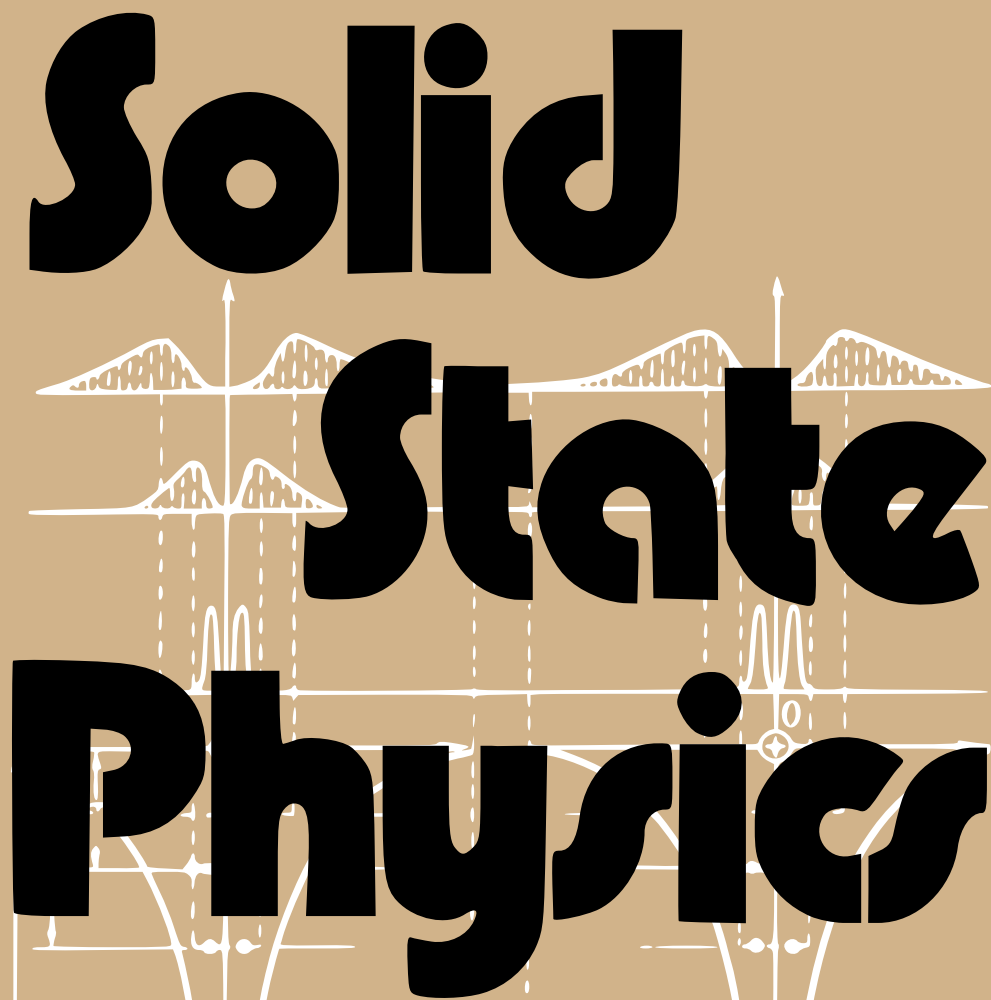
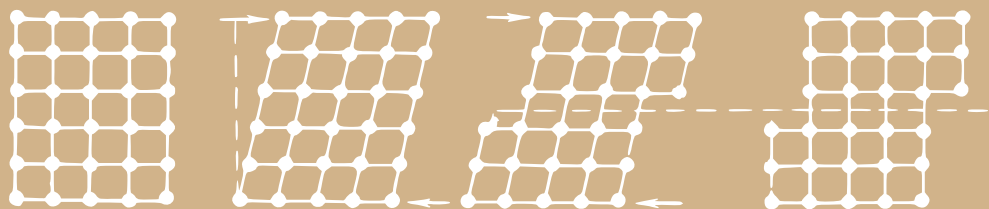


Solid State Physics

The background of the title section features several white line drawings on a tan background. At the top, there are four bell-shaped curves, each with a vertical dashed line extending downwards. Below these are two smaller bell-shaped curves. In the center, there are two oscillating waveforms, one above the other, with vertical dashed lines connecting them. At the bottom of this section, there are two parabolic curves opening upwards, with dots placed at various points along them.

G. I. Epifanov



Mir Publishers Moscow

Solid State Physics

Solid State Physics

by **G. I. Epifanov**, D.Sc.

Moscow Institute of Electronic Engineering

Translated from the Russian by

Mark Samokhvalov, Cand. Sc.



MIR PUBLISHERS
MOSCOW

Revised from the 1977 Russian edition
First published 1979

English translation, Mir Publishers, 1979

PREFACE

Ten years have passed since the first Russian edition of this textbook was published. In this time solid state physics has developed rapidly as the scientific background of numerous front-line branches of technology, absorbing new discoveries and theories. This has been considered in preparing the new edition.

At the same time college curricula have been changed to improve the basic preparation of versatile engineers, especially in physics and mathematics. This too had to be reflected in this book.

Also, the years that have elapsed since the first edition have seen much comment, some critical, and many proposals from Soviet and foreign readers—from college teachers and students, teachers of vocational and secondary schools, engineers and scientists. The author is grateful for all the comment and proposals.

There was a need therefore to revise the book completely. As in the first edition, the presentation of material has followed the aim of elucidating the physical nature of the phenomena discussed. But, where possible, the qualitative relations are also presented, often though without rigorous mathematics.

The manuscript was reviewed in detail by Prof. L. L. Dashkevich, Dr. of Technical Sciences, and Prof. I. G. Nekrashevich, Honored Scientist of the Belorussian Republic. It was perused by Prof. L. A. Gribov, Dr. of Mathematical and Physical Sciences, Assistant Prof. V. B. Zernov, and Z. S. Sazonova. The author extends sincere thanks for their efforts and criticism, which he took into account when revising the manuscript.

The author is also indebted to Senior Lecturer F. Zh. Vilf, Cand. of Technical Sciences, and Assistant Prof. Yu. A. Moma, Cand. of Technical Sciences, for manuals used in this textbook on superconductivity, Gunn effect, and principles of operation of impulse and high-frequency diodes, and to Z. I. Epifanova for all her work in preparing the manuscript.

The author will be most grateful for comment and proposals that might improve this book. They should be sent to the publishers.

G. I. E.

Contents

Preface

vii

Chapter 1. Bonding. The Internal Structure of Solids

1

- § 1. The van der Waals forces 1
- § 2. The ionic bond 4
- § 3. The covalent bond 6
- § 4. The metallic bond 11
- § 5. The hydrogen bond 12
- § 6. Comparison between bonds of various kinds 14
- § 7. Forces of repulsion 15
- § 8. Crystal lattice 16
- § 9. Notation to describe sites, directions, and planes in a crystal 19
- § 10. Classification of solids based on the nature of bonds 22
- § 11. Polymorphism 27
- § 12. Imperfections and defects of the crystal lattice 30

Chapter 2. Mechanical Properties of Solids

35

- § 13. Elastic and plastic deformations. Hooke's law 35
- § 14. Principal laws governing plastic flow in crystals 40
- § 15. Mechanical twinning 44
- § 16. Theoretical and real shear strengths of crystals 45
- § 17. The dislocation concept. Principal types of dislocations 47
- § 18. Forces needed to move dislocations 51
- § 19. Sources of dislocations. Strengthening of crystals 53
- § 20. Brittle strength of solids 57
- § 21. Time dependence of the strength of solids 63
- § 22. Methods of increasing the strength of solids 67

Chapter 3. Elements of Physical Statistics

71

- § 23. Methods used to describe the state of a macroscopic system 71
- § 24. Degenerate and nondegenerate ensembles 75
- § 25. The number of states for microscopic particles 78
- § 26. Distribution function for a nondegenerate gas 81
- § 27. Distribution function for a degenerate fermion gas 83
- § 28. Distribution function for a degenerate boson gas 90
- § 29. Rules for statistical averaging 91

Chapter 4. Thermal Properties of Solids	95
§ 30. Normal modes of a lattice	95
§ 31. Normal modes spectrum of a lattice	98
§ 32. Phonons	99
§ 33. Heat capacity of solids	103
§ 34. Heat capacity of electron gas	108
§ 35. Thermal expansion of solids	110
§ 36. Heat conductivity of solids	114
 Chapter 5. The Band Theory of Solids	 121
§ 37. Electron energy levels of a free atom	121

Chapter 1

Bonding.

The Internal Structure of Solids

Matter can exist in the solid state only because there are forces of interaction acting between the structural particles when the latter are brought sufficiently close together. For the solid to have a stable structure the forces of interaction between the particles should be of two types: of attraction, to prevent the particles from moving away from each other, and of repulsion, to prevent the particles from merging.

Let us discuss briefly the nature of these forces.

§ 1. The van der Waals forces

The most general type of bond existing between any two atoms, or molecules, is due to van der Waals forces. Those forces were first introduced to explain the equation of state of real gases, the *van der Waals equation*:

$$\left(p + \frac{a}{V^2}\right)(V - b) = RT \quad (1.1)$$

where the correction terms a/V^2 and b account, respectively, for the effect of the forces of attraction and repulsion acting between the molecules of a real gas. These forces manifest themselves in an almost ideal form in the interaction between the molecules with saturated chemical bonds (O_2 , H_2 , N_2 , CH_4 , etc.), as well as between the atoms of inert gases, making possible their existence in the liquid and solid states.

As a general case, the van der Waals bond is made up of the dispersion, orientational and induction interactions. Let's consider them separately.

Dispersion interaction. Consider the simplest example of two interacting helium atoms shown in Figure 1.1. The electron density distribution in a helium

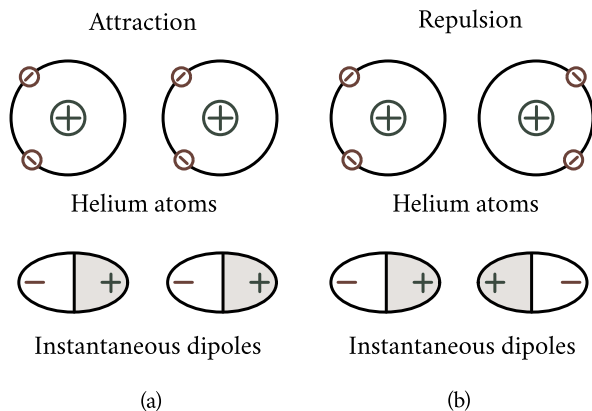


Figure 1.1: Dispersion interaction. The interaction between helium atoms is due to the correlation in the motion of electrons resulting in the appearance of instantaneous dipoles: (a) — correlation resulting in attraction; (b) — correlation resulting in repulsion.

atom is spherically symmetrical and for this reason its electric moment is zero. But this means only that the average electric moment of the atom is zero. At every moment of time the electrons occupy particular points in space, thereby creating instantaneous rapidly changing electric dipoles. When two helium atoms are brought together, the motion of their electrons becomes correlated and this leads to the forces of interaction. The forces can be of two types. If the motion of the electrons is correlated, as shown in Figure 1.1(a), the instantaneous dipoles attract each other; if the correlation is as shown in Figure 1.1(b), the resulting interaction is repulsion. Since the realization of the arrangement of Figure 1.1(a) leads to a reduction in the energy of the system, this arrangement is more probable and is realized more frequently. This is in effect the cause of the constantly existing force of attraction between helium atoms.

The bonds discussed above that owe their existence to a correlation in motion of the electrons in adjacent atoms are termed *dispersion forces*. They were first calculated by F. London in 1930. The calculation was based on the following model: the instantaneous electric dipole of one atom causes the other atom to be polarized and it becomes an induced dipole leading to the realization of the arrangement in Figure 1.1(a), which corresponds to attraction. The calculation had as its final result the following expression for the energy of the dispersion interaction of two particles:

$$U_d = -\frac{3}{4} \frac{\alpha^2 E_{\text{exc}}}{r^6} \quad (1.2)$$

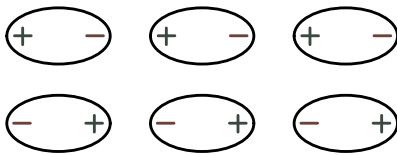


Figure 1.2: Orientational interaction of polar molecules.

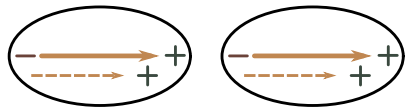


Figure 1.3: Induction interaction of molecules (dashed lines show the induced dipoles).

where α is the polarizability of the particles¹, E_{exc} their energy of excitation, and r the distance between them.

Orientational interaction. Should the molecules possess a constant dipole moment M , that is, should they be polar molecules, an electrostatic interaction would be established between them tending to arrange the molecules in a strict order (Figure 1.2), since that order corresponds to the minimum energy of the system. The correct orientation of the molecules is disturbed by thermal motion. Therefore, the energy of the system due to the mutual orientation of the molecules is strongly dependent on temperature. At low temperatures, when the orientation of the molecules is perfect, the interaction energy is determined from the expression

$$U_{\text{or}} = -\frac{M^2}{2\pi\epsilon_0 r^3} \quad (1.3)$$

where r is the distance between the molecules, and ϵ_0 the permittivity of free space.

In the high temperature range the energy of interaction of polar molecules, as had been demonstrated by W. H. Keesom, is of the following form:

$$U_{\text{or}} = -\frac{M^4}{24\pi^2\epsilon_0^2 k_B T r^6}. \quad (1.4)$$

where k_B is the Boltzmann constant and T the temperature.

The type of interaction discussed above is termed *orientational interaction*.

Induction interaction. Lastly, in case of polar molecules of high polarizability an induced moment due to the action of the field of constant dipoles may be established (Figure 1.3; the induced dipoles are shown by dashed lines). The energy of mutual attraction due to the interaction of the rigid dipole of the first molecule and the induced dipole of the second molecule, as has been shown by Debye, is

¹Let us recall the physical meaning of α . The charges in the molecule are displaced under the action of an external field of intensity \mathcal{E} . This leads to a dipole moment M proportional to \mathcal{E} : $M = \alpha\mathcal{E}$, the proportionality factor α being the *polarizability* of the molecule.

independent of temperature and is given by the expression

$$U_{\text{in}} = -\frac{\alpha M^2}{8\pi\epsilon_0^2} \frac{1}{r^6} \quad (1.5)$$

where, as before, M is the constant dipole moment of the molecule, and α its polarizability.

Such interaction is termed *induction*, or *deformation*, interaction.

In general, when two molecules are brought close together all three types of interaction may be established, the interaction energy being the sum of the energies of the dispersion (U_{d}), orientational (U_{or}), and induction (U_{in}) types of interaction:

$$U = U_{\text{d}} + U_{\text{or}} + U_{\text{in}}.$$

Table 1.1 shows the relative magnitude (in percent) of each of those components of the total bonding energy for water, ammonia, hydrogen chloride and carbon monoxide. The data presented in Table 1.1 show the induction interaction for all the substances to be weak. Three quarters or a half of the bond energy in substances made up of polar molecules is due to the energy of orientational interaction; while in materials made up of nonpolar molecules almost all of the bond energy is due to the dispersion interaction.

Table 1.2 shows the values of the bond energy for some molecular crystals held together by van der Waals forces.

§ 2. The ionic bond

Atoms that occupy places in the Mendeleev periodic table next to inert gases tend to assume the electronic configuration of these gases either by giving away or accepting electrons. The valence electron of alkali metals, which immediately follow the inert gases, moves outside the closed shell and is only weakly connected with

Table 1.1

Substance	Type of interaction		
	Dispersion	Induction	Orientational
Water	19	4	77
Ammonia	50	5	45
Hydrogen chloride	81	4	15
Carbon monoxide	100		

the nucleus. The halides, which immediately precede the inert gases, lack one electron to complete a stable shell characteristic of an inert gas. Therefore, they exhibit high affinity to an excess electron.

Such atoms, that is, typical metals and halides, are bonded in the following way. First a recharging of the atoms takes place. The electron from the metal moves over to the halide. This turns the metal atom into a positively charged ion and the haloid atom into a negatively charged one. These ions interact according to the Coulomb law as two opposite charges. Such a bond became known as an *ionic bond*.

The energy of attraction of two ions separated by the distance r is

$$U_{\text{att}} = -\frac{q^2}{4\pi\epsilon_0 r} \quad (1.6)$$

where q is the charge of the ions.

The curve 1 in Figure 1.4 shows the dependence of U_{att} on r . As r decreases the absolute value of the energy increases monotonically, tending to infinity as $r \rightarrow 0$. The force of attraction tends to bring the ions together as close as possible. This, however, is prevented by the forces of repulsion, which begin to make themselves felt at small distances and rise very rapidly with the decrease in distance. The repulsion energy U_{rep} is shown in Figure 1.4 by the curve 2. Max Born and other scientists expressed the repulsion energy in the form

$$U_{\text{rep}} = \frac{B}{r^n} \quad (1.7)$$

where B and n are constants.

The resulting interaction energy of two ions is

$$U = U_{\text{rep}} + U_{\text{att}} = \frac{B}{r^n} - \frac{q^2}{4\pi\epsilon_0 r}. \quad (1.8)$$

Table 1.2

Substance	$U_b, [10^3 \text{ J mol}^{-1}]$
Neon (Ne)	1.90
Argon (Ar)	8.40
Nitrogen (N ₂)	6.60
Carbon monoxide (CO)	8.40
Oxygen (O ₂)	8.20
Methane (CH ₄)	10.8

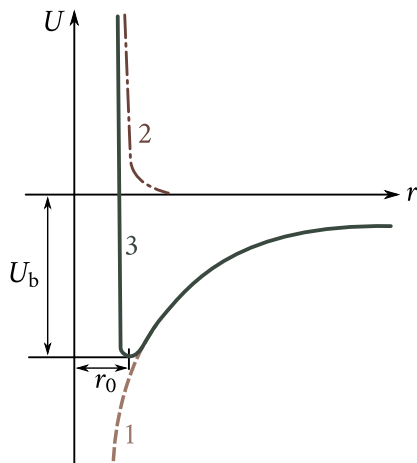


Figure 1.4: Dependence of energy of interacting ions on the distance between them: 1 — energy of attraction, 2 — energy of repulsion, 3 — total energy of interaction.

This energy is shown in Figure 1.4 by the curve 3 which has a minimum at $r = r_0$; the depth of this minimum determines the bond energy U_b , and r_0 determines the distance between the ions in the molecule. Making use of the fact that in equilibrium (at $r = r_0$) the force of attraction, $F_{\text{att}} = -(dU_{\text{att}}/dr)_{r=r_0}$, equals the force of repulsion, $F_{\text{rep}} = -(dU_{\text{rep}}/dr)_{r=r_0}$, we can easily express (1.8) as

$$U_b = -\frac{q^2}{4\pi\epsilon_0 r_0} \left(1 - \frac{1}{n}\right). \quad (1.9)$$

The energy of the lattice made up of N such molecules is

$$U_{\text{lattice}} = -NA \frac{q^2}{4\pi\epsilon_0 r_0} \left(1 - \frac{1}{n}\right) \quad (1.10)$$

where A is the *Madelung constant*, which takes account of the energy of interaction of the given molecule with all its neighbouring molecules in the crystal.

Table 1.3 shows by way of an example the experimental values of the bond energy of some ionic crystals and its values calculated with the aid of (1.10). The discrepancies do not exceed 1-2 percent, which is proof of good agreement between theory and experiment.

§ 3. The covalent bond

The ionic and van der Waals bonds are unable to account for the existence of such compounds as H_2 , O_2 , N_2 , etc., as well as for bonds in atomic crystals of the diamond type. Evidently, atoms of one kind cannot form oppositely charged ions by

changing the distribution of valence electrons, as was the case in the metal-halide interaction. On the other hand, the bond in the H_2 , O_2 and N_2 molecules is much stronger than that which could be attributed to the van der Waals forces. For such compounds the role of the van der Waals forces is that of a small correction to the bond mainly responsible for the strength of the compounds. This bond became known as the *covalent bond*.

Let us consider the nature of this type of bond using the hydrogen molecule as an example.

Suppose that two hydrogen atoms are at a rather great distance r from one another. The atom A consists of the nucleus a and the electron 1 and the atom B consists of the nucleus b and the electron 2 (Figure 1.5). Since the density of the electron cloud which describes the electron state in an atom falls off very rapidly as the distance from the nucleus increases, the probabilities to discover electron 1 near nucleus b and electron 2 near nucleus a are very small. Calculation shows that for $r \approx 50 \text{ \AA}$ each electron can visit the other nucleus on the average only once in 10^{12} years. Because of that atoms A and B may be regarded as isolated atoms and the energy of the system made up of such atoms may be taken to be equal to $2E_0$, where E_0 is the energy of the isolated atom in the ground state.

When the atoms are brought closer together, the probability of the electrons going over to nuclei other than their own increases. For $r \approx 2 \text{ \AA}$ the electron clouds begin to overlap noticeably and the transition frequency rises up to 10^{14} s^{-1} . If the atoms are brought still closer together, the frequency of the electron exchange rises so that it becomes meaningless to speak of electron 1 as belonging to the atom A and of electron 2 as belonging to atom B. This corresponds to a new state that is not characteristic for a system made up of two isolated atoms. A remarkable property of this new state is that the electrons in it belong simultaneously to both nuclei, in other words, are *collectivized*.

Table 1.3

Crystal	$U_b, [10^3 \text{ J mol}^{-1}]$	
	Experiment	Theory
Sodium chloride (NaCl)	752	754
Potassium iodine (KI)	650	630
Rubidium bromide (RbBr)	635	645
Caesium iodine (CsI)	595	585

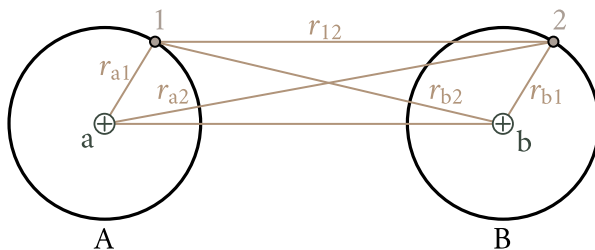


Figure 1.5: Calculating covalent bond between hydrogen atoms: A, B — hydrogen atoms; a, b — their nuclei; 1 — electron of atom A; 2 — electron of atom B; r_{a1} , r_{b2} — distances of electrons from their nuclei; r_{12} — distance between electrons; r_{a2} — distance of electron 2 from nucleus a; r_{b1} — distance of electron 1 from nucleus b; r — distance between nuclei.

The collectivization of the electrons is accompanied by a change in the electron density distribution $|\psi|^2$ and in the energy of the system as compared to the total energy $2E_0$ of the isolated atoms. The lines 1 in Figure 1.6 show the electron cloud density of the isolated atoms, the line 2 shows the total density obtained by simple superposition of the electron clouds of isolated atoms, and the line 3 the actual density distribution along the axis joining the nuclei a and b brought about by the collectivization of the electrons. The figure shows that the collectivization of the electrons results in the electron clouds being drawn into the space between the two nuclei: at a small distance from the nucleus outside this space the density of the clouds falls off, as compared with the density in isolated atoms, at the same time rising in the space between the nuclei above the sum of the densities of isolated atoms. The appearance of a state with an electron cloud of increased density that fills the space between the nuclei always results in a decrease in the system's energy and in the appearance of forces of attraction between the atoms. Speaking figuratively, we may say that the electron cloud formed in the space between the nuclei by a collectivized pair of electrons draws the nuclei together, striving to bring them as close together as possible.

Such is the qualitative picture of the origin of the covalent bond. Quantitative calculations of the hydrogen molecule were first performed by W. H. Heitler and F. London in 1927. Those calculations have shown that a system consisting of two closely spaced atoms of hydrogen can have two energy values depending on the direction of the electron spins in the atoms:

$$U_s = 2E_0 + \frac{(K + A)}{(1 + S^2)} \quad (1.11)$$

when the spins are antiparallel, and

$$U_a = 2E_0 + \frac{(K + A)}{(1 - S^2)} \quad (1.12)$$

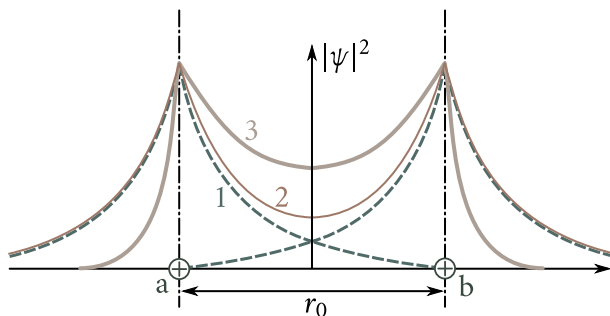


Figure 1.6: Electron density distribution in a system of two hydrogen atoms: 1 — electron density distribution in isolated hydrogen atoms, 2 — electron density resulting from simple overlapping of electron clouds of isolated atoms brought together to within a distance of r , 3 — actual electron density distribution in a hydrogen molecule.

when the spins are parallel. Here $2E_0$ is the total energy of the two isolated atoms, K is the energy of the electrostatic interaction of the electrons with the nuclei, of the electrons with one another, and of the nuclei. Another name for it is *Coulomb energy*. Its sign is negative. By A we denote the energy of exchange interaction due to the atoms exchanging electrons. This is the additional energy that appears as the result of the change in the electron density distribution in the process of the formation of the molecule. Its sign is negative and its absolute value is much larger than K ($|A| \gg |K|$). S is the *overlap integral* whose value lies within the limits $0 \leq S \leq 1$.

The state with the energy U_s is termed *symmetric* and with U_a *antisymmetric*. Since both K and A are negative and $S \leq 1$, the energy of the system in the symmetric state is less than the energy of two isolated atoms:

$$U_s < 2E_0. \quad (1.13)$$

This corresponds to the appearance of forces of attraction. Since the absolute value of the exchange energy A is considerably greater than that of the Coulomb energy K the decrease in the system's energy is mainly due to A . For this reason the force of attraction that appears between the atoms is termed the *exchange force*.

For the same reason, that is, because $|A| \gg |K|$, the formation of the antisymmetric state leads to an increase in the system's energy. This corresponds to the appearance of repulsive forces.

Figure 1.7 shows the dependence of U_s and U_a on r/a , where r is the distance between the atoms, and $a = 0.529 \text{ \AA}$ is the radius of the first Bohr orbit (the *Bohr radius*). The zeroth energy level has been fixed at $2E_0$. Figure 1.7 shows that in the antisymmetric state the system's energy rises steadily as the atoms are brought

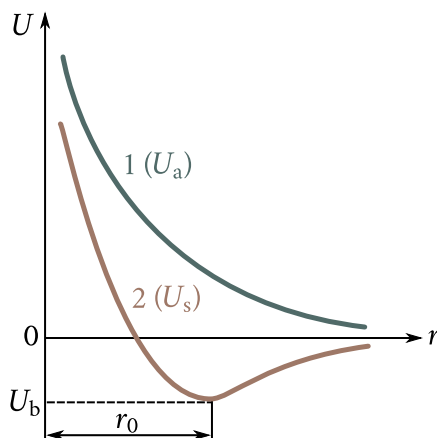


Figure 1.7: Interaction energy of two hydrogen atoms: 1 — antisymmetric state, 2 — symmetric state.

closer together (curve 1), this corresponding to the mutual repulsion of the atoms. Therefore a hydrogen molecule cannot be formed in such a state. In the symmetric state (curve 2) the system's energy at first falls as the distance r between the atoms decreases, attaining its minimum value at $r = r_0$. As the distance r decreases still further, the energy begins to rise because of the appearance of strong repulsive forces. The existence of a minimum on the potential energy curve makes the existence of a stable system of two hydrogen atoms, that is, a hydrogen molecule, possible. To destroy this system work must be performed equal to the depth of the potential well, U_b .

Calculation provides the following values of U_b and r_0 : $U_b = 4.37 \text{ eV}$, $r_0 = 0.735 \text{ \AA}$; the experimental values are $U_b = 4.38 \text{ eV}$, $r_0 = 0.753 \text{ \AA}$. The agreement between theory and experiment is quite good.

Table 1.4 shows the values of the bond energy for some covalent compounds—the molecules of H_2 , N_2 , O_2 , CO —and for diamond, silicon and germanium crystals in which the bonding is due to covalent forces.

Characteristic properties of the covalent bond, which distinguish it from the bonds of other types, are its *saturability* and *directionality*.

Saturability means that each atom can form covalent bonds only with a limited number of its neighbours. This means that each hydrogen atom can form covalent bonds only with one of its neighbours. The electron pair constituting such a bond has antiparallel spins and occupies one quantum cell. A third atom in this case instead of being attracted will be repelled.

The valence bond is formed in the direction of the greatest density of the elec-

tron cloud corresponding to the valence electrons. In this case there is maximum overlapping of the electron clouds of the bonding electrons, which implies that the valence bond is directional.

§ 4. The metallic bond

There is a special group of substances, called metals, that occupy places at the beginning of every period of the Mendeleev table. The formation of the metallic bond cannot be explained by the presence of the ionic or the covalent bond. Indeed, the ionic bond appears only between atoms having different affinities to the additional electron, for instance, between the atoms of a metal and a halide. Evidently, such bond between kindred atoms of a metal having identical affinity to the electron is impossible. On the other hand, metallic atoms do not have enough valence electrons to form valence bonds with their nearest neighbours. For instance, the copper atom has one valence electron and can form a valence bond only with a single atom. But in the copper lattice every atom is surrounded by twelve neighbours with which it must be connected by lines of force. This points to the fact that in metals there is a special type of bonding known as the *metallic bond*. Let us consider the nature of this bond.

In metallic atoms the external valence electrons are rather weakly coupled to the nucleus. In the liquid and solid states the atoms come so close together that the valence electrons are able to leave their respective atoms and wander throughout the lattice. This leads to an extremely homogeneous distribution of the negative charge in the crystal lattice. This conclusion is supported by direct experiments.

Table 1.4

Gas	$U_b, [10^5 \text{ J mol}^{-1}]$
Carbon monoxide (CO)	10.8
Nitrogen (N ₂)	9.50
Oxygen (O ₂)	5.00
Hydrogen (H ₂)	4.40
Diamond (C)	6.80
Silicon (Si)	4.40
Germanium (Ge)	3.50

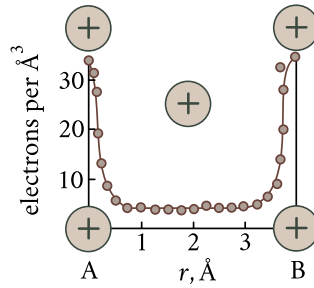


Figure 1.8: Electron density distribution in the aluminium lattice obtained by X-ray photography.

Figure 1.8 shows an experimental curve of the electron density distribution between the sites of the aluminium lattice obtained by means of X-ray photography. Most part of the distance between the sites the electron concentration remains constant. Only quite close to the sites it rises sharply because of the presence of internal shells of the aluminium atoms.

In the lattice of a metal the bond is due to the interaction of the positive ions with the electron gas. The electrons moving between the ions compensate the forces of repulsion existing between the positively charged ions and draw them closer together. As the distance between the ions becomes smaller the density of the electron gas rises and this leads to an increase in the force drawing the ions together. On the other hand, in this case the repulsive forces acting between the ions tend to move them away from each other. When the distance between the ions becomes such that the forces of attraction are compensated by the forces of repulsion, a stable lattice is formed.

It appears that the metallic bond is somewhat similar to the valence bond, since they are both based on the collectivization of external valence electrons. However, in case of the valence bond only atoms that form pairs of nearest neighbours take part in the collectivization of electrons, and the respective electrons always remain between the atoms. In case of the metallic bond all atoms of the crystal take part in the collectivization of electrons, and the collectivized electrons are no longer localized near their respective atoms but move freely inside the lattice.

§ 5. The hydrogen bond

The *hydrogen bond* is formed between an atom of hydrogen and an extremely electronegative atom, for instance, an atom of oxygen, fluorine, nitrogen, chlorine. Such an atom attracts the bonding electrons and becomes negatively charged; the

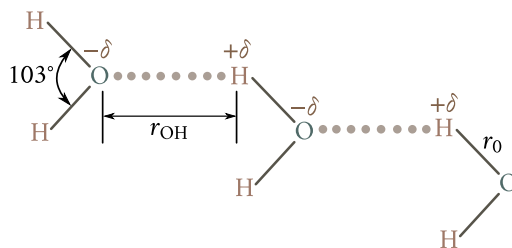


Figure 1.9: Hydrogen bond between water molecules.

hydrogen atom after losing the bonding electron assumes a positive charge. The hydrogen bond is a result of electrostatic attraction of those charges.

As a typical example we may cite the hydrogen bond between molecules of water (Figure 1.9). The O–H bond between an oxygen atom of one molecule and a hydrogen atom of another behaves as a tiny dipole with a $-\delta$ charge on the oxygen atom and a $+\delta$ charge on the hydrogen atom. The force of attraction between those charges is the cause of the hydrogen bond shown by dots in Figure 1.9. The attraction is enhanced by the small dimensions of the hydrogen atom that enable it to come close to the electronegative atom. Still, this distance $r_{OH} = 2.76 \text{ \AA}$ is much greater than the radius of the covalent bond H–O, r_0 , in the water molecule itself, which is equal to 0.96 \AA . This is quite natural since the energy of the covalent bond is about an order of magnitude higher than that of the hydrogen bond. Its value for water is $21 \times 10^3 \text{ J mol}^{-1}$ to $25 \times 10^3 \text{ J mol}^{-1}$.

The hydrogen bond is the cause of association of molecules of liquids (water, acids, spirits, etc.), which results in greater viscosity, higher boiling point, abnormal thermal expansion, etc. Water may serve best to illustrate this. Should there be no hydrogen bonds between the molecules of water, its boiling point at atmospheric pressure would be not $+100^\circ\text{C}$ but -80°C and its viscosity would be lower by almost an order of magnitude. When water is heated above 0°C , the hydrogen bond is destroyed. Since the hydrogen bond is responsible for the loose structure of associated complexes, in which the water molecules are rather far apart (2.76 \AA), the destruction of such a loose structure should result in an increase in the density of water. On the other hand, an increase in the temperature of water and a corresponding increase in the intensity of thermal motion of its molecules should lead to thermal expansion and a decrease in its density. Experiment shows that in the temperature range 0°C to 4°C the first factor—the increase in density due to the disruption of the hydrogen bonds—is the prevalent one. Because of that within this range the density of water rises upon heating. Above 4°C the other factor—thermal expansion—prevails. This is why when water is heated above 4°C its den-

sity decreases, as is the case with other (normal) liquids.

§ 6. Comparison between bonds of various kinds

The van der Waals bond is the most universal one. It exists in all cases without exception. At the same time this is the weakest, having an energy of the order of 10^4 J mol^{-1} . Ideally, it operates between neutral atoms, or molecules, with closed inner electron shells. Specifically, the van der Waals forces are responsible for the existence of the liquid and solid states of inert gases, hydrogen, oxygen, nitrogen and many other organic and inorganic compounds. They also, as we will see later, form bonds in many of the molecular valence crystals. Because of low energy values of the van der Waals bond all structures based on it are unstable, volatile and have low melting points.

The ionic bond is a typical chemical bond very frequent among the inorganic compounds such as metal-halide compounds, metallic oxides, sulfides and other polar compounds. The ionic bond is also a feature of numerous intermetallic compounds (carbides, nitrides, selenides, etc.). The energy of the ionic bond is much higher than that of the van der Waals bond and may be as high as 10^6 J mol^{-1} . Because of that solids based on the ionic bond have high sublimation heat values and high melting points.

The covalent bond is extremely widespread among organic compounds, but is also present in inorganic compounds, in some metals and in numerous intermetallic compounds. This bond is responsible for the existence of valence crystals of the diamond, germanium and other types, as will be discussed below. The energy of the covalent bond is also high ($\sim 10^6 \text{ J mol}^{-1}$), which stems from the fact that the solids with this type of bond have high melting points and high heats of sublimation.

The metallic bond formed as a result of the collectivization of the valence electrons is characteristic of typical metals and numerous intermetallic compounds. The order of magnitude of the energy of this type of bond is comparable to that of the energy of covalent bond.

Lastly, the hydrogen bond, although relatively weak, still plays an important part in nature.

It should be pointed out that in real solids no types of bonds discussed above ever exist purely by themselves. Practically, there is always a superposition of two types of bonds or more. One of them, as a rule, plays a dominant part in determining the structure and the properties of the solid.

§ 7. Forces of repulsion

For the formation of a stable system of interacting atoms or molecules, together with forces of attraction there should be forces of repulsion, which would prevent the complete merging of the particles.

The origin of the forces of repulsion is first of all the interaction of the nuclei each of which carries a considerable positive charge. The energy of this interaction, U'_{rep} depends on the distance between the nuclei and on the degree of screening by their internal electron shells.

The following expression for U'_{rep} may be obtained from quantum mechanical calculations:

$$U'_{\text{rep}} \propto e^{-r/a} \quad (1.14)$$

where r is the distance between the nuclei, and $a = 0.529 \text{ \AA}$ the Bohr radius.

This dependence of U'_{rep} on r determines the nature of the forces of repulsion: they attain enormous values at short distances and fall off abruptly as r increases. For instance, when the distance between a proton and a hydrogen atom decreases from $r = 2a$ to $r = a/2$ (4 times), the repulsive energy increases almost 300-fold.

The repulsive forces due to the interaction of the nuclei play a dominant role when light atoms, whose nuclei are rather poorly screened by the electron shells, are brought together. In all other cases the dominant part is played by repulsion due to the overlapping of closed electron shells of the atoms being brought together. Consider by way of an example the interaction between a chlorine ion with a closed 3p shell and a sodium ion with a closed 2p shell. When the ions are brought together to a distance at which the 3p and 2p shells overlap, the number of electrons in each of them begins to exceed that which is compatible with the Pauli exclusion principle. Because of that some of the electrons are forced to go to higher energy levels (for instance, 3d or 4s). This results in an increase in the system's energy and, consequently, in the appearance of forces of repulsion. Quantum mechanical calculations show the energy of such repulsion to have an exponential dependence on the distance, as well:

$$U''_{\text{rep}} \propto e^{-r/\rho} \quad (1.15)$$

where ρ is a constant usually obtained experimentally.

Often the repulsion energy is expressed in the form (1.7). This expression gives a less steep decline of U_{rep} with the increase in r and is less consistent with experiment than (1.14) or (1.15) but is nevertheless widely used by researchers.

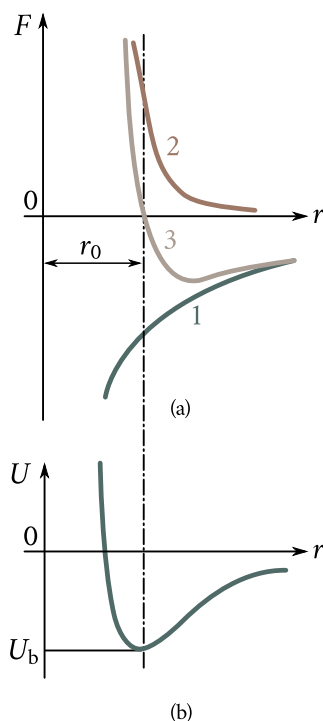


Figure 1.10: Interaction between approaching particles: (a) — interaction forces, 1 — force of attraction, 2 — repulsive force, 3 — resultant force; (b) — interaction energy.

§ 8. Crystal lattice

No matter what the origin of forces appearing when particles are brought together is, their general nature is the same [Figure 1.10(a)]: at comparatively large distances forces of attraction F_{att} increase rapidly as the distance between the particles decreases (curve 1); at small distances forces of repulsion F_{rep} come into being and with a further decrease in r they increase much more rapidly than F_{att} (curve 2). At the distance $r = r_0$ the repulsive forces counterbalance the forces of attraction, the resultant force of interaction F vanishes (curve 3), and the energy attains its minimum value U_b [Figure 1.10(b)]. Because of this the particles brought together to a distance r_0 are in a state of stable equilibrium. By the same reason the free particles should arrange themselves in a strict order at a distance r_0 from one another thus forming a body with a regular internal structure—a crystal. Such a structure will remain stable until the absolute value of the bond energy remains greater than the energy of thermal motion of the particles. The particles constituting the crystal cannot freely leave their equilibrium sites because this involves an increase in their

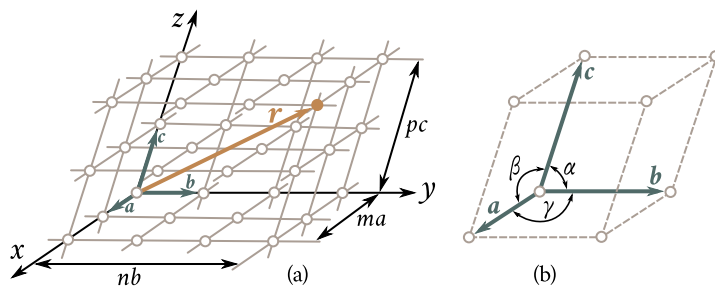


Figure 1.11: Crystal lattice: (a) — Bravais lattice, (b) — unit cell of Bravais lattice.

energy and leads to the appearance of forces tending to return them to their equilibrium sites. One may say that the particles are fixed in their equilibrium sites. The only form of motion allowed to them are random vibrations around their equilibrium sites.

To describe the regular internal structure of crystals one may conveniently use the concept of the crystal lattice. There are lattices of two types—the translational Bravais lattice and the lattice with a basis.

Bravais lattice. From the geometrical point of view a regular periodic arrangement of particles may be described with the aid of a translation. Figure 1.11(a) shows a lattice obtained with the aid of translation of a particle along the three axes: OX over the sections $a, 2a, 3a, \dots, ma, \dots$; OY over the sections $b, 2b, 3b, \dots, nb, \dots$; OZ over the sections $c, 2c, 3c, \dots, pc, \dots$ (m, n, p are integers). The position of any particle in this lattice is described by the vector

$$\mathbf{r} = m\mathbf{a} + n\mathbf{b} + p\mathbf{c}. \quad (1.16)$$

The vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ are termed the *translation vectors* and their numerical values the *translation periods*.

A lattice built with the aid of translation of any site along the three directions is termed a *Bravais lattice*. The smallest parallelepiped built on the vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$ is termed the *unit cell* of the crystal (Figure 1.11(b)). The shape and the volume of all the unit cells comprising the lattice are identical. All cell apexes are occupied by identical atoms or groups of atoms and are therefore equivalent. They are termed *lattice sites*.

To describe a unit cell, six quantities should generally be stated: three edges of the cell (a, b, c) and three angles between them (α, β, γ). Those quantities are termed the *parameters* of the unit cell. Often the sections a, b, c are used as units of length in lattices instead of the metre. They are termed *axial units*.

Unit cells with particles only at the vertices are known as *primitive cells*. There is only one particle to each such cell.

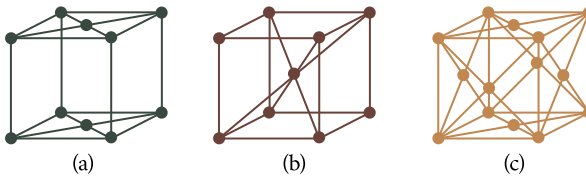


Figure 1.12: Typical crystal structures: (a) — base-centered (BaseC); (b) — body-centered (BC); (c) — face-centered (FC).

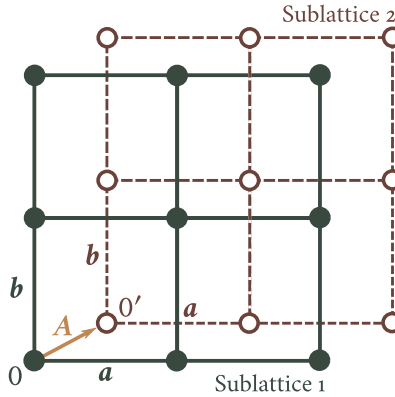


Figure 1.13: Two-dimensional lattice with a basis: \mathbf{A} — basis vector.

In some cases to express the lattice symmetry more fully the unit cells are built so that they contain particles not only in their apexes but at other points as well. Such cells are termed *complex cells*. The most widespread types of such cells are (Figure 1.12) the body-centered (BC), the face-centered (FC), and the base-centered (BaseC) cells. It may be shown that such cells may easily be reduced to primitive cells. Because of that they are Bravais-type cells.

A lattice with a basis. Not every type of lattice may be obtained by translation of a single site. Figure 1.13 shows a two-dimensional lattice with a general-type basis. It may easily be seen that it is impossible to describe the unit cell of such a lattice in terms of a single-site unit cell. Such a lattice may be imagined as consisting of two Bravais lattices, 1 and 2, each determined by the basis vectors \mathbf{a} and \mathbf{b} and inserted into each other. The relative displacement of the lattices is described by an additional basis vector \mathbf{A} . The number of such basis vectors may be arbitrary.

The lattice of this type is termed the *lattice with a basis*. It may be built with the aid of the same translations as can be used to build any of the Bravais lattices that make it up. However, in this case we shall have to translate not one site but several sites—the *basis*, defined by the totality of the basis vectors. Thus, the two-dimensional lattice shown in Figure 1.13 may be obtained by a translation of the

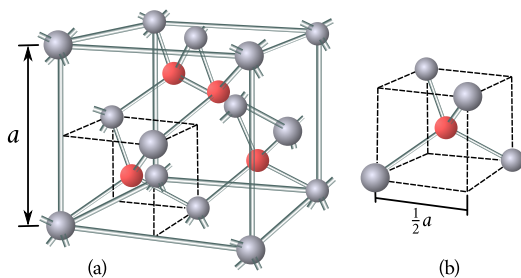


Figure 1.14: Diamond lattice: (a) — spatial arrangement of atoms in the lattice; (b) — tetrahedral pattern of atoms in the lattice.

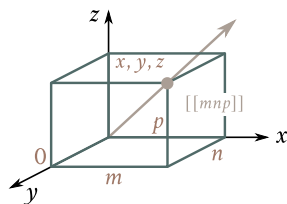


Figure 1.15: Indices of a crystal lattice site.

basis made up of two sites: 0 and 0'. An example of a three-dimensional lattice with a basis is the diamond lattice shown in Figure 1.14(a). It may be obtained by inserting one FCC (face-centered cubic) lattice into another FCC lattice displaced along the space diagonal by one-fourth of its length. Figure 1.14(b) shows a tetrahedron designated by a dotted line in Figure 1.14(a). It may be seen from Figure 1.14(b) that in the diamond lattice every atom is surrounded by four nearest neighbours in the apexes of the tetrahedron whose edge is $a/2$.

§ 9. Notation to describe sites, directions, and planes in a crystal

Let us mention briefly the notation conventionally used to describe sites, directions and planes in a lattice, the *Miller indices*.

Site indices. The position of any lattice site relative to the chosen origin of coordinates is defined by three of its coordinates (Figure 1.15): x, y, z . These coordinates may be expressed in the following form:

$$x = ma, \quad y = nb, \quad z = pc$$

where a, b, c are the lattice parameters, and m, n, p are integers.

Should we use lattice parameters as units of length along the respective axes we would obtain the lattice coordinates simply in the form of numbers m, n, p . These numbers are termed *site indices* and are written in the form $[[mnp]]$. For a negative index the minus sign is written above the index. For instance, for a site with coordinates $x = -2a, y = -lb, z = 3c$ the indices are written in the form $[2\bar{1}3]$.

Indices of direction. To describe a direction in a crystal a straight line passing through the origin is chosen. The position of this is uniquely defined by the indices of the first site through which it passes (Figure 1.15). Therefore the indices of the

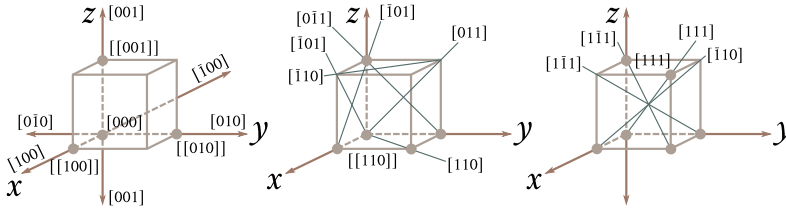


Figure 1.16: Indices of principal directions in a cubic crystal.

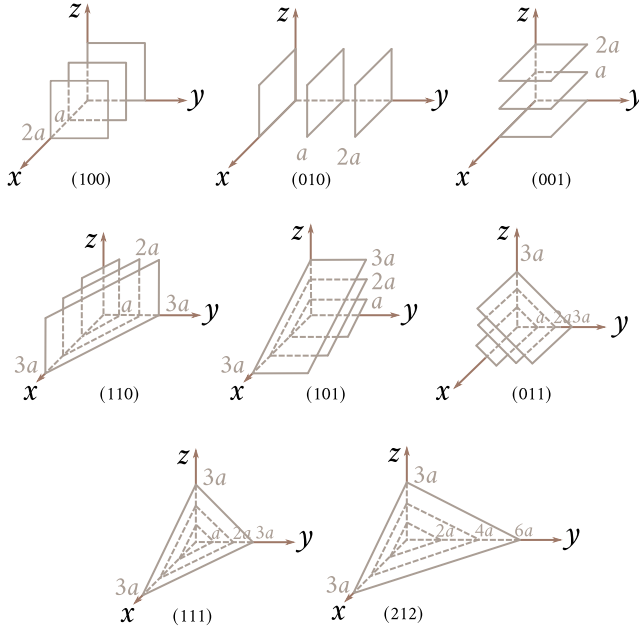


Figure 1.17: Indices of principal planes in a cubic crystal.

site are at the same time the indices of the direction. The usual notation for a direction is $[mnp]$. The indices of direction are, by definition, the three smallest integers that describe the position of the site nearest to the origin which lies on the given direction. For instance, the indices of the direction that passes through the origin and the site $[[435]]$ are $[435]$. Figure 1.16 shows the principal directions (crystallographic orientations) in a cubic crystal and their notation.

Plane indices. The position of a plane is defined by the choice of three sections A, B, C it cuts off when it intersects with the three coordinate axes. The procedure of finding the indices of such a plane is as follows.

The sections ABC are expressed in axial units and the reciprocal quantities are written as $1/A, 1/B, 1/C$. A common denominator is found for all the three

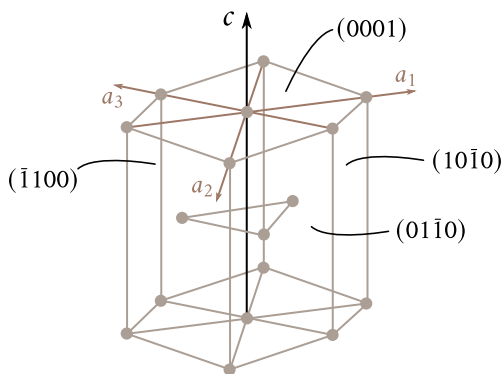


Figure 1.18: Indices of principal planes in a hexagonal crystal.

fractions. Suppose it is D . Then the integers $h = D/A$, $k = D/B$, $l = D/C$ will be the plane indices. They are written in the form (hkl) .

Determine, for example, the indices of a plane that cuts off the sections $A = 1/2$, $B = 2$, $C = 1/3$ on the x , y , z axes respectively. The ratios $1/A \div 1/B \div 1/C = 1/(1/2) \div 1/2 \div 1/(1/3) = 2 \div 1/2 \div 3$. The common denominator is $D = 2$. The indices of the plane are $h = 2/(1/2) = 4$, $k = 2/2 = 1$, $l = 3/(1/2) = 6$. The plane is denoted (416) . Figure 1.17 shows the principal planes of the cubic lattice.

It may easily be shown that in a cubic crystal the distances between the planes belonging to a given family may be expressed with the aid of the indices of these planes in the following way:

$$d = \frac{a}{\sqrt{h^2 + k^2 + l^2}} \quad (1.17)$$

where a is the lattice parameter. This formula shows that the greater are the plane indices the shorter is the distance between the planes.

To denote the planes in a hexagonal crystal a four-axis coordinate system is used (Figure 1.18): three axes (a_1 , a_2 , a_3) make angles of 120° with one another and lie in the base of a hexagonal prism, the fourth axis, c , being perpendicular to the base plane. Every plane is denoted by four indices: $hkil$. The additional label i occupies the third place and is calculated with the aid of h and k : $i = -(h + k)$. The base plane parallel to the axes a_1 , a_2 , a_3 has the indices (0001) . The planes parallel to the lateral faces of the prism have indices of the (1010) type. There are three such planes (not parallel to one another). They are termed *first-order planes*.

§ 10. Classification of solids based on the nature of bonds

The nature of the crystal structure is primarily dictated by the nature of bonding forces acting between the structural particles (atoms, ions, molecules) which make up the solid. In accordance with the five existing types of bonds there are five principal types of crystal lattices: *ionic*, or *coordination*, *lattices*, with the ionic bond playing the main part; *molecular lattices*, with the van der Waals forces mainly responsible for the bonding; *atomic lattices*, with bonds of a distinctly covalent type; *metallic lattices*, with characteristic metallic bonds; and lattices with the *hydrogen bond*.

Let us analyze from this viewpoint the crystal structure of chemical elements and of typical chemical compounds (see Appendix Sec. ??, Table ??).

The chemical elements may be roughly divided into four classes according to the type of crystal structure. The analysis may best be started with Class IV.

Class IV. This class includes all the inert gases. In the process of compression and crystallization of these gases only comparatively weak van der Waals forces act between the atoms, which have spherically symmetrical electron shells. Acted upon by these forces the symmetrical atoms draw together to form a most tightly packed face-centered cubic lattice. Every atom is surrounded by 12 nearest neighbours. The number of nearest neighbours is usually termed the *coordination number* of the lattice.

Class III. The Class III includes silicon and carbon from the short periods of the Mendeleev periodic table, germanium and tin from Group IVB, and all the elements from Groups VB, VIB, VIIB.

The crystallization of the elements of those classes proceeds in conformity with the $(8 - N)$ -rule: every atom in the lattice is surrounded by $8 - N$ nearest neighbours, N being the number of the group to which the element belongs. Thus diamond, silicon, germanium and gray tin are elements of Group IV of the Periodic Table. Therefore the coordination number of their lattices should be $8 - 4 = 4$. They all do have a tetrahedral lattice in which every atom is surrounded by 4 nearest neighbours, as is shown in Figure 1.19(a). Phosphorus, arsenic, antimony and bismuth belong to Group V. Their coordination number is $8 - 5 = 3$. Every atom has 3 nearest neighbours lying in a plane, as shown in Figure 1.19(b). Their lattice has a laminate structure, with the atomic layers bonded by van der Waals forces.

Selenium and tellurium belong to Group VI and have a coordination number 2. Their atoms form long spiral-shaped chains with each atom having two nearest neighbours [Figure 1.19(c)]. The chains are bonded by van der Waals forces. Lastly, iodine belongs to Group VII [Figure 1.19(d)] and has a coordination number 1. The atoms in the iodine lattice are arranged in pairs bonded by van der Waals forces.

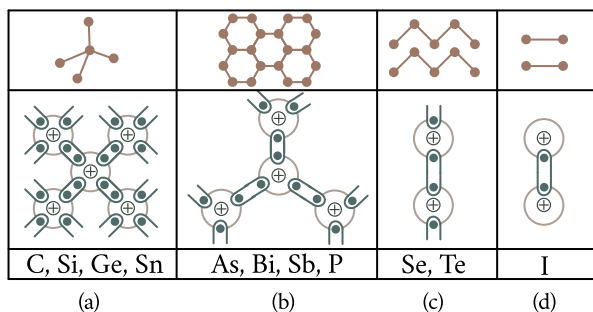


Figure 1.19: Crystal structure of chemical elements crystallizing in accordance with the $(8 - N)$ -rule: (a) — elements of Group IVB; (b) — elements of Group VB; (c) — elements of Group VIB, (d) — elements of Group VIIB.

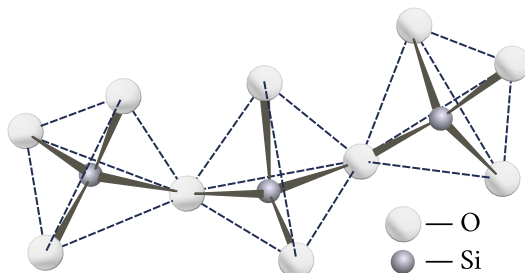


Figure 1.20: Structure of quartz SiO_2 crystal.

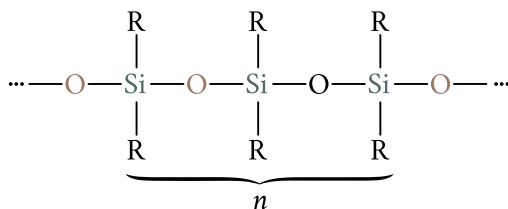
This explains the high volatility of iodine.

Such nature of the crystal structure of chemical elements whose crystallization conforms to the $(8 - N)$ -rule is quite understandable. The atoms of Group IV elements have 4 electrons in their outer shell. They lack 4 additional electrons to make up a stable 8-electron configuration. They make up the deficit by exchanging electrons with 4 nearest neighbours, as shown in Figure 1.19(a), forming a strong covalent bond with each of them. Accordingly, every atom in the crystal lattice of those elements has 4 nearest neighbours. In the same way the electron shells of Groups V, VI, VII are completed to contain 8 electrons.

Many chemical compounds crystallize in crystals with covalent bonds. Quartz SiO_2 may serve as a typical example. In the quartz crystal every silicon atom is surrounded by a tetrahedron of oxygen atoms (Figure 1.20) bonded to the silicon atom by covalent bonds. Every oxygen atom is bonded to two silicon atoms thereby joining two tetrahedrons. In this way a three-dimensional net of Si–O–Si bonds is formed, and the hardness and the melting point of the resulting crystal are high.

It may be of interest to note that the Si–O–Si bonds may be arranged into a

one-dimensional chain. Such compounds described by the common formula



where R is an arbitrary organic group, are termed *silicones*. The number n in a chain may be as high as several million. The chains may be joined together with the aid of the lateral groups R . In this way new materials, silicone resins, are formed. Because of the high strength of the Si–O–Si bonds and of the high flexibility of silicone chains such resins retain their properties at much lower and much higher temperatures than natural rubbers. This fact enables them to be used for thermal shielding of space ships and aircraft, as well as in extreme arctic conditions.

Class I. This is the most populated class which contains metals. Since metallic lattices are made up not of atoms but of ions, having the spherical symmetry of the atoms of inert gases, it is to be expected that metals too should crystallize into the same tightly packed lattices as the inert gases. Indeed, the following three types of crystal lattices are characteristic for metals: the face-centered cubic lattice with the coordination number 12 (see Figure 1.12), the hexagonal close-packed (HCP) lattice with the coordination number 12 (see Figure 1.18) and the body-centered cubic lattice with the coordination number 8 (see Figure 1.12). The latter is the least closely packed metal lattice.

Class II. The chemical elements belonging to Class II are in a sense intermediate between metals and the Class III elements, which crystallize in conformity with the $(8 - N)$ -rule. For instance, the Group IIB elements Zn, Cd and Hg are metals and one would expect them to have a typically metallic lattice with a high coordination number. Actually, Zn and Cd crystals are a special modification of the compact hexagonal lattice in which every atom has 6 nearest neighbours instead of 12, as required by the $(8 - N)$ -rule. These atoms occupy the base plane. In the case of mercury the $(8 - N)$ -rule is observed even more strictly: its crystal structure is a simple rhombohedral in which every atom is surrounded by 6 nearest neighbours. Boron—an element of Group IIIB—has a lattice that may be described as a deformed lattice with 5 nearest neighbours. This too agrees with the $(8 - N)$ -rule.

The ionic bond, as was stated above, plays one of the main parts in the world of inorganic compounds, in particular, in numerous ionic crystals typically represented by the rock salt crystal NaCl (Figure 1.21). In such crystals it is impossible to pick out single molecules. The crystal should be regarded as a closely packed

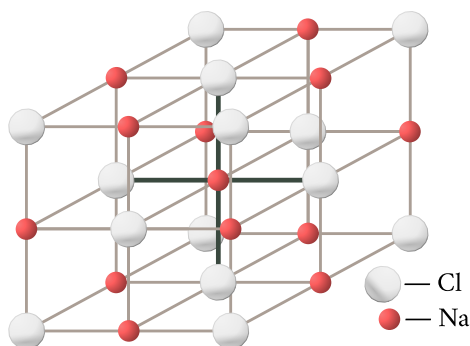


Figure 1.21: Structure of rock salt NaCl crystal.

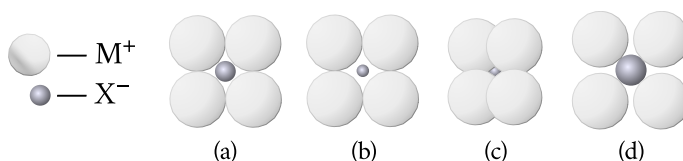


Figure 1.22: Effect of relative dimensions of ions on their packing in the lattice.

system of positive and negative ions whose positions alternate so that the electrostatic attraction between the nearest neighbours would be at its maximum. With the most favourable relation between the radii of the positive (M^+) and the negative (X^-) ions which exists in the NaCl crystal the ions “touch” one another [Figure 1.22(a)] and the closest possible packing is achieved, in which every ion is surrounded by 6 nearest neighbours of the opposite charge. When the ratio of the radii of the ions M^+ and is less favourable [Figure 1.22(b,c)] crystal structures with other coordination numbers, 4 or 8, are formed.

Ionic compounds of the MX_2 type, such as $CaCl_2$ and Na_2O , have still more complex lattices. But the principle upon which they are built remains the same: the ions are packed so as to be surrounded by ions of the opposite sign in accordance with the formula of the compound and the ratio of their radii.

Finally, let us consider crystals featuring the hydrogen bond. A typical representative of such crystals is ice. Figure 1.23(a) shows a two-dimensional diagram of the arrangement of water molecules in an ice crystal: each molecule is surrounded by four nearest neighbours a distance $r_{OH} = 2.76 \text{ \AA}$ away from it with whom it forms hydrogen bonds. In space the molecules occupy the vertices of a regular tetrahedron [Figure 1.23(b)]. The combination of such tetrahedra forms the regular crystal structure of ice [Figure 1.23(c)]. The structure is very loose and this is the cause of the abnormally low density of ice. Upon melting, some ($\sim 15\%$) of the hydrogen bonds are disrupted and the packing density of the water molecules rises

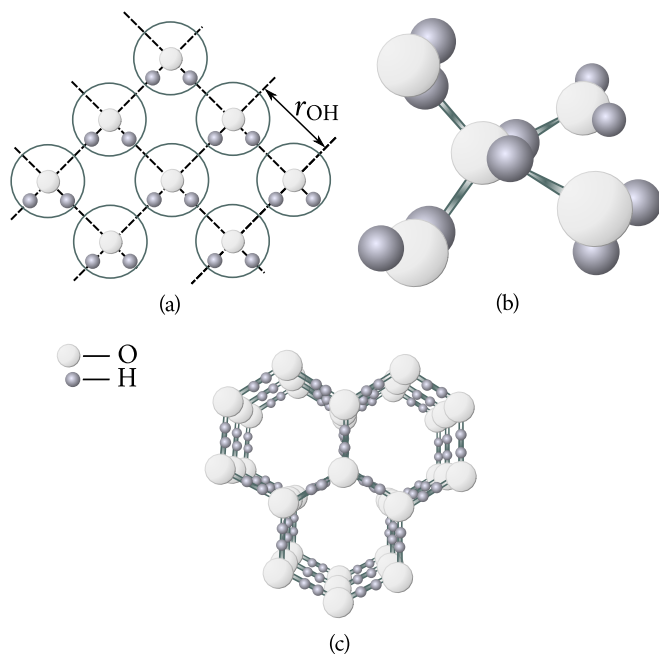


Figure 1.23: Crystals with hydrogen bond: (a) — plane diagram of arrangement of water molecules in an ice crystal; (b) — spatial arrangement of water molecules in an ice crystal; (c) — crystal structure of ice.

somewhat with the resultant rise in the density of water: the density of ice at 0°C is 916.8 kg m^{-3} , and the density of water at this temperature is 999.87 kg m^{-3} .

It may be of interest to note that if there was no hydrogen bond the melting point of ice would be -100°C instead of 0°C .

It should finally be stressed again that the hydrogen bond plays an extremely important part in vital biological compounds: protein molecules owe their helical shape exclusively to the hydrogen bond; the same type of bonds holds together the double helixes in the DNA. “It is no exaggeration to claim that life on our planet would have assumed radically different forms—if any at all—were hydrogen bonding not present in water and in the proteins and nucleic acids that compose living cells and that transmit hereditary traits”².

Table ?? of Appendix ?? shows the general classification of solids. The upper left corner contains typical metals with collectivized electrons (silver, copper) and the upper right corner typical valence crystals with distinctly localized electron

²G. C. Pimentel and R. D. Spratley, *Chemical Bonding Clarified Through Quantum Mechanics*, Holden-Day, San Francisco (1969), p. 261.

bonds. The extreme righthand part contains crystals with van der Waals bonds. Such elements as silicon and germanium occupy an intermediate position between the metals and the valence crystals. At absolute zero they are typical valence crystals; however, as temperature rises the valence bond is gradually disrupted and they begin to exhibit metallic properties. Such solids as sulphur, phosphorus, and selenium occupy an intermediate position between the valence crystals and crystals with the van der Waals bond.

The lower left corner of the diagram contains alloys of the NiCu type with the characteristic metallic bond and the lower right corner —ionic crystals (sodium chloride). Intermediate positions between them are occupied by numerous intermetallic compounds of the Mg_3Sb_2 type featuring the ionic bond (Mg_3Sb_2 corresponds to a $\text{Mg}^{+2} - \text{Sb}^{-3}$ compound). Intermediate position between the ionic and the valence crystals is occupied by such compounds as SiO_2 and SiC , with bonds of partially ionic nature made possible because of electron displacement. Compounds of the FeS and the TiO_2 (titanium dioxide) type occupy an intermediate position between ionic crystals and crystals with the van der Waals bond.

There are a great many crystals in which ionic or covalent bonds act in atomic planes while the bonds between the planes are of the van der Waals type.

§ 11. Polymorphism

Some solids have two or more crystal structures each of which is stable in an appropriate range of temperatures and pressures. Such structures are termed *polymorphic modifications*, or *polymorphs*, and the transition from one modification to another, *polymorphic transformation*.

It is the practice to denote polymorphic modifications by Greek letters: the modification stable at normal and lower temperatures is denoted by α ; modifications stable at higher temperatures are denoted by the letters β , γ , δ , etc. The polymorphism of tin may serve as a classical example. Below 13.3°C the stable modification of tin is α -Sn, which has a tetragonal cubic lattice of the diamond type. This is the so-called gray tin. It is brittle and may easily be ground to powder. Above 13.3°C α -Sn transforms into β -Sn, which has a body-centered tetragonal lattice. This is the familiar white metallic tin, a rather ductile metal. The transformation from β -Sn to α -Sn is accompanied by a considerable increase in specific volume (by about 25%). Long ago when many things were made of tin, the perplexing phenomenon of growing bulges on them and their subsequent destruction following excessive cooling was attributed to a mysterious metal disease, the “tin plague”.

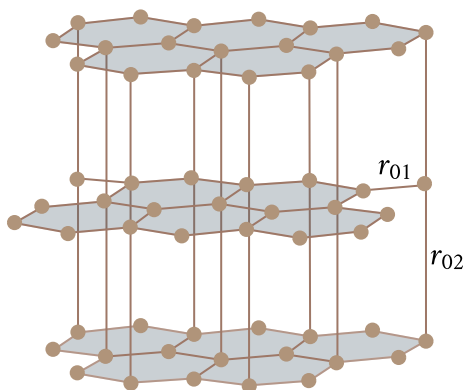


Figure 1.24: Crystal structure of graphite.

Many other chemical elements also exhibit polymorphism: carbon, iron, nickel, cobalt, tungsten, titanium, boron, berillium, and others, as well as many chemical compounds and alloys.

An interesting and a practically important case of polymorphism is that of carbon, which exists in the forms of diamond and graphite.

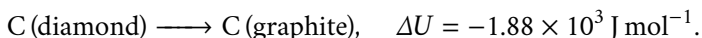
This case deserves a more detailed study. In the diamond lattice every atom is surrounded by 4 nearest neighbours occupying the vertices of a tetrahedron (see Figure 1.14) with whom it is bonded by strong covalent forces. The length of the bond is 1.544 \AA and the energy per bond is about $3.5 \times 10^5 \text{ J mol}^{-1}$.

The graphite lattice is of the pattern characteristic of the Group VB element lattices: the carbon atoms form two-dimensional layers in which every atom is surrounded by 3 nearest neighbours with whom it is bonded by covalent bonds (Figure 1.24). The length of the bond is $r_{01} = 1.42 \text{ \AA}$, that is, less than in the diamond lattice, therefore the former is stronger. The distance between the layers is much greater than the length of the C–C bond, being equal to $r_{02} = 3.6 \text{ \AA}$. Only weak van der Waals forces can act at such great distances and the layers are held together by them. The energy of this bond is $4 \times 10^3 \text{ J mol}^{-1}$ to $8 \times 10^3 \text{ J mol}^{-1}$.

Such a great difference in the nature of the bonding forces in the diamond and graphite structures should evidently manifest itself in a great difference in their properties, which is actually the case. Graphite slides easily along the planes held together by weak van der Waals forces. Therefore it is used to advantage in making “lead” pencils and as dry lubricant. The electrons in diamond are held securely between the atoms forming bonds. Light of the visible part of the spectrum is unable to knock out such electrons and therefore is not absorbed in diamond. Because of this diamond is an ideal transparent crystal unable to conduct electric current (a di-

electric). In graphite one of the four valence electrons of the carbon atom is actually collectivized by the atoms forming the layer. Such electrons can easily be moved by the action of an external electric field, making graphite a two-dimensional conductor. The presence of mobile electrons explains light absorption (the gray colour of graphite) and its characteristic metallic glitter.

In normal conditions graphite is a somewhat more stable modification than diamond, although the difference in the energies of those modifications is quite small—of the order of $2 \times 10^3 \text{ J mol}^{-1}$:



Still, such a difference is enough to bring about a sufficiently rapid transformation of diamond into graphite when heated above 1000°C in the absence of air.

The density of diamond is greater than that of graphite (3500 and 2250 kg m^{-3} , respectively), which is due to a loose packing of the atomic layers in graphite. Therefore at greater pressures diamond becomes more stable and graphite less stable. At sufficiently high pressures diamond becomes more stable than graphite. In such conditions by raising the temperature to increase the mobility of the carbon atoms we may bring about the transformation of graphite into diamond.

The conditions for such transformation to proceed at a practical rate were calculated by the Soviet physicist O. I. Leipunskii. He writes: “Firstly, graphite should be heated to at least 2000°C for the carbon atoms to be able to move from place to place. Secondly, it must be subjected to very high pressure, not less than 60000 atm ”³. These conditions were first achieved by the scientists of the General Electric Research and Development Center, who in 1954 succeeded in producing the first synthetic diamonds in the form of dark unsightly crystals, the largest being 1.5 mm long. Subsequently, the synthesis of diamonds was mastered in Sweden, the Netherlands, and Japan.

In the Soviet Union the production of synthetic diamonds on a commercial scale began in 1961. The pressure in the process is of the order of 100000 atm and the temperature about 2000°C . Synthetic diamonds produced by this process are harder and stronger than natural diamonds and their industrial use is about 40% more efficient than that of natural ones.

Another material of extreme hardness had been synthesized in a process similar to that of the diamond—the cubic boron nitride BN, which became known as *borazon*. It is harder than diamond and may be heated up to 2000°C in atmospheric conditions. In its hexagonal modification boron nitride is similar to graphite—a white powder oily to the touch.

From the theoretical point of view all solids should exhibit polymorphism pro-

³Leipunskii, O. I.: Quoted from I. I. Shafranovskii, *Diamonds*, “Nauka”, Moscow (1964).

vided the range of their stability is not limited by the processes of melting and sublimation. The existence of polymorphism is a direct consequence of the variation of the strength and the nature of the bonds in the crystal lattice caused by the changes in intensity of atomic motion and in the distance between them as a result of heating or of application of external pressure to the crystal. Close to absolute zero the stable structure should be that with the strongest bonds possible for the given atomic ensemble. In the case of tin, which belongs to Group IV of the Mendeleev periodic table, such structure is the diamond structure, in which every atom is bonded to 4 nearest neighbours by strong covalent bonds. However, as the temperature is raised, those bonds, because of their strict directionality and rigidity, are easily destroyed by thermal motion, and already above 13.3 °C the flexible metallic structure formed with the aid of collectivized electrons becomes more favourable. This bond has its own stable crystal structure, the tetragonal body-centred lattice.

The transition from one modification to another is accompanied by the liberation or absorption of latent heat of transformation and is therefore a phase transition of the first kind. Such a transition involves the transformation of the crystal lattice and this fact together with a low mobility of atoms in solids makes possible a practically infinitely long existence of a modification thermodynamically unstable in particular conditions. Diamond which can exist ages without turning into graphite—the stable modification in normal conditions—is a striking example of this point.

Polymorphism is very important for practical purposes. Heat treatment of steels to obtain various properties, the production of stainless steels, the treatment of various alloys to obtain the necessary properties are all to a large extent based on the use of polymorphism.

§ 12. Imperfections and defects of the crystal lattice

Mosaic structure. Numerous data obtained in the study of the structure of real crystals point to the fact that their internal structure is essentially not the same as that of an ideal crystal. To begin with, real crystals have a *mosaic structure*: they are made up of regular blocks which are only approximately parallel to one another. The dimensions of the blocks vary from 10^{-6} m to 10^{-8} m and the angles between them from several seconds to tens of minutes. Because of the different orientation of adjacent blocks there is a transition layer between them in which the lattice changes its orientation gradually from that of the first block to that of the second. Therefore in this layer the lattice is deformed as compared with that of an ideal

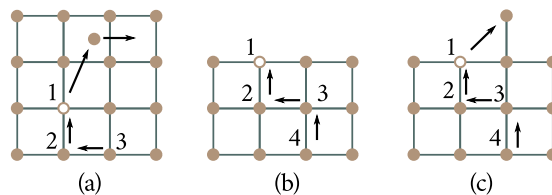


Figure 1.25: Crystal lattice defects: (a) — Frenkel defects; (b), (c) — Schottky defects.

crystal.

Lattice deformations are even greater near the grain boundaries in a polycrystal, since the orientation of adjacent grains may differ by as much as tens of degrees. The grain and block boundaries carry excess free energy, which increases the rate of chemical reactions, of polymorphic transformations, of diffusion, etc. They also serve as effective carrier scattering centres responsible for a considerable part of the solid's (metal or semiconductor) electrical resistance.

Frenkel defects. The distribution of energy among the atoms of a solid is very nonuniform, as is the case with the molecules of a gas or liquid. At any temperature there are atoms in the crystal whose energy is many times greater or less than the average energy corresponding to the law of equipartition of energy. The atoms that at a given instant of time have enough energy can not only move a considerable distance away from their position of equilibrium, but can also surmount the potential barrier set up by the neighbouring atoms and move over to new neighbours, to a new cell. Such atoms acquire the capability, so to say, to “evaporate” from their lattice sites and to “condense” in its internal cavities, or interstitials [Figure 1.25(a)]. This process results in the creation of a vacant site (a *vacancy*) and of an atom in the interstitial position (a *displaced atom*). Such lattice defects are termed *Frenkel defects*.

Calculation shows the equilibrium concentration of interstitial atoms n_F at a given temperature to be

$$n_F = AN \exp\left(-\frac{E_F}{k_B T}\right) \quad (1.18)$$

where E_F is the formation energy of the interstitial whose order of magnitude is several electron volts, N is the number of sites in the given volume, A is an integer (usually about 1) indicating the number of identical interstitial positions per one lattice atom, and k_B is the Boltzmann constant.

Both the interstitial atoms and the vacancies do not remain localized in one place but diffuse through the lattice. The diffusion of a displaced atom proceeds by the motion of this atom from one interstitial position to another, and the diffusion

of a vacancy by a relay process in which the vacancy is filled by neighbouring atoms [Figure 1.25(a)]: when atom 2 moves into vacancy 1 the vacancy moves over to site 2, when atom 3 moves to the now vacant site 2 the vacancy moves to site 3, etc.

Schottky defects. In addition to internal evaporation there is also a possibility of a partial or even complete evaporation of an atom from the surface of a crystal. Complete evaporation means that the atom leaves the crystal surface and joins the vapour phase [Figure 1.25(b)]. Partial evaporation means that the atom leaves the surface layer and arranges itself on top of it [Figure 1.25(c)]. In both cases a vacancy is produced in the surface layer. But when an atom from the interior of a crystal occupies a vacancy, the latter is pulled into the crystal and diffuses there. Here there are no displaced atoms to correspond to the vacancies, since the latter are produced without the simultaneous transition of atoms to interstitial positions. Such vacancies are termed *Schottky defects*. Calculations show the equilibrium number of vacancies n_{Scn} in a crystal of N sites to be

$$n_{\text{Scn}} = N \exp \left(-\frac{E_{\text{Scn}}}{k_{\text{B}} T} \right) \quad (1.19)$$

where E_{Scn} is the energy of formation of a single vacancy. It is somewhat lower than E_{F} . For instance, for aluminium it is equal to 0.75 eV. Substituting this value into (1.19), we obtain $n_{\text{Scn}} \approx 10^{18} \text{ m}^{-3}$ at $T = 300 \text{ K}$; at $T = 923 \text{ K}$, that is, close to the melting point of aluminium ($T_{\text{m}} = 933 \text{ K}$), $n_{\text{Scn}} \approx 10^{25} \text{ m}^{-3}$. Such values are characteristic of all metals at temperatures close to their melting points.

The energy of formation of the Frenkel defects is approximately equal to the sum of formation energies of the vacancy and the interstitial atom.

The Frenkel and Schottky defects play an important part in many processes in crystals. They act as carrier scattering centres reducing their mobility. They can also act as sources of carrier production, that is, play the role of donors and acceptors (usually the latter). They can also appreciably affect optical, magnetic, mechanical, and thermodynamic properties of crystals, especially of thin semiconducting films and fine crystalline specimens (because defect concentration in them is usually much greater than in bulk specimens).

Impurities. Impurities are one of the most important and most common type of defects in the structure of real crystals. Contemporary refining methods are unable to guarantee absolute purity of materials. Even the most pure materials contain up to 10^{-9} percent of impurities, which corresponds to a concentration of about 10^{17} impurity atoms per cubic metre of the material. To illustrate this degree of purity we would like to cite an equivalent example of one grain of rye contained in about 10 tons of wheat.

The impurities contained in the crystal may, depending on their nature, be in

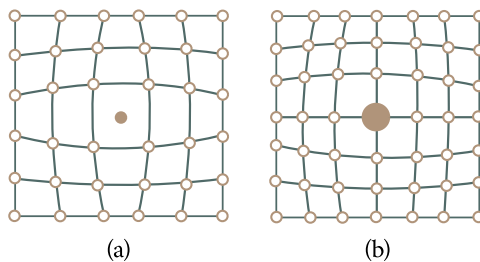


Figure 1.26: Deformations of crystal lattice of solid solutions: (a) — interstitial; (b) — substitution.

the form of dissolved atoms or in the form of inclusions of various dimensions. In the process of dissolution the impurity atoms enter the interstitial positions between the atoms of the crystal or substitute some of them in their sites. The solid solution of the first type is termed the *interstitial solution* [Figure 1.26(a)] and that of the second the *substitution solution* [Figure 1.26(b)]. Because of a difference in the physical nature and the dimensions of the impurity atoms from the atoms of the crystal, their presence results in the deformation of the crystal lattice.

Impurities may appreciably affect chemical, optical, magnetic, and mechanical properties of solids. They are effective carrier scattering centres, being the cause of electrical resistance that does not vanish at absolute zero temperature. In semiconductor crystals the impurities create new energy levels leading to the appearance of impurity conductivity. Calculations show that a perfectly pure silicon should have a specific resistance of the order of $2000\ \Omega\ \text{m}$. Active impurities contained in it in a concentration of 10^{-9} percent reduce the resistivity to several units. Technically pure germanium was for a long time regarded as a metal because its resistivity was of the same order as that of metals. Only perfect refining methods that brought impurity concentration in germanium down to 10^{-7} - 10^{-8} percent made it a typical semiconductor.

Interesting results were obtained in the course of investigations into the properties of extremely pure metals. Thus thoroughly purified iron turned out to be chemically inert and immune to corrosion even in conditions of tropical humidity. Titanium, chromium, bismuth, tungsten, molybdenum, which had a reputation for brittleness, turned out to be ductile even in conditions of extreme cooling; tin purified to contain no more than 5×10^{-6} percent impurities is so soft that it bends under its own weight like dough.

Some striking results were obtained in dehydration experiments: materials dried so as to contain negligible amounts of residual moisture change their properties in a marked degree. Thus dried oxyhydrogen gas does not explode at high

temperatures; carbon monoxide does not burn in oxygen; sulphuric acid does not react with alkali metals, etc. The English chemist H. B. Baker sealed 11 thoroughly purified individual chemical compounds in glass tubes together with phosphoric anhydride (a powerful absorber of water). The tubes were opened 9 years later in conditions that precluded the admission of moisture. The results were startling: the boiling point of all the compounds rose appreciably. For instance, the boiling point of benzol turned out to be 26 °C higher than that specified in tables, that of ethyl alcohol was 60 °C higher, that of bromine was 59 °C higher, and that of mercury was almost 100 °C higher. Subsequent experiments carried out by other investigators substantiated those results. More than that, it was established that very dry materials not only change their boiling point but melting point and other properties as well.

Despite substantial progress in the field of production of ultrapure materials there is a growing demand for better purification methods and presently there will be a need for materials with impurity concentrations of no more than 10^{-10} - 10^{-12} percent. This applies in the first instance to materials used for thermonuclear fusion apparatus, to microelectronics materials, as well as to materials used in other branches of industry. Such materials are not only difficult to produce but also difficult to keep pure, especially if they have to be processed before use. To illustrate how easy it is to make a mistake while working with such materials we would like to cite a case told by the well-known German physicist Werner Heisenberg. When a target was irradiated with a flux of neutrons in a mass spectrometer, gold nuclei were detected. This effect vanished after the experimenter took off and put away his gold-rimmed eyeglasses.

Chapter 2

Mechanical Properties of Solids

The mechanical properties—strength, hardness, ductility, wear-resistance—are the most characteristic of the properties of solids. Thanks to those properties the practical use of solids as constructional, building, electrotechnical, magnetic and other materials without which the growth of economy is impossible has become so widespread. The very names of the periods of human culture reflect the names of the solids whose mechanical properties made a qualitative leap in the process of development of human society possible—the Stone Age, the Bronze Age, the Iron Age.

This chapter deals briefly with modern physical concepts concerning the mechanical properties of solids, the laws of their plastic flow and destruction, the physical nature of strength, and prospects for the development of materials with unique mechanical properties.

§ 13. Elastic and plastic deformations. Hooke's law

When a crystal is subjected to an external extension load, the distances between the atoms become greater and the atoms are displaced from their equilibrium positions in the crystal. This destroys the equilibrium between the forces of repulsion and attraction characteristic of the equilibrium state of the atoms in the lattice and results in the appearance of internal forces tending to return the atoms to their initial equilibrium positions. The value of those forces per unit cross-sectional area of the crystal is termed *stress*. Let us calculate it.

It was shown in Chapter 1 that the energy of interaction of particles 1 and 2 in a solid is a function of the distance r between them. This can be described by the curve $U(r)$ schematically shown in Figure 2.1(a). When particle 2 is displaced from its equilibrium position to a distance x , that is, when the distance between

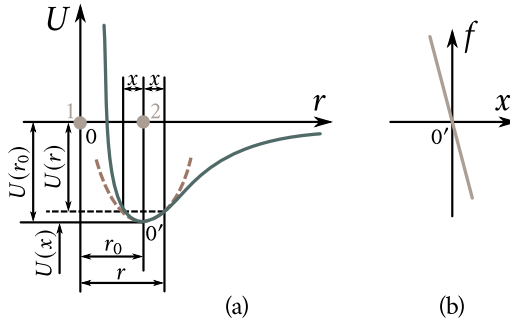


Figure 2.1: Variation of (a) interaction energy and (b) interaction force with the displacement of a particle from equilibrium position by a distance x .

the particles is increased to $r = r_0 + x$, the particle's energy grows and becomes $U(r)$. The change in energy $U(x) = U(r) - U(r_0)$ can be found if we expand $U(r)$ into a Taylor series in powers of x :

$$U(x) = \left(\frac{\partial U}{\partial r} \right)_0 x + \frac{1}{2} \left(\frac{\partial^2 U}{\partial r^2} \right)_0 x^2 + \frac{1}{6} \left(\frac{\partial^3 U}{\partial r^3} \right)_0 x^3 + \dots \quad (2.1)$$

Leaving only the quadratic term of the series and taking into account the fact that $(\partial U / \partial r)_0$ at point $0'$ is zero, we obtain

$$U(x) \approx \frac{1}{2} \left(\frac{\partial^2 U}{\partial r^2} \right)_0 x^2 = \frac{1}{2} \beta x^2 \quad (2.2)$$

where β is the rigidity of the bond.

We obtained an approximate expression for the change in energy of a particle brought about by its displacement from its equilibrium position to a distance x . This expression is an approximation because we left only the quadratic term in the expansion (2.1), neglecting higher-order terms. Graphically this dependence is expressed by a parabola shown in Figure 2.1(a) by a dotted line.

The force which appears between particles 1 and 2 when the distance between them is changed by x is equal to

$$f = - \frac{\partial U(x)}{\partial x} = -\beta x. \quad (2.3)$$

It follows from (2.3) that the force is linearly dependent on x and is directed towards the position of equilibrium, as indicated by the minus sign. It is well known that a body acted upon by such a force oscillates harmonically. Therefore such force is termed *harmonic*, the same term being applied to the approximation (2.2) (*harmonic approximation*). Figure 2.1(b) is a schematic diagram of the $f(x)$ dependence for small values of x . It is a straight line.

Now let us imagine that a tensile load F is applied to a rod with a cross-sectional

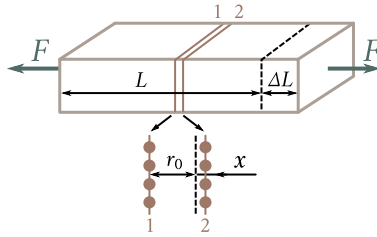


Figure 2.2: Uniaxial extension of a rod by an external force F ; 1 and 2 are the schematic representation of adjacent atomic planes.

area S and a length L . This load changes the distance between the neighbouring atomic planes 1 and 2 by the amount x causing thereby an extension of the rod by ΔL (Figure 2.2). It will be counterbalanced by the internal force F_{int} equal numerically to

$$F_{\text{int}} = fN = N\beta x \quad (2.4)$$

where N is the number of particles in the atomic layer of area S .

The stresses which appear in the extended rod will be

$$\sigma = \frac{F_{\text{int}}}{S} = \frac{N}{S}\beta x = cx \quad (2.5)$$

where $c = N\beta/S$. Multiplying and dividing the right-hand side of (2.5) by the distance between the atomic planes, r_0 , we obtain

$$\sigma = cr_0 \frac{x}{r_0} = E\varepsilon \quad (2.6)$$

where

$$E = cr_0 = \frac{N}{S}\beta r_0 \quad (2.7)$$

is termed the *elasticity modulus*, or *Young's modulus*, and

$$\varepsilon = \frac{x}{r_0} \quad (2.8)$$

is the relative change in the lattice parameter in the direction of the external force F .

Multiplying the numerator and the denominator of (2.8) by the number of atomic layers N' contained in the sample of length L , we obtain

$$\varepsilon = \frac{xN'}{r_0N'} = \frac{\Delta L}{L}. \quad (2.9)$$

Hence, ε is the relative elongation of the sample under the action of the external tensile load.

It follows from Eq. (2.6) that as long as the harmonic approximation remains valid, that is, as long as the forces acting between the particles displaced in relation

to each other as a result of the deformation of the body remain linear functions of the displacement, the stresses σ which appear in the body will remain proportional to the relative deformation of the body:

$$\sigma = E\varepsilon.$$

The elasticity modulus E serves as the proportionality factor.

Formula (2.6) expresses the well-known *Hooke's law*. It is valid only as long as the harmonic approximation is valid, that is, only for very small relative deformations ε .

The physical meaning of the elasticity modulus is quite evident from Eq. (2.6). Putting $\varepsilon = 1$, we find that $\sigma = E$. Hence, the elasticity modulus is numerically equal to the stress which is capable of causing an elongation $\Delta L = L$ of the sample, provided Hooke's law remains valid and the sample is not destroyed. No real material except rubber can stand such deformations.

Table 2.1 shows the values of the elasticity modulus of some metallic crystals.

It follows from data presented in Table 2.1 that the elasticity modulus of solids is very large (of the order of 10^{10} Pa to 10^{11} Pa), which is an indication that the bonding forces in those bodies are very strong.

For some crystals the value of the elasticity modulus depends appreciably on the direction in which the lattice is deformed. Table 2.1 shows the values of E for directions in which it is at its minimum and at its maximum. For some crystals the ratio $E_{\text{max}}/E_{\text{min}}$ may be as high as 3, pointing to a high degree of *anisotropy* of such crystals.

The elasticity modulus depends only on the nature of the atoms (molecules) making up the body and on their mutual arrangement. It can be changed only by

Table 2.1

Substance	E (GPa)		G (GPa)	
	maximum	minimum	maximum	minimum
Aluminium	77	64	29	25
Copper	194	68	77	31
Iron	200	135	118	60
Magnesium	514	437	184	171
Tungsten	400	400	155	155
Magnesium	126	65	497	278

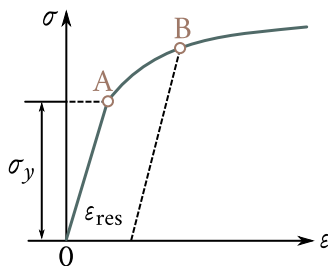


Figure 2.3: Typical extension curve of a ductile metal: σ_y — yield stress, ε_{res} — residual (plastic) deformation, OA—elastic deformation region, AB — plastic deformation region.

a substantial change in composition or internal structure of the solid. However, even in such cases the changes in E are relatively small. Thus, high concentration alloying, heat treatment, cold rolling, etc. of steel result in great improvement in its hardness and in other mechanical properties but only in negligible (up to 10%) changes in its elasticity modulus; alloying copper with zinc up to 40% leaves the elasticity modulus practically unchanged, although other properties experience a profound change.

We have discussed the tensile stress. However, all the considerations and the results obtained remain valid for other types of deformation—compression and shear—as well. In the latter case one should make use of the shear modulus G , whose values are also presented in Table 2.1.

When the external load is steadily increased, stress σ and deformation ε increase steadily too (Figure 2.3). At some stress σ_y , characteristic of the specific crystal, the crystal is either destroyed or the direct proportionality between σ and ε ceases and a residual (plastic) deformation ε_{res} sets in which remains after the external load has been removed. The first case is that of a brittle material and the second of a ductile one. The stress σ_y at which a noticeable plastic flow in the body sets in is termed the *yield stress* and OA and AB are the regions of the elastic and plastic deformations, respectively.

In brittle materials the elastic limit coincides with the tensile strength, and their destruction begins before a noticeable plastic flow sets in. In ductile metals, on the other hand, the elastic limit and the yield stress are, as a rule, much lower than the tensile strength, and these materials are destroyed only after a substantial plastic deformation has taken place.

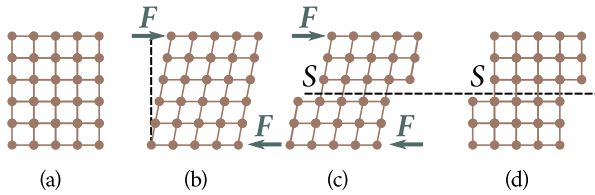


Figure 2.4: Crystal deformation by a shear force F : (a) — initial unstressed crystal; (b) — elastic deformation caused by shearing stress not exceeding elastic limit; (c) — early stages of plastic shear (slip) in the S plane caused by stress exceeding elastic limit; (d) — external force is removed, residual deformation (residual shift of one part of the lattice in relation to another) remains.

§ 14. Principal laws governing plastic flow in crystals

Residual deformations occur in all cases when the stress in ductile crystals tested for extension and compression exceeds the yield stress. However, neither extension nor compression can by themselves be the causes of such deformations. An increase in the length of the crystal can only result in an increase in the distance between the atomic planes perpendicular to the acting force. When these planes are drawn far enough apart, it may be that the forces of attraction shall no longer be able to compensate for the external load and the crystal will break. Compression can only draw the atomic planes closer together until the repulsive forces appearing between the atoms are able to counterbalance the external load. Deformation in this case is ideally elastic and cannot lead to irreversible displacement of parts of the lattice.

Plastic deformation may only be the result of shearing stresses, which are able to shift some parts of the crystal in relation to the others without destroying the bonds between them. Such displacement is termed *slipping*. It lies at the basis of the plastic flow process in crystalline materials. Figure 2.4 shows how residual deformations originate and develop in crystals [Figure 2.4(a)] acted upon by a shearing force F . As long as the elastic limit is not reached the crystal experiences elastic deformations [Figure 2.4(b)] with the tangential stresses growing in proportion to the relative shearing deformation γ (*Hooke's law*):

$$\tau = G\gamma \quad (2.10)$$

where G is the shear modulus. After the crystal is relieved from external load the atoms return to their initial positions. When the elastic limit is exceeded, one part of the crystal shifts in relation to another [Figure 2.4(c)] by one or more atomic distances along definite planes S termed *slip planes*. When the external load is withdrawn, the elastic stresses in the lattice vanish. However, one part of the crystal

remains displaced in relation to another [Figure 2.4(d)]. Such small irreversible displacements that proceed along numerous slip planes sum up to produce the residual deformation of the crystal as a whole.

The degree to which a crystal can be subjected to plastic deformations is determined, first of all, by the nature of the bonding forces acting between its structural elements.

The covalent bond with its rigorous directionality is appreciably weakened already by small relative displacements of the atoms. Shear destroys such bonds even before the atoms are able to establish them with other neighbouring atoms. On account of this the valence type crystals (such as diamond, silicon, germanium, antimony, bismuth, and arsenic) are incapable of plastic deformation. Outside the elastic deformation range they experience brittle destruction.

The metallic bond, which does not exhibit any directionality, on the other hand, remains practically unchanged as a result of relative tangential displacements of the atoms. This makes very great (some thousand atomic distances) relative displacements of some parts of the lattice possible, resulting in a high degree of ductility of crystals of this type.

The ionic bond occupies an intermediate position between the metallic and covalent bonds. It is less directional than the covalent bond but not so flexible as the metallic bond. Typical ionic crystals such as NaCl, CaF₂, and KCl are almost as brittle as the valence type crystals. At the same time silver chloride crystals are rather ductile.

Slipping takes place in crystals along definite crystallographic planes and directions, usually along the closest-packed planes and directions. This is because the closest-packed planes and directions are the strongest since the interatomic distances in them are the shortest and bonding is at its maximum. On the other hand, the distance between such planes is the greatest [see (1.17)]; on account of this the bonding between them is at its minimum. Slipping along such planes and directions results in the minimum disarrangement in atomic order and is therefore the easiest to accomplish.

The combination of the slip plane and the slip direction, which lies in it, forms the *slip system*. In the face-centered cubic lattice the slip plane coincides with the octahedral plane (111) and the slip direction with the direction of the body diagonal [111]. In hexagonal crystals the *SS* slip plane coincides with the base plane (0001) and the *X* slip direction with one of the three axes lying in the base plane (see Figure 2.5, where *P* is the external deforming load).

Numerous experiments have shown that the crystal begins to “slip” in the given slip system only after the shearing stress τ acting in this system reaches the critical

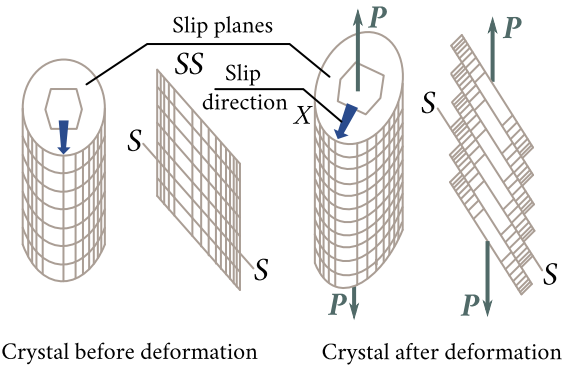


Figure 2.5: Slip planes and directions in a crystal.

value τ_{cr} termed the *critical shearing stress*. Table 2.2 shows the values of critical shearing stresses for some pure metallic single crystals.

It follows from the data of Table 2.2 that for the most ductile single crystals the critical shearing stress does not exceed 10^6 Pa.

The critical shearing stress depends to a large extent on the prior deformation of the crystal rising with the increase in the latter. This phenomenon became known as strengthening, or *cold working*. Thus a 350% preliminary deformation of the magnesium single crystal increases τ_{cr} nearly 25 times. Even greater is the effect of cold working on the cubic crystals—aluminium, copper, nickel, etc.

The strengthening of crystals is a witness to the fact that irreversible processes involving the relative displacement of atoms and of parts of the crystal take place. This results in changes of the internal energy of the crystals. Experimental study of this phenomenon has proved that the changes in the internal energy of solids in the process of their plastic deformation do, indeed, take place. Table 2.3 shows

Table 2.2

Metal	Impurity content (10^{-4})	Slip plane	Slip direction	τ_{cr} (10^7 Pa)
Cadmium	0.40	(0001)	[100]	0.058
Copper	10.0	(111)	[101]	0.100
Magnesium	5.00	(0001)	[100]	0.083
Nickel	20.0	(111)	[101]	0.580
Silver	1.00	(111)	[101]	0.060
Zinc	4.00	(0001)	[100]	0.094

the maximum amounts of energy that are accumulated by various metals in the process of their plastic deformation.

Should this energy be transformed into heat it would suffice to heat the metal by several degrees.

Since the accumulation of energy in the crystal in the process of its plastic deformation involves irreversible displacements of the atoms and of parts of the crystal, this energy is, in effect, the energy of *residual stresses* remaining in elastically deformed parts of the crystal lattice.

Because of a higher value of internal energy in a cold worked crystal it is less thermodynamically stable than the annealed crystal. This gives rise to processes that tend to bring the crystal to the equilibrium state. Relaxation and recrystallization are two such processes.

Relaxation consists in the dissipation of internal stresses, with the atoms of the deformed parts of the lattice returning to their regular positions. This process does not involve visible changes in the crystal structure and results in a partial or complete removal of the strengthening obtained as a result of plastic deformation. Being a diffusion-controlled process relaxation proceeds at a rate that strongly depends on temperature and on the latent heat of defect formation. Metals with a low melting point (such as tin, lead, cadmium, zinc) have comparatively high self-diffusion rates already at room temperatures. Accordingly, their relaxation rates at room temperatures are quite noticeable. At the same time there is practically no relaxation at room temperature in the metals with a high melting point; but the relaxation rate rises sharply as the temperature is increased (the relaxation process goes as far in 1 minute at 315 °C as it would in a hundred years at room temperature).

Another process that also results in the disappearance of the hardening in a cold worked crystal—the *recrystallization* process—proceeds intensely at temper-

Table 2.3

Metal	$Q \text{ (J kg}^{-1}\text{)}$
Aluminium	4400
Brass	2000
Copper	2000
Iron	4800
Nickel	3120

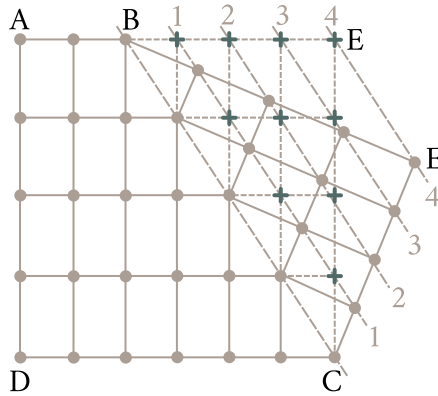


Figure 2.6: Twinning in a crystal: sign “+” denotes initial atomic positions in the twinning region.

atures of the order of one quarter of the melting temperature of the metal (on the absolute scale). In contrast to relaxation, which produces no visible changes in the crystal structure, recrystallization involves nucleation and growth of new crystals free from internal stresses. The nucleation of such crystals takes place in the first instance in the most deformed parts of the lattice, where much of the excess free energy is concentrated. In this way, a complete change in the microscopic structure of the crystal takes place with the crystal generally going over from the single state to the polycrystalline one. In the process of recrystallization the latent heat accumulated in the deformed crystal is given off in the form of heat.

§ 15. Mechanical twinning

Plastic deformation may also take the form of *twinning*, which is a process of step-by-step relative displacement of atomic planes parallel to the twinning plane by a fixed distance equal to a fraction of the lattice parameter. Figure 2.6 shows the diagram of twinning of the crystal AECDA. The area ABCDA is the undeformed part of the crystal, BECB is the part where twinning has taken place, and BC is the *twinning axis*. The positions of atoms before twinning are denoted by crosses. The plane passing through the twinning axis and separating the region of twinning from the undeformed part of the crystal is termed *twinning plane*.

Figure 2.6 shows that twinning results in the displacement of the atoms of the plane 11 relative to the twinning plane BC by a fraction of interatomic distance in the twinning direction. The plane 22 is displaced relative to the plane 11 by the same fraction of interatomic distance, the displacement relative to the twinning

plane being twice as great. In other words, every atomic plane parallel to the twinning plane is displaced in itself by a distance proportional to its distance from the twinning plane. As a result, the atoms in the twinned region assume positions that are mirror reflections of the positions in the undeformed part of the crystal in the twinning plane.

Twinning, in the same way as slipping, may take place only along specific crystallographic planes. In case of a face-centered cubic crystal this is the (112) plane, in case of a hexagonal close-packed crystal this is the (1012) plane, etc. For twinning to take place the tangential stresses should exceed some critical value. The process is a very rapid one and is usually accompanied by a characteristic crackle.

Because only negligible relative displacements of the neighbouring atomic planes are involved in the process of twinning it cannot result in a great residual deformation. For instance, a complete transition of a zinc crystal to the twinned form brings about only a 7.39% elongation. For this reason in crystals capable of plastic flow by means of slipping, twinning is responsible only for a negligible fraction of the total plastic deformation. In contrast to that, negligible deformation that precedes destruction of the valence crystals, in which slipping cannot take place, is due to twinning. In hexagonal crystals unfavourably oriented in relation to the external force twinning and subsequent reorientation of the crystal may result in appreciable residual deformations produced by the normal slipping process.

§ 16. Theoretical and real shear strengths of crystals

Shear is the principle mechanism of plastic flow in crystals. For a long time it was presumed that such shear takes place by means of a rigid displacement of one part of the crystal in relation to another simultaneously along the entire slip plane SS (Figure 2.7).

Let us make a rough estimate of the tangential stress needed to produce such shear.

The atoms of two adjacent parallel planes in an undeformed lattice occupy equilibrium sites corresponding to the minimum of the potential energy [Figure 2.7(a)]. The forces acting between them are zero. As one atomic plane is displaced relative to the other tangential stresses τ appear. They resist the shear and tend to bring back the original equilibrium state [Figure 2.7(b)]. If we assume the dependence of those stresses on the displacement to be sinusoidal (Figure 2.8), we shall be able to express the resistance to shear in the form

$$\tau = A \sin \left(\frac{2\pi x}{b} \right) \quad (2.11)$$

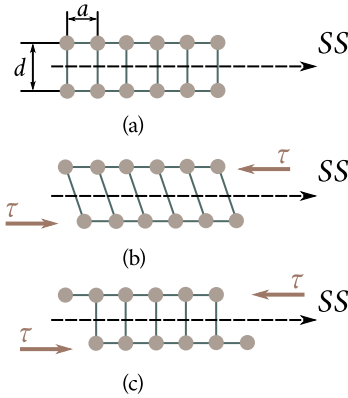


Figure 2.7: Diagram of rigid shear: (a) — equilibrium position of atoms in atomic planes adjoining the slip plane; (b) — gradual shift of one plane in relation to another caused by external stress τ ; (c) — lower atomic plane as a whole displaced by one interatomic distance in relation to the upper plane.

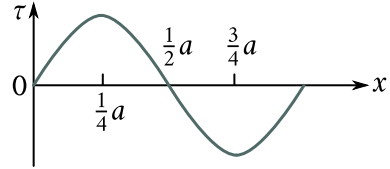


Figure 2.8: Variation of resistance to shear in the process of displacement of one part of the lattice in relation to another.

where x is the displacement of the atoms from their equilibrium positions, $b = a$ the interatomic distance in the slip plane, and A is a constant. For small displacements $\sin(2\pi x/b) \approx 2\pi x/b$ and therefore

$$\tau = A \left(\frac{2\pi x}{b} \right). \quad (2.12)$$

On the other hand, for small displacements Hooke's law is valid:

$$\tau = G \frac{x}{d} \quad (2.13)$$

where G is the shear modulus, and d the distance between the planes. From (2.12) and (2.13) we obtain $A = (b/d)(G/2\pi)$. Therefore

$$\tau = \frac{b}{d} \frac{G}{2\pi} \sin \left(\frac{2\pi x}{b} \right). \quad (2.14)$$

The maximum Value τ_{cr} of the tangential stress τ is attained for $x = b/4$ and this is assumed to be the theoretical strength:

$$\tau_{cr} = \frac{b}{d} \frac{G}{2\pi}. \quad (2.15)$$

Setting $b = d$, we obtain

$$\tau_{cr} = \frac{G}{2\pi}. \quad (2.16)$$

Hence, the critical shearing stress should be equal to about one tenth of the

shear modulus. A more rigorous consideration of the nature of the bonding forces acting between the atoms leads to a negligible correction factor. The minimum value that was obtained for τ_{cr} is $G/30$. Table 2.4 shows experimental and theoretical values of τ_{cr} for several metals.

A comparison of these figures shows that the real shear strength of crystals is some 3-4 orders of magnitude less than the theoretical value. This points to the fact that shear in crystals does not take place by means of a rigid relative displacement of atomic planes but by means of a mechanism involving the displacement of a comparatively small number of atoms at a time. The understanding of this fact led to the evolution of the dislocation theory of plastic flow of crystals.

§ 17. The dislocation concept. Principal types of dislocations

The dislocation theory of plastic flow assumes that the slipping process starts always at imperfections in the crystal structure and develops along the shear plane by means of a gradual motion of this imperfection which at a time involves only a limited number of atoms. Such imperfections are termed *dislocations*.

Edge dislocations. Suppose gliding took place in the crystal K in the plane ABCD in the direction of the vector \mathbf{b} involving the area AHED (Figure 2.9). The atomic planes on both sides from the slip plane AHED are displaced in relation to one another by the distance b in the slip direction. The boundary HE separating the area AHED, where slipping took place, from the area HBCE, where slipping has not yet taken place, constitutes an edge dislocation and the vector \mathbf{b} is termed

Table 2.4

Metal	τ_{cr} (10^7 Pa), experiment	G (10^7 Pa)	τ_{cr} (10^7 Pa), theory	
			$G/(2\pi)$	$G/30$
Cadmium	0.06	2640	420	88
Copper	0.10	4620	735	154
Iron	2.90	6900	1100	230
Magnesium	0.08	1770	280	59
Nickel	0.58	7800	1240	260
Silver	0.06	2910	459	97
Zinc	0.09	3780	600	126

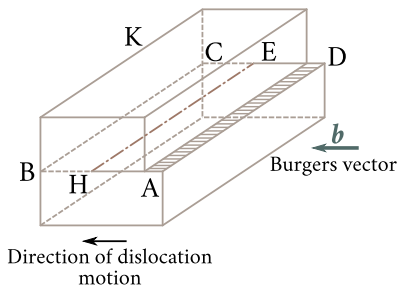


Figure 2.9: Shear that creates an edge dislocation. Shear took place only in region AHED of slip plane ABCD. Boundary HE is the edge dislocation.

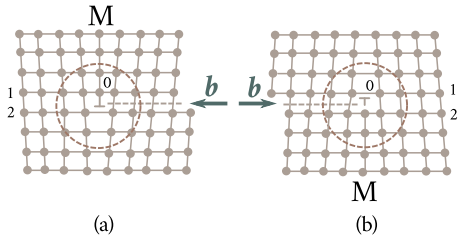


Figure 2.10: Arrangement of atoms in the plane perpendicular to dislocation line HE (see Figure 2.9). Dislocation occupies the region in which lattice atoms are displaced from their equilibrium sites (bounded by a circle): 0 — dislocation centre; (a) — positive dislocation; (b) negative dislocation.

the *Burgers vector*. It describes how far slipping has proceeded in the area AHED.

Figure 2.10 shows the arrangement of atoms in a plane perpendicular to the dislocation line. As a result of the shift which took place over the area AHED the upper part of the lattice contains one atomic plane (plane OM) more than the lower. Because of that the atomic row 1 lying above the shear plane contains one atom more than the row 2 below this plane. The interatomic distances in the upper row near the point 0 (the dislocation centre) will accordingly be shorter than the normal value (the lattice is contracted), while the interatomic distances in the lower row near the point 0 will be longer (the lattice is extended). As the distance to the left or to the right, and up or down, from the dislocation centre 0 increases, the deformation of the lattice gradually subsides and at an appropriate distance from 0 in the crystal normal disposition of atoms is restored. However, in the direction perpendicular to the plane of the diagram the dislocation passes through the entire crystal or through a considerable part of it.

Thus, a feature of the edge dislocation is the existence of an “excess” atomic plane OM in some part of the crystal. Therefore the process of formation of such a dislocation may be imagined as that of pulling the lattice apart and inserting an additional atomic plane in it. Such plane is termed *extra plane*. If the plane is inserted into the upper part of the lattice, the edge dislocation is assumed to be positive [Figure 2.10(a)]. But if the extra plane is inserted into the lower part of the lattice, the dislocation is assumed to be negative [Figure 2.10(b)]. A dislocation whose Burgers vector is equal to the lattice parameter is called the *unit dislocation*. When a unit dislocation passes through a cross section of the crystal, one part of it shifts in re-

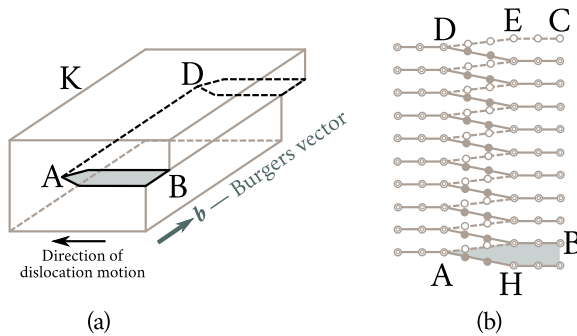


Figure 2.11: Formation of a screw dislocation: (a) — shear which produces the screw dislocation. It took place in the ABCD plane. Boundary AD is a screw dislocation; (b) — arrangement of atoms around the screw dislocation. Plane of drawing is parallel to slip plane. White circles denote atoms of the plane lying immediately above the slip plane, black circles denote atoms of the plane lying under the slip plane.

lation to the other by a distance b . The motion of a positive dislocation to the left causes the same shift of parts of the lattice as a motion of a negative dislocation to the right [Figure 2.10(a,b)].

Screw dislocations. Suppose an incomplete unit shift is made in the crystal K in the direction of the vector \mathbf{b} over the area ABCD, as shown in Figure 2.11(a); AD is the boundary of the area that experienced the shift. In Figure 2.11(b) the white circles denote the atoms of the plane immediately above the slip plane and black circles the atoms of the plane below the slip plane. In the undeformed part of the crystal to the left of AD the atoms of those planes are arranged one on top of the other; therefore the black circles coincide with the white (this is shown by white circles with circles in the centre). In the right-hand part of the crystal, where the shift has covered one interatomic distance, that is, to the right of EH, the atoms of the planes discussed above are also arranged one on top of the other. However, in the narrow strip ADEH the atoms of the upper plane are displaced in relation to those of the lower plane the more the farther away they are from the boundary AD. This displacement results in a local deformation of the lattice, which became known as the *screw dislocation*; the boundary AD is termed *dislocation axis*. The origin of the term screw dislocation may be easily understood from Figure 2.12: the motion of the atom “a” towards the atoms “b, c, d, e”, etc. [Figure 2.12(a)] lying in the plane of the screw dislocation proceeds, as may be seen from Figure 2.12(b), along a spiral. A distinction is made between right and left screw dislocations (Figure 2.13); the motion of both in opposite directions results in a shift in one direction.

Comparing Figures 2.9 and 2.11(a), we see that in contrast to the edge dislo-

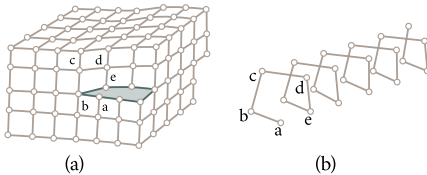


Figure 2.12: Explaining the origin of the “screw dislocation”: (a) — arrangement of atoms in a screw dislocation; (b) — atom “a” moves towards atoms “b, c, d, e”, etc. constituting the screw dislocation along a spiral.

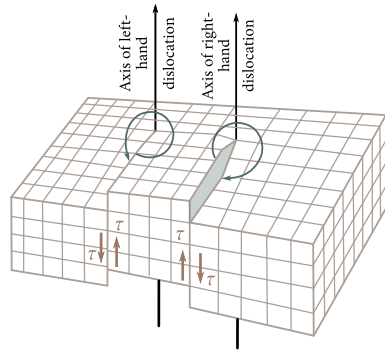


Figure 2.13: Arrangement of atoms in the plane perpendicular to dislocation line HE (see Figure 2.9). Dislocation occupies the region in which lattice atoms are displaced from their equilibrium sites (bounded by a circle): 0 — dislocation centre; (a) — positive dislocation; (b) negative dislocation.

cation, which is perpendicular to the Burgers vector **b**, the screw dislocation is parallel to it. The motion of the edge dislocation is in the direction of the Burgers vector **b**, and the motion of the screw dislocation is in the direction perpendicular to it.

Recently, experimental methods for direct observation of dislocations have been developed. Figure 2.14(a) shows a schematic diagram of an electron micrograph of a thin film of platinum phthalocyanine and Figure 2.14(b,c) an electron micrograph of a copper sulphide crystal obtained with the aid of a special procedure. Dark stripes on the micrographs are the traces of the atomic planes, which in platinum phthalocyanine are arranged at a distance of 12 Å from one another and in copper sulphide at a distance of 1.88 Å. The micrographs distinctly show the extra planes which terminate inside the crystal and form edge dislocations.

Figure 2.14(d) shows an optical micrograph of a decorated screw dislocation in a CaF_2 crystal. The method of decoration as used for transparent crystals consists in the precipitation along their dislocation cores of impurity atoms, which make the dislocation visible in an optical microscope. The striking agreement between those pictures and the theoretical concepts as set out in Figures 2.10 and 2.12 is worthy of admiration. Points of exit of dislocations on the crystal surface may be detected with the aid of etching. When a crystal is etched in a specially selected etch, the parts of the crystal where the lattice is most deformed dissolve more readily be-

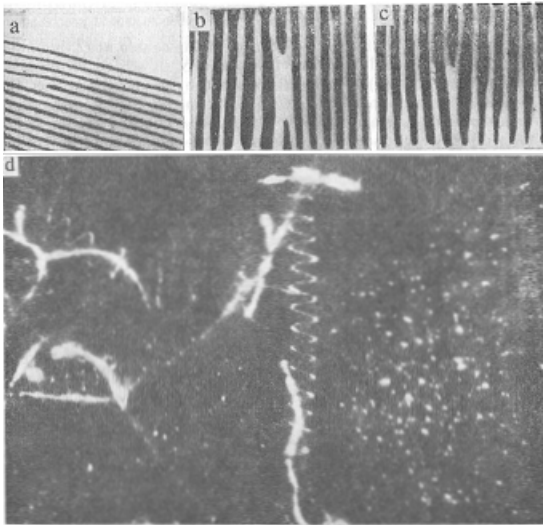


Figure 2.14: Observation of dislocations in an electron microscope: (a) — schematic diagram of an electron micrograph of a thin platinum phthalocyanine film (dark lines are atomic traces); (b), (c) — electron micrograph of a copper sulfide crystal (dark lines are traces of atomic planes); (d) — screw dislocation in a CaF_2 crystal obtained by decoration method.

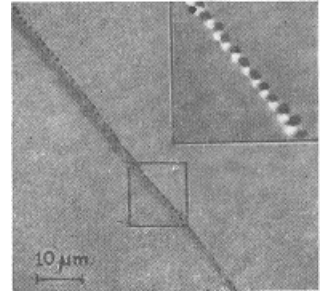


Figure 2.15: Etch pits on germanium surface. Dark points along the grain boundary are points of exit of dislocations.

cause the atoms in those parts possess an excess energy and are chemically more active. The places of exit of dislocations on the crystal surface are just such parts. Figure 2.15 shows a photograph of an etched germanium surface. The dark patches are the points of exit of dislocations.

§ 18. Forces needed to move dislocations

Suppose there is a positive dislocation with the centre at 0, bounded at points “a” and “k” and lying in the plane S in which slipping is possible [Figure 2.16(a)]. In equilibrium the force with which the lattice acts on the dislocation is zero. This may easily be seen from the roller model shown in Figure 2.17. The structure of the upper row of rollers which normally occupy recesses between the rollers of the lower row was deformed so that the section AB which previously contained 6 rollers now contains only 5. Such deformation gives rise to forces which tend to return the rollers 1, 2, 4, 5 to their stable equilibrium positions (the forces F_1, F_2, F_3, F_4, F_5). The forces applied to rollers 7, 5 and 2, 4 are equal in magnitude and opposite in

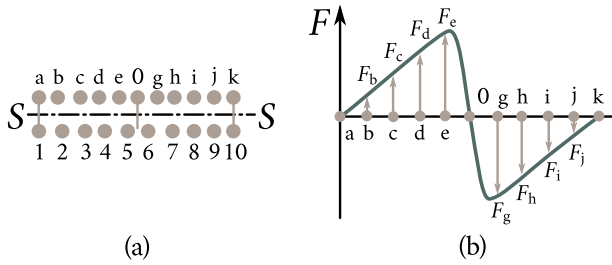


Figure 2.16: Calculating the force needed to move a dislocation: (a) — region of positive dislocation in crystal; 0 is dislocation centre, “a” and “k” are dislocation boundaries, S is the slip plane; (b) — forces needed to move an atom in the dislocation region (forces applied to atoms equidistant from the dislocation centre are equal in magnitude and opposite in direction).

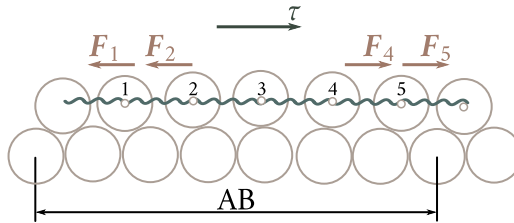


Figure 2.17: Roller model of an edge dislocation. Forces applied to “atoms” 1, 5 and 2, 4 are equal in magnitude and opposite in direction.

direction. Therefore, if the rollers of the upper row are interconnected by means of an elastic spring acting as a bond between them, the forces F_1 and F_5 , F_2 , F_4 will be mutually compensated and the system will be in a state of equilibrium.

The same situation occurs in the case of a dislocation schematically shown in Figure 2.16(b); the forces acting on atoms of the upper row occupying positions symmetrical with respect of the dislocation centre 0 are equal in magnitude but opposite in direction (the forces $F_b = F_j$, $F_c = F_i$, $F_d = F_h$, $F_e = F_g$). Therefore the resultant force is zero and the dislocation is in a state of equilibrium. If, however, the dislocation moves a little in the slip plane the symmetrical arrangement of the atoms in respect of the dislocation centre will be disturbed giving rise to a force which resists the motion of the dislocation. It is evident from Figure 2.17 that this force cannot be great since the movement of the rollers 1 and 2 to their new equilibrium position is to a large extent the result of the action of the forces exercised on them by the rollers 4 and 5, which also strive to occupy positions of stable equilibrium. Calculations show the tangential stress needed to move the dislocation to

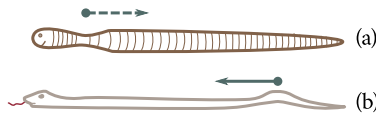


Figure 2.18: Dislocation mechanism of motion of (a) an earth-worm and (b) a snake.

be equal to

$$\tau_0 = \frac{2G}{(1-\nu)} \exp \left[-\frac{2\pi b}{d(1-\nu)} \right] \quad (2.17)$$

where G is the shear modulus, ν the Poisson ratio, b the interatomic distance, and d the distance between adjacent slip planes. The stress τ_0 is the theoretical value of the critical shearing stress. Setting $b = d$ and $\nu = 0.3$, we obtain $\tau_0 = 3 \times 10^{-4} G$. Within an order of magnitude this coincides with the experimental values of τ_{cr} . Thus, the theory of dislocations resolves the contradiction between the theoretical and the experimental values of shear strength of crystals.

The mechanism of motion by means of dislocations is quite frequent in nature. Snakes, worms, and shellfish move, because they generate dislocations. The motion of an earth-worm begins with the formation of an “extension” dislocation near the neck. The dislocation subsequently spreads along the body to the tail [Figure 2.18(a)]. In contrast, the motion of most snakes involves the formation of a “contraction” dislocation near the tail and its motion towards the head [Figure 2.18(b)].

§ 19. Sources of dislocations. Strengthening of crystals

The dislocations in a real crystal are formed in the process of its growth from the melt or from a solution. Figure 2.19(a) shows the boundaries of two blocks growing towards each other. The blocks make a small angle φ between themselves. As the blocks fuse together, some of the atomic planes do not spread through the entire crystal but terminate at block boundaries. Those are the places where dislocations are formed [Figure 2.19(b)]. The same situation occurs in the process of fusion of differently oriented grains in a polycrystalline sample. Since the block and grain boundaries in real solids are very extensive, the number of dislocations in them is enormous—as many as 10^{12} dislocations per square metre can be counted in well annealed metals. After cold working (rolling, drawing, etc.) dislocation densities rise to 10^{15} m^{-2} to 10^{16} m^{-2} . Those dislocations accumulate almost the entire energy absorbed by the metal in the process of plastic deformation.

Vacancy clusters may also serve as sources of dislocations in an undeformed crystal. Figure 2.20 shows an example of the formation of a positive and a negative dislocation from a cluster of vacancies.

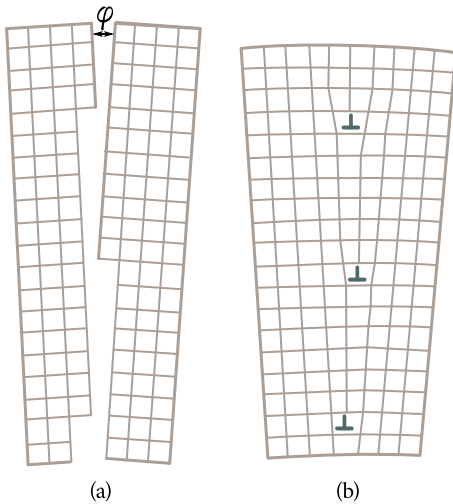


Figure 2.19: Formation of dislocations at block boundaries: (a) — two blocks growing towards each other at an angle φ ; (b) — dislocations appearing when the blocks fuse together.

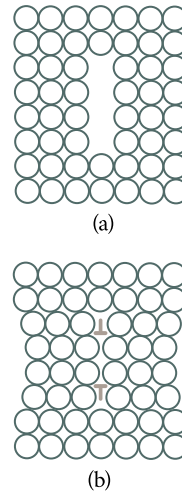


Figure 2.20: Dislocations formed from vacancy clusters: (a) — vacancy cluster in crystal; (b) — positive and negative dislocations formed from this cluster.

The shear process in a crystal in response to an applied external force is, in effect, the motion of dislocations in the slip planes and their emergence through the crystal surface. Should only the dislocations already present in the crystal be responsible for this process, plastic deformation would lead to their exhaustion and to the perfection of the crystal. This is in contradiction with experience, which demonstrates that as deformation grows the lattice does not become more perfect. In fact, just the opposite is true: the density of dislocations grows in the process. It is an established fact that dislocations responsible for plastic deformation are generated in the shear process itself by the action of the external force applied to the crystal. One such generation mechanism was discovered in 1950 by F. C. Frank and W. T. Read. For the purpose of better understanding this mechanism let us consider soap bubble formation with the aid of a tube (Figure 2.21). After the end of the tube has been immersed in a soap solution a flat film remains that closes the tube's orifice. As the air pressure in the tube is increased, the film swells and passes through the stages 1, 2, 3, 4, etc. Until it assumes the shape of a hemisphere (stage 3) its state is unstable: as pressure falls the film contracts striving to return to the original state 1. After the bubble has passed stage 3 the state of the bubble changes: it is now capable of evolution not only at a constant but also at a gradually decreasing pressure until it leaves the end of the tube. After the first bubble the second begins to be formed, followed by the third, etc.

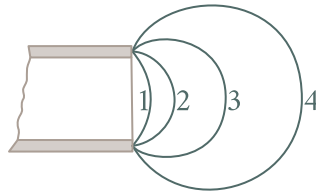


Figure 2.21: Process of formation of a soap bubble.

Now let us discuss the operation of the Frank-Read source. Figure 2.22(a) shows an edge dislocation DD' in a slip plane; points D and D' are fixed and do not take part in the motion of the dislocation. Dislocations may be anchored at the points of intersection with other dislocations, at impurity atoms, etc. Under the action of an external stress τ the dislocation starts bending in the same way as was the case with the soap film and at some time assumes the shape of a semicircle [Figure 2.22(b)]. Just like the soap film the dislocation can continue to bend only if τ grows steadily until it assumes the shape of a semicircle. Its subsequent evolution takes place by itself and results in the formation of two loops [Figure 2.22(c)], which after meeting at point C [Figure 2.22(d)] divide the dislocation in two: an external one, which closes forming an external circle [Figure 2.22(e)], and an internal one, which returns to the original position DD' . The external dislocation grows until it reaches the surface of the crystal and results in an elementary shift; the internal dislocation having returned to the initial position DD' begins again to bend under the action of the applied force and to grow in the manner described above. Such process may be repeated any number of times eventually leading to a noticeable shift of one part of the crystal in relation to another in a particular slip plane.

Low shear strength of crystals is due to the presence of innate dislocations and to the generation of others in the process of Shear. On the other hand, it is an established fact that the crystal is strengthened in the process of plastic deformation accompanied by the growth in the number of defects. The essence of such strengthening is the interaction of dislocations with each other and with other types of lattice defects causing their motion in the lattice to be obstructed.

Interaction of dislocations. Every dislocation, being the cause of elastic stresses of the lattice, creates a force field around itself which may be described by the values of the tangential τ and normal σ stresses at every point. When another dislocation enters this field, forces begin to act which strive to bring the dislocations together or to move them apart. Dislocations of like signs lying in one plane are repelled, while those of opposite signs are attracted. This is the reason why, as dislocations are accumulated in a definite plane, the crystal's resistance to shear is increased and the crystal is strengthened.

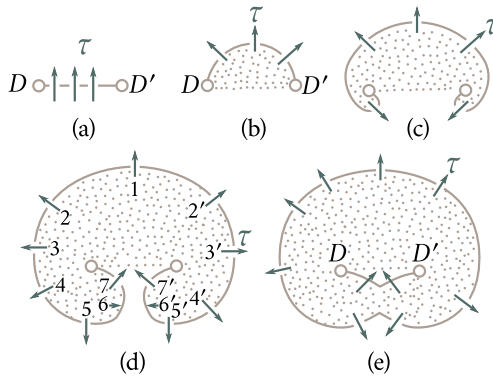


Figure 2.22: Operating sequence of a Frank-Read source: (a) — initial position of dislocation DD' , (b) — acted upon by external force the dislocation bends and assumes the shape of a semicircle; (c), (d) — further symmetric development of the dislocation loop; (e) — formation of external closed dislocation loop spreading across the crystal and of internal dislocation DD' returning to the original position.

Surmounting of obstacles. Suppose a dislocation when moving in a slip plane under the action of tangential stresses τ runs into a stationary obstacle D , for instance, an intersection with some other dislocation, an impurity atom, or some other type of defect. Figure 2.23 shows the method by means of which dislocation AB could, theoretically, surmount obstacle D : as the dislocation approaches D (positions 1, 2, 3) it gradually bends forming a loop that envelops the obstacle. Behind the obstacle the loop closes and the dislocation $A'B'$ again becomes a straight line. Figure 2.24 shows a photograph illustrating a case when a dislocation runs into a stationary obstacle (dark lines represent dislocations decorated by etching). The similarity in the pictures leaves not a trace of doubt as to the validity of the theoretical pattern shown in Figure 2.23.

In passing around the obstacle, the length of the dislocation and the deformation of the lattice are increased, which requires additional work to be performed. Therefore, the resistance to the motion of the dislocation in the interval where it has to surmount a defect is much greater than in other parts of the lattice. This is the essence of the fact that defects strengthen a crystal. The growth in the number of dislocations in the crystal with greater plastic deformation increases the number of obstacles at points of their intersection, which is the cause of strengthening brought about by plastic deformation. Impurity atoms have a similar effect: they create local lattice imperfections and thereby hinder the motion of the dislocations, with the result that the crystal's resistance to shear is increased. Block and grain boundaries and foreign inclusions in the lattice are especially effective

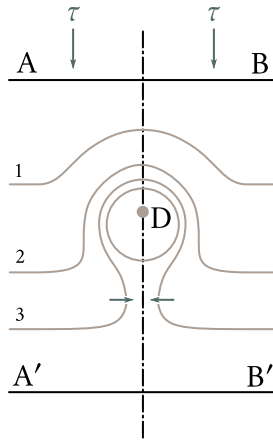


Figure 2.23: Schematic representation of an edge dislocation surmounting an obstacle: AB — shape of edge dislocation away from the obstacle D; 1, 2, 3 — gradual bending of the dislocation as it approaches D and closing of the newly formed loop behind the obstacle; A'B' — straightening of the dislocation far away from the obstacle.



Figure 2.24: Microphotographs of a chromium grain. Dark lines are etched dislocations ($\times 2000$).

in hindering the motion of the dislocations. They sharply increase the resistance to the motion of dislocations and greater stresses are required to overcome their effect. The phenomenon of strengthening in the process of cold working, in the process of introducing impurity atoms (doping), and in the process of formation of inclusions (tempering, aging, etc.) is widely used in practice to improve mechanical properties of engineering materials. This method enabled the strength of the materials to be increased from 6 to 8 times in the last 40 years.

§ 20. Brittle strength of solids

The destruction of solids may be of one of two principal types: of the *brittle* and of the *plastic*, or *viscous*, types.

Brittle destruction takes place if the tensile strength of the material is below the elastic limit. Such a material experiences only elastic deformation prior to destruction. No irreversible changes take place in such a material before it breaks down.

In the ductile materials the elastic limit is not only below the tensile strength but also below the yield stress. Because of that the destruction process is preceded by an appreciable plastic deformation, which prepares the subsequent destruction

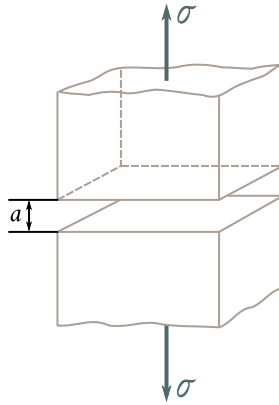


Figure 2.25: Calculating theoretical strength of solids after Polanyi (explanation in text).

process. In this case strength, being a typical kinetic parameter, is strongly dependent on the time the destructive stress is applied.

To begin with, let us discuss brittle strength of solids.

Theoretical strength of solids. There have been numerous attempts to calculate the strength of solids on the basis of molecular interaction in them. The strength σ_0 thus calculated is termed *theoretical strength*.

Here is a glance at some of the methods of estimating σ_0 .

Polanyi's method. The simplest method of estimating the strength of solids theoretically is due to M. Polanyi. Its essence is as follows.

Suppose a tensile stress σ is applied to a rod of a cross-sectional area of 1 m^2 (Figure 2.25). This stress increases the distance between the atomic planes. It is assumed that for destruction to take place a stress σ_0 able to increase the distance between the atomic planes by a value of the order of the lattice parameter a should be applied. The work needed to move an atomic plane a distance a away from the neighbouring plane is assumed to be equal to $\sigma_0 a$. It is further assumed that this work is transformed into the free energy of two new surfaces with a total area of 2 m^2 formed as a result of the breakup, the free energy being equal to 2α , where α is the surface energy ("surface tension") of the solid. Hence,

$$\sigma_0 a = 2\alpha$$

and the theoretical strength is

$$\sigma_0 = \frac{2\alpha}{a}. \quad (2.18)$$

For copper $\alpha \approx 1.7 \text{ J m}^{-2}$, $a = 3.6 \times 10^{-10} \text{ m}$, and $\sigma_0 \approx 10^{10} \text{ Pa}$, for silver $\alpha \approx 1.14 \text{ J m}^{-2}$, $a = 4 \times 10^{-10} \text{ m}$, and $\sigma_0 = 0.6 \times 10^{10} \text{ Pa}$.

Determination of σ_0 from the heat of sublimation. The energy equal to the heat of sublimation Q_s is required for the evaporation of a mole of a solid. For the evaporation of one molecular layer of the area of 1 m^2 the required energy W is a fraction of Q_s equal to the ratio of the mass of this layer m to the molar mass M :

$$W = \frac{Q_s}{m}.$$

But

$$m = N_s \mu, \quad M = N_A \mu$$

where μ is the molecular weight, $N_A = 6.023 \times 10^{23} \text{ mol}^{-1}$ the Avogadro number, and N_s the number of molecules per square metre of the solid's surface.

For an intermolecular distance of a the area per molecule is approximately equal to a^2 and the number of molecules per square metre $N_s \approx a^{-2}$. Therefore,

$$W = Q_s \frac{N_s}{N_A} \approx \frac{Q_s}{N_A a^2}.$$

Should the assumption be made that the evaporating molecules loose contact with the solid's surface when they are a distance of the order of the lattice parameter a away from it, we would obtain for the force needed to tear away an entire surface layer as a whole

$$\sigma_0 \approx \frac{W}{a} = \frac{Q_s}{N_A} \frac{1}{a^2} \quad (2.19)$$

σ_0 is assumed to be the theoretical strength of the solid.

For copper $Q_s = 3 \times 10^5 \text{ J mol}^{-1}$, $a = 3.6 \times 10^{-10} \text{ m}$, and $\sigma_0 \approx 10^{10} \text{ Pa}$. Similar calculations lead to the following results: for iron $\sigma_0 \approx 2.3 \times 10^{10} \text{ Pa}$, for aluminium $\sigma_0 \approx 0.6 \times 10^{10} \text{ Pa}$, and for silver $\sigma_0 \approx 0.6 \times 10^{10} \text{ Pa}$.

Calculating σ_0 from the forces of molecular interaction. Finally, let us discuss the method of calculating theoretical strength of solids from the forces of molecular interaction. Figure 2.26 shows the dependence of the potential energy $U(x)$ and the force of interaction between the particles $f(x)$ on the distance x between them. Since it is not easy to determine the exact law governing $f(x)$, the practice is to approximate this dependence by various functions. For instance, M. Polanyi and E. Orowan used the approximation in the form of a half of a sinusoid:

$$f(x) = f_{\max} \sin \left(\frac{2\pi x}{c} \right). \quad (2.20)$$

When a body of cross-sectional area of 1 m^2 is slowly torn in two, the required force is $\sigma = f N_s$, where N_s is the number of particles per square metre of the cross

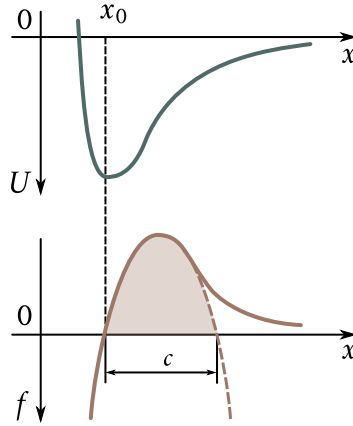


Figure 2.26: Calculating theoretical strength from forces of molecular interaction (explanation in text).

section. Substituting f from (2.20) we obtain

$$\sigma = \sigma_0 \sin\left(\frac{2\pi x}{c}\right) \quad (2.21)$$

where $\sigma_0 = f_{\max} N_s$ is the theoretical strength of the body.

For small displacements relation (2.21) may be rewritten in the form

$$\sigma = \frac{\sigma_0 2\pi x}{c}.$$

On the other hand, for small displacements Hooke's law is valid:

$$\sigma = \frac{Ex}{c}.$$

Equating the right-hand sides of these equations, we obtain

$$\sigma_0 \approx \frac{E}{2\pi} \approx 0.1E. \quad (2.22)$$

Calculations show that a more accurate estimate of the nature of the bonding in solids results only in a negligible correction to (2.22).

Comparing the values of theoretical strength σ_0 calculated with the aid of various methods we see that all of them yield nearly the same result whose order of magnitude is $0.1E$. Therefore it may be legitimately assumed that

$$\sigma_0 \approx 0.1E. \quad (2.23)$$

This is an enormous figure of the order of 10^9 Pa to 10^{10} Pa.

Real (technical) strength of solids. The strength of real crystals and solids used for technical purposes is termed *real*, or *technical*, strength σ_r . Table 2.5 shows the values of the elasticity modulus E , of theoretical strength $\sigma_0 \approx 0.1E$, of technical

strength σ_r , and of the ratio σ_0/σ_r for some industrial materials.

It follows from the data of Table 2.5 that the technical strength of solids is from 2 to 3 orders of magnitude less than their theoretical strength.

At present there is a general agreement that such discrepancy between σ_0 and σ_r is due to the presence of defects in real solids of various types, in particular of microscopic cracks which reduce the strength of solids. This is accounted for by the so-called *Griffith theory*. Let us calculate the technical strength using this theory.

We take a sample in the form of a thin plate and apply a tensile stress σ to it [Figure 2.27(a)]. The density of elastic energy in such an elastically extended sample is $\sigma^2/(2E)$ ¹.

Now let us imagine that a transverse microscopic crack of the length l running through the entire thickness δ of the sample has developed in it. The appearance of the crack is accompanied by the formation of a free surface $S \approx 2l\delta$ inside the sample and by an increase in the sample's energy by the amount $\Delta U_1 \approx 2l\delta\alpha$ (α is the free surface energy of the sample per unit area). On the other hand, the formation of a crack relieves the elastic stress from the volume $V \approx l^2\delta$ of the sample, whereby its elastic energy is reduced by the amount $\Delta U_2 \approx l^2\delta\sigma^2/(2E)$. The total change in the energy of the sample $W(l)$ brought about by the appearance

¹Indeed, the relative deformation in a sample under stress σ is $\epsilon = \sigma/E$, the absolute deformation $\Delta L = \epsilon L$ (L being the length of the sample). The work performed by the stress σ to extend the sample by ΔL is $(1/2)\sigma S \Delta L = \sigma^2 S L/(2E) = \sigma^2 V/(2E)$ (S is the cross-sectional area and V the volume of the sample). This work is transformed into the elastic energy of the sample of volume V . Therefore, the specific volume density of the elastic energy is $\sigma^2/(2E)$.

Table 2.5

Substance	Elasticity modulus	Theoretical strength	Technical strength	σ_0/σ_r
	E (10^7 Pa)	$\sigma_0 \approx 0.1E$ (10^7 Pa)	σ_r (10^7 Pa)	
Aluminium	6000	600	9.0	65
Copper	12000	1200	23	50
Glass	8000	800	8.0	100
Iron	21000	2100	30	70
Rock salt	4000	400	0.5	800
Silver	8000	800	18	45

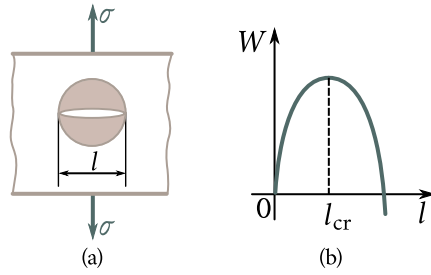


Figure 2.27: The Griffith theory of calculating the real strength of solids (explanation in text).

of a crack in it is

$$W(l) = 2l\delta\alpha - l^2\delta\frac{\sigma^2}{2E}. \quad (2.24)$$

Figure 2.27(b) shows the dependence of W on the length l of the crack. It has a maximum where its derivative vanishes: $dW/dl = 2\delta\alpha - l\delta\sigma^2/E$. Denote the length of the crack corresponding to the maximum energy by l_{cr} . We obtain from the last relation

$$l_{cr} = \frac{2\alpha E}{\sigma^2}. \quad (2.25)$$

It may be seen from Figure 2.27(b) that as long as the length l of the crack remains below the critical value l_{cr} , energy is needed for it to develop. On the other hand, starting with $l = l_{cr}$ further extension of the sample results in a reduction in its energy. Therefore, it takes place spontaneously with the brittle destruction of the sample as the final result.

Hence, the technical strength of solids having microscopic cracks should be calculated according to the Griffith theory from relation (2.25):

$$\sigma_r \approx \left(\frac{2\alpha E}{l}\right)^{1/2} \approx \beta \left(\frac{\alpha E}{l}\right)^{1/2}. \quad (2.26)$$

This result was subsequently verified by many investigators for various quite different methods of applying loads to the sample. A negligible difference was observed only in case of the numerical coefficient β .

Should we substitute the values of α, E, σ_r for copper ($\alpha \approx 1.7 \text{ J m}^{-2}$, $E = 1.2 \times 10^{11} \text{ Pa}$, $\sigma_r = 1.8 \times 10^8 \text{ Pa}$) into (2.26), we would obtain $l \approx 8 \times 10^{-6} \text{ m}$. Approximately the same values of l may be obtained for other solids.

It follows that for the strength of the solids to be reduced from the theoretical value to the value of the technical strength microscopic cracks of the order of several micrometers in length should develop in them up to the moment preceding their destruction. Many factors may be the cause of such cracks.

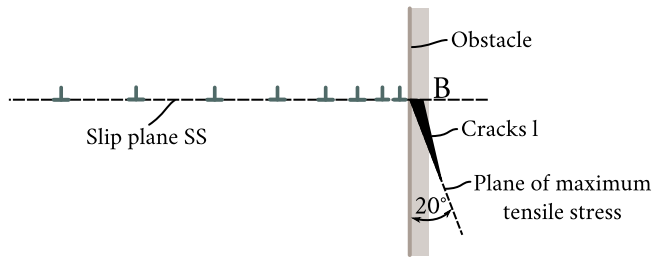


Figure 2.28: Formation of a crack near a dislocation pile-up.

The cracks may be produced in the course of the production of the solid, especially in the course of its mechanical processing. A proof of this is, in particular, a significant dependence of the strength of the sample on its dimensions, especially in the small dimensions range. Thus the strength of a glass filament of $2.5\ \mu\text{m}$ in diameter is almost 100 times that of a massive sample. The explanation is that as the dimensions of the sample are reduced so too is the probability of a large crack responsible for low strength appearing in it. Such dependence of the strength on the dimensions of the sample became known as the *scale factor*. The cracks may be the result of a large number of vacancies merging together.

Figure 2.28 shows a dislocation mechanism of crack production. Dislocations of a similar sign move in slip plane SS and meeting obstacle B begin to accumulate in its vicinity. Large stresses able to produce cracks l may develop at the head of this dislocation pile-up.

§ 21. Time dependence of the strength of solids

The theory of strength based on the condition (2.26) and discussed above describes actually the final stage of the destructive process when the body already contains cracks able to cause brittle rupture.

However, the initial stages of the destructive process during which the cracks originate and grow to attain critical dimensions l_{cr} are also important. This process is a gradual one and takes time τ to be completed. The time τ that it takes for the destructive process to develop from the moment the load is applied to the body to the moment of rupture is termed the *durability* of the material.

The firsts experiments aimed at investigating durability were carried out by S. N. Zhurkov and G. M. Bartenev with coworkers. They also developed modern notions about the physical nature of durability.

It was established by experiments that the durability τ , the tensile stress σ , and

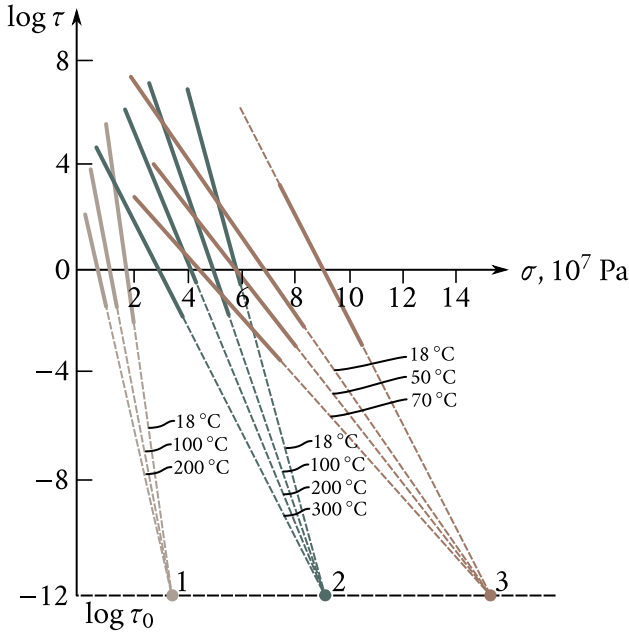


Figure 2.29: Durability versus stress for aluminium (1), Plexiglas (2), and silver chloride (3).

the absolute temperature T are related by the expression:

$$\tau = \tau_0 e^{(U_0 - \gamma\sigma)/(k_B T)} \quad (2.27)$$

where τ_0 , U_0 , and γ are constants dependent on the nature and structure of the body.

For $T = \text{constant}$, formula (2.27) may be rewritten in the form

$$\tau = A e^{-\beta\sigma} \quad (2.28)$$

where $A = \tau_0 e^{U_0/(k_B T)}$, and $\beta = \gamma/(k_B T)$.

Formulae (2.27) and (2.28) were tested on a great number of different materials (metals, polymers, haloid compounds, etc.) in an 8 to 10 order of magnitude range of the values of τ and in a wide range of the values of T .

Figure 2.29 shows the dependence of the durabilities τ of aluminium (1), Plexiglas (2), and silver chloride (3) on the applied stress σ at various temperatures expressed in the $\log \tau$ versus σ coordinate system. It may be seen from Figure 2.29 that the dependence $\tau(\sigma)$ in semilogarithmic coordinates is well represented by a straight line. A family of such straight lines obtained for a given material at different temperatures resembles a fan with the apex at some point called pole. It follows from Eq. (2.27) that τ will be independent of T and that the straight lines $\log \tau(\sigma)$ at different temperatures will intersect at one point (at the pole) only if $U_0 - \gamma\sigma = 0$;

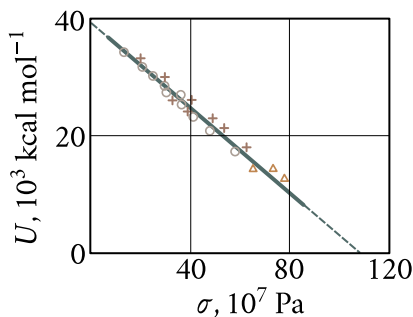


Figure 2.30: Activation energy of rupture of viscose fibre at different temperatures (triangles, $-76\text{ }^{\circ}\text{C}$; circles, $+20\text{ }^{\circ}\text{C}$; crosses, $+80\text{ }^{\circ}\text{C}$) versus stress.

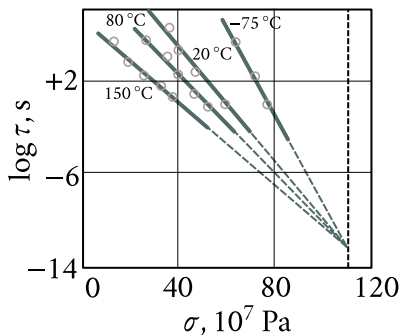


Figure 2.31: Durability of viscose fibre at different temperatures versus stress.

but in that case $\log \tau = \log \tau_0$. Hence, the pole should be at a distance $\log \tau_0$ below the σ -axis.

It is evident from Figure 2.29 that the poles for all the materials tested lie practically on the same straight line parallel to the σ -axis. This means that τ_0 is approximately the same for all the materials. Experiments show it to be of the order of 10^{-12} s to 10^{-13} s, that is, close to the period of oscillations of atoms in solids.

Let us take the logarithm of (2.27):

$$\log \tau = \log \tau_0 + \frac{(U_0 - \gamma\sigma)}{k_B T} = \log \tau_0 + \frac{U}{k_B T}, \quad U = U_0 - \gamma\sigma. \quad (2.29)$$

Measuring the dependence of $\log \tau$ on $1/T$ for constant values of σ , we can determine U for various values of the stress σ experimentally; the dimensions of U are that of energy and because of this it is termed *activation energy of the destructive process*. Figure 2.30 shows the dependence of the activation energy of rupture of viscose fibre on stress for various temperatures. It may be seen that U is independent of T and is determined solely by σ ; for $\sigma = 0$ the maximum value of U is $U_0 \approx 40 \text{ k/m}\cdot\text{o}\ell$; for a stress $\sigma \approx 107 \times 10^7 \text{ Pa}$ we see that $U = 0$. Figure 2.31 shows that for $\sigma \approx 107 \times 10^7 \text{ Pa}$ a practically instantaneous rupture of viscous fibre takes place (during the time τ_0), no matter what its temperature is.

Meticulous experiments carried out by S. N. Zhurkov with coworkers and by other investigators on a variety of materials demonstrated that for metals U_0 is quite close to the sublimation energy Q_s and for polymers to the thermal destruction energy Q_d . Table 2.6 shows the values of U_0 , Q_s , and Q_d for some materials. It may be seen that U_0 coincides either with Q_s or with Q_d with a high degree of accuracy.

Universal validity of the dependence thus obtained merits the conclusion that the process of destruction of a solid is one of a kinetic nature (that is, develops in time) and its origin is the same for all solids. Modern notions of the physical mechanism of this process are set out below.

The atoms in a solid take part in thermal oscillations with a period of $\tau_0 \approx 10^{-12}$ s to 10^{-13} s. Thermal fluctuations from time to time result in the rupture of chemical bonds. The probability of this process depends on the height of the potential barrier of destruction U and on the temperature T . This probability increases with the rise in T and the decrease in U . In the absence of external stress σ the energy required to break a bond is equal to the energy of the bond itself. This is the reason why the height of the potential barrier U_0 obtained from experiments in mechanical destruction of solids turned out to be equal to the sublimation heat of metals and to the thermal destruction energy of polymers.

The stresses induced in a body reduce the height of the potential barrier from U_0 to $U_0 - \gamma\sigma$ and thus increase the probability of rupture of the bonds and, consequently, the number of ruptured bonds per unit volume.

The formation of submicroscopic volumes in which the bonds have been broken and their merging results eventually in the nucleation and development of cracks. When the length of the cracks attains a critical value, the body breaks up under the applied stress. The higher is the stress the lower the activation barrier $U_0 - \gamma\sigma$ and the greater the rate of bond rupture; therefore it takes less time for the destructive process to develop, that is the less should be the durability of the body. This is exactly what is observed in practice.

From the above point of view the destruction of solids should take place at any stresses provided the time they act is long enough. But in that case it is not easy

Table 2.6

Substance	Activation energy of destruction, U_0 (10^5 J mol $^{-1}$)	Sublimation energy Q_s (10^5 J mol $^{-1}$)	Thermal destruction energy, Q_d (10^5 J mol $^{-1}$)
Aluminium	2.16	2.2	
Nickel	3.48	3.4	
Nylon	1.8		1.72
Platinum	4.8	5.1	
Polymethyl methacrylate	2.16		2.1-2.2
Polyvinyl chloride	1.4		1.28
Silver	2.56	2.72	
Teflon	3.0		3.0-3.1
Zinc	1.0	1.08	

to understand why bridges and other installations built many centuries ago and carrying loads all that time still remain intact.

To explain this fact we again turn to Figure 2.29. We see that the lower the temperature the weaker the load dependence of durability is. This dependence is practically nill at sufficiently low temperatures. For glasses and metals with a high melting point already room temperatures are low enough. Because of that their strength is actually a unique characteristic of the material. In all other conditions it is not justified to speak of strength without mentioning the time during which the material is to work under load. Thus, industrial products made of Plexiglas during a year's service can endure loads not exceeding 30% of their short-time strength; steam turbine blades working at high temperatures are calculated for strength always with account taken of their durability.

§ 22. Methods of increasing the strength of solids

The nucleation mechanism of breaks in continuity and the mechanism of crack growth are both greatly influenced by the atomic structure of solids. Therefore, strength is a structure-sensitive characteristic of such bodies.

Stresses in crystals occasion the production of dislocations and their motion in slip planes. In this way, plastic shifts resulting in plastic deformations are realized. Meeting impurities, grain and block boundaries, interceptions of slip planes, etc., the dislocations lose their mobility and the crystal is hardened. As was mentioned above, stresses may develop at the head of a dislocation pile-up capable of causing cracks.

To increase the strength of such bodies it is necessary to retard the production of dislocations and the nucleation and growth of cracks.

This can be done by two methods.

(i) By producing imperfection-free crystals free from internal stress sources, which in the long run cause the nucleation of cracks.

This method has up to the present been realized only in the filament type crystals known by the name of "whiskers". They are single crystals grown under special conditions using the method of decomposition or reduction of appropriate chemical compounds, the method of condensation of vapours of pure metals at an appropriate temperature in hydrogen or in an inert gas, and the method of electroplating metals from a solution onto electrodes of extremely small dimensions. The filament-type crystals are usually 2 mm to 10 mm long and 5 μm to 50 μm thick.

A striking property of such crystals is that their mechanical parameters are extremely high. Their strength turned out to be close to the theoretical strength

of solids. Thus, the strength of iron whiskers is about 1.34×10^{10} Pa, of copper whiskers about 3×10^9 Pa, and of zinc whiskers 2.3×10^9 Pa, while the strength of normal samples made from those metals is 3×10^8 Pa, 2×10^8 Pa, and 1.8×10^8 Pa, respectively.

Filament-type crystals of iron experience only elastic deformation reaching an enormous figure of the order of 5%-6%, after which brittle destruction occurs. Note that in normal iron noticeable plastic flow starts already at a deformation of $\varepsilon \approx 0.01\%$.

The unusually high mechanical parameters of the filament crystals are due to their ideal internal structure. Such crystals contain practically no dislocations, are exceptionally pure and their surface is so perfect that even a magnification of 40000 times fails to reveal any traces of roughness. Such perfection is mainly due to the condition of growth of small-size crystals, in which the freezing-in of lattice imperfections is less probable because it is easier for them to leave the crystal through a nearby surface.

Because of the absence of dislocations and of other defects in filament crystals a shift in a slip plane can only take the form of a rigid shift, in which the bonds of all the atoms in the slip plane are simultaneously broken. Stresses close to the theoretical stress limit of the crystals are needed to effect such a shift and this is what is observed in practice.

An unnaturally great elastic deformation of the whiskers is due to the absence of mobile dislocations, which in normal crystals are responsible for the plastic deformations occurring already at very low stresses.

Hence, the first method, the method of producing imperfection-free (in particular, dislocation-free) crystals, holds out a promise of producing materials of extreme strength close to the theoretical strength of solids.

(2) The second method is a direct opposite of the first. It consists in the maximum deformation of the internal structure of a crystal through the introduction of impurities, precipitation of dispersed phases, great plastic deformation, etc. Such defects hinder the motion of the dislocations and the growth of cracks and thus increase the strength of the material, as was already discussed in detail above. Science and industry have up to now made use only of this method and succeeded in attaining with it a strength of the order of 4×10^9 Pa. The effect this had on technology may be inferred from the following example. The specific weight of a modern aircraft engine is about 1 kgf per hp; at the turn of the century it was about 250 kgf per hp.

The recent times have witnessed the appearance of composite materials consisting of a matrix filled with filament crystals. Stainless steel, nickel, titanium and

other materials are used for the matrix. The matrices are filled with tungsten, aluminium oxide, etc. filaments. The results obtained so far hold out a promise of obtaining by this method in the near future materials of 5 to 10 times the strength (especially at elevated temperatures) of the best steels and of 1.5-2 times lighter weight.

The strength of amorphous bodies and glass polymers is no less sensitive to internal structure. The strength of glass and quartz filaments newly extruded at a high temperature and practically free from defects² is 100 times as high as that of normal specimens and quite close to the theoretical value.

The room temperature strength of unoriented glass polymers is of the order of 10^8 Pa. Films and fibres made of them having an oriented structure have a strength of the order of 10^9 Pa comparable to that of high quality steels. With a perfect orientation of the polymer molecule chain the strength of the needlelike crystals of the polymer may be as high as 3×10^{10} Pa. If one takes into account that the density of the polymers is close to unity, one can imagine how great their value for technology may be.

There is a rapidly growing demand on the quality of the materials for modern science and technology. Already now there is a need for materials able to withstand temperatures of several thousand degrees with the necessary strength characteristics at such temperatures and without any noticeable plastic deformation at normal loads.

What are the prospects for such materials?

One of the feasible methods for producing such extra strong and extra heat-resistant materials was proposed by the Soviet physicist A. V. Stepanov who pointed to a particular property of such molecular crystals as sulfur. The crystal of sulfur is constructed of molecules bonded by relatively weak molecular forces. Because of that the strength of the crystal and its melting point are low (115°C). The atoms in the sulfur molecule itself, on the other hand, are held together by powerful chemical bonds. If one would be able to construct a sulfur lattice with the atoms retaining the same bonds that act in the molecule, the result would be an extremely strong crystal with the melting point of about 34700°C . Similar modifications could be introduced into other molecular crystals as well. Are there any real grounds for such projects? The fact that we were able to transform soft graphite and hexagonal boron nitride into extra strong, hard, and high melting point diamond and borazon crystals by substituting powerful covalent bonds for weak van der Waals forces

²Since the atomic structure of amorphous bodies is irregular, the term defect may apply only to inclusions (clusters of foreign atoms, cracks, inhomogeneities) large if compared with atomic dimensions.

lends ground to such hopes. The prospects that will be opened by such materials are so enormous that any work, no matter how great, put into their production shall be generously rewarded.

Chapter 3

Elements of Physical Statistics

Every solid is a system, or an ensemble, consisting of an enormous number of microscopic particles. Such systems obey specific *statistical laws*, which are the subject of statistical physics, or physical statistics.

The present chapter deals briefly with the principal elements of physical statistics needed to describe the properties of solids.

§ 23. Methods used to describe the state of a macroscopic system

There are two methods of describing the state of a system consisting of a great number of microscopic particles, the *thermodynamic* and the *statistical* method. Let us discuss them.

Thermodynamic description of a system. In the thermodynamic approach to the description of the state of a system consisting of an enormous number of particles the latter is regarded as a macroscopic system, it being of no interest of what type of particles it consists. Such a system is termed a *thermodynamic system*.

A thermodynamic system may be either *closed* or *open*. A closed system does not interact in any way with the surroundings, and an open system can exchange heat and/or work with the surroundings.

The state of a system in which it can remain infinitely long is termed the *equilibrium state*. It is uniquely determined by a set of independent physical parameters, the *state parameters*. The principal state parameters are the *volume* of the system V , the *pressure* p , and the *temperature* T . However, often those parameters are inadequate for a complete characteristic of the system. For a system made up of several substances one has also to know their concentrations; for a system in an electric or a magnetic field the intensities of these fields should be specified; etc.

Any change in a thermodynamic system involving the variation of at least one

state parameter is termed a *thermodynamic process*.

The sum of all types of energy of a closed system is termed the *internal energy* (E) of the system. It is made up of the kinetic energy of the particles constituting the system, of the potential energy of the interaction between the particles, and of the internal energy of the particles themselves (which shall not be considered here since it is not subject to change in usual processes).

The internal energy is a *function of state* of the system. This means that there is one and only one definite value of internal energy that corresponds to each state no matter how the system arrived at this state.

Interacting with the surroundings a thermodynamic system may receive or reject some amounts ΔQ of heat, may perform work ΔA or have work performed on it. In all cases the variation in internal energy of the system, dE , should be equal to the difference in the amount of heat received from outside, ΔQ , and the work ΔA performed by the system against external forces:

$$dE = \Delta Q - \Delta A. \quad (3.1)$$

This is the first law of thermodynamics.

It should be pointed out that in contrast to the internal energy the work ΔA and the amount of heat ΔQ depend not only on the initial and the final states of the system but on the way the state is changed as well. Since

$$\Delta A = p dV \quad (3.2)$$

where dV is the variation of the volume of the system the pressure in which is p , we may write (3.1) in the form

$$dE = \Delta Q - p dV. \quad (3.3)$$

The second law of thermodynamics maintains that the amount of heat ΔQ received by the system in a reversible process results in the increase of the entropy of the system by

$$dS = \frac{\Delta Q}{T} \quad (3.4)$$

where T is the temperature at which the heat is received. Substituting ΔQ from (3.4) into (3.3), we obtain

$$dE = T dS - p dV. \quad (3.5)$$

It follows from (3.5) that the system's internal energy can be changed at the expense of work performed or heat exchanged.

However, the system's energy may also change with the change in the number of particles it contains, for every particle leaving the system takes away a definite amount of energy with it. Therefore, the general expression for the law of conser-

variation of energy (3.5) should be written in the form

$$dE = T dS - p dV + \mu dN \quad (3.6)$$

where dN is the variation of the number of particles in the system. Parameter μ is termed the *chemical potential* of the system. Its physical meaning is as follows. For an isolated system of constant volume which neither receives nor gives away heat, $dS = \Delta Q/T = 0$ and $dV = 0$. For such a system

$$dE = \mu dN. \quad (3.7)$$

Whence

$$\mu = \frac{dE}{dN}. \quad (3.8)$$

Hence, the chemical potential expresses the variation of the energy of an isolated system of a constant volume brought about by a unit variation in the number of particles it contains.

Let us consider the conditions of equilibrium of a system whose total number of particles remains constant but the particles can go over from one body belonging to the system to another. Two electron conductors, for instance, two metals, in contact with each other at a constant temperature may serve as an example of such a system. Denote the chemical potential of the electron gas in the first metal by μ_1 and in the second by μ_2 . Suppose dN electrons flow from one metal to another. According to (3.7) this will reduce the energy of the first metal by $dE_1 = \mu_1 dN$ and increase the energy of the second by $dE_2 = \mu_2 dN$. For the metals to be in a state of equilibrium the necessary condition is

$$dE_1 = dE_2, \quad \text{or} \quad \mu_1 dN = \mu_2 dN.$$

Hence the condition of equilibrium is

$$\mu_1 = \mu_2. \quad (3.9)$$

This condition is valid not only in the case of two electron conductors in contact with each other but for any phases in contact with each other: the solid and the liquid, the liquid and the gaseous, etc. *In all cases the condition of equilibrium is the equality of the chemical potentials.*

Statistical method of describing a system. To describe the state of every particle one should specify its three coordinates and three components of the momentum. Apparently, if one was to write the equations of motions of the particles and solve them, he would be able to obtain complete information on the behaviour of the system and to predict its state at any moment of time. Such calculations, however, are not only extremely tedious but, in fact, useless. The complexity of the problem stems from the fact that to describe the behaviour of the gas molecules normally contained in 1 m^3 one would have to solve about 10^{26} interconnected

equations of motion and also take into account the initial conditions, which is practically impossible. Should such calculations be carried out, they would be of no value since the properties of a system in the state of equilibrium not only are independent of the initial values of the coordinates and of the momentum components but generally remain constant in time, although the coordinates and the momenta of the particles do change. It follows from here that there is a qualitative distinction between the system and the individual particles and that the behaviour of the former is governed by laws different from those that govern the behaviour of individual particles. These laws are the statistical laws. The following examples are proof of their existence.

The velocity of an individual gas molecule is a random quantity, which is impossible to predict. Despite this fact, in a gas with a very large number of particles, on the average a distinct velocity distribution of its molecules may be observed. In other words, on the average a quite definite fraction of the molecules has a speed of, say, from 100 m s^{-1} to 200 m s^{-1} , from 400 m s^{-1} to 500 m s^{-1} , etc.

It is a matter of chance whether or not a given molecule shall enter a specified volume of the gas. Despite this fact there is a definite regularity in the distribution of the molecules over the volume: equal elements of volume contain, on the average, equal numbers of molecules.

The situation here is similar to that when a coin is tossed. The landing of the coin heads or tails up is a random event. Nevertheless, when the number of times the coin is tossed is very great, a quite definite regularity may be observed: on the average, the coin lands heads up half the number of times.

Such regularities are termed *statistical*. The principal feature of statistical laws is that they deal with *probabilities*. They enable predictions to be made only as to the probability of some event occurring or some result being realized. In the example with the coin the predicted probability of the coin landing one or the other side up is $1/2$. The results of individual tests may, and undoubtedly shall, deviate from those values the more the less the number of tests is. If we toss a coin five times, the head may fall out any number of times from 0 to 5. But the greater the number of tosses, that is, the more numerous the ensemble, the more accurate the statistical predictions are. Calculations show the relative deviation of an observed physical quantity (for instance, of the number of particles per unit volume) from the average value \bar{M} in a system of N noninteracting particles to be

$$\frac{\sqrt{\Delta M^2}}{\bar{M}} \propto \frac{1}{\sqrt{N}}$$

or inversely proportional to \sqrt{N} .

As N is increased, the ratio $\Delta M/\bar{M} \rightarrow 0$. For very great N we have that

$M/\overline{M} \approx 1$. Thus 1 m^3 of air normally contains on the average 2.7×10^{25} molecules. The relative deviation from this number is on the average equal to

$$\frac{100\%}{\sqrt{N}} \approx 2 \times 10^{-11}\%.$$

This deviation is so negligible that there are no instruments capable of detecting it. Therefore when dealing with large volumes it is always reasonable to assume that the distribution of molecules over the volume is uniform.

It should, however, be pointed out that deviations from the average values are not a possibility but a necessity. Such deviations are termed *fluctuations*.

§ 24. Degenerate and nondegenerate ensembles

Microscopic particles and the ensemble. All microscopic particles making up an ensemble may be subdivided into two classes according to their behaviour: fermions and bosons.

Fermions include electrons, protons, neutrons and other particles with a half-odd integral values of spin: $\hbar/2, 3\hbar/2, \dots$. Bosons include photons, phonons and other particles with integral values of spin: $0, \hbar, 2\hbar, \dots$

The fermions in an ensemble exhibit marked “individualistic” tendencies. If some quantum state is already occupied by a fermion, no other fermion shall settle in it. This is the essence of the well-known *Pauli exclusion principle*, which governs the behaviour of fermions. Bosons, on the other hand, strive for “unification”. They can settle in the same state in any numbers and do it the more readily the more populated the state already is.

Degenerate and nondegenerate ensembles. Let us discuss the possible effects of the nature of the particles (their fermionic or bosonic character) on the properties of the ensemble as a whole.

For the nature to be felt the particles must “meet” often enough. This means that they must occupy the same state or at least sufficiently closely-spaced states.

Suppose that there are G different states which any one of N similar particles can occupy. The ratio N/G may serve as a measure of the “meeting” frequency. The meetings will be rare if

$$\frac{N}{G} \gg 1. \quad (3.10)$$

In this case the number of different vacant states is much larger than the number of particles: $G \gg N$. Evidently, in such circumstances the specific nature of fermions and bosons shall not be felt, since every particle has at its disposal a large number of different free states and the problem of several particles occupying the same

state actually does not arise. Therefore, the properties of the ensemble as a whole shall not depend on the nature of the particles that make it up. Such ensembles are termed *nondegenerate*, and condition (3.10) is the *condition of nondegeneracy*.

If, however, the number of states G is of the same order of magnitude as the number of particles N , that is if

$$\frac{N}{G} \approx 1 \tag{3.11}$$

then the problem of how the states should be occupied, whether individually or collectively, indeed assumes much importance. In this case the nature of the microscopic particle is fully revealed in its effect on the properties of the ensemble as a whole. Such ensembles are termed *degenerate*.

Degenerate ensembles are a unique property of quantum objects, since the parameters of state of such objects only change discretely with the result that the number of possible states G can be finite. The number of states for classical objects whose parameters change continuously is always infinite and they can form only nondegenerate ensembles.

It should be pointed out that quantum mechanical objects too may form nondegenerate ensembles provided condition (3.10) is fulfilled (see Table 3.1).

Classical and quantum statistics. Physical statistics that studies nondegenerate ensembles is termed *classical statistics*. It owes much to J. C. Maxwell and L. E. Boltzmann (the *Maxwell-Boltzmann statistics*).

Physical statistics that studies degenerate ensembles is termed *quantum statistics*. Owing to the effect of the particles' nature on the properties of a degenerate ensemble, degenerate ensembles of fermions and bosons behave in essentially different ways. On this ground a distinction is made between two quantum statistics.

Quantum statistics of fermions owes much to E. Fermi and P. A. M. Dirac (this, by the way, explains the origin of the term "fermion"). It is termed the *Fermi-Dirac statistics*.

Quantum statistics of bosons owes much to S. N. Bose and A. Einstein (hence

Table 3.1

Object	Ensembles	
	Degenerate	Nondegenerate
Classical	No	Yes
Quantum	Yes	Yes

the term “boson”). It is termed the *Bose-Einstein statistics*.

It follows then that quantum statistics deals only with quantum objects while classical statistics may deal both with the classical and the quantum objects. If we reduce the number of particles in an ensemble or increase the number of states, we shall eventually turn a degenerate ensemble into a nondegenerate one. In that case the ensemble shall be described by the Maxwell-Boltzmann statistics no matter whether it contains fermions or bosons.

Distribution function. What is the connection between the distribution of the particles over particular states and the state of the ensemble as a whole? To specify the state of an ensemble, for instance, of a gas of particles, one should specify its state parameters. To specify the state of each particle one should specify its coordinates and momentum components or its energy, which is a function of coordinates and momentum.

The two types of quantities are connected by the *statistical distribution function*

$$N_{\mu,T}(E) dE \quad (3.12)$$

which specifies the number of particles having an energy from E to $E + dE$ in the system described by the state parameters μ and T .

This function is termed *complete statistical distribution function*. To simplify notation the indices denoting the state parameters are usually omitted.

The complete distribution function may be represented by the product of the number of states $g(E) dE$ per energy interval dE and the probability of occupation of those states by the particles. Let us denote the latter by $f(E)$. Then

$$N(E) dE = f(E)g(E) dE. \quad (3.13)$$

The function $f(E)$ is termed simply the *distribution function*. As was stated before it signifies the probability of the occupation of the respective states by the particles. If, for instance, 10 of 100 closely spaced states are occupied by particles (the total number of particles in the system being much greater than 100), the probability of occupation of such states will be equal to 0.1. Since on the average there is 0.1 of a particle per each state, we may take $f(E)$ to be the average number of particles in a given state.

Hence, the problem of finding the complete distribution function of particles over the states is reduced to that of finding the function $g(E) dE$, which describes the energy distribution of the states, and the function $f(E)$ which determines the probability of their occupation.

We start by determining the function $g(E) dE$.

§ 25. The number of states for microscopic particles

Concept of phase space of a microscopic particle and quantization. In classical mechanics the state of a particle is determined if its three coordinates (x, y, z) and three components of its momentum (p_x, p_y, p_z) are specified. Let us imagine a six-dimensional space with the coordinate axes x, y, z, p_x, p_y, p_z . The state of the particle at every moment of time will be described by a point (x, y, z, p_x, p_y, p_z). Such space is termed *phase space* and the points (x, y, z, p_x, p_y, p_z) are termed *phase points*. The quantity

$$\Delta\Gamma = \Delta\Gamma_V \Delta\Gamma_p = dx dy dz dp_x dp_y dp_z \quad (3.14)$$

is termed an element of the phase space. Here, $\Delta\Gamma_V = dx dy dz$ is an element of volume in coordinate space and $\Delta\Gamma_p = dp_x dp_y dp_z$ an element of volume in momentum space.

Since the coordinates and the momentum components of a classical particle may change continuously, the elements $\Delta\Gamma_V$, $\Delta\Gamma_p$ and with them the element $\Delta\Gamma$ as well can be chosen as small as desired.

The potential energy of a system of noninteracting particles not acted upon by an external field is zero. Such particles are termed *free*. For such particles it is convenient to use a three-dimensional momentum space instead of the six-dimensional phase space. In this case, the element $\Delta\Gamma_V$ is simply equal to the volume V in which the particles move, because no additional restrictions are placed on them.

The division of the phase space into elements of volume is not quite so simple if the particle in question is an electron or some other microscopic object possessing wave properties. The wave properties of such particles make it impossible, in accordance with the uncertainty principle, to distinguish between two states, (x, y, z, p_x, p_y, p_z) and ($x + dx, y + dy, z + dz, p_x + dp_x, p_y + dp_y, p_z + dz$), if the product $dx dy dz dp_x dp_y dp_z$ is less than h^3 . Since this product represents an element of volume in six-dimensional phase space, it follows from the above that different quantum states shall correspond to different elements of volume in this space only if the size of those elements is no less than h^3 . Therefore quantum statistics makes use of an elementary cell in the six-dimensional space with the volume

$$\Delta\Gamma = \Delta\Gamma_V \Delta\Gamma_p = h^3. \quad (3.15)$$

The element of the three-dimensional momentum space for free particles for which $\Delta\Gamma_V = V$ is

$$\Delta\Gamma_p = \frac{h^3}{V}. \quad (3.16)$$

Each element of this kind has corresponding to it a definite quantum state.

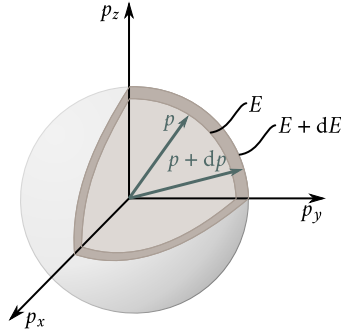


Figure 3.1: Calculating the number of states of a microscopic particle.

The process of dividing the phase space into cells of finite size (h^3 or h^3/V) is termed *quantization of phase space*.

Density of states. We wish to calculate the number of states of a free particle in the energy interval from E to $E + dE$. To this end draw two spheres of the radii p and $p + dp$ in the momentum space (Figure 3.1). There is a spherical layer with the volume of $4\pi p^2 dp$ contained between the spheres. The number of phase cells contained in this layer is

$$\frac{4\pi p^2 dp}{\Delta\Gamma_p} = \frac{4\pi V}{h^3} p^2 dp. \quad (3.17)$$

Since there is one particle state to correspond to every cell the number of states in the interval dp between p and $p + dp$ is

$$g(p) dp = \frac{4\pi V}{h^3} p^2 dp. \quad (3.18)$$

For free (noninteracting) particles

$$E = \frac{p^2}{2m}, \quad dE = \frac{p}{m} dp.$$

Using these relations to express p and dp and substituting the results into (3.18), we obtain

$$g(E) dE = \frac{2\pi V}{h^3} (2m)^{3/2} E^{1/2} dE. \quad (3.19)$$

This is the number of states of a free particle in the energy interval $(E, E + dE)$. Dividing the right- and the left-hand sides of (3.19) by dE , we obtain the density of states, $g(E)$, which specifies the number of states of a microscopic particle per unit energy interval:

$$g(E) = \frac{2\pi V}{h^3} (2m)^{3/2} E^{1/2}. \quad (3.20)$$

It follows from (3.20) that as E increases the density of states rises in proportion

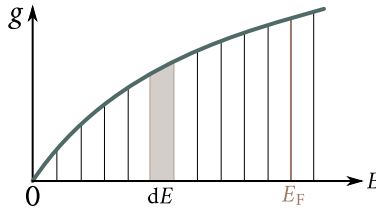


Figure 3.2: Energy dependence of density of states.

to $E^{1/2}$ (Figure 3.2). The density of states depends, besides, on the particle's mass and increases with m .

In case of the electrons each phase cell corresponds, to be exact, not to one but to two states, each distinguished by its spin. They are termed *spin states*. Therefore, in case of the electrons the number of states (3.18) and (3.19) and the density (3.20) should be doubled:

$$g(p) dp = \frac{8\pi V}{h^3} p^2 dp, \quad (3.21)$$

$$g(E) dE = \frac{4\pi V}{h^3} (2m)^{3/2} E^{1/2} dE, \quad (3.22)$$

$$g(E) = \frac{4\pi V}{h^3} (2m)^{3/2} E^{1/2}. \quad (3.23)$$

Condition of nondegeneracy for an ideal gas. Integrating (3.20) with respect to energy from 0 to E , we obtain the number of particle states contained within the energy interval $(0, E)$:

$$G = \frac{2\pi V}{h^3} (2m)^{3/2} \frac{2}{3} E^{3/2}.$$

Setting $E = 3k_B T/2$, we obtain

$$G \approx V \left(\frac{2\pi m k_B T}{h^2} \right)^{3/2}.$$

Substituting this expression into (3.10), we obtain the condition for nondegeneracy:

$$n \left(\frac{h^2}{2\pi m k_B T} \right)^{3/2} \ll 1 \quad (3.24)$$

where $n = N/V$ is the number of particles per unit volume.

Consider some molecular gas, for instance, nitrogen in normal conditions. For it $n \approx 10^{26} \text{ m}^{-3}$, $m = 4.5 \times 10^{-26} \text{ kg}$, and $k_B T = 4 \times 10^{-21} \text{ J}$. Substituting the figures into the left-hand side of (3.24), we obtain $nh^3(2\pi m k_B T)^{-3/2} \approx 10^{-6}$, which is much less than unity. Accordingly, the molecular gases are normally nondegenerate and must be described with the aid of the Maxwell-Boltzmann classical statistics.

Consider now the electron gas in metals. For it we have that $n \approx 5 \times 10^{24} \text{ m}^{-3}$ and $m = 9 \times 10^{-31} \text{ kg}$. For such values of n and m the electron gas turns out to be nondegenerate only at temperatures above 105 K; the left-hand side of (3.24) for such temperatures diminishes to less than unity (at $T = 105 \text{ K}$ it is approximately 0.5). Therefore, in practice the electron gas in metals is always degenerate and on account of this should be described with the aid of the Fermi-Dirac statistics.

It follows from (3.24) that a nondegenerate state of a gas can be realized not only by raising its temperature but by reducing its concentration n as well. For $n \approx 10^{22} \text{ m}^{-3}$ the left-hand side of (3.24) for electrons at normal temperatures is approximately 10^{-3} and the electron gas becomes nondegenerate. Such (and smaller) concentrations of the electron gas are found in some semiconductors. In such semiconductors termed *nondegenerate*, the electron gas is nondegenerate and is described by the classical Maxwell-Boltzmann statistics.

Let us try now to find the distribution function $f(E)$. The form of this function depends in the first instance on whether the gas is degenerate or nondegenerate. In the case of a degenerate gas the important point is whether the gas consists of fermions or bosons.

Let us start with a nondegenerate gas whose distribution function $f(E)$ is independent of the particles' nature.

§ 26. Distribution function for a nondegenerate gas

Appendix ?? contains an elementary derivation of the distribution function for a nondegenerate gas. It is of the following form:

$$f(E) = e^{(\mu-E)/k_B T} \quad (3.25)$$

where k_B is Boltzmann's constant, and μ the chemical potential. Calculations give the following expression for μ of a nondegenerate gas:

$$\mu = k_B T \ln \left[\frac{N}{V} \left(\frac{h^2}{2\pi m k_B T} \right)^{3/2} \right]. \quad (3.26)$$

Substituting it into (3.25) we obtain:

$$f_M(E) = \frac{N}{V} \left(\frac{h^2}{2\pi m k_B T} \right)^{3/2} e^{-E/(k_B T)}. \quad (3.27)$$

We would like to remind again that $f_M(E) dE$ expresses the probability of occupation of the states in the energy interval $(E, E + dE)$; the term for it is the *Maxwell-Boltzmann distribution function*.

Figure Figure 3.3(a) shows a graph of the function $f_M(E)$. It has a maximum at $E = 0$ and asymptotically approaches zero as $E \rightarrow \infty$. This means that the lower

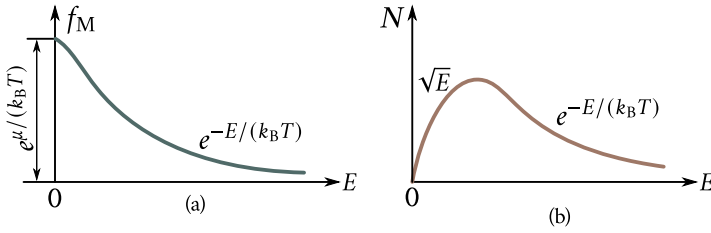


Figure 3.3: Distribution functions for nondegenerate gas: (a) —the Maxwell-Boltzmann distribution function expressing the average density of state occupation by particles; (b) — the complete Maxwell-Boltzmann distribution function.

energy states have the greatest probability of occupation. As the energy of a state increases its probability of occupation diminishes steadily.

Multiplying $f_M(E)$ by the number of states $g(E) dE$ [see Eq. (3.22)], we obtain the complete distribution function of particles over the energy

$$N(E) dE = \frac{4\pi V}{h^3} (2m)^{3/2} e^{\mu/(k_B T)} e^{-E/(k_B T)} E^{1/2} dE, \quad \text{or} \quad (3.28)$$

$$N(E) dE = \frac{2N}{\sqrt{\pi} (k_B T)^{3/2}} e^{-E/(k_B T)} E^{1/2} dE.$$

It is termed the *complete Maxwell-Boltzmann distribution function*. Figure 3.3(b) shows the graph of this function. Because of the factor $E^{1/2}$ its maximum is displaced to the right of the origin.

Knowing the distribution function $f_M(E)$ we may easily find the laws of distribution of the particles over the momentum, $N(p) dp$, and over its components, $N(p_x, p_y, p_z) dp_x dp_y dp_z$. We may also find them over the velocity, $N(v) dv$, and over its components $N(v_x, v_y, v_z) dv_x dv_y dv_z$ over one of the components of velocity, say $N(v_x) dv_x$, etc. Those distributions are shown below

$$N(p) = \frac{4\pi N}{(2\pi m k_B T)^{3/2}} e^{-p^2/(2m k_B T)} p^2, \quad (3.29)$$

$$N(v) = 4\pi N \left(\frac{m}{2\pi k_B T} \right)^{3/2} e^{-mv^2/(k_B T)} v^2, \quad (3.30)$$

$$N(p_x, p_y, p_z) = N \left(\frac{N}{2\pi m k_B T} \right)^{3/2} e^{-(p_x^2 + p_y^2 + p_z^2)/(2m k_B T)}, \quad (3.31)$$

$$N(v_x, v_y, v_z) = N \left(\frac{m}{2\pi k_B T} \right)^{3/2} e^{-m(v_x^2 + v_y^2 + v_z^2)/(k_B T)}, \quad (3.32)$$

$$N(v_x) = N \left(\frac{m}{2\pi k_B T} \right)^{1/2} e^{-mv_x^2/(k_B T)}. \quad (3.33)$$

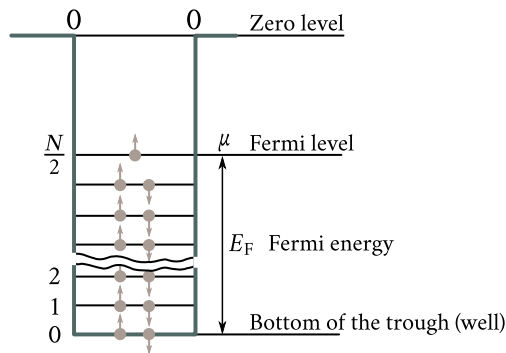


Figure 3.4: Schematic representation of a metal as a potential well for free electrons.

The reader is requested to obtain those results himself.

§ 27. Distribution function for a degenerate fermion gas

The distribution function for a degenerate fermion gas was first obtained by Fermi and Dirac. It is of the following form:

$$f_{\text{F}}(E) = \frac{1}{e^{(E-\mu)/(k_{\text{B}}T)} + 1}. \quad (3.34)$$

An elementary derivation of this expression is presented in Appendix ?? . Here, as before, μ , denotes the chemical potential, which in the case of a degenerate fermion gas is termed the *Fermi level*.

Equation (3.34) shows that for $E = \mu$, the distribution function $f(E) = 1/2$ at any temperature $T \neq 0$. Therefore, from the statistical point of view the Fermi level is a state whose probability of occupation is $1/2$.

The function (3.34) is termed the *Fermi-Dirac function*. To obtain a clear picture of the nature of this function one should consider the degenerate electron gas in metals at absolute zero.

Electron distribution in a metal at absolute zero. Fermi energy. The metal is a sort of a potential well (trough) for free electrons. To leave it the electron should have work performed on it against the forces retaining it in the metal. Figure 3.4 shows the diagram of such a potential well. The horizontal lines denote energy levels which the electrons may occupy. In compliance with the Pauli exclusion principle there may be two electrons with opposite spins on each such level. For an electron gas of N electrons the last occupied level will be the $N/2$ level. This level is termed the *Fermi level* for a degenerate electron gas. It corresponds to the maximum kinetic energy E_F an electron in a metal may possess at absolute

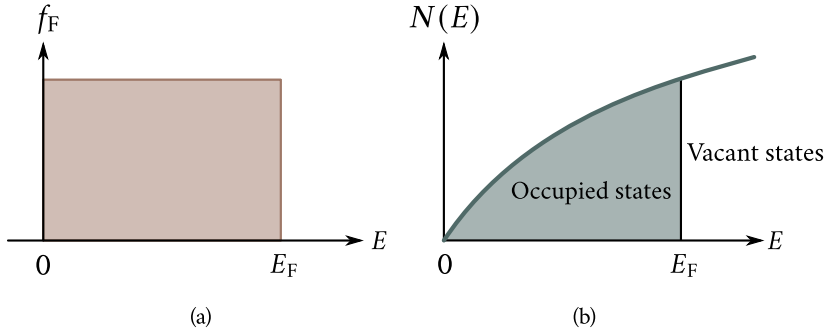


Figure 3.5: The distribution function for degenerate fermion gas: (a) — the Fermi-Dirac distribution function for $T = 0$ K; (b) — the complete Fermi-Dirac distribution function for $T = 0$ K.

zero. This energy is termed the *Fermi energy*.

Thus, at absolute zero all states with the energy $E < E_F$ are occupied by electrons and all states with the energy $E > E_F$ are free. In other words, the probability of occupation of a state with the energy $E < E_F$ at $T = 0$ K is unity and the probability of occupation of a state with the energy $E > E_F$ is zero:

$$f_F(E) = \begin{cases} 1, & \text{for } E < E_F, \\ 0, & \text{for } E > E_F. \end{cases} \quad (3.35)$$

To obtain this result from (3.34) one should assume that at $T = 0$ K the chemical potential of the electron gas measured from the bottom of the potential well is equal to the Fermi energy E_F :

$$\mu = E_F. \quad (3.36)$$

Indeed, setting in (3.34) $\mu = E_F$ we obtain

$$f_F(E) = \frac{1}{e^{(E-E_F)/(k_B T)} + 1}. \quad (3.37)$$

If $E < E_F$, then $e^{(E-E_F)/(k_B T)} \rightarrow 0$ at $T = 0$ K and $f_F = 1$; if $E > E_F$, then $e^{(E-E_F)/(k_B T)} \rightarrow \infty$ at $T = 0$ K and $f_F = 0$.

Figure 3.5(a) shows the graph of the Fermi-Dirac distribution function at absolute zero. It has the shape of a step terminating at $E = E_F$.

Multiplying (3.35) by the number of states $g(E) dE$ [see Eq. (3.22)], we obtain the *complete Fermi-Dirac distribution function at absolute zero*:

$$N(E) dE = \frac{4\pi V}{h^3} (2m)^{3/2} E^{1/2} dE \quad (3.38)$$

because $f_F = 1$ in the energy interval $(0, E_F)$. The graph of the function is presented in Figure 3.5(b), where the area of the occupied states is shaded.

Integrating expression (3.38) from 0 to E_F , we obtain

$$N = \frac{8\pi V}{3h^3} E_F^{3/2} (2m)^{3/2}$$

whence the Fermi energy may be easily obtained

$$E_F = \frac{h^2}{2m} \left(\frac{3n}{8\pi} \right)^{2/3} \quad (3.39)$$

where $n = N/V$ is the concentration of electron gas in the metal.

Knowing the energy distribution function of the electrons, we may easily calculate the average energy of the electrons at absolute zero, \bar{E}_0 :

$$\bar{E}_0 = \frac{3}{5} E_F = \frac{3h^2}{10m} \left(\frac{3n}{8\pi} \right)^{2/3}. \quad (3.40)$$

Lastly, knowing E_F and \bar{E}_0 , we can calculate the maximum velocity v_F and the effective velocity v_{eff} (corresponding to average energy) of free electrons in a metal at absolute zero:

$$v_F = \left(\frac{2E_F}{m} \right)^{1/2}, \quad v_{\text{eff}} = \left(\frac{2\bar{E}_0}{m} \right)^{1/2}. \quad (3.41)$$

Table 3.2 shows the Fermi energy E_F , the average energy \bar{E}_0 , the maximum and effective velocities of free electrons at absolute zero for some metals. The last column contains the Fermi temperature determined from the relation

$$T_F = \frac{E_F}{k_B}. \quad (3.42)$$

This is the temperature at which a molecule in a normal nondegenerate gas would have the energy of thermal motion $3k_B T/2$ equal to the Fermi energy E_F multiplied by 3/2.

It may be seen from Table 3.2 that Fermi temperatures are so high that no metal can exist in a condensed state.

Table 3.2

Metal	E_F (eV)	\bar{E}_0 (eV)	v_F (10^6 m s $^{-1}$)	v_{eff} (10^6 m s $^{-1}$)	T_F (10^4 K)
Copper	7.10	4.30	1.60	1.25	8.20
Lithium	4.72	2.80	1.30	1.00	5.50
Silver	5.50	3.30	1.40	1.10	6.40
Sodium	3.12	1.90	1.10	0.85	3.70

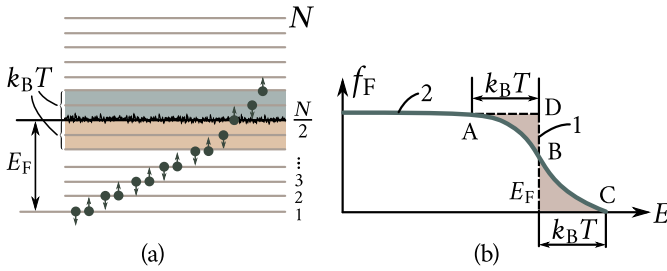


Figure 3.6: Temperature dependence of the Fermi-Dirac distribution function: (a) — thermal excitation of electrons; (b) — the Fermi-Dirac distribution function for $T > 0$ K.

It should be stressed that, although the Fermi energy represents the kinetic energy of translational motion of free electrons, it is not the energy of their thermal motion. Its nature is purely quantum mechanical and is due to the fact that electrons are fermions satisfying the Pauli exclusion principle.

Temperature dependence of the Fermi-Dirac distribution. When the temperature is raised, the electrons become thermally excited and go over to higher energy levels. This causes a change in their distribution over the states. However, in the range of temperatures in which the energy of their thermal motion, $k_B T$, remains much less than E_F only the electrons in a narrow band about approximately $k_B T$ wide adjoining the Fermi level may be thermally excited [Figure 3.6(a); the excited states are shaded]. The electrons of the lower levels remain practically unaffected because the energy of thermal excitation $k_B T$ is not enough to excite them (to transfer them to levels above the Fermi level).

As a result of thermal excitation some of the electrons with an energy less than E_F are transferred to the levels with energies greater than E_F and a new distribution of electrons over the states is established. Figure 3.6(b) shows the curves of the electron distribution over the states for $T = 0$ K (curve 1) and for $T > 0$ K (curve 2). It can be seen that the rise in temperature causes the original distribution to smear to a depth of $k_B T$ with a “tail” BC appearing to the right of E_F . The higher the temperature the greater the change in the distribution function. The tail BC itself is described by the Maxwellian distribution function.

The shaded areas in Figure 3.6(b) are proportional to the number of electrons transferred from the states with $E < E_F$ (the area ADB) to the states with $E > E_F$ (the area $E_F BC$). Those areas are equal since they represent the same number of electrons.

Let us make an approximate estimate of this number, ΔN . There are $N/2$ energy levels inside the interval $(0, E_F)$, where N is the number of free electrons in the metal. To simplify the problem we may assume that these levels are equidistant,

the separation being $\Delta\varepsilon = E_F/(N/2) = 2E_F/N$. Only the electrons in the band $k_B T$ wide just below E_F [Figure 3.6(a)] are thermally excited. There are $k_B T/\Delta\varepsilon = E_F N/(2E_F)$ levels inside this band occupied by $2k_B T N/(2E_F) = k_B T N/E_F$ electrons. Assuming that not more than a half of those electrons go over the Fermi level, we obtain the following approximate relation for ΔN :

$$\Delta N \approx \frac{k_B T}{2E_F} N. \quad (3.43)$$

At room temperature $k_B T \approx 0.025$ eV, $E_F = 3$ eV to 10 eV, therefore, $\Delta N/N < 1\%$; at $T = 1000$ K we find that $\Delta N/N \approx 1\%$ to 2% .

Hence, in all the temperature range in which the electron gas in a metal is degenerate its distribution is close to that at absolute zero.

Accordingly, only a negligible fraction of the electrons close to the Fermi level are thermally excited. At room temperature this fraction is less than 1% of the total number of conduction electrons. The laws governing the distribution of the electrons in metals discussed above remain valid practically in all cases, because in all the temperature range in which the existence of metals in the condensed state is possible the electron gas remains degenerate.

Consider the temperature dependence of the chemical potential. Integrating the complete Fermi-Dirac distribution function $f_F(E)g(E) dE$ over energy, we obtain the total number N of free electrons in a metal:

$$N = \int_0^\infty f_F(E)g(E) dE = \frac{4\pi V}{h^3} (2m)^{3/2} \int_0^\infty \frac{E^{1/2}}{[e^{(E-\mu)/(k_B T)} + 1]} dE.$$

Generally, there is no analytic expression for this integral. Approximate calculation in the temperature range in which the electron gas remains strongly degenerate yields the following temperature dependence of μ :

$$\mu = E_F \left[1 - \frac{\pi^2}{12} \left(\frac{k_B T}{E_F} \right)^2 \right]. \quad (3.44)$$

As $k_B T$ remains much less than E_F up to the melting point of a metal, the decrease in μ with the rise in T turns out to be so small that it can often be neglected and the Fermi level can be assumed to coincide with E_F at any temperature.

One can also calculate the average energy of the electrons \bar{E} in a degenerate electron gas dividing its total energy, $E_t = \int_0^\infty E f_F(E)g(E) dE$, by the number of the electrons, N :

$$\bar{E} = \frac{E_t}{N} = \frac{\int_0^\infty \frac{E^{3/2}}{[e^{(E-\mu)/(k_B T)} + 1]} dE}{\int_0^\infty \frac{E^{1/2}}{[e^{(E-\mu)/(k_B T)} + 1]} dE}.$$

An approximate calculation of these integrals yields

$$\bar{E} = \frac{3}{5} E_F \left[1 + \frac{5\pi^2}{12} \left(\frac{k_B T}{E_F} \right)^2 \right]. \quad (3.45)$$

At $T = 0$ K turns into (3.40).

Lifting of degeneracy. Nondegenerate electron gas. When the condition for nondegeneracy (3.10) is fulfilled, every gas including the electron gas must become nondegenerate. Let us discuss this in more detail.

According to (3.10) the gas is nondegenerate if the average occupancy of the states by the particles is much less than unity. Since the distribution function $f(E)$ represents the occupancy of the states, the condition of nondegeneracy (3.10) may be written in the form

$$f(E) \ll 1. \quad (3.46)$$

The Fermi-Dirac function (3.34) will be much less than unity if the exponential term, $e^{(E-\mu)/(k_B T)}$, will be much greater than unity:

$$e^{(E-\mu)/(k_B T)} \gg 1. \quad (3.46')$$

This inequality should hold for all the states including that with $E = 0$:

$$e^{(E-\mu)/(k_B T)} \gg 1. \quad (3.47)$$

It follows from (3.47) that for a nondegenerate electron gas in which condition (3.46) is satisfied, $-\mu$ should be a positive quantity considerably greater than $k_B T$:

$$-\mu \gg k_B T. \quad (3.48)$$

The chemical potential μ should be negative and greater than $k_B T$ in absolute value.

If the condition (3.46') is fulfilled, unity in the denominator of the Fermi-Dirac function can be neglected and the following expression for the distribution function of a nondegenerate electron gas may be obtained:

$$f(E) = e^{\mu/(k_B T)} e^{-E/(k_B T)}. \quad (3.49)$$

Comparing (3.49) with (3.25), we see that a nondegenerate electron gas like every other nondegenerate gas is described by the Maxwell-Boltzmann distribution function.

The electron gas in metals, where the free electron concentration is always very high ($\approx 10^{28} \text{ m}^{-3}$), is always in a degenerate state described by the Fermi-Dirac distribution function.

A nondegenerate electron gas is a feature of the *intrinsic* (pure) and weakly doped semiconductors, which are the mainstay of modern semiconductor electronics. The concentration of free electrons in such semiconductors is substantially less than in metals, varying from 10^{16} - 10^{19} m^{-3} to 10^{23} - 10^{24} m^{-3} depending

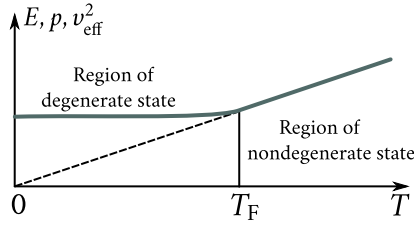


Figure 3.7: Temperature dependence of energy, pressure, and the square of effective velocity of electrons in a metal.

upon the concentration of electrically active impurities. The nondegeneracy condition (3.10) remains valid for such concentrations and the electron gas is nondegenerate.

In conclusion, to illustrate the drastic difference in the behaviour of the ideal nondegenerate gas obeying the Maxwell-Boltzmann statistics and of the degenerate electron gas described by the Fermi-Dirac statistics, we would like to cite some of their properties (Table 3.3).

It follows that average energy \bar{E} , effective velocity v_{eff} and pressure p of a nondegenerate ideal gas are functions of temperature that vanish at absolute zero. At the same time \bar{E} , and p of a degenerate electron gas are very large already at absolute zero and are practically independent of temperature (Figure 3.7). This points to the fact that, as noted above, \bar{E} , v_{eff} , and p of a degenerate electron gas are for the most part not of thermal origin, the contribution of the thermal electron motion

Table 3.3

Parameters	Gas	
	Nondegenerate	Degenerate
\bar{E} at 0 K	0	$\bar{E}_0 = \frac{3}{5} E_F$
\bar{E} at T K	$\bar{E} = \frac{3}{2} k_B T$	$\bar{E} = \frac{3}{5} E_F \left[1 + \frac{5\pi^2}{12} \left(\frac{k_B T}{E_F} \right)^2 \right] \approx \frac{3}{5} E_F$
v_{eff} at 0 K	0	$v_{\text{eff},0} = \left(\frac{E_F}{m} \right)^{1/2} \approx 10^6 \text{ m s}^{-1}$
v_{eff} at T K	$v_{\text{eff}} = \left(\frac{3k_B T}{m} \right)^{1/2}$	$v_{\text{eff}} = v_{\text{eff},0}$
p at 0 K	0	$p_0 = \frac{2}{3} \bar{E}_0 \approx 10^{10} \text{ Pa}$
p at T K	$p = \frac{RT}{V}$	$p \approx p_0$

to these quantities being negligible.

§ 28. Distribution function for a degenerate boson gas

In contrast to the electrons, which satisfy the Pauli exclusion principle, the bosons can occupy both the free states and the states already occupied by other bosons the more readily the greater the occupancy of the latter.

The distribution function of bosons over the states was first obtained by Bose and Einstein. It is of the following form:

$$f_{\text{Bose}}(E) = \frac{1}{[e^{(E-\mu)/(k_B T)} - 1]}. \quad (3.50)$$

(An elementary derivation of this expression is given in Appendix ??). It is termed the *Bose-Einstein distribution function*. Let us use it to describe the state of a photon gas.

Suppose that a cavity inside a black body at a temperature T is filled with equilibrium thermal radiation. From the quantum mechanical point of view this radiation may be regarded as consisting of an enormous number of photons constituting a photon gas. The photon's spin is $s = 1$. Therefore, photons are bosons, which implies that the photon gas should satisfy the Bose-Einstein distribution.

The photons have some peculiarities as compared with other bosons, for instance, the helium nucleus, $\frac{4}{2}\text{He}$:

- (1) The rest mass of a photon is zero.
- (2) All photons move with the same speed equal to that of light, c , but can have different energy E and momentum p ; E and p depend on the photon's frequency:

$$E = h\nu = \hbar\omega, \quad p = \frac{h\nu}{c} = \frac{\hbar\omega}{c} \quad (3.51)$$

where $\hbar = h/2\pi$ and $\omega = 2\pi\nu$. It follows from (3.51) that

$$E = pc. \quad (3.52)$$

- (3) The photons do not collide with one another. Therefore, an equilibrium distribution in a photon gas can be established only in the presence of a body capable of absorbing and emitting photons. The walls of the cavity in which the radiation is contained may serve as an example of such a body. The transformation of a photon of one frequency into a photon of another frequency takes place in the processes of absorption and subsequent emission.
- (4) The photons may be generated (in the act of emission) and annihilated (in the act of absorption) in any numbers. Therefore, the number of photons in a photon gas does not remain fixed but depends on the state of the gas. For

specified values of V and T the photon gas in a state of equilibrium contains so many photons N_0 as are needed for the energy of the gas to be at its minimum. This makes it possible to express the condition for the equilibrium of the photon gas in the form:

$$\left(\frac{dE}{dN} \right)_{V,T} = 0. \quad (3.53)$$

Since according to (3.8) $(dE/dN)_{V,T} = \omega$ the equilibrium condition (3.53) means that $\mu = 0$. Hence, the chemical potential of an equilibrium photon gas is zero.

For a nondegenerate gas the chemical potential is negative and has a relatively great absolute value. The fact that for the photon gas $\mu = 0$ means that such gas is always degenerate.

Setting $\mu = 0$ in (3.50), we obtain the distribution function for the photon gas:

$$f_p(E) = \frac{1}{[e^{E/(k_B T)} - 1]} = \frac{1}{[e^{(\hbar\omega)/(k_B T)} - 1]}. \quad (3.54)$$

This formula was first obtained by Max Planck and is termed the *Planck formula*. It represents the average fraction of photons having the energy $E = \hbar\omega$. Using this formula, we may easily formulate the law for the energy distribution in the spectrum of a black body. The following expression can be obtained for the energy density of the radiation of such a body:

$$\rho(\omega) = \frac{2\hbar}{\pi c^3} \left[\frac{\omega^3}{e^{(\hbar\omega)/(k_B T)} - 1} \right]. \quad (3.55)$$

The readers are requested to derive this formula themselves making use of (3.54) and (3.18).

§ 29. Rules for statistical averaging

As was already stated, to specify the state of an ensemble one should specify its state parameters. To specify the state of a particle one should specify the values of its coordinates and momentum components.

The problem of going over from the parameters of the individual particles to the state parameters characterizing the ensemble involves the problem of transition from the dynamical laws describing the behaviour of the individual particles to the statistical laws describing the behaviour of the ensemble. To effect such a transition it is necessary to perform the averaging of the characteristics of motion of the individual particles assuming the chances of all the particles belonging to the ensemble to be identical. The state parameters of the ensemble are expressed

in terms of the averaged parameters of the individual particles belonging to the ensemble.

To make the rules of averaging apparent, let us consider an ensemble of N identical particles each of which is capable of assuming one of the discrete set of energy values: E_1, E_2, \dots, E_m . Choose an arbitrary moment of time and note the energy every particle has at that moment. We obtain as a result a set of numbers $N(E_i)$ expressing the number of particles having the energy E_i . To determine the average energy \bar{E} of the particles we add up the energies of all of them and divide the sum by the number of particles. The total number of particles is $N = \sum_{i=1}^m N(E_i)$, and the sum of their energies is $\sum_{i=1}^m E_i N(E_i)$. Therefore,

$$\bar{E} = \frac{\sum_{i=1}^m E_i N(E_i)}{\sum_{i=1}^m N(E_i)}. \quad (3.56)$$

The average value E obtained in this fashion is termed an *ensemble average*.

If the particle's energy assumes a continuous set of values, the practice is to count the number of particles having an energy lying within an interval $(E, E + dE)$ instead of the number of particles having an exact value of energy. The average energy will then be

$$\bar{E} = \frac{\int_0^\infty E N(E) dE}{\int_0^\infty N(E) dE}. \quad (3.57)$$

Such averaging may be performed for any physical quantity M that is a function of the coordinates and the momenta of the particles making up the ensemble. If M is continuous,

$$\bar{M} = \frac{\int_0^\infty M N(M) dM}{\int_0^\infty N(M) dM}. \quad (3.58)$$

Let us determine the average energy of the particles of an ideal nondegenerate gas. According to (3.57) and (3.28), we have

$$\bar{E} = \frac{\int_0^\infty E N(E) dE}{\int_0^\infty N(E) dE} = \frac{2}{\sqrt{\pi}(k_B T)^3} \int_0^\infty e^{-E/(k_B T)} E^{3/2} dE = \frac{3}{2} k_B T. \quad (3.59)$$

The results of the calculations of the average values of the velocity, of a velocity component, of the effective velocity, and of its component for the particles of an ideal gas are presented thus:

$$\bar{v} = \left(\frac{8k_{\text{B}}T}{\pi m} \right)^{1/2}, \quad \bar{v}_x = \left(\frac{k_{\text{B}}T}{2\pi m} \right)^{1/2},$$
$$v_{\text{eff}} = \left(\frac{3k_{\text{B}}T}{m} \right)^{1/2}, \quad v_{\text{eff},x} = \left(\frac{k_{\text{B}}T}{m} \right)^{1/2}.$$

The reader may for the sake of practice perform these calculations himself.

Chapter 4

Thermal Properties of Solids

§ 30. Normal modes of a lattice

The atoms of a solid take part in thermal vibrations around their equilibrium positions. Because of a strong interaction between them, the nature of those vibrations turns out to be extremely complex and an accurate description of it presents enormous difficulties. Therefore, approximate methods and various simplifications are used to solve this problem.

Instead of describing the individual vibrations of the particles the practice is to consider their collective motion in the crystal, which is a spatially ordered structure. This simplification is based on the fact that powerful bonds immediately transmit the vibrations of one particle to other particles and a collective motion in the form of an elastic wave involving all the particles of the crystal is excited in it. Such collective motion is called the *normal mode of a lattice*. The number of normal modes coincides with the number of degrees of freedom, which is $3N$ if N is the number of particles constituting the crystal.

Figure 4.1(a) represents a one-dimensional model of a solid—a linear chain of atoms separated by a distance a and able to vibrate in the direction perpendicular to the chain. Such a chain may be regarded as a string. If the ends of the chain are fixed, the fundamental mode corresponding to the lowest frequency ω_{\min} is represented by the standing wave with a node at each end [Figure 4.1(b), curve 1]. The second mode is represented by the standing wave with an additional node in the centre of the chain (curve 2). The third mode, or third harmonic, is represented by the standing wave with two additional nodes that divide the chain in three equal parts (curve 3), etc. The length of the shortest wave in such a chain is evidently equal to twice the distance between the atoms of the chain [Figure 4.1(c)]:

$$\lambda_{\min} = 2a. \quad (4.1)$$

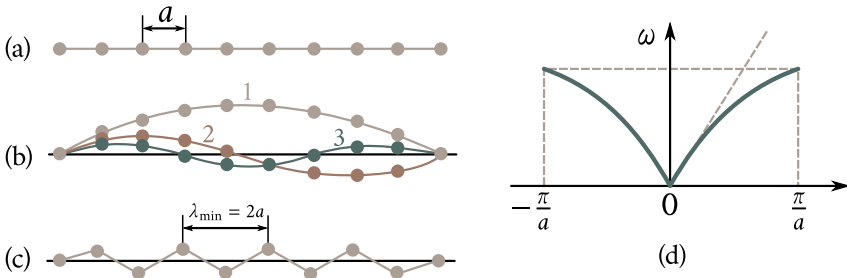


Figure 4.1: Normal modes of a linear chain made up of identical atoms: (a)—linear chain; (b)—normal modes of the chain; (c)—normal modes of the chain corresponding to shortest wavelength (to highest frequency); (d)—dispersion curves expressing dependence of normal mode frequency on wave vector.

The corresponding maximum frequency ω_{\max} is

$$\lambda_{\max} = \frac{2\pi v}{\lambda_{\min}} = \frac{\pi v}{a} \quad (4.2)$$

where v is the velocity of wave propagation (of sound) along the chain.

This maximum frequency is a parameter of the chain's material and is determined by the interatomic distance and the velocity of wave propagation. Should we set $a = 3.6 \times 10^{-10}$ m (the lattice parameter of copper) and $v = 3550$ m s⁻¹ (the velocity of sound in copper) we would obtain $\omega_{\max} \approx 3 \times 10^{13}$ s⁻¹, which corresponds to the frequency of atomic vibrations in a solid.

To describe wave processes one usually uses the wave vector \mathbf{q} whose direction coincides with that of wave propagation and whose absolute value is

$$q = \frac{2\pi}{\lambda}. \quad (4.3)$$

It follows from Eq. (4.2) that $2\pi/\lambda = \omega/v$. Therefore¹,

$$q = \frac{\omega}{v} \quad \Rightarrow \quad \omega = qv. \quad (4.4)$$

¹The phase velocity v , which enters (4.4), is itself a function of the wave vector \mathbf{q} and for a linear chain of atoms bonded by elastic forces is expressed by the following relation:

$$v = v_0 \frac{\sin(qa/2)}{(qa/2)} \quad (4.4')$$

where v_0 is the velocity of wave propagation in a continuous string for which $a = 0$. It follows from Eq. (4.4') that for a constant a , the velocity v is practically independent of \mathbf{q} and is approximately v_0 only in the range of small q 's, where $[\sin(qa/2)]/(qa/2) \ll 1$. In this range ω increases approximately in proportion to q [Figure 4.1(d)]. As q increases, the value of $[\sin(qa/2)]/(qa/2) \ll 1$ steadily diminishes and for $q = \pi/a$ tends to $2/\pi$. This causes the dispersion curve $\omega(q)$ to flatten out, so that for $q = \pi/a$ it runs parallel to q .

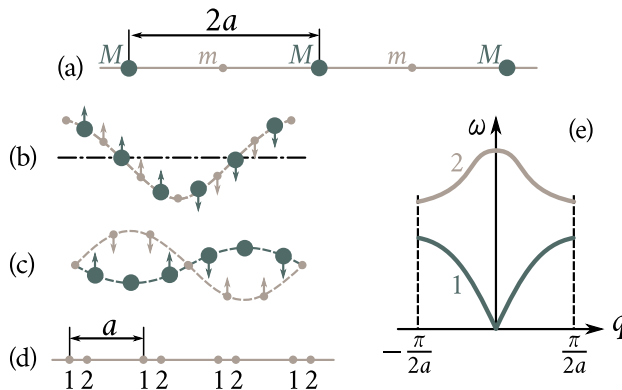


Figure 4.2: Normal modes of a chain made up of atoms of two kinds: (a)—arrangement of atoms in the chain; (b)—acoustic modes; (c)—optical modes; (d)—dispersion curves for acoustic and optical modes; (e)—linear lattice with a basis in which optical and acoustic modes are present.

Figure 4.1(d) shows the dependence on wave vector q of the frequency of normal modes in a linear chain made up of atoms of one kind. As q increases from 0 to π/a , the frequency of the normal modes rises, reaching the maximum value $\omega_{\max} = \pi v/a$ for $q = \pi/a$, that is, for $\lambda = 2a$. Curves of this type expressing the dependence of the vibration frequency on the wave vector (the wavelength) are termed *dispersion curves*.

Consider now a chain made up of atoms of two kinds placed in a regular sequence one after another [Figure 4.2(a)]. Denote the mass of the heavier atoms by M and that of the lighter atoms by m . Two types of normal modes can be present in such a chain, as is shown in Figure 4.2(b,c). The modes in Figure 4.2(b) are quite identical to the modes of a uniform chain: the phases of the neighbouring atoms are practically the same. Such vibrations are termed *acoustic modes*, since they include the entire spectrum of the acoustic modes of the chain. For them, $\omega_{\text{ac}} = 0$ for $q = 0$. They play a decisive part in determining the thermal properties of the crystals such as heat capacity, heat conductivity, thermal expansion, etc.

In case of normal modes shown in Figure 4.2(c) the phases of the neighbouring vibrating atoms are opposite. Such modes can be treated as the relative vibrations of two interpenetrating sublattices, each made up of atoms of one kind. They are termed *optical modes*, since they play a decisive part in the processes of interaction of light with crystals.

Figure 4.2(d) shows the dispersion curves for the acoustic (1) and optical (2) normal modes of a chain made up of atoms of two kinds. In contrast to an acoustic mode, whose frequency rises with the wave vector reaching the maximum value at

$q_{\max} = \pi/(2a)$, the maximum frequency of an optical mode corresponds to $q = 0$; with the increase in q the frequency of an optical mode falls off to its minimum at $q_{\max} = \pi/(2a)$.

The optical vibrations are possible not only in a chain made up of atoms of different kinds, but in a complex chain made up of two, or more, interpenetrating chains containing atoms of one kind, as shown in Figure 4.2(e). Unit cell of such a complex chain contains two atoms. The optical modes are the result of the relative vibrations of two sublattices.

§ 31. Normal modes spectrum of a lattice

One of the principal problems of the theory of lattice vibrations is the problem of the frequency distribution of normal modes. Consider now the simplest case of the normal modes in a linear atomic chain (see Figure 4.1).

The wavelengths of the normal modes in such a chain are

$$\lambda_n = \frac{2L}{n} \quad (n = 1, 2, 3, \dots, N) \quad (4.5)$$

where L is the length of the chain, and N the number of atoms in it.

The number of normal modes z having the wavelength equal to or greater than λ_n will evidently be n :

$$z = n = \frac{2L}{\lambda_n}.$$

In the same way the number of standing waves in a three-dimensional crystal of volume V (for instance, in a cube with edge L and volume L^3) having the wavelength equal to or greater than λ should be

$$z = \left(\frac{2L}{\lambda} \right)^3 = \frac{8V}{\lambda^3}.$$

A more accurate calculation yields

$$z = \frac{4\pi V}{\lambda^3}. \quad (4.6)$$

Since $\lambda = 2\pi v/\omega$, it follows that

$$z = \frac{V}{2\pi^2 v^3} \omega^3. \quad (4.7)$$

Differentiating this expression, we obtain

$$dz = g(\omega) d\omega = \frac{3V}{2\pi^2 v^3} \omega^2 d\omega. \quad (4.8)$$

Equation (4.8) expresses the number of normal modes per frequency interval

$(\omega, \omega + d\omega)$. The function

$$g(\omega) = \frac{dz}{d\omega} = \frac{3V}{2\pi^2 v^3} \omega^2 \quad (4.9)$$

determines the *density* of the normal vibrations in $d\omega$ of the spectrum, that is their spectrum. The function $g(\omega)$ is termed *spectral distribution function of normal modes*.

Since the number of normal vibrations in a lattice is $3N$, the function $g(\omega)$ should satisfy the following normalization condition:

$$\int_0^{\omega_D} g(\omega) d\omega = 3N \quad (4.10)$$

where ω_D is the maximum frequency limiting the spectrum of normal modes from above.

Substituting (4.9) into (4.10) and integrating, we obtain

$$\frac{V\omega_D^3}{2\pi^2 v^3} = 3N. \quad (4.11)$$

Hence,

$$\omega_D = v \left(6\pi^2 \frac{N}{V} \right)^{1/3}. \quad (4.12)$$

The frequency ω_D is termed the *Debye frequency* and the temperature

$$\Theta = \frac{\hbar\omega_D}{k_B} \quad (4.13)$$

the *Debye temperature*. Table 4.1 shows the Debye temperatures of some chemical elements and compounds.

At the Debye temperature the entire vibration spectrum is excited in the solid, including the one with the maximum frequency ω_D . Accordingly, any further rise in temperature (above Θ) shall not be accompanied by the appearance of new normal modes. In this case, the role of the temperature is to increase the intensity of each of the normal modes with a resultant increase in their average energy.

Temperatures $T > \Theta$ are usually referred to as *high temperatures*.

Substituting v^3 from (4.11) into (4.9), we obtain

$$g(\omega) = 9N \frac{\omega^2}{\omega_D^3}. \quad (4.14)$$

§ 32. Phonons

Each normal mode carries with it some energy and momentum. Oscillation theory contains the proof of the fact that the energy of a normal mode is equal to the

energy of an oscillator with a mass equal to the mass of the vibrating atoms and the frequency of the normal mode. Such oscillators are *termed normal*.

Denote the energy of the i -th mode characterized by the frequency ω_i by $E_{i,n.m.}$. It is equal to the energy $E_{i,n.o.}$ of a normal oscillator of the same frequency ω_i : $E_{i,n.m.} = E_{i,n.o.}$. The total energy of the crystal in which all the $3N$ normal modes have been excited is

$$E = \sum_i^{3N} E_{i,n.o.}$$

Hence, the total energy of the crystal made up of N atoms taking part in coupled vibrations is equal to the energy of $3N$ independent normal harmonic linear oscillators. In this sense, a system of N atoms whose vibrations are interconnected is equivalent to a set of $3N$ normal oscillators and the problem of calculating the average energy of such a system is reduced to a simpler problem of calculating the average energy of normal oscillators.

It should be pointed out that normal oscillators have nothing in common with real atoms except the mass. Every oscillator represents one of the normal modes of the lattice in which all the atoms of the crystal take part vibrating with the same

Table 4.1

Element	Θ , (K)	Element	Θ , (K)	Element	Θ , (K)
Al	418	Fe	467	Pb	94.5
Ag	225	Ge	366	Pd	275
Au	165	Hg	275	Pt	229
Be	1160	In	109	Si	658
Bi	117	KBr	174	Sn (gray)	212
C (diamond)	1910	KCl	227	Sn (white)	189
Ca	219	La	132	Ta	231
CaF ₂	474	Mg	406	Ti	278
Cd	300	Mo	425	Tl	89
Co	445	NaCl	320	V	273
Cr	402	Nb	252	W	(379)
Cu	339	Ni	456	Zn	308

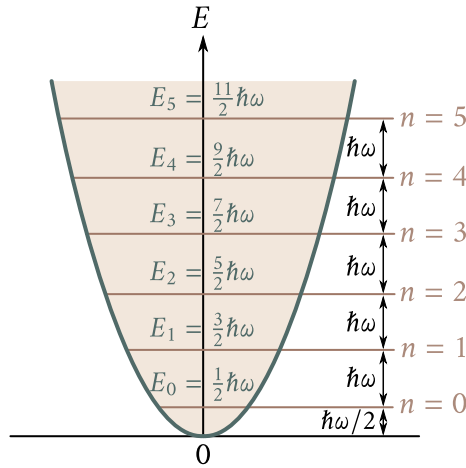


Figure 4.3: Energy spectrum of linear harmonic oscillator.

frequency ω .

The energy of a quantum oscillator is, as is well known, expressed by the relation

$$E_n = \left(n + \frac{1}{2}\right) \hbar \omega \quad (n = 0, 1, 2, \dots) \quad (4.15)$$

where ω is the oscillator's vibration frequency, and n the quantum number.

Figure 4.3 shows the energy spectrum of a linear harmonic oscillator. It consists of a set of discrete levels spaced at an interval of $\hbar \omega$.

Since $E_{n,m} = E_{n,o}$, the expression for the energy of the normal modes of a lattice should be (4.15) and the energy spectrum should coincide with that shown in Figure 4.3.

The minimum portion of energy that can be absorbed or emitted by the lattice in the process of thermal vibrations corresponds to the transition of the normal mode being excited from the given energy level to the adjacent level and is equal to

$$\varepsilon_{ph} = \hbar \omega. \quad (4.16)$$

This portion, or quantum, of energy of thermal vibrations of the lattice is termed a *phonon*.

The following analogy may help to clear up the point. The space inside a black body is filled with equilibrium thermal radiation. From the quantum mechanical point of view such radiation is treated as a gas made up of the light quanta, or photons, whose energy is $\varepsilon = \hbar \omega = h\nu$ and whose momentum is $p = \hbar \omega / c = h / \lambda$, where c is the velocity of light, and λ its wavelength.

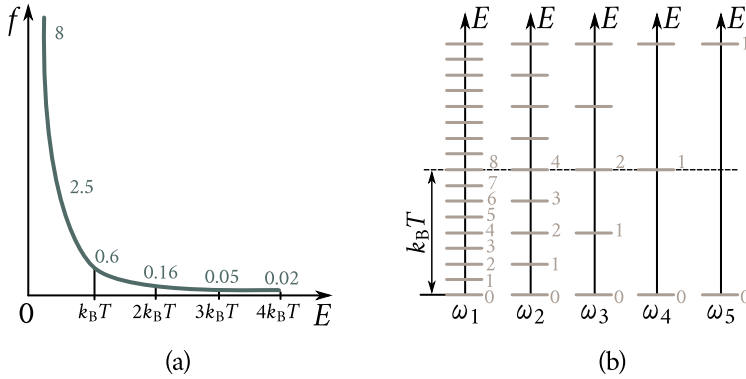


Figure 4.4: Energy spectrum of linear harmonic oscillator.

The field of elastic waves in the crystal may be treated similarly as a gas made up of quanta of the normal modes of the lattice, or of phonons having the energy $\varepsilon_{\text{ph}} = \hbar\omega = h\nu$ and momentum

$$p_{\text{ph}} = \frac{\hbar\omega}{v} = \frac{h}{\lambda} = \hbar q \quad (4.17)$$

where v is the velocity of sound, and λ the length of the elastic wave.

From this point of view a heated crystal may be likened to a box filled with phonon gas. The analogy may be extended.

Phonons are described by the same Bose-Einstein distribution function (3.54) as photons:

$$f(E) = \frac{1}{e^{\varepsilon_{\text{ph}}/(k_B T)} - 1} = \frac{1}{e^{(\hbar\omega)/(k_B T)} - 1}.$$

Depending on the intensity of excitation of the normal mode it can “emit” a definite number of phonons. Thus, if some normal mode was excited to the third level (Figure 4.3), its energy became $E_3 = (3 + 1/2)\hbar\omega$; this means that the particular normal mode has “generated” three identical phonons each with an energy of $\hbar\omega$.

Figure 4.4(a) shows the graph of the phonon energy (frequency) distribution function $f(E)$. We see that for a given temperature T , all normal modes in a lattice up to those with the energy $\hbar\omega \approx k_B T$ are excited; practically no quanta of higher frequencies with the energy $\hbar\omega > k_B T$ are excited. This is quite evident from Figure 4.4(b). Horizontal strokes here denote energy spectra of normal modes with the frequencies $\omega_1 = k_B T/(8\hbar)$, $\omega_2 = k_B T/(4\hbar)$, $\omega_3 = k_B T/(2\hbar)$, $\omega_4 = k_B T/\hbar$ and $\omega_5 = 2k_B T/\hbar$; the level corresponding to $k_B T$ is shown by a dotted line. It follows that for a given temperature T the mode with the frequency ω_1 is excited approximately to the 8th level. As was stated before, this means that this normal mode “generates” eight identical phonons with the energy $\hbar\omega_1 = k_B T$ each. The normal

mode with the frequency ω_2 is excited approximately to the 4th level, that with the frequency ω_3 to the second, and that with the frequency ω_4 (whose quantum of energy is $\hbar\omega_4 = k_B T$) to the first. At the same time, the vibration ω_5 is rarely excited at T because its excitation energy $\hbar\omega_5$ is too high. The excitation of still higher frequencies is a much more rare event. Therefore, we can say that approximately only the vibrations with frequencies not greater than ω corresponding to the energy $\hbar\omega \approx k_B T$ are excited in a solid at temperatures $T < \Theta$.

By definition, the distribution function $f(E)$ expresses the average number of phonons having the energy $\varepsilon_{ph} = \hbar\omega$. Therefore, to obtain the average energy of an excited normal mode, $\bar{E}_{n,m}$, of the frequency ω one has to multiply (3.54) by $\hbar\omega$:

$$\bar{E}_{n,m} = \frac{\hbar\omega}{e^{(\hbar\omega)/(k_B T)} - 1}. \quad (4.18)$$

§ 33. Heat capacity of solids

The thermal energy of a solid E_{lattice} is the sum of the energies of its normal modes. The number of normal modes per spectral interval $d\omega$ is $g(\omega) d\omega$ [see Eq. (4.8)]. Multiplying this number by the average energy $\bar{E}_{n,m}$ of the normal mode, (4.18), we obtain the total energy of the normal modes in the interval $d\omega$

$$dE_{\text{lattice}} = \bar{E}_{n,m} g(\omega) d\omega.$$

Integrating this expression over the entire spectrum of the normal modes, that is, from 0 to ω_D , we obtain the energy of the thermal vibrations of the lattice of a solid:

$$E_{\text{lattice}} = \int_0^{\omega_D} \bar{E}_{n,m} g(\omega) d\omega. \quad (4.19)$$

The heat capacity at constant volume of a solid, C_V , is the change in the thermal energy of a solid brought about by a one degree change in its temperature. To find it one should differentiate lattice with respect to T :

$$C_V = \frac{dE_{\text{lattice}}}{dT}. \quad (4.20)$$

The fundamental problem in the theory of heat capacity is the temperature dependence of C_V . Let us first consider it from a qualitative point of view for two temperature ranges: for the range of temperatures much below the Debye temperature

$$T \ll \Theta \quad (4.21)$$

which is termed the *low temperature range*, and for the range of temperatures above the Debye temperature

$$T > \Theta \quad (4.22)$$

the term for which is the *high temperature range*.

Low temperature range. In this range mainly the low frequency normal modes, with the energy quanta $\hbar\omega < k_B T$, are excited. The approximate value of the average energy of normal vibrations may in this case be calculated with the aid of the following method. Expand the denominator of expression (4.18) into a series leaving only two terms:

$$\bar{E}_{n.m} = \frac{\hbar\omega}{e^{(\hbar\omega)/(k_B T)} - 1} \approx \frac{\hbar\omega}{1 + \frac{\hbar\omega}{k_B T} + \dots - 1}$$

Hence, in the low temperature range the average energy of every normal mode increases in proportion to the absolute temperature T :

$$\bar{E}_{n.m} \propto T. \quad (4.23)$$

This law is due to the increase in the probability of excitation of every normal mode with the rise in temperature resulting in an increase in its average energy.

In addition to this, the rise in temperature in the low temperature range, causes new higher frequency normal modes to be excited. The approximate number of the latter, z , may be calculated with the aid of Eq. (4.8). If we assume that at a temperature T all normal modes up to the frequency $\omega \approx k_B T / \hbar$ are excited, we get

$$z = \int_0^{k_B T / \hbar} g(\omega) d\omega \approx \int_0^{k_B T / \hbar} \omega^2 d\omega \propto T^3.$$

It follows that with the rise in temperature the number of normal modes increases in proportion to the cube of the absolute temperature:

$$z \propto T^3. \quad (4.24)$$

To sum up, the crystal's energy in the low temperature range increases with the rise in temperature by means of two mechanisms: (1) the increase in the average energy of every normal mode, $\bar{E}_{n.m}$, due to the rise in the probability of its excitation, and (2) the increase in the number of the normal modes of the lattice.

The first mechanism is responsible for the increase in energy proportional to T and the second for the one proportional to T^3 .

Therefore, the total effect is an increase in the energy of the lattice proportional to T^4 :

$$\bar{E}_{\text{lattice}} \propto T^4 \quad (4.25)$$

and a rise in heat capacity proportional to T^3 :

$$C_V \propto T^3. \quad (4.26)$$

Formula (4.26) is the *Debye T^3 law*, which agrees well with experiment in the

low temperature range.

High temperature range. As has been already stated, all normal modes of a lattice are excited at the Debye temperature, so that a further rise in temperature cannot increase their number. Therefore, the variation in energy of a solid in the high temperature range may only be due to the rise in intensity of the normal modes, resulting in an increase in their average energy $\bar{E}_{n.m.}$. Since $\bar{E}_{n.m.} \propto T$, the variation of the energy of the body as a whole, too, should be proportional to T :

$$\bar{E}_{\text{lattice}} \propto T \quad (4.27)$$

and the heat capacity must be independent of T :

$$C_V = \frac{d\bar{E}_{\text{lattice}}}{dT} = \text{constant}. \quad (4.28)$$

Relation (4.28) is the expression of the *Dulong and Petit law*, which is quite well substantiated by experiment.

A rather wide range of temperatures, the so-called *medium temperature range*, lies between the high and low temperature ranges. In this medium temperature range, gradual transition from the Debye T^3 law to the Dulong and Petit law takes place. Calculations in this range are the most difficult.

To sum up, the following physical picture of the variation of the temperature dependence of energy and of heat capacity of a solid with the rise in temperature may be presented.

In the low temperature range ($T \ll \Theta$) the solid's energy increases with the rise in temperature firstly because of the increase in the probability of excitation of every normal mode, that is because of the increase in its average energy, $\bar{E}_{n.m.}$, which is proportional to T , and secondly because new normal modes are drawn into the process causing the body's energy to increase in proportion to T^3 . The energy of the lattice, as a whole, rises in proportion to T^4 and the specific heat in proportion to T^3 .

As the temperature approaches the Debye temperature the second mechanism gradually becomes inoperative and E_{lattice} becomes less dependent on T , causing a deviation from the Debye T^3 law.

At the Debye temperature, the entire spectrum of normal modes is excited. Therefore, the second mechanism has no part to play; only the first mechanism operates here causing the energy to rise in proportion to T and the heat capacity C_V to remain independent of T (the Dulong and Petit law).

The qualitative laws of the variation of C_V with T obtained from the study of physical processes in solids may be substantiated by more rigorous quantitative calculations. To this end let us turn to (4.19) and try to calculate the lattice energy as a function of temperature more accurately.

Substituting $g(\omega)$ from Eq. (4.14), and $\bar{E}_{n.m}$ from (4.18) into (4.19), we obtain

$$E_{\text{lattice}} = \frac{9N}{\omega_D^3} \int_0^{\omega_D} \frac{(\hbar\omega)^3}{e^{(\hbar\omega)/(k_B T)} - 1} d\omega. \quad (4.29)$$

One can introduce the dimensionless quantity $x = (\hbar\omega)/(k_B T)$ and rewrite (4.29) in the form

$$E_{\text{lattice}} = 9Nk_B \Theta \left(\frac{T}{\Theta} \right)^4 \int_0^{\Theta/T} \frac{x^3}{e^x - 1} dx \quad (4.30)$$

where Θ is the Debye temperature.

We will consider the high and the low temperature ranges separately.

Low temperature range ($T \ll \Theta$). In this range we can substitute infinity for the limit of integration in (4.30). Taking into account that

$$\frac{x^3}{e^x - 1} dx = \frac{\pi^4}{15}$$

we obtain

$$E_{\text{lattice}} = \frac{3\pi^4}{5} Nk_B \Theta \left(\frac{T}{\Theta} \right)^4 \propto T^4. \quad (4.31)$$

Differentiating Eq. (4.31) with respect to temperature, we obtain

$$C_V = \frac{12\pi^4}{5} Nk_B \left(\frac{T}{\Theta} \right)^3 \propto T^3. \quad (4.32)$$

We have arrived at the Debye T^3 law in accordance with which the heat capacity of a lattice varies in the low temperature range as the cube of the temperature.

High temperature range. For such temperatures the values of x are small and, hence, it is possible to drop all but the first two terms of the expansion $e^x = 1 + x + \dots$. Then,

$$E_{\text{lattice}} = 9Nk_B \Theta \left(\frac{T}{\Theta} \right)^4 \int_0^{\Theta/T} x^2 dx = 3Nk_B T \propto T. \quad (4.33)$$

The heat capacity of the crystal is

$$C_V = \frac{dE_{\text{lattice}}}{dT} = 3Nk_B = \text{constant}. \quad (4.34)$$

For a mole of a monatomic substance $N = N_A = 6.023 \times 10^{23} \text{ mol}^{-1}$ (Avogadro's number), $N_A k_B = R \approx 8.31 \text{ J mol}^{-1} \text{ K}^{-1}$ (the gas constant) and

$$C_V \approx 3R \approx 25 \text{ J mol}^{-1} \text{ K}^{-1}. \quad (4.35)$$

Formula (4.35) expresses the Dulong and Petit law, which was formulated by them in 1819.

The solid line in Figure 4.5 shows the theoretical temperature dependence of the heat capacity of solids, the points being experimental values for silver, diamond,

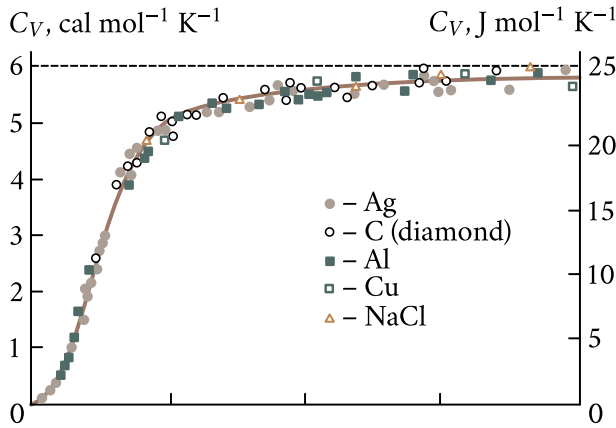


Figure 4.5: Temperature dependence of heat capacity of solids. The solid line is the theoretical Debye curve.

aluminium, copper, and rock salt. The agreement between theory and experiment is quite satisfactory not only from the qualitative but also from the quantitative point of view.

Knowing the temperature dependence of the energy of a lattice, we can easily find at least the qualitative dependence of the concentration of the phonon gas on temperature, that is, the number of phonons n_{ph} excited in a unit of volume of the crystal.

The concentration of the phonon gas in the low temperature range, in which $E_{\text{lattice}} \propto T^4$ and the phonon energy $\hbar\omega \approx k_B T \propto T$, must be proportional to T^3 :

$$n_{\text{ph}} \propto T^3. \quad (\text{low temperature range, } T \ll \Theta) \quad (4.36)$$

In the high temperature range, where $E_{\text{lattice}} \propto T$ and the phonon energy attains the maximum value of $\hbar\omega_D \approx k_B T$ independent of T , the concentration of the phonon gas should be proportional to T :

$$n_{\text{ph}} \propto T. \quad (\text{high temperature range, } T > \Theta) \quad (4.37)$$

To be exact, to calculate the concentration of the phonon gas one must know the average energy of the phonons $\bar{\epsilon}_{\text{ph}}$ both in the low and the high temperature ranges since the lattice energy is equal to the product of the average phonon energy and their concentration. The calculation of $\bar{\epsilon}_{\text{ph}}$ yields

$$\bar{\epsilon}_{\text{ph}} = \frac{\pi^2 k_B T}{5} \quad (4.38)$$

for the low temperature range, and

$$\bar{\epsilon}_{\text{ph}} = \frac{2k_B \Theta}{3} \quad (4.39)$$

for the high temperature range.

This justifies the temperature dependence of n_{ph} expressed by formulae (4.36) and (4.37) and obtained by qualitative methods.

§ 34. Heat capacity of electron gas

The metals, in addition to ions, which constitute the lattice and vibrate around their equilibrium positions, contain also free electrons, the number of which per unit volume is approximately the same as that of the ions. For this reason, the specific heat of a metal should be the sum of the heats capacity of the lattice C_{lattice} calculated in the previous paragraph and of the electron gas C_e :

$$C_V = C_{\text{lattice}} + C_e.$$

If the electron gas was a normal classical (nondegenerate) gas, every electron would have an average energy $3k_B T/2$ and the energy of the electron gas per mole of the metal would be

$$E_e^{(\text{cl})} = \frac{3}{2} N_A k_B T = \frac{3}{2} RT$$

and its heat capacity would be

$$C_e^{(\text{cl})} = \frac{3}{2} N_A k_B = \frac{3}{2} R. \quad (4.40)$$

The total heat capacity of the metal in the high temperature range would in this case be

$$C_V = C_{\text{lattice}} + C_e = \frac{9}{2} R \approx 37 \text{ J mol}^{-1} \text{ K}^{-1}.$$

Actually, the heat capacity of metals, as well as that of dielectrics, in the high temperature range, where the Dulong and Petit law is valid, is $C_V \approx 25 \text{ J mol}^{-1} \text{ K}^{-1}$, a proof that the contribution of the electron gas is negligible.

This situation incomprehensible from the point of view of classical physics found its natural explanation in quantum theory.

Indeed, as was demonstrated in Chapter 3, the electron gas in metals is a degenerate gas described by the Fermi-Dirac quantum statistics. As the temperature is raised, not all the electrons are thermally excited; only a negligible fraction of them, ΔN , occupying states close to the Fermi level (see Figure 3.6) are thermally excited. The number of such electrons is approximately expressed by the relation (3.43):

$$\Delta N \approx N \frac{k_B T}{2E_F}$$

where E_F is the Fermi energy. For copper at $T \approx 300 \text{ K}$ and $E_F \approx 7 \text{ eV}$, we have $\Delta N/N \approx 0.002$, that is, less than one percent.

Every thermally excited electron absorbs an energy of the order of $k_B T$ just as a particle of a normal gas does. The energy absorbed by the electron gas as a whole is the product of $k_B T$ and the number of thermally excited electrons ΔN :

$$E_e \approx k_B T \Delta N \approx N k_B T \frac{k_B T}{2E_F}. \quad (4.41)$$

The heat capacity of the electron gas is

$$C_e = \frac{dE_e}{dT} \approx N k_B \frac{k_B T}{E_F}. \quad (4.42)$$

A more accurate calculation yields the following expression:

$$C_e \approx \pi^2 N k_B \frac{k_B T}{2E_F}. \quad (4.43)$$

Comparing Eq. (4.40) with (4.43), we obtain

$$\frac{C_e}{C_e^{(cl)}} \approx \pi \frac{k_B T}{E_F}. \quad (4.44)$$

It follows from Eq. (4.44) that the ratio of the heat capacity of a degenerate electron gas to that of a nondegenerate monatomic gas is approximately equal to the ratio of $k_B T$ to E_F . At normal temperatures, the ratio $\pi k_B T / E_F \lesssim 1\%$. Therefore,

$$C_e \lesssim 0.01 C_e^{(cl)}. \quad (4.44')$$

Hence, because of the degeneracy of the electron gas in metals even in the high temperature range only a small portion of the free electrons (usually less than one percent) is thermally excited; the rest do not absorb heat. This is why the heat capacity of the electron gas is negligible as compared to that of the lattice, and the heat capacity of a metal as a whole is practically equal to the latter.

The situation is different in the low temperature range close to absolute zero. Here, the heat capacity decreases in proportion to T^3 , with the fall in temperature and close to absolute zero may prove to be so small that the contribution of the heat capacity of the electron gas, C_e , which decreases much more slowly than C_{lattice} ($C_e \propto T$), may become predominant. Figure 4.6 shows the temperature dependence of the lattice and electron components of specific heat of an alloy (20% vanadium and 80% chromium) whose Debye temperature is $\Theta = 500$ K. It may be seen from Figure 4.6, that close to absolute zero the heat capacity of the electron gas is much greater than that of the lattice ($C_{\text{lattice}} < C_e$), the sign of the inequality remaining the same up to $T \approx 8.5$ K. At $T > 8.5$ K, the sign is reversed, the inequality becoming stronger with the rise in T . Already at $T \approx 25$ K, the heat capacity of the alloy is mainly due to that of its lattice (at $T = 25$ K the heat capacity is $C_{\text{lattice}} \approx 10C_e$).

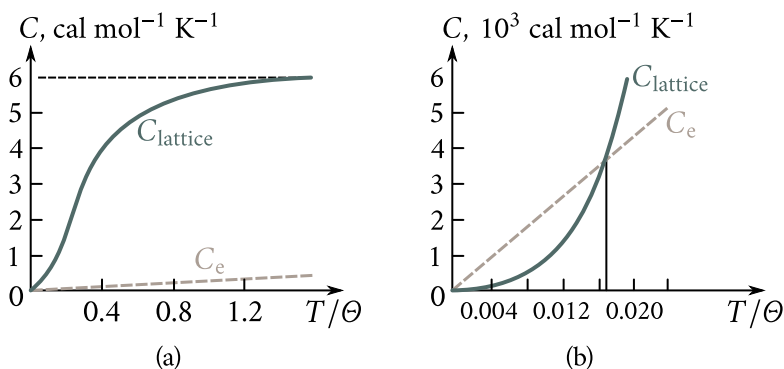


Figure 4.6: Temperature dependence of lattice and of an alloy consisting of 20% vanadium and 80% chromium.

§ 35. Thermal expansion of solids

To explain the elastic properties of solids, in Chapter 2 we have introduced the harmonic approximation according to which the elastic force acting on a particle displaced from its equilibrium position is proportional to the displacement [see Eq. (2.3)] and its potential energy is proportional to the square of the displacement [see Eq. (2.2)]; this fact, is represented by a parabola (the dotted line in Figure 2.1).

The immediate result of the harmonic approximation was the Hooke's law, which describes the elastic deformation of solids. The same approximation was used in Chapter 3 as a basis for calculating the thermal vibrations of a lattice and constructing the theory of heat capacity of a lattice, which is in fair agreement with experiment.

However, the harmonic approximation was unable to explain such well known phenomena as, for instance, the thermal expansion of solids, their heat conductivity, etc.

Indeed, let us turn to the dependence of the potential energy of interaction of the particles of a solid on the distance between them (Figure 4.7). At absolute zero, the particles occupy positions r_0 corresponding to the minimum interaction energy U_0 (at the bottom of the potential trough (well) "abc"). Those distances determine the dimensions of the body at absolute zero. As the temperature rises the particles begin to vibrate around their equilibrium positions 0. For the sake of simplicity, let us assume that particle 1 is fixed and only particle 2 is vibrating. The kinetic energy of the vibrating particle is at its maximum E_k when the particle passes its equilibrium position 0. In Figure 4.7 the energy E_k is measured upwards from the bottom of the potential trough. When particle 2 moves to the left, its kinetic

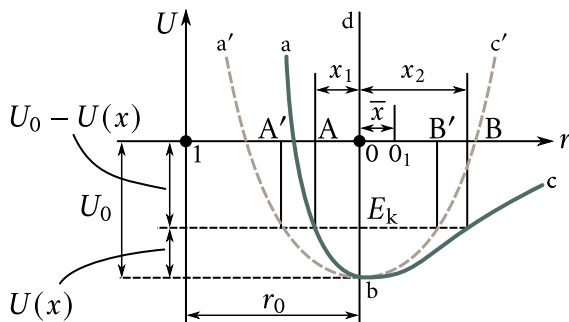


Figure 4.7: The origin of thermal expansion of solids (explanation in text).

energy is used to overcome the repulsive forces acting from particle 1 and is transformed into the potential energy of the particles' interaction. The displacement to the left stops when all the kinetic energy E_k is transformed into the potential energy. In the extreme left position of particle 2 displaced by the distance x_1 , the potential energy's increment is $U(x_1) = E_k$ and its value $-[U_0 - U(x_1)]$. When particle 2 moves to the right, its kinetic energy is spent to overcome the forces of attraction to particle 1 and is, as in the previous case, transformed into the potential energy of the particles' interaction. At point B displaced from the equilibrium position by the distance x_2 , the entire kinetic energy E_k is transformed into the potential energy, the latter increasing by $U(x_2) = E_k$ to become $-[U_0 - U(x_2)]$.

If the vibrations of particle 2 were purely harmonic, the force $f(x)$ caused by its displacement from the equilibrium position by a distance x would be strictly proportional to this displacement and directed towards the equilibrium position:

$$f = -\beta x. \quad (4.45)$$

The change in the particle's potential energy $U(x)$ would, in this case, be described by the parabola $a'bc'$ (Figure 4.7) whose equation is

$$U(x) = \frac{\beta x^2}{2}. \quad (4.46)$$

This parabola is symmetric about the straight line bd passing parallel to the axis of ordinates at a distance of r_0 from it. Therefore, the displacements x_1 and x_2 would be equal in magnitude and the centre of AB would coincide with the equilibrium position 0. In this case, heating a body would not bring about its expansion, for a rise in temperature would result only in an increase in the particles' amplitude of vibrations, the average distances between them remaining unchanged.

Actually, the potential curve abc is, as may be seen from Figure 4.7, not symmetric about the straight line bd , its left branch ba rising much more steeply than the right branch be . This means that the vibrations of the particles in a solid are

anharmonic (not harmonic). To account for the asymmetry of the potential curve an additional term $-gx^3/3$ expressing this asymmetry (g is a proportionality factor) should be introduced. Then, Eqs. (4.45) and (4.46) will assume the following form:

$$U(x) = \frac{\beta x^2}{2} - \frac{gx^3}{3}, \quad (4.45')$$

$$f(x) = -\frac{\partial U}{\partial x} = -\beta x + gx^2. \quad (4.46')$$

When the particle 2 is displaced to the right ($x > 0$), the term $gx^3/3$ is subtracted from $\beta x^2/2$ and the slope of the branch be is less than that of the branch be'; when the displacement is to the left ($x < 0$), the term $gx^3/3$ is added to $\beta x^2/2$ and the slope of ba is greater than that of ba'.

Because of the asymmetric nature of the potential curve, the right and left displacements of particle 2 turn out to be different, the former being greater than the latter (Figure 4.7). As a result, the central position of particle 2 (point 0₁) no longer coincides with its equilibrium position 0 but is displaced to the right. This corresponds to an increase in the average distance between the particles by x .

Hence, heating a body should result in an increase in the average distances between particles and the body should expand. The cause of this is the anharmonic nature of the vibrations of particles making up the solid due to the asymmetry of the dependence of the particles' interaction energy on the distance between them.

Let us estimate the value of the thermal expansion coefficient.

The average value of the force caused by the displacement of particle 2 from its equilibrium position is

$$\bar{f} = -\beta \bar{x} + g \bar{x}^2.$$

When the particle vibrates freely, $\bar{f} = 0$; therefore, $g \bar{x}^2 = \beta \bar{x}$. Hence, we obtain

$$\bar{x} = \frac{g \bar{x}^2}{\beta}. \quad (4.47)$$

The expression for the potential energy of a vibrating particle correct to the second order of magnitude is (4.45) and its average value is $\bar{U}(x) \approx \beta \bar{x}^2/2$. Hence,

$$\bar{x}^2 \approx \frac{2\bar{U}(x)}{\beta}.$$

Substituting this into Eq. (4.47), we obtain

$$\bar{x} = \frac{2g\bar{U}(x)}{\beta^2}.$$

In addition to potential energy $U(x)$, a vibrating particle has kinetic energy E_k such that $\bar{U}(x) = \bar{E}_k$. The total average energy of the particle is $\bar{E} = \bar{U}(x) + \bar{E}_k =$

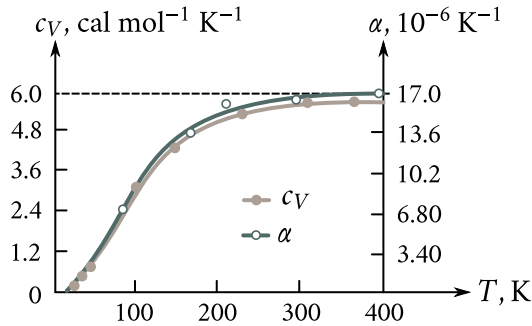


Figure 4.8: Temperature dependence of linear expansion coefficient α and of heat capacity c_V of copper.

$2\bar{U}(x)$. This fact makes it possible to rewrite the expression for x in the following form:

$$\bar{x} = \frac{g\bar{E}}{\beta^2}.$$

The relative linear expansion, that is, the ratio of the variation of the average distance between the particles, x , to the equilibrium distance between them, r_0 , is equal to

$$\frac{\bar{x}}{r_0} = \frac{g}{\beta^2 r_0} \bar{E}$$

and the *linear expansion coefficient* is

$$\alpha = \frac{1}{r_0} \frac{d\bar{x}}{dT} = \frac{g}{\beta^2 r_0} \frac{d\bar{E}}{dT} = \chi c_V \quad (4.48)$$

where

$$\chi = \frac{g}{\beta^2 r_0} \quad (4.49)$$

and c_V is heat capacity per particle.

Thus, the thermal expansion coefficient proves to be proportional to temperature. Figure 4.8 shows the temperature dependence of c_V and α . It can be easily seen that both are interrelated.

In the high temperature range the energy of a particle engaged in linear vibrations is $k_B T$ and its heat capacity is $c_V = k_B$. Therefore, the thermal expansion coefficient of a linear atomic chain will be

$$\alpha = \chi c_V = \frac{g k_B}{\beta^2 r_0}.$$

Substituting the values for g , k_B , β , and r_0 for various solids, we obtain a value of the order of 10^{-4} to 10^{-5} for α , which is in fair agreement with experiment.

Experiment also supports the conclusion that in the high temperature range α is practically independent of temperature (Figure 4.8).

In the low temperature range α behaves in a way similar to that of c_V : it decreases with the fall in temperature and tends to zero as absolute zero is approached.

In conclusion, we would like to remark that for metals a formula similar to (4.48) was first proposed by E. Grüneisen in the form

$$\alpha = \frac{\gamma\kappa}{3V}c_V \quad (4.50)$$

where κ is the metal's compressibility, V the atomic volume, and γ the Grüneisen constant, ranging from 1.5 to 2.5, depending on the metal.

§ 36. Heat conductivity of solids

Heat conductivity of dielectrics (lattice heat conductivity). The second effect caused by the anharmonic nature of atomic vibrations is the *thermal resistance* of solids. There could be no such resistance should the atomic vibrations be strictly harmonic propagating through the lattice in the form of noninteracting elastic waves. In the absence of interaction, the waves would be able to travel without scattering, that is, without meeting any resistance, like light in vacuum.

If we were to set up a temperature difference in such a crystal, the atoms of the hot end vibrating with large amplitudes would transmit their energy to the neighbouring atoms and the front of the heat wave would travel along the crystal with the velocity of sound. Because the wave would meet no resistance there would be a considerable heat flux even for an infinitesimally small temperature difference and the heat conductivity of the crystal would be infinitely great.

The nature of atomic vibrations in real crystals at temperatures not too low is anharmonic, as indicated by the second term in Eq. (4.45'). The anharmonicity destroys the independence of the normal modes of the lattice and causes them to interact, exchanging energy and changing the direction of their propagation (through mutual scattering). It is just those processes of interaction between the elastic waves that make possible the transfer of energy from the modes of one frequency to the modes of another and the establishment of thermal equilibrium in the crystal.

The process of mutual scattering of normal modes is conveniently described in terms of phonons, the thermally excited crystal being regarded as a box containing phonons. In the harmonic approximation, in which the normal modes are presumed to be independent, the phonons make up an ideal gas (a gas of noninteracting phonons). The transition to the anharmonic modes is equivalent to the introduction of an interaction between phonons, which may result in the splitting

of a phonon into two or more phonons and in the formation of a phonon from two other phonons. Such processes are termed *phonon-phonon scattering*. Their probability, as is the case of all scattering processes, is characterized by the *effective scattering cross-section* σ_{ph} . Should the phonon be, from the point of view of the scattering processes, represented by a sphere of the radius r_{ph} then, $\sigma_{\text{ph}} = \pi r_{\text{ph}}^2$. The phonon-phonon scattering may take place only if the phonons approach to within a distance at which their effective cross sections begin to overlap. Since the scattering is due to the anharmonicity of the atomic vibrations, numerically described by the coefficient of anharmonicity g , it would be natural to assume that the phonon effective cross section radius is proportional to g and $\sigma_{\text{ph}} \propto g^2$.

Knowing the effective scattering cross section σ_{ph} , we can calculate the mean free path λ_{ph} of the phonons, that is, the average distance the phonons travel between two consecutive scattering acts. Calculations show that

$$\lambda_{\text{ph}} = \frac{1}{n_{\text{ph}}\sigma_{\text{ph}}} \propto \frac{1}{n_{\text{ph}}g^2} \quad (4.51)$$

where n_{ph} is the phonon concentration.

In the kinetic theory of gases it is proved that for gases the *heat conductivity* is

$$\mathcal{K} = \frac{\lambda \nu C_V}{3} \quad (4.52)$$

where λ is the mean free path of the molecules, ν their thermal velocity, and C_V the heat capacity of the gas.

Let us apply this formula to the phonon gas substituting for C_V the specific heat of the crystal (the phonon gas), for $\lambda = \lambda_{\text{ph}}$ the mean free path of the phonons, and for ν the velocity of sound (the phonon velocity). We obtain the following expression for the lattice heat conductivity:

$$\mathcal{K}_{\text{lattice}} = \frac{\nu \lambda_{\text{ph}} C_V}{3}. \quad (4.53)$$

Substituting λ_{ph} into (4.53) from (4.51), we obtain

$$\mathcal{K}_{\text{lattice}} \propto \frac{\nu C_V}{n_{\text{ph}} g^2}. \quad (4.54)$$

In the high temperature range, in accordance with (4.37), $n_{\text{ph}} \propto T$; hence,

$$\mathcal{K}_{\text{lattice}} \propto \frac{\nu C_V}{T g^2}. \quad (4.55)$$

Since C_V in this range is practically independent of T , the lattice thermal conductivity should be inversely proportional to the absolute temperature, which is in qualitative agreement with experiment. Equation (4.55) also includes the anharmonicity factor g and the sound velocity ν , which depend substantially on the rigidity of the bonds between the particles of the solid. Bonds of lesser rigidity cor-

respond to lower ν 's and to greater g 's, since the weakening of the bonds makes for greater thermal vibration amplitudes (for a specified temperature) and for greater anharmonicity. Both those factors should, according to (4.55), bring about a reduction in $\mathcal{K}_{\text{lattice}}$. This conclusion is supported by experiment. Table 4.2 presents the values of sublimation heat Q_s , which is a measure of bonding energy, and of the lattice heat conductivity lattice for some covalent crystals with the diamond lattice: diamond, silicon, and germanium.

We see that the decrease in the bond energy from the value of diamond to that of silicon and, finally, germanium is accompanied by a noticeable decrease in the lattice heat conductivity. A more detailed analysis shows that $\mathcal{K}_{\text{lattice}}$ is also strongly dependent of the mass M of the particles, being less for greater M 's. This, to a large extent, accounts for the fact that the lattice heat conductivity of the lighter elements occupying the upper part of the Mendeleev periodic table (B, C, Si) is of the order of tens or even hundreds of watts per metre per kelvin, the corresponding values for the elements of the middle part of the Mendeleev table being several watts per metre per kelvin, and that for the heavier elements occupying the lower part of the periodic table even to several tenths of a watt per metre per kelvin.

A striking feature is that the lattice heat conductivity of crystals with light particles and rigid bonds may be very high. Thus, at room temperature $\mathcal{K}_{\text{lattice}}$ of diamond is greater than the total heat conductivity of the best heat conductive metal, silver: $\mathcal{K}_{\text{Ag}} = 407 \text{ W m}^{-1} \text{ K}^{-1}$.

At temperatures below the Debye temperature there is a sharp drop in phonon concentration with a fall in temperature leading to a sharp increase in their mean free path, so that at $T \ll \Theta/20$ it becomes comparable with the dimensions of the crystal. Since the crystal surface usually is a poor reflector of phonons, any further decrease in temperature does not lead to an increase in λ_{ph} , for the latter is determined by the crystal's dimensions only. The temperature dependence of the lattice heat conductivity within this range of temperatures is determined by the temperature dependence of the heat capacity C_V . Since $C_V \propto T^3$ in the low

Table 4.2

Substance	$Q_s, (10^5 \text{ J mol}^{-1})$	$\mathcal{K}_{\text{lattice}}, (\text{W m}^{-1} \text{ K}^{-1})$
Diamond	71.23	550.0
Silicon	46.09	137.0
Germanium	37.00	54.00

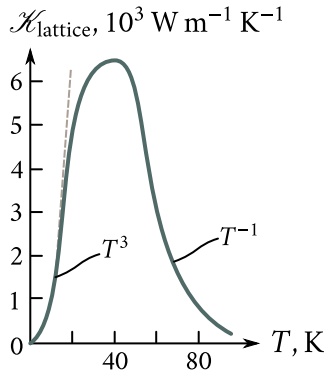


Figure 4.9: Temperature dependence of heat conductivity of synthetic sapphire.

temperature range, $\mathcal{K}_{\text{lattice}}$ too should be proportional to T^3 :

$$\mathcal{K}_{\text{lattice}} \propto T^3. \quad (4.56)$$

This result is also in qualitative agreement with experiment. Figure 4.9 shows the temperature dependence of thermal conductivity of synthetic sapphire. In the low temperature range $\mathcal{K}_{\text{lattice}}$ is indeed approximately proportional to T^3 .

As the temperature rises so does the concentration of phonons n_{ph} , and this should *per se* cause $\mathcal{K}_{\text{lattice}}$ to rise. However, an increase in n_{ph} is accompanied by an increase in the phonon-phonon scattering and a consequent decrease in the mean free path of phonons λ_{ph} , which should result in a decrease in $\mathcal{K}_{\text{lattice}}$. For low n_{ph} , the first factor should be the predominant one and $\mathcal{K}_{\text{lattice}}$ should rise with T . However, starting with a definite concentration n_{ph} , the second factor should assume primary importance and $\mathcal{K}_{\text{lattice}}$ after passing through a maximum should fall with the rise in T . This decrease in the high temperature range is approximately of the $1/T$ type.

Amorphous dielectrics in which the size of regions with a regular structure is of the order of interatomic distances should present a similar picture. Phonon scattering on the boundaries of such regions should be the dominant factor at all T 's and, therefore, λ_{ph} should be independent of T . Because of that the heat conductivity of such dielectrics should be proportional to T^3 in the low temperature range and independent of T in the high temperature range. This is just what is observed in experiment.

However, at present the theory is unable to predict not only the exact values of lattice heat conductivity, $\mathcal{K}_{\text{lattice}}$, but even its order of magnitude.

Heat conductivity of metals. In metals, in contrast to dielectrics, heat is transported not only by phonons but by electrons as well. Therefore, generally

the heat conductivity of metals is the sum of the lattice heat conductivity $\mathcal{K}_{\text{lattice}}$ (conductivity due to phonons) and the heat conductivity \mathcal{K}_e of the free electrons:

$$\mathcal{K} = \mathcal{K}_{\text{lattice}} + \mathcal{K}_e.$$

The heat conductivity of the electron gas, can be calculated with the aid of Eq. (4.52). Substituting into this formula the heat capacity of the electron gas, C_e , the electron velocity, v_F , and their mean free path, λ_e , we obtain

$$\mathcal{K}_e = \frac{1}{3} C_e v_F \lambda_e. \quad (4.57)$$

Substituting C_e from Eq. (4.43) into Eq. (4.57), we have

$$\mathcal{K}_e = \frac{\pi^2}{3} \frac{N k_B^2}{m_n v_F} \lambda_e T. \quad (4.58)$$

Let us make a qualitative estimate of the temperature dependence of heat conductivity of pure metals.

High temperature range. Practically, of all the quantities contained in the right-hand side of Eq. (4.58), only λ_e depends on T . For pure metals at temperatures not too low, λ_e is determined by electron-phonon scattering and, all other conditions being equal, is inversely proportional to phonon concentration: $\lambda_e \propto 1/n_{\text{ph}}$. In the high temperature range, according to Eq. (4.37), $n_{\text{ph}} \propto T$. Substituting this into Eq. (4.58), we obtain

$$\mathcal{K}_e = \text{constant}. \quad (4.59)$$

Hence, the heat conductivity of pure metals in the high temperature range should be independent of temperature. This is an experimental fact. Figure 4.10 shows the experimental curve depicting the temperature dependence \mathcal{K} for copper. It follows that above 80-100 K the heat conductivity of copper is practically independent of temperature.

Low temperature range. The phonon concentration in this range is $n_{\text{ph}} \propto T$, therefore, $\lambda_e \propto 1/T^3$. Substituting into Eq. (4.58), we obtain

$$\mathcal{K}_e \propto T^{-2}. \quad (4.60)$$

Consequently, in the low temperature range where the Debye T^3 law is true, the heat conductivity of metals should be inversely proportional to the square of the absolute temperature. This conclusion too is in general supported by experiment (Figure 4.10).

Very low temperature range. Close to absolute zero the phonon concentration in a metal becomes so small that the main part in electron scattering processes is taken over by the impurity atoms, which are always present in a metal no matter how pure it is. In this case, the mean free path of an electron, $\lambda_e \propto 1/N_i$ (N_i is the concentration of impurity atoms), is no longer dependent on temperature and the

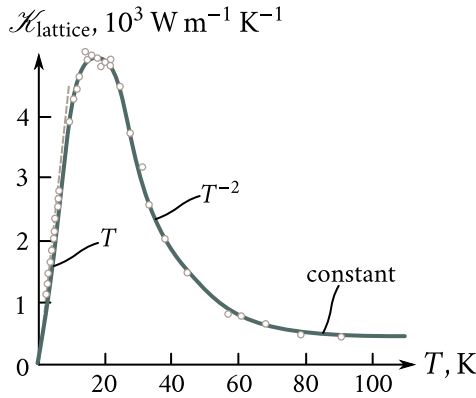


Figure 4.10: Temperature dependence of heat conductivity of copper.

heat conductivity of a metal becomes proportional to T :

$$\mathcal{K}_e \propto T \tag{4.61}$$

which is an experimental fact.

Let us estimate the magnitude of \mathcal{K}_e for metals, making use of Eq. (4.57). For typical metals, we have $C_e \approx 0.01 C_V \approx 3 \times 10^4 \text{ J m}^{-3} \text{ K}^{-1}$, $v_F \approx 10^6 \text{ m s}^{-1}$, and $\lambda_e \approx 10^{-8} \text{ m}$. Substituting into Eq. (4.57), we obtain $\mathcal{K}_e \approx 10^2 \text{ W m}^{-1} \text{ K}^{-1}$. Thus, \mathcal{K}_e for metals may be as high as hundreds of watts per metre per kelvin. This is substantiated by experiment. Table 4.3 shows the room temperature heat conductivities for some typical metals and for one alloy, constantan, which consists of 60% copper and 40% nickel.

It follows that for pure metals \mathcal{K} can indeed be as high as hundreds of watts per metre per kelvin.

Let us also estimate the contribution of the lattice heat conductivity of a metal,

Table 4.3

Metal	$\mathcal{K}, (\text{W m}^{-1} \text{ K}^{-1})$	Metal	$\mathcal{K}, (\text{W m}^{-1} \text{ K}^{-1})$
Silver	403	Aluminium	210
Copper	384	Nickel	60.0
Gold	296	Constantan	23.0

making use of Eqs. (4.53) and (4.57):

$$\frac{\mathcal{K}_{\text{lattice}}}{\mathcal{K}_e} = \frac{C_V v \lambda_{\text{ph}}}{C_e v_F \lambda_e}$$

v being the phonon (sound) velocity. For pure metals, we have $C_e/C_V \approx 0.01$, $v = 5 \times 10^3 \text{ m s}^{-1}$, $\lambda_{\text{ph}} \approx 10^{-9} \text{ metre}$, $v_F \approx 10^6 \text{ m s}^{-1}$, and $\lambda_e \approx 10^{-8} \text{ m}$. Hence, $\mathcal{K}_{\text{lattice}}/\mathcal{K}_e \approx 5 \times 10^{-2}$.

It follows then that the heat conductivity of typical metals is almost entirely due to the heat conductivity of their electron gas, the contribution of lattice heat conductivity being a few percent. This picture may, however, totally change when we go over to metallic alloys, in which impurity scattering is the principal electron scattering mechanism. The electron mean free path for which such scattering is responsible, λ_e , is inversely proportional to the impurity concentration N_i ($\lambda_e \propto 1/N_i$) and for high N_i 's may become comparable with the phonon mean free path ($\lambda_{\text{ph}} \propto \lambda_e$). Naturally, in such a case the electron contribution to heat conductivity may become of the same order of magnitude as that of the phonon contribution: $\mathcal{K}_e \approx \mathcal{K}_{\text{lattice}}$. This too is supported by experiment. Table 4.3 gives the heat conductivity of constantan. It is much less than that of nickel or copper. This proves the fact that electron scattering in constantan is mainly due to lattice defects caused by impurity atoms. We also note that \mathcal{K}_e and $\mathcal{K}_{\text{lattice}}$ measured by R. Berman in constantan proved to be of the same order of magnitude.

Chapter 5

The Band Theory of Solids

The theory of free electrons was the first successful attempt to explain the electric and magnetic properties of solids (primarily of metals). It was based on the assumption that metals contain free electrons capable of moving around the metal like gas molecules in a vessel. The theory of free electrons was successful in explaining such phenomena as the electric and the heat conductivities, thermionic emission, thermoelectric and galvanomagnetic effects, etc. However, this theory proved incapable of dealing with such properties of solids as are determined by their internal structure. It could not even explain why some bodies are conductors and some insulators.

The next stage in the progress of the electron theory has been the band theory of solids, which is outlined in this chapter.

§ 37. Electron energy levels of a free atom

