кр.-отт. 20,64. Уч.-изд. л. 20,75. Изд. № 10/6221. Тираж 8000 экз. Зак. 7797. Цена 2 р. 20 к.

ИЗДАТЕЛЬСТВО «МИР»
В/О «Совэкспорткнига» Государственного комитета СССР по делам издательств, полиграфии и книжной торговли
129820, ГСП, Москва, И-110, 1-й Рижский пер., 2

Экспериментальная типография ВНИИ полиграфии Государственного комитета СССР по делам издательств, полиграфии и книжной торговли
103051, Москва, Цветной бульвар, 30

С $\frac{1604010000\text{-}086}{056(01)\text{-}89}$ 22-89, ч. 2