Co-funded by the
Erasmus+ Programme
of the European Union

# FUNCTIONAL NANOMATERIALS

## (Electronic book)

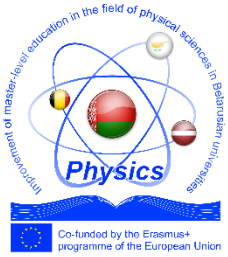
Scientific advisers of the project: Leonids Ribickis, Nadezhda Kunicina (RTU, Latvia)

Project Manager: Anatoly Zabashta (RTU, Latvia)

Editor: Alexander Fedotov (BSU, Belarus)

Riga, 2019

Co-funded by the
Erasmus+ Programme
of the European Union




This document was prepared with the financial support of the European Union. The authors of the Belarusian State University, the Grodno State University, the Gomel State University, the Belarusian State Technological University and the Catholic University of Leuven (Ostend, Belgium) are responsible for the content of this document and can not be considered as official position of the European Union.

The book was developed within the framework of the project "Improvement of Master's Education in the Physical Sciences in Belarusian Universities" (acronym PHYSICS), Nr. 561525-EPP-1-2015-1-LV-EPPKA2-CBHE-JP.

The project was approved by the European Commission within the framework the EAC/A04/2014 project proposal competition by the ERASMUS + Program.
Sub-programme: Joint projects
Activity: Curriculum reform
Final result: 2.1. Developed and translated to teaching language master-level study programs and courses for specialities «Functional nanomaterials» and «Photonics».

This project is financed with the support of the European Commission Programme. This publication reflects only the views of the authors, and the Commission can not be held responsible for any use that may be made based on the information contained therein.

**Athors: A. Fedotov, V. Odzhaev, A. Mazanik, M. Tivanov, N. Poklonski, V. Ksenevich, N. Gorbachuk, J. Fedotova, J. Kasiuk, M. Lukashevich, N. Strekal, N.R. Prokopchuk, Zh.S. Shashok, K.V. Vishnevskii, N.R. Prokopchuk, A.L. Shutova, A.V. Rogachev, A.A. Rogachev, M. Yarmolenko, D.Pilipcov, N. Fedosenko, D. Kovalenko, V. Gaishun, A. Semchenko, J. Peuteman**

**This document is a translation of the electronic book “Functional nanomaterials” prepared under the editorship of Professor Alexander Fedotov. (BSU)**

ISBN 978-9934-22-224-5 (pdf)

Co-funded by the Erasmus+ Programme of the European Union

## Content

## Executive summary

Recently, more and more attention has been attracted to research in the field of functional nanomaterials and nanotechnologies, which work with objects whose sizes are at least in one dimension close to nanometers. On the recommendation of the International Union of Pure and Applied Chemistry (IUPAC) for "nanocriterium" we take the value of 100 nm. With this criterion, the properties of substances in the nanostructured state are often very different from those of bulk phases. That is why sometimes we speak about a nanostructured state as a speciefic state of matter.

The use of nanostructured materials is gaining ambitious prospects not only because of the unique possibility of devices miniaturization previously unseen, but also because of their fundamental difference from conventional materials and the acquisition of new properties. The latter includes the occurrence of quantum and tunnel effects, and also very high role of surface area, excess surface energy, and the reactivity of nanoparticles.

The most important parameters of nanosystems are becoming size (diameter), dimensionality (0D, 1D, 2D), ordering and functionality. Taking into account all these characteristics results in a new generation of materials that have the highest innovative potential and determine all further progress in nanotechnology

The book consists of 20 sections (chapters) devoted to the most actual problems of nanotechnology and nanostructured materials. It can be used to develop lecture courses for master-level students with a 2-year cycle of study by the specialties "1-31 81 03 Functional nanomaterials" and "1-31 81 02 Photonics".

The content of the book will contribute to the competitiveness of higher education in the Republic of Belarus in the field of applied physics, photonics and functional nanomaterials in the world educational space, as well as further development of international cooperation of Belarusian universities with educational and scientific institutions of foreign countries. Participation of the four leading Belarusian universities in this project will contribute to achieving one of the goals for the universities of the Republic of Belarus - ensuring a high level of preparation and stable quality of the educational process in the field of nanotechnology to meet the requirements of external and internal consumers of the specialists.

## Introduction

The electronic book «Functional Nanomaterials» is written to provide training for master-level students in Belarusian universities in the framework of the 2-year master's tarining program by the specialty «1-31 81 03 Functional Nanomaterials».

The book was developed within the framework of the grant 561525-EPP-1-2015-1-LV-EPPKA2-CBHE-JP "Improvement of Master's Education in the Physical Sciences in Belarusian Universities" of the international educational program ERASMUS + of the European Union

Taking into account the presence in the world of a large number of textbooks and handbooks on nanotechnology and nanomaterials, the book is not built as a textbook, but rather as a monograph consisting of 20 Chapters.

The Chapters of the book were written by leading specialists of four Belarusian universities (Belarusian State University, F. Skorina Homel State University, I Kupala Grodno State University and the Belarusian State Technological University), and the Catholic University of Bruges-Ostend (Belgium), for a number of important directions of functional nanomaterials. It is within the framework of these problems that the themes of lecture courses taught for students of Belarusian universities are formulated, and the themes of master's theses will be offered starting from 2017-2018 academic year.

The head of the team of the book authors is professor of the BSU physics faculty, doctor of phys.-math. sciences Alexander Fedotov as one of the leading specialists of Belarus in the field of nanostructured materials.

## 1. The concept of low-dimensional systems

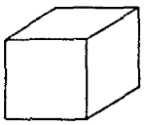
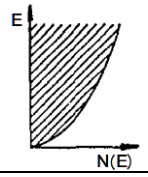
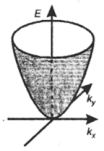
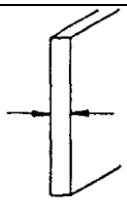
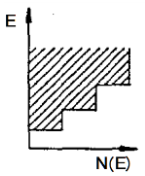
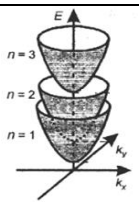
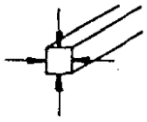
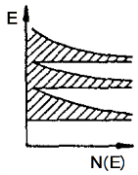
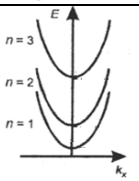
Recently, more and more attention is attracted to research in the field of nanomaterials and nanotechnologies. In this case, we work with objects, the size of which in at least one dimension should be commensurable with the *correlation radius* (or *the characteristic length*) of a physical phenomenon for a particular quasiparticle system. In this sense, as will be shown below, the transition to a *nanoscale* object of investigation can not always be regarded as a feature (or driving force!) of the transition to a new state of physical phenomena or process. That is why, the title of this chapter uses words "Concept of low-dimensional system", rather than nanoscale.

Let us enumerate the most important correlation parameters in the physics of low-dimensional systems (LDS):

- *the dephasing length* $L\varphi$ of an electron in scattering theory,
- *length of thermal dephasing* $L_T$,
- *the mean free path* $l$ of quasiparticles (electrons, phonons, etc.)
- *the coherence length* $\xi$ of Cooper electron pairs (in a superconductors),
- *the de Broglie wavelength* $\lambda_{dB}$ (or *the Fermi wavelength* $\lambda_F$),
- *the Debye shielding length* $L_D$,
- *the magnetic length* $L_m$ (the thickness of the domain wall),
- *the size of the magnetic domain*,
- *diameter of the solid phase nucleus*,
- and etc.

Depending on the ratio between the dimensions of the object (its length *d*, width *a* and height (thickness, diameter) *h*) and the correlation radii mentioned above, which are characteristic for describing certain quasiparticle systems (electrons, phonons, photons, cooperons, plasmons, etc.), we can talk about three-dimensional (3D), two-dimensional (2D), one-dimensional (1D) or zero-dimensional (0D) objects. Examples of such objects and their most important parameters are shown in Figs 1 and 2.

Fig. 1.1. Examples of LDS, their band diagrams and energy distributions of the density of states $N_i(E)$: $i$ = 0, 1, 2, 3 - dimensionality of the object

| Dimension | Form | The density of states of $N_i(E)$ | $N_i(E)$ | Band energy diagrams | Objects |
|---|---|---|---|---|---|
| **3D** |  | $N_{3D}(E) = \frac{1}{2\pi^2}(\frac{2m^*}{\hbar^2})^{3/2}\sqrt{E - E_{c0}}$ |   |   | Bulk, 3D lattices |
| **2D** |  | $N_{2D}(E) = \frac{m^*}{\pi\hbar^2}\sum_{n_z}\theta(E - E_{n_z})$ |   |   | 2D electron gas AlGaAs / AsGa |
| **1D** |  | $N_{1D}(E) = \frac{1}{\pi}\sqrt{\frac{2m^*}{\hbar^2}}\sum_{n_x n_y}\frac{1}{\sqrt{E - E(n_x, n_y)}}$ |   |   | 1D wires or tubes |

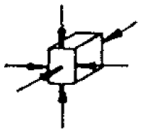
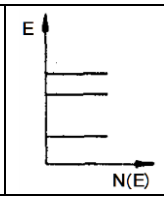
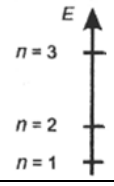
| | | | | | |
|---|---|---|---|---|---|
| **0D** |  | $N_{0D} = \sum_{n_x,n_y,n_z} \delta(E - E_{n_x,n_y,n_z})$ |   |   | Quantum dots, antidots or molecules |

Fig. 1. 2. Examples of the most important parameters of LDS

| **The most important of LDS parameters** | **3 D** | **2 D** | **1 D** |
|---|---|---|---|
| **Fermi energy** $E_F$ | $\frac{\hbar^2}{2m^*}(3\pi^2 n_e)^{3/2}$ | $\frac{\pi\hbar^2}{m^*}n_e$ | $\frac{\pi^2}{4m^*}{n_e}^2$ |
| **The Fermi wave vector** $k_F$ | $(3\pi^2 n_e)^{1/3}$ | $\sqrt{2\pi n_e}$ | $\frac{1}{2}\pi n_e$ |
| **Quantum corrections to conductivity** $\Delta\sigma_i$ | $\Delta\sigma_3 = \frac{e^2}{\hbar}\left(\frac{1}{L_\phi} - \frac{1}{l}\right)$ | $\Delta\sigma_2 = -\frac{e^2}{\hbar}\ln L_\phi / l$ | $\Delta\sigma_1 = \frac{e^2}{\hbar}(l - L_\phi)$ |

Let us consider in more detail the concept of LDS and its criteria, taking in the beginning as an illustrative example the motion of electrons in an infinite thin metal film, reducing its thickness *h* to one of the correlation radii. For 2D systems we denote the plane of the system as *xy*, and *z* - as the direction perpendicular to it. The wave function of an electron moving freely in such film can be represented as a de Broglie wave with a definite energy and momentum. When moving through a film, electrons undergo various scattering processes on defects - impurities, vibrating atoms, etc. The scattering of an electron by impurities with good accuracy is elastic. This means that the energy and momentum modulus is preserving, but the direction of the momentum is changed. However, any inelastic scattering event leads to dephasing of the wave function. Therefore, the motion of scattered electrons can be represented qualitatively as diffusion with conservation of the phase between the scattering events with dephasing of the wave function. Thus, as long as the distance to which the electron diffuses (the dephasing length) becomes less than the film thickness, the system is considered 3D. In the opposite limiting case, the system behaves as 2D for effects in which interference between electron waves is significant.

Let us consider what characteristic length scales allow us to say that some properties of the film can be described as 2D.

***The dephasing length.*** The largest scale characterizing the electron scattering processes is *the dephasing length* (or the *Thouless length* $L_{th}$ or *phase coherence length*). At low temperatures, the processes of electron collisions with defects and smooth surfaces of films are elastic: although the direction of propagation of the de Broglie wave of an electron changes during scattering, the phase of the wave function is conserved. At the same time, inelastic electron-electron and electron-phonon scattering processes knock down the phase of the wave function.

As was mentioned above, the electron motion can be regarded as a diffusion process with conservation of phase when it is subjected by multiple elastic scattering events. Due to the phase conservation, this electron can participate in various interference phenomena and give the so-called interference quantum corrections to electric conductivity (see Fig. 2). A criterion-condition when we need to take into account interference of an electron with the scattering amplitude $A_i$ at various impurity centers is relation

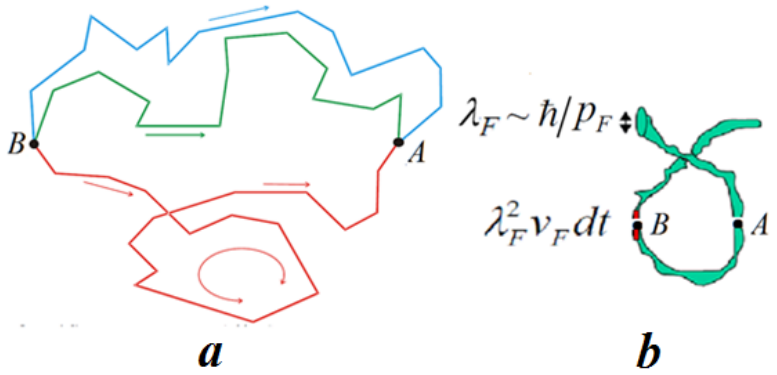
$$l << \lambda_F \sim \hbar/p_F. \qquad (1.1)$$

Let us elucidate this with the example of a 2D electron gas in Fig. 3. When taking into account the quantum effects for calculating the probability of the electron route from point *A* to point *B*, it is necessary to sum the amplitudes $A_i$ of de Broglie waves traveling along all possible trajectories connecting these points. To find the probability of the appearance of a trajectory with self-intersection (i.e. return to the starting point *A*), in the case of diffuse motion of an electron from point *A* to point *B*, we can use

$$W = W_{Cl} + W_Q = |\Sigma A_i|^2 = \Sigma_i |A_i|^2 + \Sigma\, i \neq j\, A_i\, A^*. \qquad (1.2)$$

Here $W_{Cl}$ is a classical contribution, whereas the second term $W_Q$ is the result of the interference of de Broglie waves, because of which the total probability can differ from the classical sum of the probabilities along different trajectories.

Fig. 1.3. Scheme for estimating the probability of the appearance of a trajectory with self-crossing at the diffuse motion of an electron from point A to point B (a) and a separate image of a self-crossing trajectory (b)




The key here are the possible self-crossing trajectories, because electron can pass by such "loops" in two directions – clockwise and anticlockwise. Since the phase change for both routes turns out to be the same, these two (coherent) waves interfere effectively. At the same time, "ordinary" trajectories without self-crossing come to the end point with random dephasing and the interference correction for them is zero.

We denote the characteristic time between the elastic processes $\tau$, and the time between inelastic processes leading to dephasing, $\tau_\phi$. Then the mean free path of the electron with the Fermi velocity $v_F$ in the electric field between the elastic collisions will be $l \sim v_F \cdot \tau$, and its characteristic displacement to the moment of dephasing is

$$L_\phi \approx \sqrt{D\tau_\phi} = l(\tau_\phi/\tau)^{1/2}, \qquad (1.3)$$

where $D \sim l^2/\tau = v_F^2 \cdot \tau$ is the diffusion coefficient.

Let us define more precisely how the approaching of the LDS size to the dephasing length leads to a change in the role of interference effects. For the interference of de Broglie waves, it is important that their phases were coherent. This property is automatically satisfied by de Broglie waves propagating along trajectories with self-intersections: the resulting loop can be traversed in two directions, therefore each such trajectory actually represents two different possible paths for an electron. Since the change in the phase of the de Broglie wave for an electron moving with the Fermi velocity is determined only by the path traveled, then for the wave packets circumventing the loop in different directions, the phases automatically coincide and these waves are interfered.

Thus, the magnitude of the interference correction will be proportional to the probability that the electron will cross its trajectory during the diffuse motion. This interference contribution reduces the classical Drude conduction (see Fig. 2).

Here it is necessary to make an important remark. Due to the uncertainty relation, the trajectory of the electron in the process of diffuse motion is not infinitely thin - it is actually a tube with a diameter of the order of the de Broglie (Fermi) wavelength $\lambda_F = h/\sqrt{2mE_F}$. During the time $t \gg \tau$, the electron can get to any of the points at the characteristic distance $\sqrt{Dt}$ from the starting point, that is, the volume $\sim(Dt)^{3/2}$ is accessible to it. In order for the endpoint of the trajectory to close the loop (coinciding with the starting point) during the time interval $dt$, it must fall into the volume $\sim v_F \cdot \lambda^2 dt$ around the starting point (Fig. 2b). To find the total probability, it is necessary to sum over all possible loop lengths, that is, to integrate with respect to time from time $\tau$ (it is senseless to consider the smaller times) to $\tau_\phi$ (at large lengths, the interference loops will not be due to a dephasing). That is, in the 3D case the probability of closing the electron trajectory is

$$w_{3D} \sim \int_{\tau}^{\tau_\varphi} \frac{v_F \lambda^2}{D^{3/2} t^{3/2}}\, dt \sim -\frac{\lambda^2}{v_F^2} \frac{1}{\tau^{\frac{3}{2}}} \frac{1}{\sqrt{t}}\Big|_{\tau}^{\tau_\varphi} = \frac{\lambda^2}{l^2} - \frac{\lambda^2}{l L_\varphi}. \tag{1.4}$$

In the 3D case, the finiteness of the dephasing length is only a small correction. If the film thickness $h < L_\phi$, then we can assume that across the film the probability of finding an electron at some point is constant and then the available volume will be $\sim h \cdot Dt$ and the probability of trajectory closure will be equal to

$$w_{2D} \sim \frac{v_F \lambda^2}{D} \frac{1}{d} ln \frac{\tau_\varphi}{\tau} \sim \frac{\lambda^2}{ld} ln \frac{L_\varphi}{l}. \tag{1.5}$$

As a result, we obtain a logarithmic divergence with increasing $L_\phi$, which indicates an increasing role of interference corrections.

***The mean free path.*** The next specific scale of the electron system is the mean free path $l \sim v_F \cdot \tau$, which characterizes the distance traveled by an electron between two scattering events. If $l$ becomes comparable with some of the dimensions of the system (for example, with wire diameter or film thickness $h$), then two consequences in the behavior of electrons are possible. For the first case, if in this direction the electron "flies" through the whole sample without scattering (in the ballistic regime), then this can lead, for example, to the quantization of the conductivity for 1D or 2D conductor (see below).

The second consequence is that electrons at low temperatures can be scattered by the surface, leading to a decrease in conductivity. The magnitude of this decrease depends on the surface roughness of the film or wire: when specular reflection of electrons, the conductivity remains practically unchanged, whereas for diffuse scattering, it decreases by several orders of magnitude as the indicated size of the object decreases. This question will be considered in detail in Chapter 2, paragraph 2.7.1. However, we note that, according to the Fuchs model, the size effect over the mean free path is most pronounced if the film thickness $d$ is less than or on the order of the mean free path $l$, but it can be very significant even for thicker films. The value of the relative decrease in the conductivity of thin films under the condition $h \le l$ can be estimated with the help of the relation derived by Fuchs

$$\frac{\sigma}{\sigma_0} \cong 1 - 3(1-P)/8(h/l_0), \tag{1.6}$$

where $\sigma_0$ is the specific conductivity of the bulk material, and $P$ is the specularity coefficient. According to this model, after specular collision of electrons by a surface ($P = 1$), its energy and the parallel to the surface momentum component are conserved. When diffuse scattering ($P = 0$), after each collision with the surface, the electrons begin to move in a longitudinal electric field with an average thermal energy and a randomly directed momentum.

***Debye shielding length.*** The concept of LDS with respect to the Debye screening length is considered in detail in Chapter 2, Section 2.7.7. Here we are talking about the classical dimensional effect in the form of a change in the electrical conductivity of a semiconductor film whose thickness $h$ is close to the width of the space-charge region (the depth of penetration of the electric field into the semiconductor), which, in turn, depends on the Debye length of the screening $L_D$. In considering this effect, it is assumed that for $h < L_D$, identical concentrations of surface electronic states are localized on both surfaces of the crystal with coordinates ($z = 0$, $h$), so that the same bending of the energy bands occurs near both surfaces. Since the crystal is thin, the electric fields of these charges "permeate" the entire thickness of the film (their space-charge regions overlap), so that as a result of their superposition, the boundaries of the energy bands in the field of their bending are also overlapped and lowered in the middle part of the film. The latter effect is called "dimensional flatness" of energy bands). Its physical meaning is that there are not enough free electrons in a thin crystal to shield the surface electronic states, and therefore the necessary number of electrons enters to the crystal from the external circuit.

***Dimensional quantization*** occurs when the mean free path exceeds the thickness of the film or the diameter of the object. If the film is ideally mirror and plane-parallel, then when moving in the z-direction, the electron will feel both boundaries, so that due to reflection on the surface, an interference picture of de Broglie waves is obtained, analogous to the optical interference pattern in thin transparent films. In the language of quantum physics, such a thin film corresponds to the presence of infinitely high potential walls at the boundaries of the film, and then for the component of the wave vector $k_Z$ perpendicular to the film, the obvious quantization condition $k_Z h = \pi n$ appears. It means that in the potential well with infinite walls, an integer number of de Broglie half-waves must be kept on the width of the well. In this case, there arises a certain amount (generally speaking, which can be large) of the size-quantized subbands, since the electron energy is also quantized, i.e. discrete energy levels arise. For a thin film, these energy levels have the form

$$E(n_x, n_y, k_z) = \left[\frac{\hbar^2}{2m_{xx}}\left(\frac{\pi}{L_x}n_x\right)^2 + \frac{\hbar^2}{2m_{yy}}\left(\frac{\pi}{L_y}n_y\right)^2\right] + \frac{\hbar^2 k_z^2}{2m_{zz}^*}, \tag{1.7}$$

which differs significantly from the dispersion law for three-dimensional objects

$$E(k) = E_{c0}(r) + \frac{\hbar^2}{2}\left(\frac{k_x^2}{m_{xx}^*} + \frac{k_y^2}{m_{yy}^*} + \frac{k_z^2}{m_{zz}^*}\right). \tag{1.8}$$

If in a film the bulk concentration of carriers is $N$, then they must be distributed over several subzones with different values of the quantum number $n$. In the film plane, the free motion of electrons and the filling of states of each subband is a two-dimensional analogue of the Fermi gas problem: states with a wave vector below the Fermi gas will be filled. The next subband will be filled only if the electron concentration is so high that the Fermi energy in the previous subband is compared with the distance between the subbands. By minimizing the energy under the condition that the total number of particles is conserved, it is possible to obtain the distribution of Fermi momentums over the size-quantized subbands. The exact solution of this problem is considered in one of the chapters below. We note only one obvious conclusion: the Fermi momentum will decrease with increasing $n$ and beginning with some quantum number the subzones will not be filled at all.

Thus, if the electron concentration is sufficiently small, or the distance between the subbands is sufficiently large, then all the electrons will gather in the lower subband. Only the $x$- and $y$-components of the wave vector remain in the particle energy: the particle, as it were, moves only in the $xy$ plane. It is precisely these systems that are spoken of as a 2D electron gas.

However, it must be borne in mind that they also remain quasi-2D, since the wave function depends on the coordinate $z$ (the particle is delocalized in the $z$ direction at the scale of the thickness

of the film). Therefore, in some phenomena (for example, when calculating the interaction energy of electrons at a sufficiently high surface density of particles, when the distance between electrons becomes of the order of the film thickness), it may be necessary to take into account this degree of freedom.

Similarly, if we restrict the motion of electrons in two directions ($z$ and $y$), 1D conducting systems (quantum wires) are obtained. With restriction in all three directions (that is, when a completely closed potential well for an electron is created), by analogy we speak of a 0D electron system (a quantum dot). In this case the dispersion law will have the form

$$E(n_x, n_y, n_z) = \frac{\hbar^2}{2m_{xx}}\left(\frac{\pi}{L_x}n_x\right)^2 + \frac{\hbar^2}{2m_{yy}}\left(\frac{\pi}{L_y}n_y\right)^2 + \frac{\hbar^2}{2m_{zz}}\left(\frac{\pi}{L_z}n_z\right)^2. \qquad (1.9)$$

We also remind that the scale of the energy distance between the size-quantized subbands also imposes the necessary condition on the temperature at which such effects can be observed: it is necessary that $T << \Delta E/k$. In practice, in order to reliably observe the effects associated with size quantization, a temperature is required at which $kT$ is an order of magnitude smaller than the distance between the subbands $\Delta E$.

Thus, it can be seen that the low-dimensional criteria in the theory of electron scattering and in describing the behavior of electronic systems of different dimensions are determined by a specific physical problem, and, accordingly, can be very different. As can be seen from Fig. 1, the dimensionality of the object has a significant influence on the dispersion laws and the energy dependence of the density of states $N_i(E)$, where $i = 0, 1, 2, 3$ is the dimensionality of the object. Typical characteristic scales in the physics of LDSs are shown in Fig. 4.

Fig. 1.4. Typical values of the characteristic dimensions in the physics of LDS

| | | |
|---|---|---|
| Size of the object | $L$ |  |
| Fermi wavelength | $\lambda_F \sim \hbar/p_F$ | <1 nm (metals)<br>~ 10 nm (semiconductors) |
| Mean free path | $l \sim v_F \cdot \tau$ | 10 - 100 nm (metals)<br>0.1 - 1.0 μm (semiconductors) |
| The dephasing (phase coherence) length | $L_\phi \approx \sqrt{D\tau_\phi} = l(\tau_\phi/\tau)^{1/2}$ | $L_\phi > l$, usually 0.1 to 10 μm |
| Thermal dephasing length | $L_T \approx \sqrt{\hbar D / k_B T}$ | $L_T > l$, usually 0.03 - 1 μm |
| Magnetic Length | $L_H = \sqrt{\Phi_O / 2\pi H}$ | 1.0 μm @ 20 Gs |

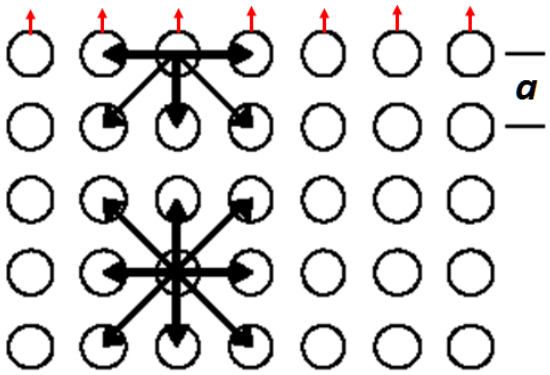
## 2. Introduction to Physics of Surface/Interface

## 2.1. Introduction to the subject

### 2.1.1. External and internal surfaces in solids

The objects of research in this Chapter of the book are the *external* and *internal surfaces*. The *external surface* is understood as the interface between a solid and an external medium (vacuum, gas, liquid), Fig. 2.1.1. Under the inner surfaces (they are often called the *interface*) are meant interfaces between two condensed bodies (interphase or intergranular boundaries), Fig. 2.1.2.

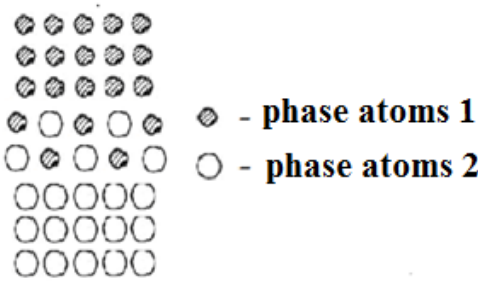
Fig. 2.1.1. Anisotropy of the forces acting on atoms on the surface and in the volume of a crystal




The surface as the interface between the crystal and vacuum is of interest for fundamental physics because the atoms in its near-surface layers are in special conditions in comparison with the atoms in the crystal bulk. These special conditions are associated with a violation of the strict periodicity of the crystal lattice in one of the directions: it consists of the formation of dangling bonds (arrows in Fig. 2.1.1) at the surface at the initial moment of the surface formation due to an abrupt breaking of the translational symmetry in the crystal.

As shown below, the presence of dangling bonds at the surface of a crystal in vacuum leads to a redistribution of the electron density and a change in the atomic positions in several near-surface atomic layers. As a result, electrons and lattice waves moving near the surface "feel" this break, and therefore their behavior near the surface of a solid body is not quite the same as in its volume.

Fig. 2.1.2. Interface between two crystalline phases




Internal surfaces (interfaces) are formed either between the same type of crystals (for example, in the form of grain boundaries) or between different phases (see Figure 2.1.2). In the case of intergrain boundaries, the interface consists of identical atoms and is formed as a matching region of identical but disoriented crystal lattices of neighboring grains. In the case of interphase boundaries, the coordination of neighboring phases (as a transition from one type of periodicity to another) is achieved by transition structures consisting of both types of atoms that form a chemical compound or consist of randomly mixed atoms of two sorts within several atomic layers (Fig. 2.1.2).

Due to the violation of the periodicity in the arrangement of atoms, the surface is one of the main defects of the three-dimensional crystalline structure, being a classical object in solid state physics. The breakage of chemical bonds when the surface formation (for example, due to rapture of crystal on two pieces) leads to a change in the coordination sphere of surface atoms and the re-hybridization of their valence orbitals (redistribution of the electron density) in the near-surface region of the raptured crystal. This leads to the appearance of own localized surface electronic states (SES); the change in the effective charges of surface atoms; the change of atoms arrangement due to reconstruction or relaxation; the appearance of additional deformation due to a change in the interatomic distances; change in atomic vibrations (the phonon spectrum).

### 2.1.2. Choice of the physical model of the crystal surface. Homogeneous and inhomogeneous surfaces.

When a surface is formed (for example, by breaking a crystal into two parts), chemical bonds are broken between atoms located on opposite sides of the crystal being broken. As can be seen from Fig. 2.1.1, the breaking of bonds leads to the appearance of anisotropy of the forces acting on the surface atoms (compared to what happened when these atoms were in the volume of a solid). The latter will lead to the redistribution of electrons near the surface (the re-hybridization of their valence orbitals) and, as a consequence, the immediate change in the mutual arrangement of surface atoms.

The structure of the surface, which has been altered by breaking the crystal, can not be changed abruptly to a three-dimensionally ordered structure of the volume of the crystal. Consequently, a finite (quasi-two-dimensional) transition region should always appear near the surface. Therefore, the surface should be treated not as a geometric plane ($x$, $y$) with zero thickness, but as a three-dimensional surface phase whose structure and physical properties differ from those of a bulk crystal.

The anisotropy of the interatomic forces mentioned above, leading to a restructuring of the atomic positions, arises not only on the crystal-vacuum boundary, but also at the interface (Fig. 2.1.2). Therefore, in the future, the concept of "*surface*" will be used not only for a free surface, but also for interphase and grain boundaries, which separate contacting crystals.

As a result of these processes of the atomic rearrangement in the near-surface region or in the vicinity of the interface, the interatomic distances will change. On the one hand, the change in the interatomic distances (distortion of the crystal lattice) can cause the appearance of additional strains (deformations) near the surface. On the other hand, as noted above, effective charges of surface atoms will change, that is especially important for semiconductors, when intrinsic SES in the band energy spectrum of electrons are created which act as capture and recombination centers of free charge carriers.

In addition, all the above-mentioned near-surface changes in the crystal will necessarily lead to a change in the vibrational spectrum of the atoms of the crystal lattice (the phonon spectrum).

Since the above-mentioned redistribution of the near-surface atoms in a crystal does not always go uniformly (including, due to the influence of atoms of the external medium), any surface phase is more or less inhomogeneous. Its heterogeneity is due to a number of interrelated physical, chemical, technological, and other factors. Consequently, the classification of surface defects on the real surface of crystalline solids is often based on the nature of their introduction during growth of crystal with a free surface or when a surface is formed due to breaking of already prepared crystal. In this case, the concepts of physical, chemical, and induced inhomogeneities are introduced, the names of which characterize the nature of the appeared defects.

Under *physical inhomogeneities* we mean any deviations in the surface structure of real crystals from the exact periodicity of an ideal crystal described by the Bravais lattice. Such inhomogeneities are usually referred to as surface roughnesses of a different nature and crystal lattice irregularities in the periodicity of the crystal lattice in the form of crystal lattice defects (point defects, dislocations, etc.).

*Surface roughness* can be either atomic (nanometer) scale, or rather large (micrometer and even more scale). Usually they have the form of steps on a more or less smooth surface, which arise both during the crystallization process (due to local disturbances in crystal growth conditions) and due to subsequent surface treatments (cutting, grinding, polishing, etching, etc.). In the case where the dimensions of the steps on the crystalline surface do not exceed the interatomic distance (usually the lattice parameter is characterized by values of the order of 0.3-0.5 nm), such surfaces are usually called atomically smooth.

Point defects, dislocations and other *crystallographic defects* are formed both during the crystallization process and under the influence of thermal, mechanical, electrical, radiation and other effects on the crystal.

The physical inhomogeneities of the real surface play an important role in a wide variety of

processes like friction of surfaces, adsorption of atoms and molecules on the surfaces, and so on.

Experiments on the study of such processes as friction, field-emission and thermionic emission, thermal accommodation of gases, physical and chemical adsorption confirm that the interaction of foreign atoms with surfaces results in different types of physical inhomogeneities and proceed in different ways. The work function of an electron from a solid also depends on the dimensions of the roughnesses and the orientation of the surface relative to the crystallographic planes of the atomic lattice.

*The chemical heterogeneity of the crystal surface* is understood as the defects caused by the presence of single foreign atoms in the crystal lattice (they are often called *adatoms*). Even in small amounts, these extraneous atoms on the surface can serve as "active centers", affecting the processes of adsorption and catalysis, oxidation and corrosion, etc. Foreign atoms, for example, form dipole moments on the surface, which in magnitude and direction can be very different from those on an atomically clean substrate. The problem of obtaining a chemically homogeneous surface is much more complicated than the problem of obtaining chemically pure materials in general. For example, the presence of only $10^{-6}$ volume fractions per 1 $cm^3$ of the sample can yield up to 20 monolayers of foreign atoms on the surface, leading to a significant change in its properties. At the same time, the distribution of the same amount of impurity atoms more or less uniformly over the volume of the crystal practically does not change its properties.

*The induced inhomogeneities* arise from adsorption processes when the surface is covered by a layer of adsorbed atoms from the surrounding medium. The essence of the phenomenon of adsorption is that molecules or atoms of the external medium interact with the surface and are there for some time, which leads to an increase in the concentration of foreign atoms or molecules on the surface in comparison with the bulk of the crystal.

If there is a gaseous medium above the surface of the crystal, then two cases of interaction of gas molecules with the surface are possible - elastic reflection or capture and stay on the surface for some time *t*, where *t* is the average lifetime of the molecule in the adsorbed state.

From what has been said, it follows that the existence of a surface introduces enormous complications in the physics of crystals. In order to find some general patterns in this complicated situation, it is necessary to turn to models that correctly reflect the regularities common to the surface phases, neglecting more particular details in the first approximation.

Naturally, the simplest is the model of a homogeneous surface, which is often used in all practical aspects of surface physics. In this case, we assume that all the physical properties - the arrangement of atoms, the distribution of their effective charges and the charges of SES - are homogeneous around the surface plane (*x, y*). For such a model, surface is characterized by the average effective values of the macroscopic electric field, surface charge, potential, etc. Only changes in these physical parameters along the *z* axis (that is, along the normal to the surface) are considered, which reduces the three-dimensional problem to the one-dimensional one. The energy spectrum of SES in this case is characterized by such phenomenological parameters as concentration, energy, and electron and hole capture cross sections.

Such a surface model is often used in the analysis of surface electronic phenomena. Phenomenological approach to the problem made it possible to describe in a consistent and sufficiently complete manner the main regularities in various electronic processes that are raged at the solid surface and at interphase boundaries. Within the framework of theoretical concepts, it is possible to establish important relationships between the macroscopic properties of the surface and the SES parameters, which determine such fundamental processes as the capture and recombination of charge carriers, and their transport in surface phases. It is shown that all these interconnections are changed significantly at transition from bulk single crystals to systems with reduced dimensionality, when the sizes of the objects themselves begin to approach the characteristic lengths of the electron and phonon subsystems of the crystal. Such systems include zero-dimensional (*0D*) atomic-molecular clusters of substances, one-dimensional (*1D*) filamentary structures and two-dimensional (*2D*) thin films.

### 2.1.3. Ideal and nonideal surfaces in solids

As shown above, any surface is always heterogeneous to some extent, and this heterogeneity includes structural and chemical heterogeneity, heterogeneity of electronic properties, fields of deformations, etc. Heterogeneity has a cooperative character - a change in one kind of heterogeneity, as we shall see below, is always accompanied by a change in its other types. Precisely for the theoretical models to be at least qualitatively compatible with the experimental data, we separate surfaces on two types: "atomically pure" and "real". Let us consider the conditions, when studying the surface of crystals, one can speak of the so-called "pure" (or ideal) and "real" (i.e. nonideal) surface.

*An atomically pure* surface can be obtained only under special, carefully controlled conditions, for example, when a crystal is broken in an ultrahigh vacuum or in an absolutely neutral medium (for example, in liquid helium). In this case, we minimize only chemical inhomogeneities of the surface.

*The real* or, as it is often called, "*coated*" surface can be obtained either by growing the crystal or by further laboratory processing (for example, by cutting, mechanical grinding, polishing, etching, etc.). In this case, all types of heterogeneities are present, and the surface treatment conditions must be strictly standardized.

Experimental study of the surface is associated with great difficulties. The main of them is a small number of atoms in the surface phases in comparison with the volume. If there are $n_v = 10^{22}$ atoms·cm$^{-3}$ in a unit volume of the substance, then on the surface unit, $n_s = n_v^{2/3} = 10^{15}$ atoms·cm$^{-2}$. Therefore, the signals from the surface in the measuring devices are often observed against a background of an intense signal from the bulk of the crystal.

Naturally, the "pure" surface is of more interest for physicists. They tend, first of all, to understand how atoms are arranged on it and how surface electrons and vibrating atoms waves behave. At the same time, it is important to investigate real surfaces. In their study, two approaches are possible. The first is based on comprehensive studies of atomically clean surfaces with the following creating of these or other defects (chemical complexes, new phases, roughnesses, etc.) to some extent modeling the real surface in them under controlled conditions. Another approach consists in a direct analysis of the state of the real surface by attracting all available spectroscopic, electrophysical, and adsorption methods.

The development of special methods for surface cleaning, analysis of the structure and chemical composition of the surface phases in the 60-70s of the last century gave a serious push to the development of physics and chemistry of pure or ideal surfaces. This happened as a result of the coincidence of three circumstances.

*The first circumstance* was the development of new highly sensitive diffraction, spectroscopic and probe methods for studying the structure, roughness, and chemical composition of the crystal surface.

*The second favorable* circumstance was the space programs that led to the creation of ultrahigh-vacuum set-ups, allowing to keep the sample atomically pure (with sensitivity to small fractions of the atomic monolayer) for a considerable period of time.

*The third circumstance* is associated with the emergence and wide dissemination of high-speed computers, which led to the rapid development of methods for theoretical modeling of the ideal surfaces structure and properties.

The appearance of reliable experimental methods for obtaining atomically pure, clean and smooth surfaces, the possibility of their conservation in this state for a long time after its formation, and also the determination of their chemical composition sharply increased the reproducibility of the experimental data. These circumstances made it possible to perform experiments under reliably controlled conditions and on reliably characterized surfaces of solids, which, in turn, made it possible to compare the experimental results on real surfaces with theoretical models of ideal surfaces.

As a result of these three circumstances, the birth of a new field of surface science - the physics of atomically pure, clean and smooth surfaces - occurred. The latter, naturally, did not lowered the actuality of physical and chemical studies of real surfaces and, especially, of interphase boundaries, because the atomic and electronic processes on them in many cases determine the functioning of integrated circuits, converters of solar energy, sensors and many other devices and materials using in modern technologies.

### 2.1.4. The role of external and internal surfaces and the formation of properties of nanosystems

The external surfaces and interfaces of crystalline solids, as well as the processes occurring on them, are of interest to physics, technology and technics insofar as they determine many important practical properties of materials. Moreover, it turns out that in practice either surface or interfacial properties of materials, especially for nanostructured materials, are more often used.

For example, in the cutting tool (knives, drills, cutters, etc.), the mechanical properties of the surface are usually used. On the other hand, the corrosive destruction of various parts and structures begins most often just from the surface and interphase boundaries. Radiation destruction of crystals, caused by the action of charged particles (ions, electrons), also first occurs since their surface.

On the surface of certain metals (nickel, platinum, etc.) many chemical reactions are accelerating, while they are going with low rates in the liquid or gas phases. This phenomenon, called heterogeneous catalysis, is the basis of a large number of modern chemical technologies.

Processes occurring in the near-surface layers (in particular, the nature of roughness and other defects in them) determine the frictional forces in the friction parts. The same processes determine the strength of the connection of solids (for example, when soldering or gluing structural parts), as well as the wetting and spreading of liquids over the surface.

Defectiveness (for example, grain boundaries in polycrystals or interphase boundaries in steels and constructional alloys) determines the corrosion resistance of parts and structures, their fragility and strength, and also many of electrical, magnetic and other properties. Defects at the interfaces significantly affect the efficiency of semiconductor solar cells and other semiconductor devices.

The quality and condition of the surface completely determine the properties of many optical elements and devices.

Of great importance are the technical aspects of surface physics in electronic engineering, which uses such modern surface transformation technologies as molecular epitaxy, ion doping, laser processing of materials, etc. Because of this, for semiconductor devices the bulk of material often plays the role of a substrate, which can be replaced by some other one. In the same cases, when the choice of material is dictated by its bulk properties, it often turns out that the properties of the natural surface are not suitable for the intended use, and it becomes necessary to modify them appropriately or to provide a sample of the selected material with an artificially created surface possessing the required properties.

The list of examples demonstrating the great role of the surface in engineering and technology and, accordingly, the applied value of studies of the solid surface, it is not difficult to continue.

However, the structure of external and internal surfaces, as well as the processes occurring on them, are of considerable interest for fundamental sciences, especially such as physics, chemistry and biology. When atoms or molecules are adsorbed on an atomically smooth and clean surface of single crystals, two-dimensional systems of quasiparticles arise which, depending on the temperature, concentration, and singularities of their interaction, can exist in different "aggregate" states - in the form of a two-dimensional "gas", or “liquid” or a two-dimensional "crystal". The study of such two-dimensional phases and phase transitions between them is the content of an important and beautiful section of modern physics of the condensed matter. The information obtained is important for understanding the properties of complex self-organizing systems

consisting of a large number of interacting particles, which can also be used in chemistry and biology.

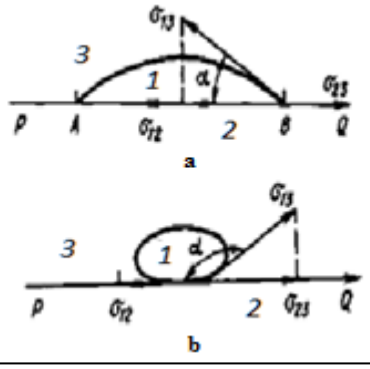
It is now difficult to name any division of the natural sciences, which to some extent does not concern phenomena at the interfaces of solid phases. Not to mention electronics and catalytic chemistry, the problem of the surface is of keen interest to specialists in the field of structural materials (powder metallurgy), specialists in magnetic and optici materials and radiophysicists (film-layered structures). Even experts in the field of nuclear physics are forced to deal with phenomena on the surface (the problem of the second wall of a thermonuclear reactor). Biologists, geophysicists and geochemists are intensively studying complex interphase processes in cell membranes, in porous inorganic and organic substances, etc.

As noted earlier in the preface, electronic processes and their interrelation with atomic/molecular processes on the surface occupy a big place in physics and chemistry surface. This is due to the transition from micro- to nanoelectronics, which tends to increase the degree of integration of solid-state electronic structures and to reduce the size of their individual components, which began to approach nanoscale. At the same time, the contribution to all the processes that occur between the different interfaces greatly increases. The physics of these systems ceases to be described by standard models used in the description of massive rigid bodies, so that the problems of surface phenomena are becoming increasingly important. Undoubtedly, progress in the field of nanoelectronics is closely related to the successes of physics and chemistry of surface.

### 2.1.5. The phenomena of wetting and spreading of liquids over the surface of solids

In many technological processes, an important role is played by phenomena occurring at the contact of liquid and solid surfaces. For example, this refers to the technology of making electrical contacts, where the method of alloying a metal drop of solder with a crystal is widely used. The spreading of such a droplet upon heating largely determines the fusion area, as well as the process of formation and the quality of the contact (its ohmicity, mechanical strength, etc.).

Fig. 2.1.3. Spreading a liquid drop 1 along the surface of a solid 2 in a gaseous medium 3 with a sharp (a) and blunt (b) wetting contact angle




A schematic representation of the spreading of a drop on a solid surface is shown in Fig. 2.1.1*a*. Here, a drop of liquid *1* of *AB* diameter lies on the surface *PQ* of solid body *2* and is also in contact with the gas phase 3 (for example, air). The surface tension $\sigma_{23}$ at the crystal-gas interface 2-3 tends to stretch the drop along the surface of the crystal. At the same time, the surface tension of $\sigma_{12}$ drops at the liquid-crystal interface 1-2, on the contrary, tends to pull the drop to its center. The surface tension $\sigma_{13}$ at the liquid-gas interface 1-3 is directed at the boundary point B of the drop to its surface tangentially at an angle $\alpha$ with the surface *PQ*. This angle is called *the contact angle*. In Fig. 2.1.3*a*, this angle is less than 90° (acute). In this case, the projection of the tension force $\sigma_{13}$ on the plane *PQ* is $\sigma_{13}\cos\alpha$ and coincides in direction with the direction of the action of the tension force $\sigma_{12}$, so that the action of these forces is added. Consequently, the equilibrium condition for the drop on the surface is determined by the equality of forces:

$$\sigma_{12} + \sigma_{13}\cos\alpha = \sigma_{23}. \qquad (2.1.1)$$

The case of a drop with a blunt contact angle of wetting is shown in Fig. 2.1.3b. In this case, the contribution $\sigma_{13}\cos\alpha$ in the relation (2.1.1) is negative.

In the absence of equilibrium of a drop on the surface, the relation (2.1.1) is not observed and the deviation from it can be in any direction. If $\sigma_{23}$ is greater than the sum of the forces on the left side of the equation, then the force $\sigma_{23}$ stretches the drop in all directions from the center. As a result, the droplet height and the angle $\alpha$ decrease, while the value of $\cos\alpha$ increases. Consequently, the projection of the force $\sigma_{13}$ on the plane *PQ* increases. With such an increase, there may come a time when the left-hand side of formula (2.1.1) is compared with the right-hand side and the extended drop appears in the equilibrium state. But it may happen that even with the maximum value of $\cos\alpha = 1$, the value of $\sigma_{23}$ will be greater than the sum on the left-hand side of the relation (2.1.1)

$$\sigma_{23} > \sigma_{12} + \sigma_{13}. \tag{2.1.2}$$

In this case, there will be no equilibrium of the droplet, and it will spread over the surface of the *PQ* of the solid, forming a liquid film. Such a typical case is the spreading of a water or alcohol drops on the glass.

If the contact angle of wetting is acute, then we say that the liquid *moistens* the solid and, conversely, at the blunt edge angle - the liquid does not moistens it. The case $\alpha = 0$ is called *absolute wetting*, and the case $\alpha = 180°$ is *called absolute non-wetting*.

## 2.2. Atomically clean and smooth surfaces

### 2.2.1. Obtaining of atomically pure and smooth surfaces

A surface is considered *atomically pure* if the concentration of foreign atoms on it is below the sensitivity limit of the chemical analysis method used, and *atomically smooth* if the characteristic scale of the irregularities is about one atomic diameter.

The creation of surfaces with high atomic purity and smoothness, as well as their long-term maintenance under ideal conditions, involves the development of methods for obtaining and methods for analyzing the chemical composition and structure of the surface at the atomic level. To maintain an atomically clean surface during the experiment, it is necessary to maintain a vacuum no worse than $10^{-9}$-$10^{-10}$ Pa.

Let us estimate the maximum pressure in a vacuum chamber, in which the surface purity will be kept as long time as needed for the experiments to take place. It follows from the elementary kinetic theory of gases that when pressure P and temperature T are applied, the number of molecules with molecular mass M, striking about 1 $cm^2$ per 1 sec, equals

$$N_S = P \cdot N_A \cdot (2\pi \cdot MRT)^{-1/2},$$

where $N_A$ is the Avogadro number and R - the universal gas constant. In the least favorable case, when each molecule striking the surface is irreversibly bound to it, the time required to cover the entire surface with a monomolecular layer for simple diatomic molecules (like $N_2$, $O_2$, etc.) will amount to 24 hours at $P = 10^{-8}$ Pa. Such a vacuum is considered to be the lower limit, from which it is possible to confidently conduct experiments with atomically clean surfaces. It should be noted that in modern facilities it is possible to maintain a vacuum with a residual pressure not exceeding $10^{-12}$ Pa.

It should be noted that real surfaces are too complex to study, and therefore atomically clean and smooth surfaces can be considered as simplified models of real systems. The use of special analytical methods makes it possible to register chemical substances on the solid surface with sensitivity up to small fractions of a monolayer.

As a criterion for the purity of atomically clean surfaces, a *pollution factor* is sometimes used

that is equal to the ratio of the number of molecules in the gas phase to the total surface area of the sample. When working with the surface of single crystals in a vacuum of $10^{-8}$ Pa, this factor is usually $10^8$ - $10^9$ $cm^{-2}$. Reduction can be achieved either by improving the vacuum, or by increasing the surface of the tested samples, that is often used in spectroscopy. When passing to nanodispersed particles, it is possible to reduce the contamination factor to $10^5$ $cm^{-2}$.

Atomically smooth and absolutely clean surface is an atomic plane of the crystal filled with only the host atoms. Obviously, we can not obtain such surface in an ideal form, but it is possible only to approach these conditions to a certain extent.

The use of ultrahigh vacuum to some extent guarantees the absence of extraneous atoms on the resulting layer, but in addition, the ordering of the surface at the atomic level is necessary.

To achieve cleanliness and smoothness of the surface, four basic methods are used at the atomic level: cleavage, heating, ion bombardment (usually by Ar ions), chemical treatment.

The method of *cleavage* in ultrahigh vacuum is used to obtain an atomically clean surface in the case of materials that are readily dissociated in certain crystallographic planes (oxides, alkali halides, semiconductors, layered compounds). For example, single crystals of germanium and silicon split along the (111) plane, and type $A^{III}B^{V}$ (InSb, InAs, GaAs, GaSb) semiconductors split along the (110) plane.

The method of *thermal surface cleaning* consists in heating the crystal in an ultrahigh vacuum, which leads to the desorption of adsorbed particles. For many materials, the desorption temperature is lower than the melting point of the material, but in a number of cases, due to the very strong coupling of the adsorbate to the surface, it can not be completely removed by heating at temperatures below the melting points of the sample. Therefore, the method of thermal cleaning is used mainly for refractory materials (tungsten and the like).

The use of thermal surface cleaning is ineffective if carbon atoms are present on the substrate: they form compounds with a very strong bonds with the material. For example, to remove the oxide from the silicon surface, it is necessary to heat the sample up to 1170 K. However, even at such temperatures, carbides are formed on the silicon surface, which are much more difficult to remove than the compounds of silicon with oxygen. Therefore, heating is usually used for the regeneration of a surface that has been cleaned beforehand in some other way.

The most universal method of surface cleaning is the bombarding of surface with inert gas ions (*ion etching*). Argon ions with energy (0.5 - 5) keV interact with the solid surface, leading to the knockout of surface atoms, both sample material and impurities. The method of ion bombardment is effective when removing a large number of atomic layers from the surface, even if the rate of removal of the particles of the mixture is much less than that of the material atoms.

The disadvantage of the ion etching method is a strong surface damage by the introduction of bombarding ions into the target, so that the surface must be annealed to restore the ordered structure. Usually in practice it is necessary to conduct several cycles of bombardment and annealing.

Methods of chemical surface cleaning in air do not make it possible to obtain atomically clean surfaces. However, for example, treatment of Si (111) in $NH_4F$ leads to the formation of atomically smooth areas of up to 1 $\mu m^2$.

The chemical method of purification is also the discharge at low pressure into the vacuum system of gases reacting with impurities on the surface to form loosely bound adsorbates, which can then be thermally desorbed. This method is most widely used to remove carbon from refractory metals, such as tungsten, which can be purified from most other impurities simply by heating.

In conclusion, it should be noted that there is no universal purification procedure. For each material, an individual technique with certain parameters is required, or even a combination of several techniques. Sometimes the cleaning technique can be quite complex and include additional procedures.

### 2.2.2. Thermodynamics of Atomically-Clean Surfaces

The basic theses of finite volumes matter thermodynamics can be briefly formulated as follows.

In equilibrium, a single-component system (for example, a crystal) is completely characterized by the value of the internal energy $U$, which is a single-valued function of the entropy ($S$), the volume ($V$) and the number of particles ($N$) in the crystal:

$$\begin{gathered} U = U(S, V, N), \\ dE = \left.\frac{\partial E}{\partial S}\right|_{V,N} dS + \left.\frac{\partial E}{\partial V}\right|_{S,N} dV + \left.\frac{\partial E}{\partial N}\right|_{S,V} dN \\ dU = TdS\text{-}PdV + \mu dN \end{gathered}, \qquad (2.2.1)$$

These equations, representing the first and second laws of thermodynamics, determine the temperature ($T$), pressure ($P$) and the chemical potential ($\mu$) of a bulk crystal. Using the property of extensiveness of internal energy

$$U(\lambda S, \lambda V, \lambda N) = \lambda U(S, V, N), \qquad (2.2.2)$$

and the first and second laws of thermodynamics, using Euler's theorem, we obtain the equation

$$U = TS - PV - \mu N. \qquad (2.2.3)$$

Differentiating (2.2.3) and using (2.2.1), we come to the relation between the intensive variables, known as the Gibbs-Duhem equation:

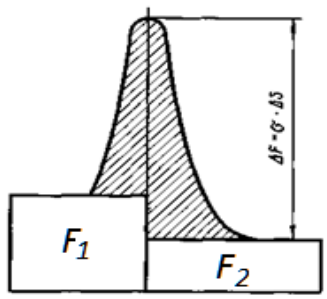
$$TdS - VdP - Nd\mu = 0. \qquad (2.2.4)$$

Let us consider the change in the above equations after the formation of a free surface with area $A$ by splitting of an infinite crystal into two parts. Since in the bulk phase spontaneous splitting should not occur, the total energy of the system should increase by an amount proportional to $A$, as follows from the modification of equation (2.2.1):

$$U = TS - PV - \mu N + \sigma A. \qquad (2.2.5)$$

Thus, from the point of view of thermodynamics, the quantitative characteristic of the surface is the proportionality coefficient σ in the additional fourth term of equation (2.2.5), which is the surface (or interfacial γ) energy and is called the *surface tension*. The quantity σ is understood as the work to be done to increase the surface per unit length of the line limiting the surface. For a better understanding of this definition, we refer to Fig. 2.2.1. Remind that the boundary (surface) is formed by splitting the crystal into two parts by breaking chemical bonds between the atoms of the crystal.

Fig. 2.2.1. Energy enhancement at the interphase boundary




Inside the crystal, the force acted to any of the atoms are compensated by the force fields of the surrounding atoms (see, Fig. 2.1.1). For atoms belonging to the formed surface there is no such compensation and the symmetry of the force interactions is broken. That is why the energy spent on the work to break interatomic bonds is concentrated on the surface.

We denote by $F_1$ and $F_2$ the energies of both parts of the crystal without taking into account the surface energy. Then the total energy of the two components of the crystal with the surface consists of the bulk energy ($F_1 + F_2$) and the surface component σΔA

$$F = F_1 + F_2 + \sigma \Delta A, \tag{2.2.6}$$

where ΔA is the contact surface. Consequently,

$$\sigma = \frac{F - (F_1 + F_2)}{\Delta A} = \frac{\Delta F}{\Delta A} \tag{2.2.7}$$

Hence, the excess energy of the crystal-vacuum interface is

$$\Delta F = \sigma \Delta A. \tag{2.2.8}$$

In the general case, the quantity σ (or interfacial energy γ) is determined by the change (excess) of the free Gibbs or Helmholtz energy per unit area when the interface is created

$$\sigma = \left( \frac{\partial G}{\partial A} \right)_{P,T,N_i} \text{ or } \sigma = \left( \frac{\partial F}{\partial A} \right)_{V,T,N_i}, \tag{2.2.9}$$

where $N_i$ is the number of moles of the i-th component. In accordance with (2.2.9), the quantity σ has the dimension J / $cm^2$. In this case, we can write the Gibbs-Duhem equation (2.2.4) for the excess of energy in the form

$$S \cdot dT - V \cdot dP - N \cdot dm - \sigma \cdot dA = 0. \tag{2.2.10}$$

Since $\Delta F > 0$, i.e. atoms on the surface of the crystal have more energy than that inside the crystal, the surface energy of the crystal can be defined as the free energy that must be expended to create a surface.

Note that the identity of σ and γ is only apparent (due to the same dimension). In fact, these quantities are numerically identical only for the interfaces between liquid-gas and liquid-liquid. For solid-solid interfaces, the difference between σ and γ can be significant. The reason for this difference is that in the liquid there is only a short-range order in the arrangement of the atoms. Therefore, if shear forces act on the liquid, the stresses arising in it are removed by rearrangement of small groups of atoms or molecules. In a solid, shear stresses are removed only by moving the constituent atoms, so that a critical shear stress is required to create residual deformation not equal zero. In addition, dislocations are present in the solid, the displacement of which also leads to a shear. Therefore, the change in the energy of the dislocations also contributes to the total surface energy of the solid body.

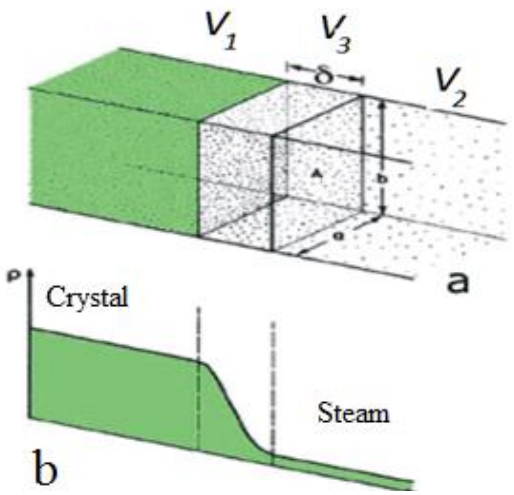
The equilibrium condition for which the surface free energy is minimal is given by equation

$$dG = \int \sigma dA \to 0. \tag{2.2.11}$$

For a liquid, a stable state is reached when the surface obtains a spherical shape, but for a solid, the change in the free energy *dG* (or *dF*) depends on the type of the crystalline faces and becomes minimal for plane faces joining the sharp edges. Therefore, surface tension leads to stresses associated with the expansion of the entire mass of the crystal, and to not very large shear stresses at its edges.

In equilibrium, at any finite temperature and pressure, the crystal with the surface coexists with its own vapor. The dependence of the number of particles on the coordinate z, which varies in the normal direction to the surface, is shown in Fig. 2.2.2 and 2.2.3.

Fig. 2.2.2. The distribution of the density of atoms $\rho$ (z) in a one-component system depending on the distance to the surface




Gibbs have shown that it is convenient to refer certain values of extensional variables to the surface area A = ab (see Figure 2.2.2a). In accordance with the Gibbs approach, the vertical lines in Fig. 2.2.3b divide the space into the solid body region $V_1$, the vapors region $V_3$ and the transition region $V_2$ pertaining to the surface. Then all the extensive Gibbs parameters can be represented as follows:

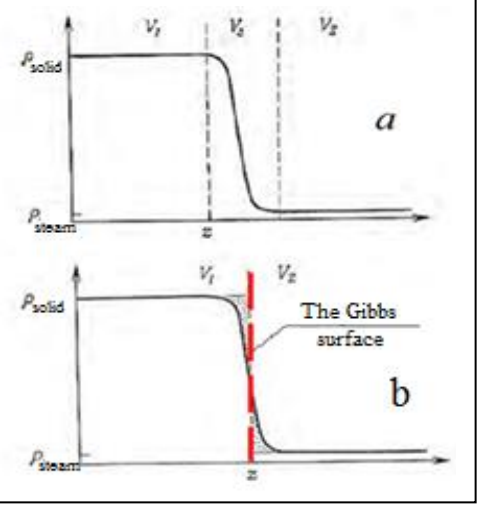
$$S = S_1 + S_2 + S_3;$$
$$V = V_1 + V_2 + V_3;$$
$$N = N_1 + N_2 + N_3; \qquad (2.2.12)$$

In these formulas, bulk parameters are defined as

$$S_i = s_i V_i,$$
$$N_i = \rho_i V_i, \qquad (2.2.13)$$

where i = 1,2, and the entropy $S_i$ and the density $\rho_i$ characterize the volumes of homogeneous phases - the crystal (1) and the vapor (2). In accordance with (2.2.12), if the surface of the phase volume $V_1$ is defined (the position of the vertical dashed line in Fig. 2.2.3b is chosen), then the remaining values relating to the interface as the transition region $V_3$ between $V_1$ and $V_2$, are determined as *thermodynamic excesses*, see the shaded regions in Fig. 2.2.3b. We note that changes in the surface excesses of the quantities in (2.2.12) are completely determined by changes of the quantities in the volumes $V_1$ and $V_2$:

Fig. 2.2.3. The choice of the phase interface by Gibbs in accordance with the condition (2.2.15).




$$\Delta S_s = -\Delta S_1 - \Delta S_2,$$
$$\Delta V_s = -\Delta V_1 - \Delta V_2,$$
$$\Delta N_s = -\Delta N_1 - \Delta N_2, \qquad (2.2.14)$$

Obviously, the specific choice of the location of the crystal-vapor interface (as a symbol of the surface) in Fig. 2.2.2 and 2.2.3a is ambiguous. Nevertheless, it turns out that one can always choose a subset of surface excesses between the vertical dashed lines in Fig. 2.2.3a. With reference to the volume and concentration of particles, Gibbs have chosen the condition when the total excesses are equal to zero in the form of relations

$$\Delta V_s = 0 \text{ и } \Delta N_s = 0. \qquad (2.2.15)$$

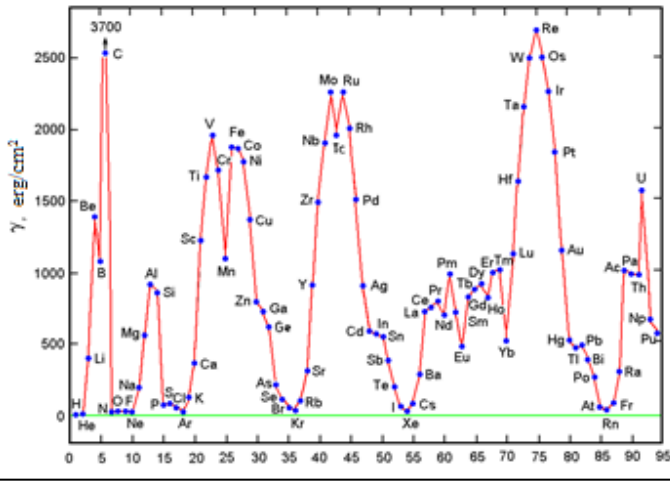
Basing on these relations, the position of the surface is chosen so that the sum of the positive and negative excesses characterizing the surface (the shaded areas to the left and to the right of the vertical dashed line in Fig. 2.2.3b) becomes zero.

By the definition of the tension force, $\sigma$ represents the energy per unit surface, which is necessary for breaking bonds to form the surface when the crystal has split. Then for this parameter we can write the following relation

$$\gamma(\sigma) = E_b \cdot (Z_s/Z) \cdot N_s, \qquad (2.2.16)$$

where $E_b$ is the bulk binding energy, $Z_s/Z$ is the relative number of dangling bonds (per surface atom), after the crystal has split, and $N_s$ is the surface density of the atoms. Estimates show that typical values of surface tension are $\gamma(\sigma) \approx 0.00012$ J/cm$^2$. The following values of the parameters entering into the parity (2.2.16) are used in the calculation: $E_b \approx 3$ eV, $Z_s/Z = 0.25$, $N_s \approx 10^{15}$ atom/cm$^2$.

Fig. 2.2.4. The experimental values of the surface tension $\gamma(\sigma)$ for various elements of the periodic table reflecting the variations in the binding energies $E_b$ in crystals

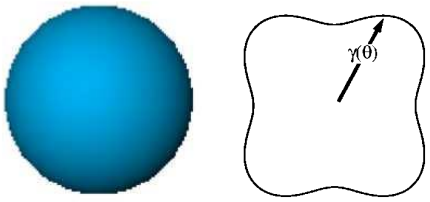



The distribution of experimentally determined surface tension values according to Mendeleyev's table has a periodic character (Figure 2.2.4), following the distribution of valence electrons in chemical elements that determine the energy of the interatomic bonds $E_b$.

### 2.2.3. Surface tension anisotropy

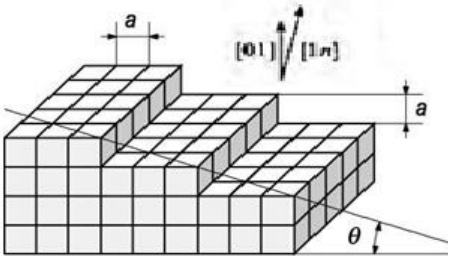
The surface tension σ (or γ) of even a flat solid surface depends on its crystallographic orientation. If the surface tension of the crystal were isotropic, then from the point of minimization of the surface energy, it should have assumed the shape of a sphere (Fig. 2.2.5a). If the surface tension is anisotropic, the crystal should assume a non-spherical shape (Fig. 2.2.5b).

Fig. 2.2.5. A crystal with an isotropic (a) and anisotropic (b) surface tension




To understand this, consider a two-dimensional solid body, with surface normal vector ***n*** of which is slightly deflected from the plane [01] direction (Fig. 2.2.6). Such surface is called *a vicinal surface*. It consists of a large number of monatomic steps separated by terraces of width *a* equal to the lattice constant. If the number n is large, then the small angle between the directions [01] and [1n] is equal to θ ≈ 1/n. The superextension along the direction [1n], which we denote by γ(0), contains the contribution of the surface tension of the (01) plane, equal to $\sigma(0)$, and the contributions from each individual step. If we denote by $\beta$ the energy per step, then the total surface tension of the surface (01) will be equal to

Fig. 2.2.6. Illustration for the determination of the energy of the vicinal surface, misoriented on the angle θ relative to the crystalline face (100).




$$\sigma(\theta) = \sigma(0) + (\beta/a)|\theta|. \qquad (2.2.17)$$

The change in the sign of the second term reflects the fact that to create a step on a flat surface we always need to spend energy. It follows from (2.2.17) that the function $\sigma(0)$ is continuous in a neighborhood of the point $\theta = 0$, but its derivative has a

discontinuity at this point. More accurately,

$$\Delta\left(\frac{d\gamma}{d\Theta}\right) \sim \frac{2\beta}{a}. \qquad (2.2.18)$$

Now consider the case of large angles $\theta$ when density of the steps will increase. In this case, to calculate the surface tension, we need to take into account the energy of interaction between the steps. Landau showed that in this case the function $\gamma(0)$ has a return point for each value of the angle that corresponds to a rational Miller index *n*. The change in the derivative of the function $\gamma(0)$ (that is, the sharpness of the vertex) at the point under consideration decreases rapidly with *n* increasing:

$$\Delta\left(\frac{d\gamma}{d\Theta}\right) \sim \frac{1}{n^4}. \qquad (2.2.19)$$

Consequently, the dependence of the surface tension on the angle $\theta$ in polar coordinates at $T = 0$ will have the form shown in Fig. 2.2.5b. We note that the growth of small crystals confirms their equilibrium non-spherical shape. For example, for lead crystals of micron sizes, it was found that the equilibrium shape represents a cubic octahedron with a number of well-developed faces.

### 2.2.4. Ideal surfaces in crystals

The *ideal surface* is an atomically pure, clean and smooth surface. In this context, according to above said, it can be described as a sharp transition from a crystal to an ideal vacuum. If the surrounding medium is not a vacuum, but another solid body, then we have an internal surface between solid bodies (*interface*). This transition between two solid bodies can be sharp or not sharp. In the latter case, *a blurred interface* with the transition region is obtained.

### 2.2.5. The atomic structure of an ideal crystal surface

In the simplest case, *an ideal crystal surface* can be obtained by breaking a crystal along some crystallographic plane, i.e. by removing all atoms located on one side of a certain crystallographic plane of a three-dimensional crystal with an unchanged arrangement of atoms on the other side of this crystallographic plane (see Figure 2.1.1). In this case, as can be seen from

Fig. 2.2.7. The form of the unit cell for a fcc lattice (a) and ideal surfaces of a crystal broken along different crystallographic planes: b - (100), c - (110), d - (111)

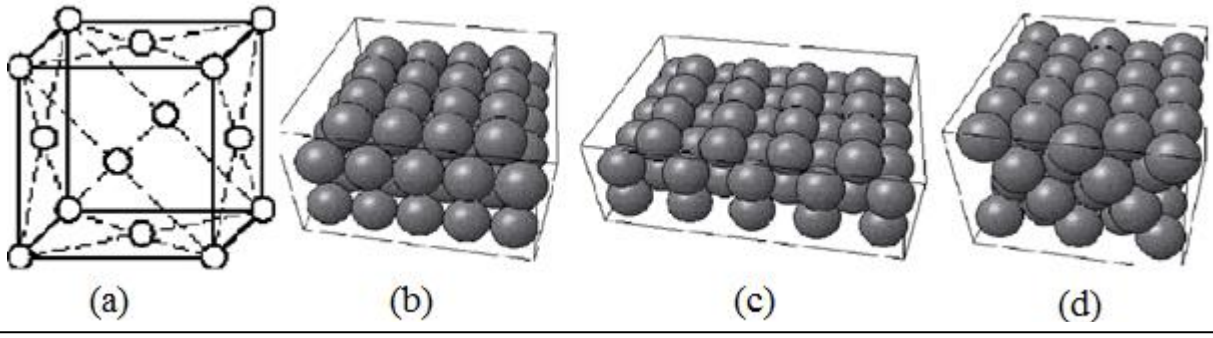



Fig. 2.2.7 - 2.2.9, an ideal crystalline surface preserves the symmetry of the atomic plane along which a discontinuity occurs. However, this situation turns out to be (thermodynamically) disadvantageous. Therefore, as noted earlier, in reality the arrangement of atoms on the surface and in some undersurface layers will differ from their location in the crystal bulk, because the anisotropy of the interatomic forces at its surface (Fig. 2.1.1) will lead to a change in the location of surface atoms. In other words, unsaturated, dangling bonds tend to form new interatomic bonds.

Fig. 2.2.8. The form of the unit cell for the bcc lattice (a) and ideal surfaces of the crystal broken along the planes: b - (100), c - (110), d - (111)

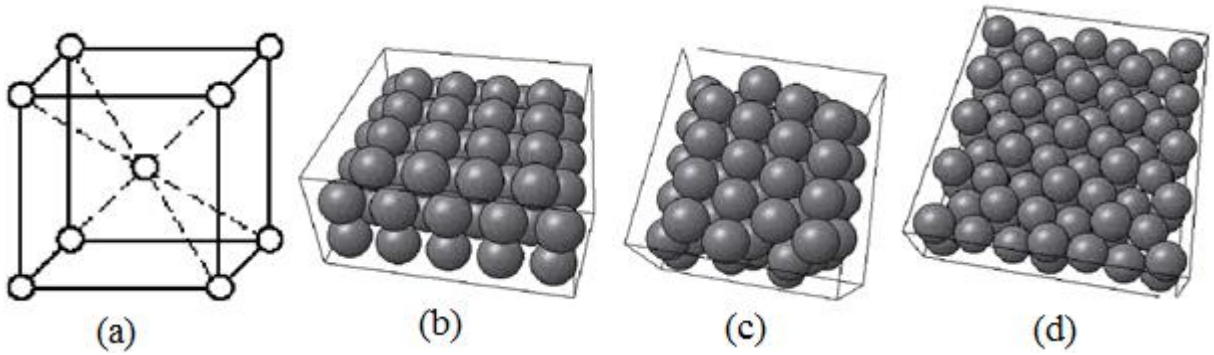
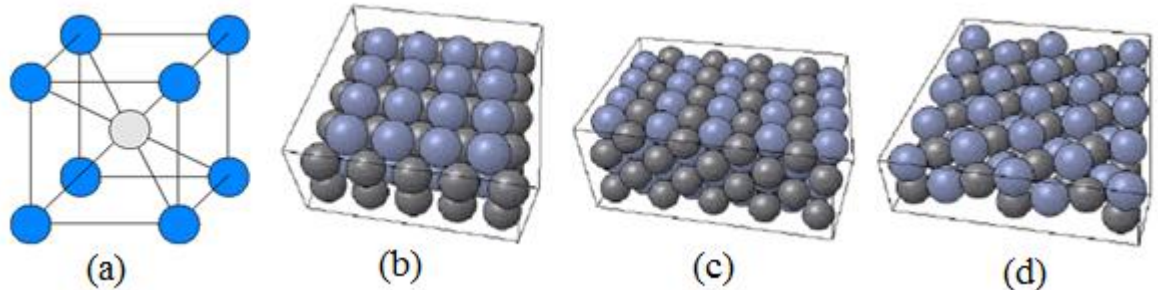



Fig. 2.2.9. The form of an elementary cell for a lattice of the CsCl (a) type and ideal surfaces of a crystal broken along the planes: b - (100), c - (110), d - (111)




As a result of this, striving for equilibrium, surface electrons will be redistributed, so that the number of dangling bonds (red arrows in Fig. 2.1.1) will be significantly reduced. In other words, the formation of dangling bonds means the readiness of unpaired electrons to form a bond either with some foreign atoms adsorbed on the surface or with atoms of the host lattice. But if the surface is pure (there are no foreign atoms on it) and smooth, then there is only one possibility for surface electrons for pairing: to form additional bonds between the surface atoms themselves. In particular, in a number of cases, neighboring atoms, forming additional bonds among themselves, can be combined into dimer pairs, so that the atoms of each dimer approach each other while simultaneously moving away from neighboring atoms on the surface that are included in other dimers (see below).

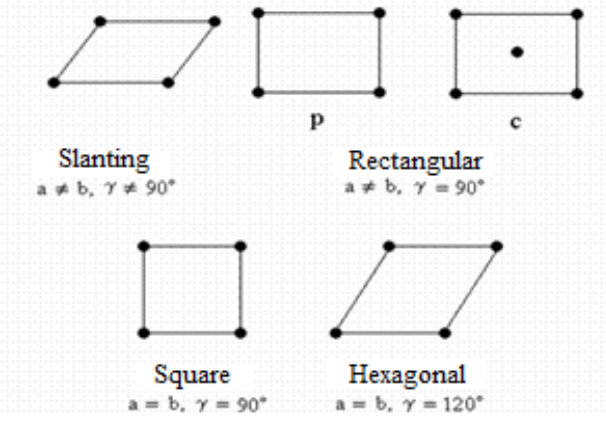
So, when an ideal surface emerges, the surface atoms form a two-dimensional periodic structure, which in general differs from the crystal structure at the initial discontinuity plane. As in the case of a three-dimensional crystal, such an ordered two-dimensional periodic surface structure of atoms can be associated with a two-dimensional spatial lattice, each node of which is given by a translation vector

$$\mathbf{R} = m\mathbf{a} + n\mathbf{b}, \qquad (2.2.20)$$

where **a** and **b** are the elementary translation vectors that define an elementary two-dimensional or surface lattice, *m* and *n* are integers.

The two-dimensional periodicity of the spatial lattice for the crystal surface allows us to classify the possible atomic structure of the surface by symmetry and arrangement of the nodes. The symmetry of tree-dimensional crystal lattices is described by 10 two-dimensional crystallographic point groups (*1, 2, 3, 4, 6, m, 2mm, 3m, 4mm, 6mm*), united in 4 two-dimensional crystalline systems (*oblique*, *hexagonal*, *rectangular* and *square*). For two-dimensional crystal lattices formed on the surface, only 5 Bravais lattices are possible (Fig. 2.2.10).

Fig. 2.2.10. Two-dimensional Bravais lattice. The indices p and c denote primitive and centered rectangular lattices, respectively




The oblique lattice with the unequal edges of the Bravais cell ($a \neq b$) and the indirect angle between the elementary translation vectors ($\gamma \neq 90^{o}$) corresponds to the point groups 1 and 2. The rectangular lattice corresponds to the point groups *m* and *2mm*. To these groups there corresponds a centered rectangular Bravais lattice, in the center of which there is a node corresponding to translation $(a + b)/2$. For such a lattice, it would be possible to choose a rhombus as an elementary primitive cell, but in this case the symmetry of the unit cell would not correspond to the complete set of crystal symmetry elements, which is the main condition for choosing the Bravais lattice. The square lattice corresponds to point groups *4* and *4mm*. The hexagonal lattice with a unit cell in the form of a rhombus corresponds to the point groups *3*, *3m*, *6* and *6mm*.

### 2.2.8. Relaxation and surface reconstruction

In the ideal case, when the crystal split does not disturb the remaining semi-infinite crystal structure, the arrangement of the atoms in the first and subsequent surface layers would have to be exactly the same as it was in the crystal bulk. However, such a situation close to ideal is realized, apparently, only for the surfaces of nonpolar dielectric compounds with a cubic structure (such as rock salt). In such compounds between the neutral atomic planes (where the anions and cations alternate) there is only a very weak Coulomb interaction and the removal of half the crystal has almost no effect on the position of the ions in the plane of this surface.

However, such surfaces are the exception rather than the rule. As mentioned above, the formation of dangling chemical bonds in the formation of a surface usually leads to a significant change in the atomic positions not only in the first surface layer, but also in several subsequent ones (see Figure 2.2.11). These changes decrease from layer to layer deep into the crystal, and at some distance from the surface the arrangement of atoms remains exactly the same as in the bulk. This applies primarily to the surfaces of crystals with a covalent bond between atoms (primarily semiconductors), due to the strict directivity of the interatomic interaction in these crystals.

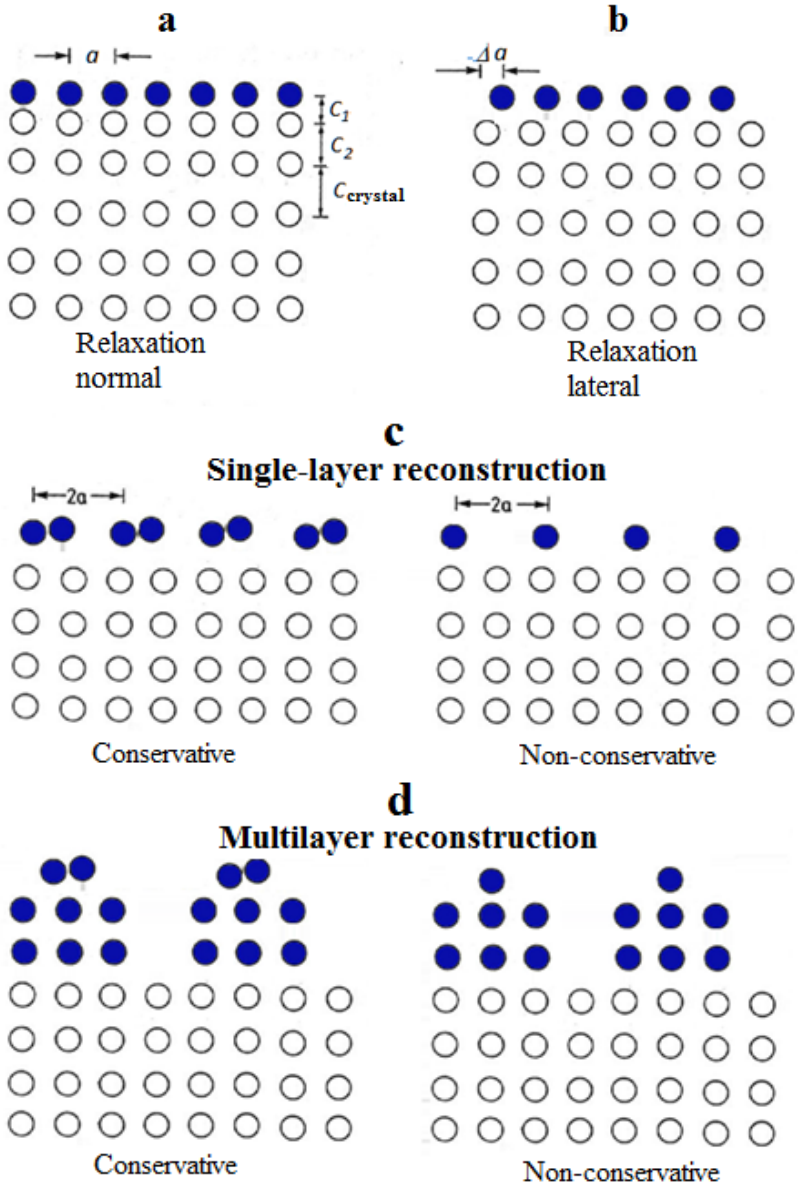
Thus, as follows from Fig. 2.2.11, in the case of rearrangement of surface atoms, it is convenient to assign a composite form to the real crystal with surface. In this case, the crystal, as it was, consists of *a substrate* having the same three-dimensional periodic structure as its bulk part and some thin layer that contains several atomic layers near the surface of the breaking and has a different structure than the crystal substrate. If the surface layer that has changed the structure consists of the same atoms as the substrate, then this layer can be called *an selvage*. If, on the other hand, the surface layer includes atoms of the external medium in which the crystal has broken, this surface layer will be called *an adsorbate*. Ideally, with all rearrangements, the selvage or adsorbate must retain a two-dimensional periodicity, along the surface.

As can be seen from Fig. 2.2.11a,b, two fundamentally different types of rearrangement of

atoms in the near-surface layer are possible: *relaxation* and *reconstruction*.

In the case of *relaxation*, the change in the edge structure can occur either due to a change in the interatomic distances only in the direction of the normal to the surface or synchronously in each of the near-surface atomic layers. Relaxation is divided into *normal* and *lateral*. *Normal*

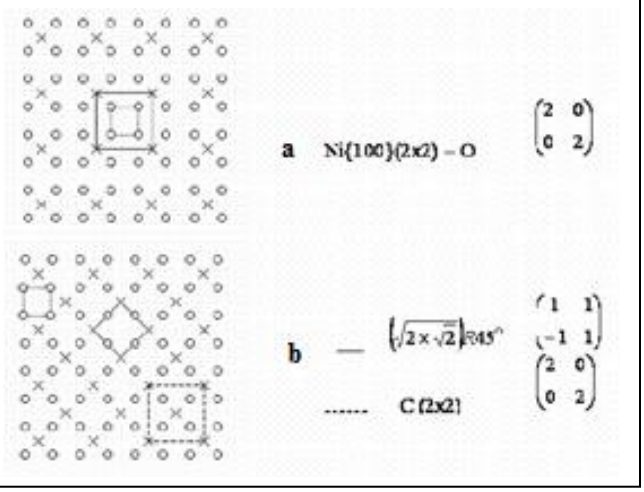
Fig. 2.2.11. Schematic representation of the rearrangement of atoms at the surface (in the cross section) in the case of relaxation (a, b) and reconstruction (c, d)




*relaxation* corresponds to the case when the atomic structure of the upper layer is the same as in the bulk, but the distance between the upper and the second layer differs from the distance between the planes in the bulk (Fig. 2.2.11a). In its pure form, normal relaxation is observed in metals. In most cases, the first interlayer distances decrease. For deeper layers, the deviation from the bulk value of the interlayer distance decreases with depth, often oscillating. On the high-index planes with a low concentration of atoms, *lateral relaxation* can be observed in the form of a uniform displacement of the upper layer parallel to the surface (Fig. 2.2.11b)

*Reconstruction of the surface* is the formation of altered translational symmetry on the pure surfaces of single crystals. In this case, the unit cell has at least one period that exceeds several times the period in the crystal bulk (in parallel planes). For this case, the displacements of atoms in the near-surface atomic layers can occur both in the direction of the normal to the surface and

Fig. 2.2.12. Examples of relaxed (a) and reconstructed (b) surfaces for the cubic lattice of nickel




in the plane of the surface edge (Fig. 2.2.11c,d). As can be seen, in the case of reconstruction, the atomic structure of the surface becomes more friable.

Theoretical consideration of reconstructed surfaces is based on quantum-chemical calculations. Calculations show that, if the free surface of homeopolar crystals is obtained, unsaturated dangling covalent bonds establish a new equilibrium configuration of surface atoms with a decreased energy of the surface. When calculating the total energy of a crystal, the dimensions of the surface unit cell are usually taken from the experiment, and the character of the bond closure is chosen by the model method.

When *conservative reconstruction*, the number of atoms is conserved and reconstruction consists only in the displacement of surface atoms from their ideal positions in the bulk. A simple example of conservative reconstruction is the above-mentioned pairing of atoms of the upper layers (the formation of dimers). This reconstruction is denoted by a 2×1 symbol, which indicates that the period along one of the directions on the surface is doubled, and along the other, it remains the same.

For *nonconservative reconstruction*, the number of atoms in the reconstructed layers differ from the bulk. A good example is a reconstruction with missing rows, in which every second atomic row is removed, in result of which half of the atoms remain in the upper layer.

Reconstruction of the 2×1 type is the simplest and often encountered in the reconstruction of the surface. In silicon and germanium, it inevitably occurs on clean, freshly prepared surfaces of crystals. Experiments show that the symmetry of the surface can change dramatically with a change in temperature, i.e. structural phase transformations occur on the surface. If such a transformation is of the type of a second-order phase transition, then one can investigate the stability of an ideal surface relative to different types of displacements of surface atoms from equilibrium positions. As a result, instead of a 1×1 or 2×1 lattice, much more complex structures such as 7×7 in silicon or 2×8 in germanium can appear.

### 2.2.9. Crystallographic methods for describing relaxed and reconstructed surfaces

As a rule, the structure of the selvage or adsorbate formed during the relaxation or reconstruction of the surface is coherent with the substrate structure. Therefore, it is convenient to describe the structures of the upper near-surface layers by associating the structure of the adsorbate (selvage) with the Bravais lattice of the substrate.

The most universal is the matrix method for designating the structures of the upper atomic layers. It is based on the fact, that the adsorbate (selvage) lattice with basic translation vectors **a'** and **b'** can always be linear combination of basic translation vectors **a** and **b** of the substrate lattice:

$$\begin{aligned} \mathbf{a}' &= G_{11}\mathbf{a} + G_{12}\mathbf{b} \\ \mathbf{b}' &= G_{21}\mathbf{a} + G_{22}\mathbf{b} \end{aligned}. \qquad (2.2.20)$$

In this case, the coefficient matrix

$$G = \begin{pmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{pmatrix}. \qquad (2.2.21)$$

can serve to indicate the structure of the upper layer (here the indices i,j = 1,2 for $G_{ij}$)

In the second method, the ratio of the *p* and *q* periods of the cells of the adsorbate (selvage) and the substrate, as well as the angle $\varphi$, to which one of the cells must be rotated, is used to designate the reconstructed surfaces, so that both pairs of basis translation vectors are aligned along one line. Thus, if the adsorbate *A* on a surface with the type {hkl} of material *X* forms a reconstructed structure whose modulus of basis vectors *a'= pa* and *b'= qb* (*a* and *b* are the moduli of the basic vectors of the substrate), and $\varphi$ is the rotation angle of the elementary cells, then the designation ("passport") of this structure has the form

$$X\{hkl\}(p \times q)R\varphi^0 - A \qquad (2.2.22)$$

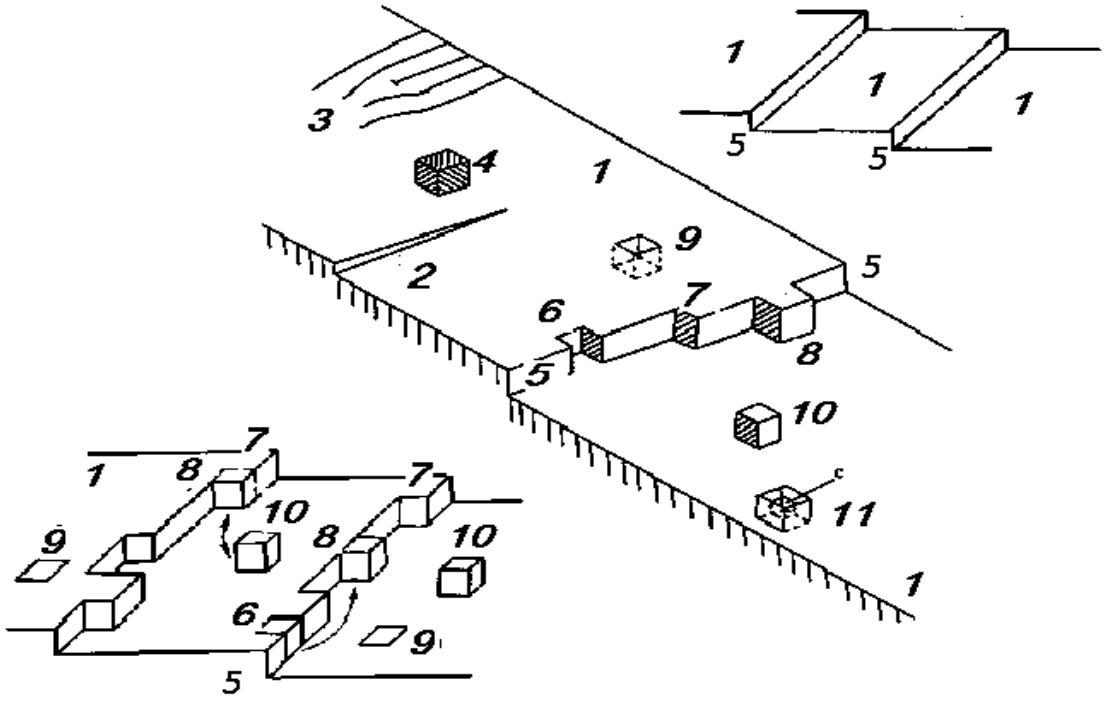
This method of designations is less universal than the first because it requires that the angles of rotation of both pairs of elementary cells basic vectors of the adsorbate (selvage) and the substrate are the same. Examples for this type of structure designation are shown in Fig. 2.2.12. Atoms of the substrate are shown in circles, and the adsorbate (selvage) - with crosses. Dashed and solid lines denote the Bravais lattices of the substrate and adsorbate (selvage), respectively.

### 2.2.10. Structural defects on ideal surfaces

It is clear from the previous presentation, that under the surface we understand a certain, even atomically thin, but still a three-dimensional layer. The thicknesses of such layers can reach even a dozen periods of the crystal lattice. Therefore, the surface structure of a real crystal is very complicated. In addition, since the surface is in contact with the external atmosphere, oxygen plays a significant role in the formation of its structure. Therefore, in the surface layers of solids, there are almost always natural oxide layers 5-200 nm in thickness. In addition to them, because of the phenomenon of adsorption on the surface and in the near-surface layers, extraneous impurity atoms of nitrogen, carbon, water molecules, etc. are present.

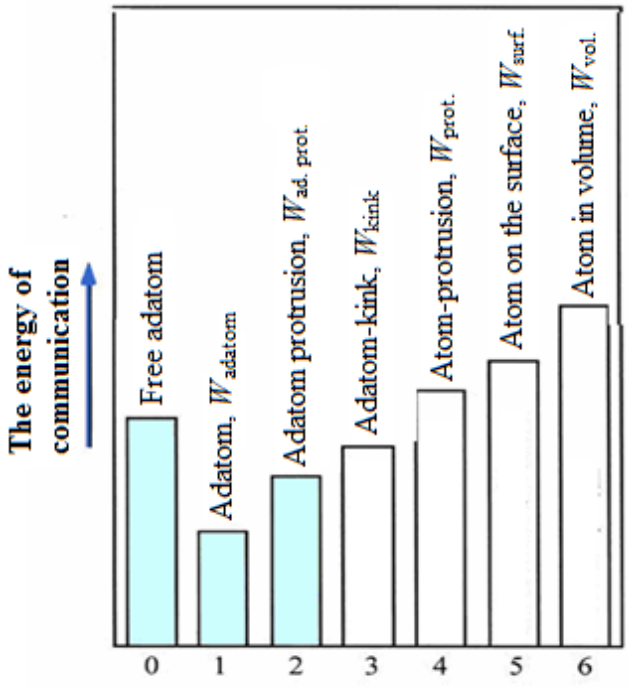
Consequently, an ideal surface crystal structure, as well as an ideal bulk crystal lattice, is not realized in practice. In the structures of real surfaces, there are always defects. As a result, the surface of a solid is essentially heterogeneous by its structure.

Fig. 2.2.13. The main types and processes of formation of simple defects on real surfaces with low values of Miller indices: 1 - terrace (the perfect flat part of the surface of the crystal face); 2 - the output of the screw dislocation on the terrace surface; 3 - intersection of an edge dislocation with a terrace; 4 – adatom of impurity on the terrace; 5 - monatomic step; 6 - vacancy in the step; 7 - ledge in the step (kink); 8 - adatom on the step; 9 - vacancy in the terrace; 10 - adatom on the terrace; 11 - vacancy in the terrace with captured electron (the color center in alkali-halide crystals); 6-8 – fluctuation-like mechanism of kinks formation; 8-10 - exchange of atoms between the kink and the adsorbed atom.




*Surface defects*, as well as bulk defects, are classified according to their dimensions. A schematic representation of the main types of surface defects is shown in Fig. 2.2.13.

Fig. 2.2.14. Relative binding energies for atoms in different positions of the crystal lattice within the framework of the nearest-neighbor model for the (100) surface of the crystal with a simple cubic lattice




To the surface defects, we should first of all refer to the *terraces* 1, which are part of the flat face of the crystal. Two terraces are separated by *steps* 5. On them, due to fluctuations 6-8, *kinks* can be formed (shown by an arrow) at $T > 0$. As can be seen, monatomic steps 5 and kinks 7predominate in most cases. Important linear surface defects are also *outputs of edge* 3 and *screw* 2 *dislocations*.

*Adatoms* are the most important point defects of the surface. The adsorbed atom can be attached to an empty place (node) on the terrace 4, as well as a kink or free step 8. *Surface vacancies* can also appear on the stap 6.

In non-single-atomic crystals (GaAs, ZnO, NaCl, etc.) adatoms can be of various types (in the form of Ga or As atoms on the GaAs surface). On the surface of GaAs, there can be two types of vacancies - Ga and As, which can be found on terraces (11) and on steps (6). Since the charges of vacancies are different, they significantly affect the electrical and optical properties of the crystal surface, especially if they can themselves capture electrons (such as color centers 11 in alkali-galloid crystals).

The type of defects that are formed in the surface is determined by the difference in the binding energies of atoms with the surface. This is due, first of all, to the fact that the number of neighboring atoms with which this atom interacts in different places of the surface is different.

Let us consider the energetics of the formation of some simple defects on an atomically clean and smooth surface in a state of thermodynamic equilibrium above the absolute zero of temperature in the presence of terraces on the surface. Surface defects are similar in terms of energy to such defects observed in bulk crystalline solids, like interstitial atoms, vacancies and Schottky defects.

We will proceed from the following statements:

1. To add or remove an atom from a crystal lattice site and move it to another location, an energy equal to the difference in the number of bonds present in the initial state and arising in the final state is required.

2. Atoms in a bulk crystal with a simple cubic lattice can always be connected with six nearest neighbors.

3. It is believed that all bonds in the crystal are the same.

In accordance with the statement 1, the change in the free Gibbs energy upon the transition of the atom from the terrace to another site of the crystal lattice can be described by the relation

$$\Delta G = W_{defect} - W_{terrace}. \qquad (2.2.22)$$

Let us analyse the dependence of the binding energies on the type of surface defects formed shown in Fig. 2.2.14. Since all bonds in the crystal are considered to be the same (statement 2), to move the atom from the position on the terrace to, for example, a position on the step, it is required to break 5 bonds and then form 3 bonds. It follows from (2.2.22), that the change in the free Gibbs

energy corresponds to the energy of the two bonds.

Fig. 2.2.15. Stepped structures of a real atomically smooth surface (cross section):
a - monatomic step, b - diatomic step, c - faceted surface

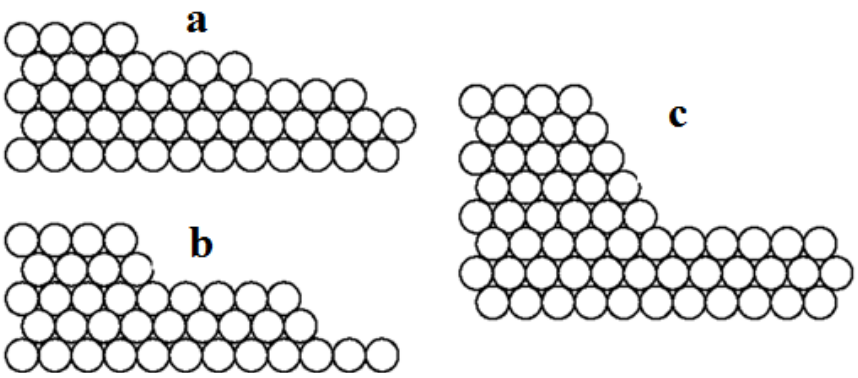
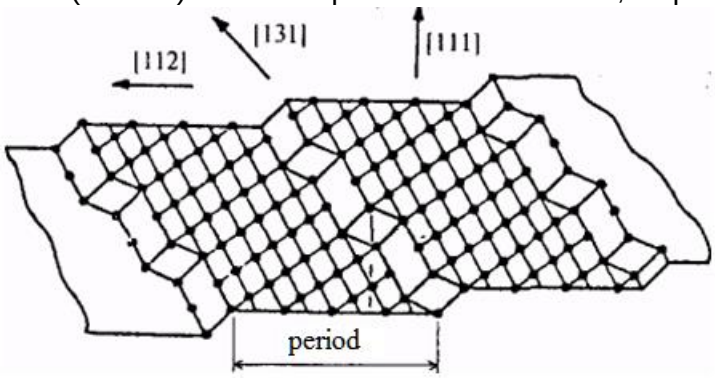



Fig. 2.2.16. The vicinal (faceted) surface of platinum with terraces, steps and kinks




Studies have shown that the surface selvage is often not two-dimensional (2D), but three-dimensional (3D). In this case, its structure, as a rule, is inhomogeneous and strongly depends on the method of obtaining the surface and its purity.

As noted above, one of the ways to obtain an atomically pure and atomically smooth surface is to split the crystal in an ultrahigh vacuum or a chemically neutral medium along cleavage planes. However, a small (by several degrees) deviation from this plane by an angle leads to the formation of so-called vicinal surfaces containing steps, dislocation outputs, and other defects. In this case, the height of the steps increases with increasing deflection angle $\theta$, leading to the formation of not only monatomic but also two-, three- and multiatomic steps (Figures 2.2.15 and 2.2.16).

### 2.2.11. Experimental data on the structure of surfaces of semiconductors and metals

have shown that the surfaces of most The atomic structure of atomically clean and smooth surfaces is established by the method of diffraction of slow electrons (DSE), scanning tunneling microscopy (STM), and electron microscopy of atomic resolution. With the help of these studies, the features of the atomic structure of the relaxed or reconstructed surfaces of a number of single crystals were established.

Studies metals are susceptible to relaxation and can not be reconstructed. In this case, normal surface relaxation prevails for metals with low index planes. In Fig. 2.2.17, an example of a low-index metal surface after normal relaxation is shown for aluminum surface (110) with a face-centered lattice. Like most metals, the interatomic distance $d_{12}$ in the first layer of atoms is less than the period $a$ in the depth of the crystal lattice. The magnitude of this compression, normalized

by the value of the interlayer distance in the bulk of the crystal, $\Delta_{12} = (d_{12}-a)/a$, is -8.6 %. In the general case, for fcc metals this value is in the range from zero to several tens of percent, and it is

Fig. 2.2.17. A schematic top and side views of the relaxed surface (110) of Al single crystal

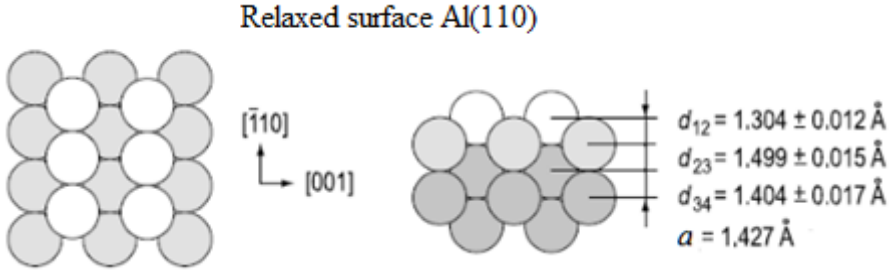



Fig. 2.2.18. Schematic representation of the hexagonal packing of the upper layer atoms in platinum single crystal (shown by gray circles) on the Pt (100) atomic plane with the original square lattice (shown by white circles).

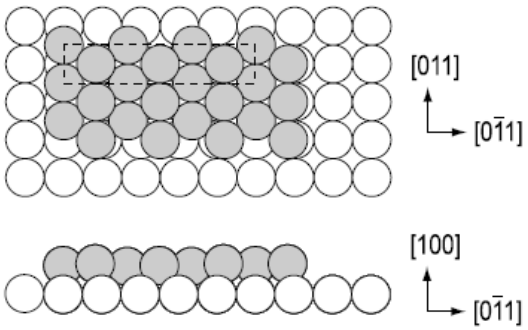
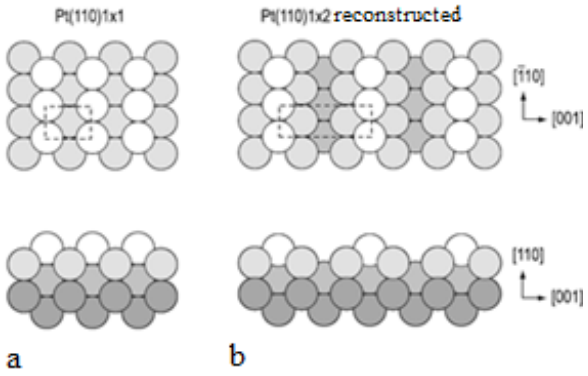



Fig. 2.2.19. Schematic representation of the atomic structure of the Pt (110) surface:
a - ideal unreconstructed; b - reconstructed with the structure of missing rows (missing-row structure).

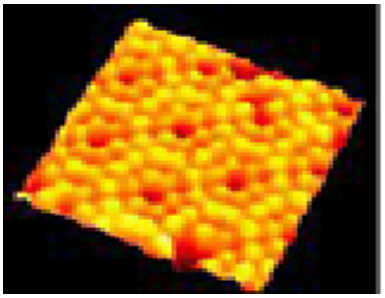



higher for surfaces with a low packing density of atoms. The deviation of the interlayer distance from the bulk value decreases with depth, often oscillating. In the case of the A1 (110) surface, the second interlayer distance is stretched by +5.0 %, and the third is again compressed, albeit slightly, by -1.6 %.

An example of reconstructed platinum surfaces with a face-centered lattice is shown in Fig. 2.2.18 and 2.2.19. On an ideal unreconstructed surface (100) a untight quadrant lattice is formed, which is under the influence of tensile stresses and, therefore, is unstable. Calculations show that a more dense packing of atoms is more preferable. The experiments showed that, in fact, this surface is reconstructed into almost close-packed hexagonal layer whose atomic density is ~ 20 % lower (Fig. 2.2.18). Another consequence of the reconstruction of the upper atomic layer is the modification of its interatomic bonds with the atoms of the underlying layer. Thus, the reconstruction of the platinum surface is controlled by the balance between the gain in energy due to the increase in the packing density of atoms, and the loss in energy due to mechanical stresses due to the mismatch between the lattices of the upper and lower layers.

As can be seen from Fig. 2.2.19, the ideal (110) face for a platinum single crystal consists of atomic rows along the [110] direction. For a real atomically clean Pt (110) surface (prepared, for example, by ionic spraying followed by annealing), a doubling of the period along the [001]

Fig. 2.2.20. STM image of the reconstructed Si (111) 7 × 7 surface



direction was observed. It was found that this doubling is due to the absence of every second row on the surface.

Fig. 2.2.20 shows the structure of the silicon surface obtained by the STM method in the case of its splitting along the (111) plane. It turned out that after reconstruction it is characterized by a period exceeding the period of the bulk lattice by a factor of 7. Such a surface according to the described above classification is denoted as Si (111) -7×7.

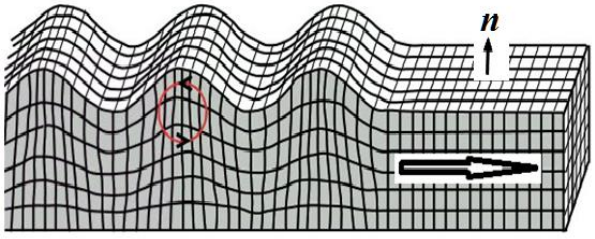
## 2.3. Vibrating of atoms on atomically clean surfaces

Since the surface of an ideal crystal can be considered as a new phase (selvage or adsorbate), its spectrum of atomic vibrations should differ from the bulk one. In this section, elementary excitations of a semi-infinite crystal, related to the correlated vibrational motion of ions, whose amplitude falls off exponentially with the deepening into the crystal, will be considered. This problem will first be considered on the basis of the Rayleigh approach in the approximation of a semi-infinite continuous medium.

### 2.3.1. Elementary excitations of a semi-infinite crystal in the approximation of a continuous medium (Rayleigh waves)

Rayleigh, considering a solid as a continuous medium, predicted the appearance of elastic waves on the surface as the boundary of a semi-infinite crystal with a vacuum or a rarefied gas.)

Fig. 2.3.1. Image of Rayleigh wave




The phase velocity of the Rayleigh waves is parallel to the surface (Figю 2.3.1), and the particles of the medium vibrating near it have both transversal (normal to the surface), and longitudinal components of the displacement vector. As a result, the particles of the medium in such a wave make elliptical trajectories in the plane in which the phase velocity vector and the normal ***n*** to the surface lie (this plane is called the *sagittal* one).

The amplitudes of longitudinal and transversal oscillations decrease the depth according to exponential laws with different attenuation coefficients. Therefore, the energy of the Rayleigh wave is concentrated in the region at a distance of the order of the wavelength from the surface. These surface density waves are naturally divided into 4 categories: a) long-wave elastic waves, b) long-wave optical waves, c) short-wave acoustic waves and d) short-wave optical waves.

Let us consider long-wave elastic surface waves within the framework of the classical theory of elasticity for a continuous medium. We denote by $u = x - x_o$ the displacement from the position of equilibrium for some element of a semi-infinite isotropic continuous medium. In general, this vector can be represented as the sum of the transversal ($\nabla \cdot u_t = 0$) and longitudinal ($\nabla \times u_l = 0$) displacements. The Cartesian components of the vectors $u_t$ and $u_l$ satisfy the usual wave equation taking into account the corresponding longitudinal $c_t$ and $c_l$ and the transverse component of the sound velocities:

$$\frac{\partial^2 \mathbf{u}_t}{\partial t^2} - c_t^2 \nabla^2 \mathbf{u}_t = 0,$$

$$\frac{\partial^2 \mathbf{u}_l}{\partial t^2} - c_l^2 \nabla^2 \mathbf{u}_l = 0. \qquad (2.3.1)$$

In the bulk, these mutually perpendicular displacements are not connected with each other. Near the surface, we seek the solution of equations (2.3.1) in the form

$$\mathbf{u}_l = \mathbf{a}_l e^{i(q_\parallel x - \omega t)} e^{-\varkappa z}, \qquad (2.3.2)$$

where $\kappa = (q_{II}^2 - \omega^2/c^2)^{1/2}$ is the damping constant of the surface wave along the surface plane (positive magnitude). We note that these waves are macroscopic in the sense that in the limit $q_{II} \to 0$ the displacement continues to exist in the bulk at a sufficiently large depth.

As always, the nature of the solution is determined by the boundary conditions. In an infinite medium, equilibrium forces are emitted by forces that balance each other on each side of an infinitesimal volume element. On the surface, all the components of the forces crossing the boundary plane must vanish:

$$dF_i = \sigma_{iz} dA_z = 0. \qquad (2.3.3)$$

The relation (2.3.3) is the *1st boundary condition*, which reflects the requirement that the displacement of the surface volume elements occurs only in the *xy* plane containing the direction of propagation and the normal to the surface.

We assume that the components of the stress tensor $\sigma_{ij}$ are related to the components of the strain tensor $\varepsilon_{ij} = (\partial u_i/x_j + \partial u_j/x_i)/2$ by Hooke's law. In an isotropic medium, the elastic constants that play the role of the proportionality coefficients can be written in terms of Young's modulus *Y* and Poisson's ratio *P*. Thus, taking into account the condition (2.3.3), we obtain *the 2nd boundary condition*

$$\sigma_{yz} = \frac{Y}{1+P}\delta_{yz} = 0 \to \frac{\partial u_y}{\partial z} = 0 \to u_y = 0. \qquad (2.3.4)$$

It reflects the requirement that the displacement of surface modes occurs only in the sagittal plane, i.e. the plane containing the direction of propagation and the normal to the surface.

Since, by definition, it follows from (2.3.1) that

$$\nabla \cdot \mathbf{u}_t = iq_\parallel u_{tx} - \varkappa_t u_{tz} = 0,$$
$$(\nabla \times \mathbf{u}_l)_y = \varkappa_l u_{lx} + iq_\parallel u_{lz}, \qquad (2.3.5)$$

After simple arguments it follows that the components of the total displacement for the Rayleigh wave are equal to:

$$u_x = \left(a\varkappa_t e^{-\varkappa_t z} + bq_\parallel e^{-\varkappa_l z}\right) e^{i(q_\parallel x - \omega t)},$$
$$u_z = \left(aq_\parallel e^{-\varkappa_t z} + b\varkappa_l e^{-\varkappa_l z}\right) i e^{i(q_\parallel x - \omega t)}. \qquad (2.3.6)$$

Using *the third* and *fourth boundary conditions* $\sigma_{xz} = \sigma_{zz} = 0$, for solving the equations (2.3.1), it is easy to show that the Rayleigh wave of type (2.3.5) is acoustic with the linear dispersion law

$\omega = c_R \cdot q_{II}$, where the sound speed $c_R$ is determined by bulk elastic constants. From the fact that $\kappa > 0$, it follows unambiguously that $c_R$ has a smaller value than $c_t$ and $c_l$. Such a "slow" surface acoustic mode is called a *Rayleigh wave*.

With a more complete consideration, it should be borne in mind that in real substances, elastic constants are usually anisotropic. In this case, we obtain either the exactly monotonically decreasing solution, like (2.3.6) or a generalized Rayleigh mode with subcritical damping, the amplitude of which falls off as we penetrate deep into the volume according to the sinusoidal law.

Another class of long-wave surface modes exists in ionic compounds, where the transversal (TO) and longitudinal (LO) optical waves known from the atomic dynamics of the bulk crystal lattice cause the appearance of an oscillating dipole moment in each unit cell of the crystal. For the optical mode, any cell can be characterized by an effective charge, which describes the change in the dipole moment with the displacement of the atoms

$$e_{ij}^* = \frac{\partial M_i}{\partial u_j}. \tag{2.3.7}$$

In this case, the optical properties of a solid are well described through dielectric permittivity:

$$\varepsilon(\omega) = \varepsilon(\infty) + \frac{\Omega^2}{\omega_{TO}^2 - \omega^2}, \qquad \text{где} \quad \Omega^2 = \varepsilon(\infty)\frac{4\pi e^{*2}}{M a_0^2}. \tag{2.3.8}$$

There is another analog of bulk optical modes, which are concentrated in the near-surface region. The macroscopic fields generated by this mode must be specified by the electrostatic potential exponentially decreasing with distance from the surface into the depth.

Carrying out the same arguments as before, we find that the surface mode satisfies the equality $\varepsilon(\omega) = -1$, where the dielectric constant is given by the formula (2.3.8). As a result, we have

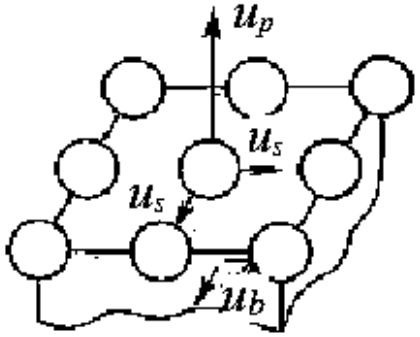
$$\omega_s = \sqrt{\frac{\varepsilon(0)+1}{\varepsilon(\infty)+1}}\,\omega_{TO}. \tag{2.3.9}$$

The existence of surface elastic waves in the ultra- and hypersonic range is not only proved experimentally, they are widely used in acoustoelectronics.

### 2.3.2. The vibrations of atoms near the surface

As a result of the formation of dangling bonds and the absence of "neighbors" from the vacuum side of the surface atoms (see, Fig, 2.1.1), the harmonic and anharmonic components of interatomic forces of interaction are changed.

Fig. 2.3.2. Scheme of the relationship between the root-mean-square values of different composing vibrational displacements of atoms




The harmonic forces of interaction between atoms on the surface when they make thermal oscillations can be characterized by the mean square displacement of the atoms $\langle u^2 \rangle$ from their nodes. The values of $\langle u^2 \rangle$ and the anisotropy of the surface atoms vibrations can be estimated from the DSE data. Fig. 2.3.2 shows schematically the relationship between the mean-square values of the various components of the displacements of atoms from the equilibrium position at the surface ($u_p$ and $u_s$) and in the bulk of the crystal ($u_b$).

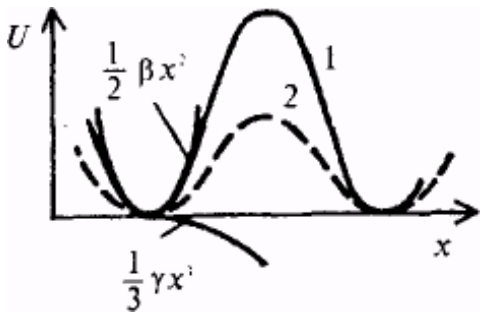
In the harmonic approximation, the displacement is related to the force acting on the atom by the relation $F = -\beta u$, where $\beta$ is the elastic constant. As will be shown below, at temperatures above the Debye $\langle u^2 \rangle \sim T/\beta$. Hence, from the experimentally measured values of $\langle u^2 \rangle$, it is possible to estimate the differences between the elastic forces on the surface and in the bulk of the crystal.

We elucidate the foregoing with the simplest example of the dynamics for a one-dimensional monatomic lattice (a chain of atoms). As is well known, the dependence of the potential energy on the displacement of atoms from the equilibrium position $x$ at the absolute zero temperature (0 K) can be expressed as a series

$$U(x) = U_0(x) + \frac{1}{2}\beta x^2 - \frac{1}{3}\gamma x^3 + \frac{1}{4}\delta x^4 - \cdots \qquad (2.3.10)$$

where the first term $U_0(x)$ is the energy of interatomic interactions, the second term characterizes the harmonic vibrations of atoms, the third term takes into account the asymmetry of the repulsive potential of neighboring atoms, etc. The minimum of the function $U(x)$ at $x = 0$ is well described by the interatomic Lennard-Jones potential. The contributions of the harmonic and first anharmonic terms in $U(x)$ for the crystal volume are shown in Fig. 2.3.3 with curves 1 and 2, respectively.

Fig. 2.3.3. The dependence of the potential energy $U$ of atoms on the surface (1) and in the bulk (2) of the crystal




We firstly consider the harmonic approximation. Using the Debye model of the phonon spectrum of a bulk crystal $g(\omega_q)$, the mean-square amplitudes of the displacements of atoms with thermal vibrations $\langle u^2 \rangle$ are represented in

$$\langle u^2 \rangle = \int_0^{\omega_D} \langle u_q^2 \rangle n_q g(\omega_q)\, d\omega_q \,, \qquad (2.3.11)$$

where $\omega_D$ is the limiting frequency of the Debye phonon, $\langle u^2 \rangle$ is the mean square displacement of atoms for normal modes with a frequency $\omega_q$ and momentum $q$; $n_q = n + 1/2$, and

$$n = exp[(\hbar\omega_q)/(kT - 1)]^{-1} \qquad (2.3.12)$$

is the number of phonons with energy $\hbar\omega_q$ in accordance with Bose-Einstein statistics and $g(\omega_q)$ is the Debye density of vibrational states of atoms (phonon spectrum). Since the average energy of a classical oscillator with frequency $\omega_q$ equals

$$\hbar\omega_q = M\omega_q^2\langle x^2\rangle, \tag{2.3.13}$$

where *M* is the mass of the atom, for the mean square displacements of atoms from the position of equilibrium, we obtain

$$\langle u^2\rangle = \frac{\hbar}{M\omega_q}, \tag{2.3.14}$$

At high temperatures $T > \theta_D$, where

$$\Theta_D = \frac{\hbar\omega_D}{k} \tag{2.3.15}$$

is Debye temperature, we get from (2.3.12) the value

$$n_q \cong kT/\hbar\omega_q. \tag{2.3.16}$$

As a result, the integrating of (2.3.11) for $T > \theta_D$ leads to the well-known relation

$$\langle u^2\rangle = \frac{3kT}{M\omega_q^2} = \frac{3\hbar^2 T}{Mk\Theta_D^2}. \tag{2.3.17}$$

As noted above, in the harmonic approximation, $\omega^2 = (\beta/M)$. Therefore

$$\langle u^2\rangle = \frac{3kT}{\beta} \tag{2.3.18}$$

The amplitudes of zero vibrations of atoms at low ($T < \theta_D$) temperatures do not depend on temperature

$$\langle u^2\rangle \approx (M\beta)^{1/2} \tag{2.3.19}$$

According to the theory of DSE, the intensity of the diffraction lines is proportional to the square of the number of ordered surface atoms. Since thermal vibrations disorder the surface, the intensity *I* of the reflexes should be weakened. According to the Debye-Waller model, at high temperatures ($T > \theta_D$), for specular reflection, the intensity is weakened by the law

$$I = I_o \exp[-\beta \langle u^2\rangle], \tag{2.3.20}$$

where exp {-β $\langle x^2\rangle$} is the Debye-Waller factor. Thus, by determining experimentally the attenuation factor, we can use the relations (2.2.20) and (2.3.17) to estimate the values of $\langle u^2\rangle$ and the Debye surface temperature $\theta_D$.

The root-mean-square displacement of the atom on the surface for different scattering angles can be decomposed into normal (*s*) and tangential (*p*) components:

$$\langle u^2\rangle = \langle u_s^2\rangle \cos^2\xi + \langle u_p^2\rangle \sin^2\xi, \tag{2.3.21}$$

with $\langle u^2\rangle = \langle A^2\rangle/2$, where *A* is the amplitude of the oscillations.

A large number of data on the estimates of $<u_s^2>$, $<u_p>$ and $\theta_D$ by the DSE method for a large number of objects have been published in the literature: on a variety of faces of crystals of the 4-th group of the periodic table, some metals, complex semiconductor compounds of the type, $A^3B^5$, $A^2B^6$, $A^4B^6$ etc.). The overall result of these calculations is as follows. For surface lattices, there is an anisotropy of oscillations of the type $<u_з^2> > <u_p^2> > <u_b^2>$. In this case, the values of $<u_s^2>$ are 2-4 times larger than displacements $<u_b^2>$ in the bulk, whereas $<u_p^2>$ are 1.5 times larger than $<u_b^2>$. On the other hand, $\theta_D^s \approx 0{,}5\theta_D^b$, where the $\theta_D^b$ values for diamond are 1860 K, for Si - 625 K, Ge - 360 K, W - 310 K, Pt - 230 K). Qualitatively, these results indicate a weakening of the elastic harmonic forces on the lattice surface in comparison with the bulk (due to the greater loosenes of the crystal lattice at the surface due to relaxation or reconstruction).

As can be seen, the greatest influence of the surface is manifested in the variation of the normal components. The fact of growth of the atomic vibrations amplitude at the surface indicates the need to take into account nonlinear anharmonic terms in the expression for the potential energy (2.3.10). And the point here is not only in quantitative corrections, which are estimated by 20-30 % to increase the values of $<u^2>$ and $\theta_D^s$ calculated from formulas (2.2.20) and (2.3.17). It is much more important to establish at least qualitatively the mechanism of phonon exchange between the surface phase (selvage or adsorbate) and the bulk of the crystal (substrate), which determines many of the thermal and electrophysical properties of the surfaces.

Let us dwell for simplicity on the first anharmonic term $\gamma x^3$ in the expansion (2.3.10). In the Boltzmann approximation, the probability of the atom deviating from the equilibrium position at the lattice site

$$W(x) = A \cdot exp(-U(x)/kT). \tag{2.3.22}$$

Substituting the first two terms of the expansion (2.3.17) instead of $U(x)$, we can estimate the average deviation of the atom from the equilibrium position

$$\langle x \rangle = \int_0^\infty x W(x) dx \cong \frac{\gamma k T}{\beta^2}. \tag{2.3.23}$$

Comparing with the relation (2.3.17), we see that the anharmonicity coefficient $\gamma$ renders influences the value $<u>$, and consequently also $<u^2>$.

### 2.3.3. Influence of surface atoms vibrations on the thermal properties of the crystal surface

The anharmonicity of lattice vibrations is decisive in such phenomena as thermal expansion and thermal conductivity. Naturally, all this applies to the surface phases (selvage or adsorbate). In accordance with the relation (2.3.23), the coefficient of thermal expansion

$$\alpha^s = \frac{\langle x \rangle}{d_0 T} = \frac{\gamma k}{a \beta^2}, \tag{2.3.24}$$

where $d_o$ is the lattice period. As follows from (2.3.24), the quantity $\alpha^s$ increases with the anharmonic contribution ($\gamma$).

The thermal conductivity of solids is determined by the contribution of the electron $\chi_e$ and lattice $\chi_{latt}$ components. For metals, $\chi_e >> \chi_{latt}$ and $\chi$ is calculated in the free-electron approximation by the Wiedemann-Franz formula. The lattice component $\chi_{latt}$ depends in a complex manner on the temperature $T$, passing through a maximum at a temperature much lower

than the Debye temperature (for example, for Ge it is observed at 20 K). Such a course of temperature dependence is due to two competing processes. At low temperatures (below the Debye temperature), the thermal conductivity of the lattice grows because of the increase in the phonons number (with increasing heat capacity), and at higher temperatures, $\chi_{latt}$ falls as a result of inelastic phonon-phonon scattering (*umklapp processes*). In theory, such processes are described by the anharmonic term $\gamma \cdot x^3$. The calculation shows that the value of the lattice component of the thermal conductivity depends not only on the elastic lattice constant ($\beta$), but also on the anharmonicity coefficient of the vibrations of surface atoms ($\gamma$):

$$\chi_{latt} = \frac{\beta^{7/2}}{\gamma^2 T}. \qquad (2.3.25)$$

Estimates of the coefficient of thermal expansion ($\alpha^s$), based on measurements of the intensity of DSE reflexes (relation (2.2.20)), show that the normal component of $\alpha^s$ is 5-20 times larger than the corresponding value for volume ($\alpha^b$). The coefficient of thermal conductivity in the surface phases of $\chi^s_{lat}$ is at least two or more times less than for the volume one. Naturally, the data on $\alpha^s$ and $\chi^s_{lat}$ are values averaged over a thin near-surface layer (~ 1 nm). The values of $\alpha^s$ grow, and $\chi^s_{lat}$ decrease when approaching the surface. The hindered exchange of vibrational energy between the surface and the volume leads to the accumulation of energy in the surface phase, which is crucial for the flow of atomic and electronic processes, which will be discussed below.

Optical phonons play an important role in the surface dynamics of ionic crystals, leading to the appearance of oscillations of dipole moments associated with the polarization of the medium. Indeed, surface optical phonons with an energy of tens of meV have been detected by some methods of electron spectroscopy for some oxides (for example, ZnO). Optical phonons were also found on the surface of Si (111) 2 x 1 (covalent crystal).

It is shown that the zone of surface phonons enters the forbidden region between the acoustic and optical branches of bulk phonons and partially is embedded in them. As a result, resonance states are formed that change the phase of plane waves of normal lattice vibrations. The calculations lead to the conclusion that the local modes of a partially disordered surface are weakly related to the "thermal phonon bath" of the crystal. The theory directly indicates the possibility of the appearance of excess energy in the surface phases.

As a result, due to reconstruction of near-surface atomic layers, large amplitudes of vibrations of atoms on the surface, and a difficult exchange of vibrational energy with volume create favorable conditions for all kinds of structural rearrangements. One type of such rearrangement with increasing temperature is melting (order-disorder phase transition), which starts from the surface at a temperature $T^s_m < T^b_m$, where $T^s_m$ and $T^b_m$ are the melting points in the surface or in the bulk correspondingly. For a rough estimate of the onset of mobility of lattice structural elements, the Tamman temperature $T_T = (0.3\text{-}0.5)\ T_m$ is often used.

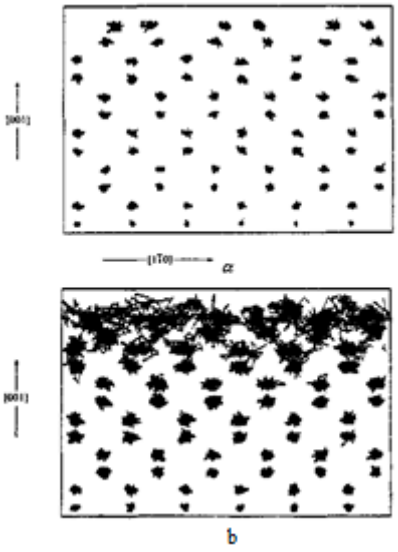
Possible reasons for the earlier melting of the crystal surface were explained by Lindemann. He suggested that the crystal begins to melt when the rms displacement of its atoms $\langle u^2 \rangle$ reaches a certain critical value

$$\sqrt{\langle u^2 \rangle} = c \cdot d_o \ , \qquad (2.3.26)$$

where $c \approx (0.1 \div 0.2)$ depending on the type of lattice. In the Debye approximation, according to (2.3.11), the melting temperature of the near-surface layers of atoms is

$$T_m^s \cong c d_0^2 \frac{Mk\Theta_D}{3\hbar^2}. \qquad (2.3.27)$$

Fig. 2.3.4. A model picture of the onset of near-surface melting of silicon, obtained by the molecular dynamics method. The particle trajectories (side view) are shown at T = 1003 K (a) and at T = 1683 K (b).



As noted above, the Debye temperature of the surface phase is almost half that of the bulk phase, and therefore, $T^s_m < T^b_m$.

Thus, the simplified application of the Lindemann criterion allows us to conclude that the first layer of atoms in the near-surface region of the crystal lattice is disordered ("melts") at a much lower temperature than the bulk phase. Accordingly, the second layer melts at a slightly higher temperature than the first surface layer, but still less than the bulk melting point. A similar reasoning applies to the third layer, and so on. Each layer is melted unevenly, as soon as the local Lindemann conditions are realized for it. As follows from the results of numerical simulation of this process (see Fig. 2.3.4.), the fusion front extends into the depth of the crystal, and the melting point with each layer increases until the process terminates at $T^b_m$.

A clear presentation of the atomic vibrations trajectories for a free Si (100) 1×1 surface (Fig. 2.3.4) is given by the molecular dynamics method for such a phase transition. It is clearly seen from the figure that at a certain higher temperature, which is below the bulk melting temperature, the first surface layer of atoms is disordered more strongly than the deeper layers. At the same time, in accordance with the considerations given above, the motion of the fusion front to the depth of the crystal is actually observed with increasing temperature.

The dependence of the thickness d of the molten layer on temperature is described for metals by the relation

$$d \sim ln(T_m - T)^{-1}, \qquad (2.3.28)$$

but for molecular crystals

$$d \sim ln(T_m - T)^{-1/2}. \qquad (2.3.29)$$

## 2.4. Electronic structure of atomically pure surfaces in crystals

In this section, we will consider those changes in the electronic structure of solids, which are associated with the appearance of special charged states named Surface Electronic States (SES) either due to the discontinuity of the three-dimensional periodicity of the bulk crystal or additional modifications due to reconstruction and relaxation of the surface. As a result, the electrical properties (including zone structure) near the surface are very different from those in the bulk. The above changes in the electronic surface structure also affect many optical, magnetic, diffusion, catalytic and other properties of the surface and crystal as a whole.

In this regard, to begin with, we confine ourselves here to the most important physical questions:

- How are chemical bonds rearranged in the first few atomic planes after the crystal split?
- What is the charge density in the neighborhood of the crystal boundary with vacuum?
- How does the energy spectrum of electrons change in a crystal that is confined by a surface?
- What is the difference between SES and bulk states due to defects?
- What is the electrostatic potential in near-surface atomic layers?

On the one hand, this is necessary for a fundamental understanding of electronic processes and phenomena on surfaces and internal interfaces of solids, which is important for modern micro- and nanoelectronics. On the other hand, consideration of the surface electronics is essential for a consistent analysis of surface phenomena such as surface conductivity, work function, oxidation, heterogeneous catalysis, crystal growth, brittle fracture, etc.

### 2.4.1. Origin and basic properties of surface states

The surface of any real crystal can be considered as another disturbance of periodicity (in the form of break of the periodic potential in the crystal), along with impurities and defects. I.E. Tamm first showed in 1932 that this circumstance allows additional solutions of the Schrödinger equation for an electron in a crystal (in comparison with the confined crystal), which rapidly decay when deepening from the surface. In other words, in the band model confined by the crystal surface, in addition to the "band" and "impurity" states, it is necessary to take into account the SES. They have a discrete energy spectrum and wave functions exponentially damped as they move both to the crystal depth and toward the vacuum (see below).

An especially important role is played by the appearance of SES in the change of properties in semiconductors. The presence of local surface energy levels leads to the fact that electrons and holes can "stick" to the SES, forming a surface electric charge. In this case, a charge equal to the magnitude and opposite in sign of the SES appears under the surface, induced by the redistribution of free electrons and holes in the bulk. As a result, near-surface layer appears in the crystal enriched or depleted by electrons or holes. The appearance of such layers explains the influence of the surface on many of equilibrium properties of semiconductors (electrical conductivity, work function, contact potential difference, optical properties, etc.).

Surface energy levels can significantly change the kinetics of electronic processes, since they create additional centers of recombination and generation of charge carriers. Therefore, all phenomena associated with nonequilibrium electrons and holes (photoconductivity, photo-emf, processes in injection semiconductor devices, etc.) also depend on the state of the surface.

To explain the nature of the SES appearance, we consider a one-dimensional crystal in which only neighboring atoms interact (Figure 2.4.1). This allows maintain a fundamental feature of the energy spectrum in a bulk (infinite) crystal in the form of alternating allowed and forbidden energy bands (Figure 2.4.1a). It is known from the band theory of crystals that in the presence of a defect in a crystal lattice (for example, by removing one atom from a linear chain that disturbes interaction of atoms), the latter, like the impurity atom, can create a local energy level $E_t$ lying in the forbidden band (Figure 2.4.1b). On the other hand, in the approximation of close interaction, when a half of atoms are removed, an infinite crystal breaks up into two noninteracting parts that have a confining

"surface". Then for a crystal with a surface, we come to the energy scheme of Fig. 2.4.1c, in which, in addition to the bulk bands we have a surface levels $E_s$. If the level enters the forbidden zone, it corresponds to the local charged state (SES).

Fig. 2.4.1. Model representation of the band energy scheme explaining the appearance of SES in a crystal confined by surface: a - ideal crystal, b - crystal with crystal lattice defect, c - crystal with SESs

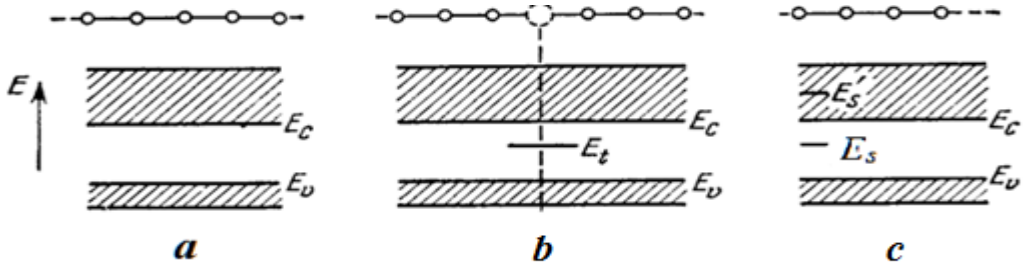



Fig. 2.4.2. The potential energy of the electron (a) and the wave functions of the bulk (b) and surface (c) states in a confined linear chain of atoms in the Tamm model

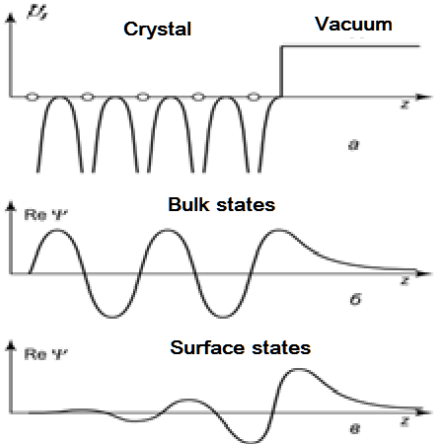
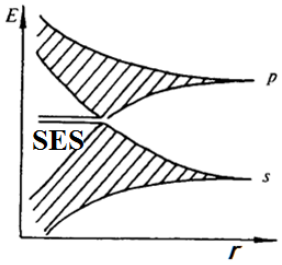



Fig. 2.4.3. Dependence of the energy of electrons in a confined crystal on the interatomic distance in the Shockley model. The shaded regions correspond to allowed states in an infinite crystal (without a surface), and the region between two horizontal lines is SES area in a crystal confined by a surface.

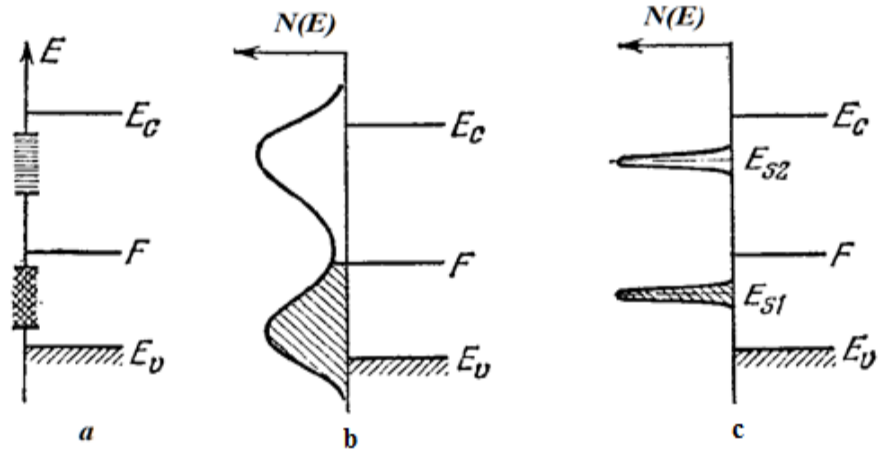



In more precise arguments, one should start from the form of the potential energy of an electron in a confined crystal, which enters into the Schrödinger equation. The Tamm model considers the simplest one-dimensional model of a semi-infinite crystal as a sequence of delta-like potential barriers, confined by a potential "walls" (Figure 2.4.2a). In contrast to the case of an infinite crystal, the function $U_1(x)$ is not periodic, so that it can be represented as the sum of the periodic part U(x), characteristic of an infinite crystal, and the additive W(x) describing the effect of the surface. We can see from the course of this curve $U_1(x)$ in Fig. 2.4.2a that in order to go to vacuum, the electron should overcome the potential barrier by penetrating the surface.

Thus, the direct cause of the SES appearance for the Tamm model is the distortion of the potential *U(x)* in the boundary cell due to the breakage of the periodic crystal lattice and the formation of a potential "wall". This is the so-called Tamm SESs, whose wave functions are localized near the crystal surface (Figure 2.4.2b), in contrast to the bulk ones (Figure 2.4.2c). This approach is applicable for d-electrons in transition metals, as well as for semiconductors and insulators. The concentration of Tamm SESs is equal to the surface concentration of atoms in the crystal, i.e. of the order of $10^{15}$ $cm^{-2}$.

We note that the existence of surface energy levels based on the Tamm model, which was stated above, can be explained analyzing chemical bonds in crystals. Shockley, analyzing a similar problem, came to the conclusion about the possibility of the SES appearance even in the case when the potential in the boundary cell remains undistorted. Calculating the energy spectrum of an atomic chain of finite dimensions, he showed that the presence of boundaries (surface) leads to the splitting off from the allowed bands by one bulk state and the appearance of states in the forbidden band localized near the surface. From the point of view of the Shockley model, this is possible only in those materials whose bulk energy bands arise due to the "mixing" of states genetically originating from different atomic orbitals - Fig. 2.4.3. This means that the surface levels arise due to the appearance of unpaired electrons, the number of which is equal to the number of dangling bonds on the surface (their density is also close to $10^{15}$ $cm^{-2}$). This approach is mainly applicable to normal metals and some narrow-band semiconductors.

Fig. 2.4.4. Surface energy bands (a) and SES densities for a continuous (b) and quasi-discrete (c) distribution of surface states




In the three-dimensional crystals, instead of individual surface levels, the whole zone of surface energy states arises (Fig. 2.4.4a). Each zone consists of closely spaced energy levels corresponding to various possible components of the quasimomentum $p_y$ and $p_z$, parallel to the surface. In reality, however, the crystal surface is always inhomogeneous and the potential distribution along it is far from periodicity. For this reason, the idea of surface zones as an exact two-dimensional analogue of the bulk energy bands of a crystal is an idealization. However, the very concept of SES zones is not canceled by the fact of the inhomogeneity of the real surface.

Surface energy zones can overlap with each other, as well as with bulk bands. Therefore, it is often more correct to talk about the continuous distribution of surface levels and describe them by specifying the density of surface levels *N(E)* (the number of levels per unit surface and unit energy interval). An example of such a distribution is shown in Fig. 2.4.4b. In this case, the surface concentration of trapped electrons (up to the thermal smearing of the Fermi distribution) is expressed by an area confined by the corresponding portion of the *N(E)* curve lower than Fermi level F ($E < F$) and the E axis.

Sometimes, when interpreting the experimental data, it can be assumed that the *N(E)* distribution has one or several sharp and narrow maxima (Fig. 2.4.4c). In this case, one speaks of quasi-discrete (or simply discrete) surface levels, which are to a certain extent analogous to the levels of one-dimensional crystals. Such an approximation is often used for qualitative (or semi-quantitative) analysis of phenomena.

The separation of SESs into "Tamm" and "Shockley" states has become quite widely used in solid-state physics, despite the certain conventionality of such a separation. It can be shown that the displacement of the split plane by half the unit cell can transform one kind of states into another.

For crystals with the predominant ionic type of bonds (like NaCl, ZnO, ZnS), the Madelung ion electrostatic potentials on the surface and in the bulk can differ very strongly. Such a situation can be fully described within the framework of the Tamm model, so that electronic states on the

surface of ionic crystals are often referred to as Tamm SESs. It follows from simple electrostatic considerations that SES, are associated with surface cations, the energy levels of which are located slightly below the bottom of the conduction band for a bulk crystal and form the so-called "cation zone". As to SES connected With surface anions, they are located above the top of the valence band in the bulk and form the "anion zone". Sometimes SES in ionic crystals are called “ionic”.

For diamond type crystals (like Ge, Si) with strong covalent bonds, the distortions of the electrostatic potential in elementary cells near the crystal surface are small, as a rule. At the same time, the very formation of strong covalent bonds is associated with mixing (in form of *sp*-hybridization) of wave functions of different types. In this case, Shockley-type SES arise, which, as noted above, are identified with dangling bonds. However, as noted earlier, the surface structure with a large number of unsaturated electron orbitals is extremely disadvantageous from the energy point of view. Therefore, on the atomically pure surfaces of covalent compounds, the orbital regib-ridization and the associated relaxation or reconstruction of the under-surface region of the crystal usually occur (see above). At the same time, the total number of SES and their energy spectrum undergo significant changes.

### 2.4.2. The band model of the surface

In order to characterize the electronic structure in a semi-infinite crystal within the framework of the band model, it is necessary to resolve the Schrödinger equation for all electrons in a solid. This is obviously an unrealistic problem, since a solid body contains about $10^{23}$ electrons per $cm^3$, while an exact solution of the Schrödinger equation is possible only for a very small number of electrons. Therefore, it requires significantly simplifying assumptions, for example, the use of the one-electron adiabatic approximation. On the other hand, we would like to preserve sufficient accuracy and reliability of theoretical predictions. To achieve this goal, the *theory of the density functional* (*DFT*) turned out to be especially successful. There is an alternative approach, namely the *Hartree-Fock theory*, which makes it possible to calculate the electronic structure of molecules and small clusters, but it turns out to be too rough in the case of large systems, like crystal surfaces, and therefore will not be discussed here.

DFT is based on the theorem formulated by Hohenberg and Kohn: the total energy of the system (for example, the crystal or its surface) is completely determined by the distribution of the electronic density $n(r)$ in its ground state. Moreover, it is possible to determine the energy functional

$$E = \boldsymbol{E}\,[n(r)], \tag{2.4.1}$$

possessing the property that it has a minimum when $n(r)$ corresponds to the density distribution in the ground state.

Usually, $E[n(r)]$ is represented as the sum of three terms: the kinetic energy $T$, the electrostatic (or Coulomb) energy $U$, and the exchange-correlation term $E_{xc}$:

$$E[n(r)] = T + U + E_{xc} \tag{2.4.2}$$

The term $T$ corresponds to the kinetic energy of a noninteracting inhomogeneous electron gas in its ground state. The Coulomb term $U$ is purely classical, which describes the electrostatic energy of the Coulomb attraction between the valence electrons and the ions of the core nuclei, the repulsion of electrons between themselves and repulsion between the ions of the core. The exchange-correlation term $E_{xc}$ combines the contributions from the quantum-mechanical effects of many bodies. The most important of them corresponds to the exchange term, which is associated with the action of the Pauli principle, as a result of which the electron-electronic Coulomb repulsion decreases. The corresponding gain in energy is called *the exchange energy*. The contribution of additional terms describing the interaction between electrons with oppositely

directed spins is defined as *the correlation energy*. Usually, the kinetic energy and the Coulomb term have close values, while the exchange-correlation part is approximately 10 % of this value, with the exchange energy greater than the correlation energy.

The electron density distribution $n(r)$, which minimizes the energy functional $E[n(r)]$ of the form (2.4.2), is found as a self-consistent solution of the system of one-electron Schrödinger-type equations:

$$E[n(r)] = T[n(r)] - \sum_R Ze \int dr \frac{n(r)}{|R-r|} + \frac{1}{2}\int\int dr dr' \frac{n(r)n(r')}{|r-r'|} + E_{\text{exc.c}}|n(r)| \tag{2.4.3}$$

In this expression, the term $T[n(r)]$ describes the kinetic energy of a *noninteracting* inhomogeneous electron gas in its ground state with the density distribution $n(r)$. The second term again represents the potential energy of the ion-electron interaction, and the third term is the average potential energy of the electrostatic interaction of the electrons. All the quantum-mechanical features of the multiparticle problem are contained in the so-called *exchange-correlation term* $E_{\text{exc-c}}[n(r)]$.

The great advantage of this formulation of the problem lies in the fact that the density $n(r)$, which minimizes the functional (2.4.3), is found by solving a system of coupled ordinary differential equations:

$$-\frac{1}{2}\nabla^2\psi_i(r) + \upsilon_{\text{eff}}(r)\psi_i(r) = \varepsilon_i\psi_i(r), \tag{2.4.4}$$

$$\upsilon_{\text{eff}}(r) = -Ze^2\sum_R \frac{1}{|r-R|} + \int dr' \frac{n(r')}{|r-r'|} + \upsilon_{\text{exc.c}}(r), \tag{2.4.5}$$

where $n(r) = \sum |\psi_i|^2$. This result is exact. Of course, the electron-electron interaction is "hidden" in the exchange-correlation potential

$$\upsilon_{\text{exc.c}}[n(r)] = \delta E_{\text{exc.c}}[n(r)]/\delta n(r), \tag{2.4.6}$$

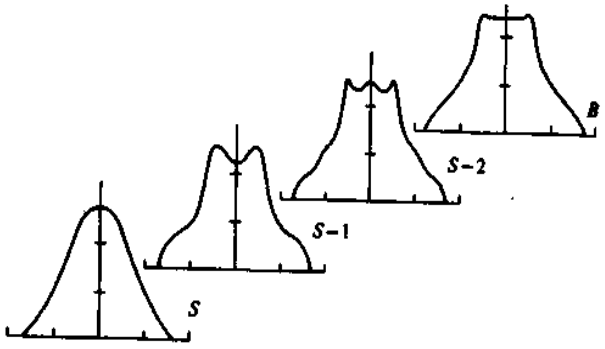
and for practical application of the method under consideration a good approximation for this quantity is necessary. The parameters $\varepsilon_i$ and $\psi_i$, which enter the equation (2.4.4), similar to the Schrödinger equation, formally have no physical meaning. Despite this, they are often successfully interpreted, respectively, as single-particle excitation energies and eigenfunctions.

When considering the band model of a three-dimensional crystal with a surface, it is useful to be able to divide the value of $n(E)$ into contributions from each atomic layer parallel to the surface. It is even better to determine the *local density of states* (*LDS*) at each point of space, taking into account the statistical weight of each eigenfunction of the system:

$$\rho(z,E) = \sum_\alpha |\psi_\alpha(z)|^2 \delta(E - E_\alpha). \tag{2.4.7}$$

In Fig. 2.4.5 we present the results of LDS calculations divided by atomic layers for a three-dimensional solid body in the framework of the strong-coupling model. Calculations indicate a narrowing of *LDS* on the surface. From Fig. 2.4.5 it also follows that *LDS* "straightens" to its bulk value at a distance of no more than two or three atomic layers below the boundary with vacuum.

Fig. 2.4.5. Allowed on LDS layers for the three upper surface planes (s, s-1, s-2) in the strong-coupling model in comparison with the bulk density of states (*B*)



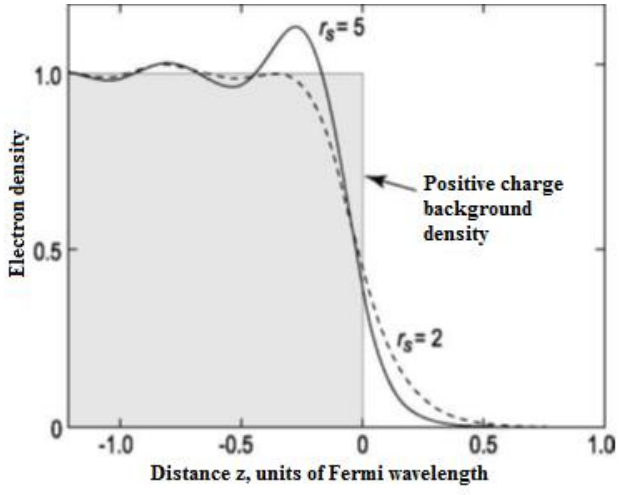

The solution of equations of the type (2.4.3) for the problem of a semi-infinite crystal lattice is rather difficult. At the same time, for example, in alkali metals, where electrons are weakly interacting with the potential of a shielded ion core, a so-called *jelly model* can be used. In this model, the discrete ion cores are replaced by a uniform positively charged background, characterized by a charge density equal to the space-averaged ion charge distribution. This simplest model provides with reasonable accuracy a description of the electronic properties of simple metals, such as Na, Mg or Al, for which only s- and p- electrons form the conduction band.

For the surface problem, the semi-infinite array of ions is replaced by a homogeneous positive charge that fills half the space

$$n_+(z) = \begin{cases} \bar{n} & z \le 0, \\ 0 & z > 0. \end{cases} \qquad (2.4.8)$$

The electrostatic potential created by such a charge distribution *n(z)* has the form of a step at z = 0 (Figure 2.4.6). Instead of the ion-electron potential, it is substituted into expression (2.4.5).

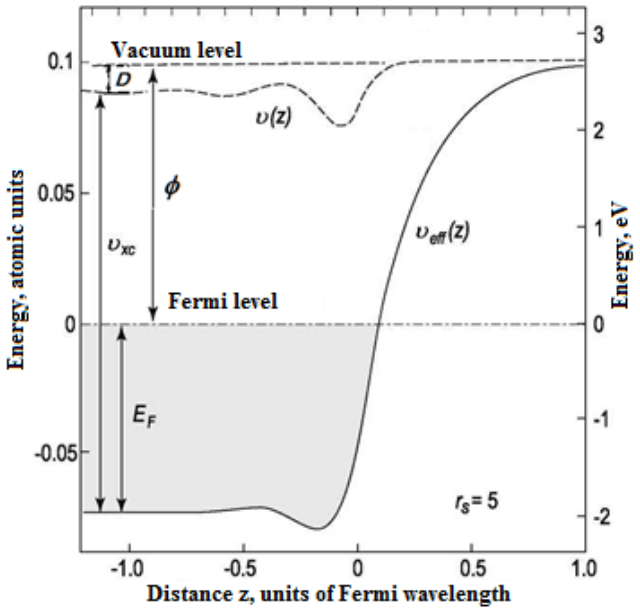
Fig. 2.4.6. The electron density profiles on the surface in *the jelly model* for the two selected values of the background density




In practice, it is customary to express the density of the positively charged background *n(z)* in terms of the dimensionless quantity $r_s$ (the average distance between electrons in Fig. 2.4.6) and the Bohr radius $a_0 = h^2/me^2 = 0{,}0529$ nm. By definition, $r_s$ is the radius of a sphere containing exactly one electron. Therefore, the density of the positive background charge *n(z)* is often expressed in terms of the inverse volume of the sphere $(4\pi/3)r_S^3 = 1/\mathrm{n(z)}$ . Typical $r_s$ values range from two to five.

The electron density profile *n(z)* for the ground state in the jelly model has translational invariance in the x-y plane on the surface. However, a change in the density *n(z)* in the direction perpendicular

Fig. 2.4.7. The electrostatic potential $\upsilon(z)$ and the total effective one-electron potential $\upsilon_{eff}(z)$ near the surface of the crystal in *the jelly model*




to the surface reveals two features that are very typical for all surface problems (Figure 2.4.7).

First, the electron distribution does not have a sharp edge, since electrons penetrate through the surface into the vacuum, so that the electron density decays to zero at a distance of 0.1-0.3 nm from the surface. In other words, electrons partially "flow out" beyond the surface ($z > 0$), creating an *electrostatic dipole* at the surface. However, we can localize the position of the effective surface by defining it as a condition

$$d_{||} = \frac{1}{n(z)} \int_{-\infty}^{\infty} dz\, z \frac{dn(z)}{dz}. \qquad (2.4.9)$$

In this case, the change of $n(z)$ in the vicinity of the surface defined in this way should be regarded as the Gibbs excesses of the electronic subsystem of the crystal at the surface.

Secondly, the function $n(z)$ oscillates (the so called *Friedel oscillations*), when its value approaches the asymptotic value of the bulk density of electron states, which exactly compensates the homogeneous (bulk)background charge. The Fermi wavelength of these oscillations is equal to $n/k_F$, where $k_F = \left(3\pi^2 n(z)\right)^{1/3}$ is the Fermi wave vector. These oscillations are due to the fact that electrons (with constant wave vectors between zero and $k_F$ try to screen the distribution of the positive background charge, which contains a step at z = 0. We note that Friedel oscillations resemble the picture of light waves diffraction on the half-plane. The similarity is not accidental: only electrons with energies close to the Fermi energy that form practically monochromatic waves can move freely in metals and degenerate semiconductors. And just the interference of these waves on the metal boundary leads to the appearance of oscillations.

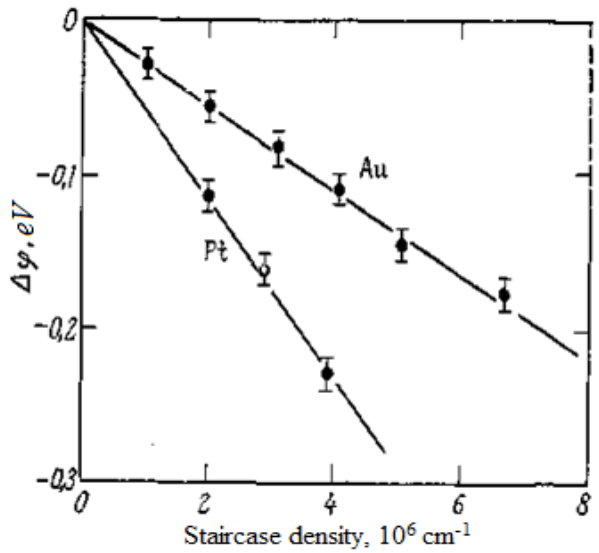
The formation of a surface dipole layer means that the electrostatic potential in a vacuum at points sufficiently far from the surface is greater than the average electrostatic potential deep in the volume of the body, i.e.

$$D = \upsilon(\infty)\, \upsilon(-\infty). \qquad (2.4.10)$$

Variations of the electrostatic potential $v(z)$ in the presence of a surface dipole are shown by the dashed line in Fig. 2.4.7. The jump in the potential partially ensures the retention of electrons inside the crystal. The remaining part of the potential barrier is due to short-range Coulomb interactions. The potential energy of each electron is lowered, because the surrounding electrons tend to retire over long distances. This is a three-dimensional effect, which is entirely due to the exchange effect and correlation.

### 2.4.3. Electronic properties of surface

Fig. 2.4.8. Change in the work function for the facetted metal surface in *the jelly model*



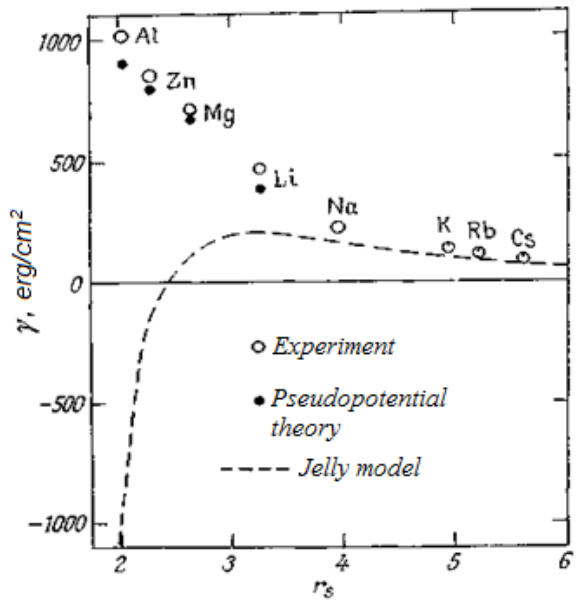

The jelly model for a semi-infinite crystal can also be used to estimate surface parameters such as electron work functions and surface tension.

The work function from the crystal surface is, by definition, equal to the minimum energy required to remove an electron from the depth of the crystal to the vacuum point located at a macroscopic distance from the surface. It can be expressed by the following relation:

$$\varphi = \upsilon(\infty) + E_{N-1} - E_N = \upsilon(\infty) - \mu = D - E_F. \quad (2.4.11)$$

In relation (2.4.11), both the electrostatic potential $\upsilon(\infty)$ and the chemical potential $\mu$ are reckoned from the average electrostatic potential deep in the crystal bulk. In this case, changes in both the surface ($D$) and bulk ($E_F$) contributions to the work function during the transition from element to element in the periodic table can be significant (~1 Rydberg). Nevertheless, the measured work functions of all elements lie in the range φ = 3,5 ± 1,5 eV.

Fig. 2.4.9. Comparison of the experimental surface tension data with the results of the jelly model and the pseudopotential theory




The specific features of the surface also contribute to the surface dipole and affect the magnitude of the work function. For example, the reconstruction of the surface, smoothing of the charge density, as well as the appearance of steps (facets) on the surface will change the parameter D. Fig. 2.4.8 demonstrates a change in the work function on metal surfaces containing regular facets (with terraces).

The jelly model for semi-infinite crystals can also be used to estimate the surface tension of simple metals. As can be seen from Fig. 2.4.9, surface tension in metals characterized by low electron density is well described by the jelly model, while for metals with high electron density it predicts spontaneous splitting. This difficulty is due to the complete neglect of the contribution of the ionic lattice to it. Using the first approximation of perturbation theory, one can take into account the linear response of the electron gas to the weak pseudopotential of the crystal lattice. When the resultant energy shifts are taken into account, the jelly model agrees much better with the experiment (see Figure 2.4.9).

Thus, as follows from Fig. 2.4.9, the results of calculations of the surface tension of metals based on the jelly model coincide with the experimental data only for alkali metals with a

sufficiently low electron density ($n < 2{,}5 \cdot 10^{22}$ cm$^{-3}$). For heavy metals with an electron concentration $n$ above $2{,}5 \cdot 10^{22}$ cm$^{-3}$, in which the average distance between free electrons and lattice ions is small, it is no longer possible to neglect the periodic potential of the crystal lattice.

## 2.5. Features of the electronic structure of the semiconductor surface under equilibrium conditions

Let us consider how the concentration of free carriers (and, correspondingly, the band diagram) in the near-surface region of the semiconductor will change when an electric field exists near this surface. This field near the surface or interface can occur under the action of:

- *Surface (interface) electronic states (SES);*
- *External electric field;*
- *"Internal" electric field* (due to the contact potential difference at the internal interfaces, and also the accumulation on the surface of not only the electron but also the ion charge, for example, due to polarization in ionic crystals).

First, consider an ideal surface (without *SES*), when an external electric field is created by an electrical contact in the form of a charged metal plane with surface charge density $\rho$. Since the lines of the electric field force must be closed, an electric charge of equal magnitude but opposite in sign must appear below the metal surface in an ideal semiconductor. This screening external field charge will arise due to the movement of free electrons and holes. Depending on the sign of the charge on the metal contact (positive or negative), the charge in the near-surface region of the semiconductor will also be of different signs. The case when the concentration of the main free carriers in the near-surface region increases is called *enrichment*, and when - decreases - *depletion*.

The appearance of a *space charge region* (*SCR*) due to a change in the concentration of free carriers in the near-surface region of a semiconductor under the action of an external electric field was called the *field effect*.

In the presence of an electric field (external or internal), the subsurface region in the semiconductor will not be electrically neutral because the screening charge in form *SCR* arises in this region. The presence of the electric field $E(z) = (\rho(z)/2\varepsilon\varepsilon_o) = (qN_M/2\varepsilon\varepsilon_o)$ (here $N_M$ is the density of charges on a metal electrode of unit area, ε is the relative permittivity of the gap between the metal contact and semiconductor) in the *SCR* changes the value of the potential energy of electrons and holes.

If the donor concentration in the bulk of the semiconductor is about $N_D \sim 10^{15}$ cm$^{-3}$, then the average distance between the free electrons (and ionized donors) in the quasi-neutral volume of the semiconductor will be $a = N_D^{-1/3} = 10^{-5}$ cm = 100 nm. At a surface charge density of $\rho = 10^{12}$ cm$^{-2}$, the thickness of the space-charge layer of ionized donors will be $10^{11}/10^{15} = 10^{-4}$ cm (of the order of 1000 nm). It follows that the electric field can penetrate into the semiconductor to a considerable depth.

We note that in metals where the concentration of free electrons per unit volume is very high ($n \approx 10^{22}$ cm$^{-3}$), such an amount of free carriers corresponds to their redistribution at a depth (from the surface) of only a few interatomic distances. In other words, the external electric field does not penetrate deep into the metals. In dielectrics, where the concentration of free carriers is less than $10^5$ cm$^{-3}$, the electric field is not screened (except by polarization processes) and penetrates any distance deep into the dielectric. Thus, in semiconductors the situation is intermediate.

If the field is directed from the surface into the semiconductor depth, then the electrons in this case will have a minimum energy at the surface, which corresponds to the presence of a potential well for the electrons in the same place. The change in the potential energy of the electrons has the form

$$\Delta U = U(z) - U(\infty) = \int_{\infty}^{z} E(z)dz\,,$$

where U(∞) is the potential energy of the electrons in the quasi-neutral volume of the semiconductor. Since the kinetic energy of the electrons ($\varepsilon = h^2k^2/2m^*$) is zero at the bottom of the conduction band, the change in the potential energy along the coordinate must also change the energy position of the bottom of the conduction band $E_c$, (and, accordingly, the top of the valence band $E_v$). On the band diagrams this is expressed in the form of *the energy bands bending*.

The magnitude of the potential difference between the quasi-neutral volume and an arbitrary point of the SCR is called the *electrostatic potential*:

$$\psi = \frac{1}{q}\int_{\infty}^{z} E(z)dz\,.$$

The value of the electrostatic potential on the surface of the semiconductor is called the *surface potential* $\psi_s$, the sign of which corresponds to the sign of the bending energy band of the charge on the metal contact: for $\psi_s > 0$, the bands are curved downward, and for $\psi_s < 0$ - upward.

### 2.5.1. The charge of surface electronic states

The band bending in the near-surface region of the semiconductor can arise even without an external electric field if the surface is not ideal and there are charged SES on it. As in the case of localized electronic states in the crystal bulk, SES is usually classified according to their possible charge states - "donor-like" or "acceptor-like". Donor SES can be either electrically neutral or positively charged ($N^0_D$, $N^+_D$), whereas acceptor ones are electrically neutral and negatively charged ($N^0_A$, $N^-_A$).

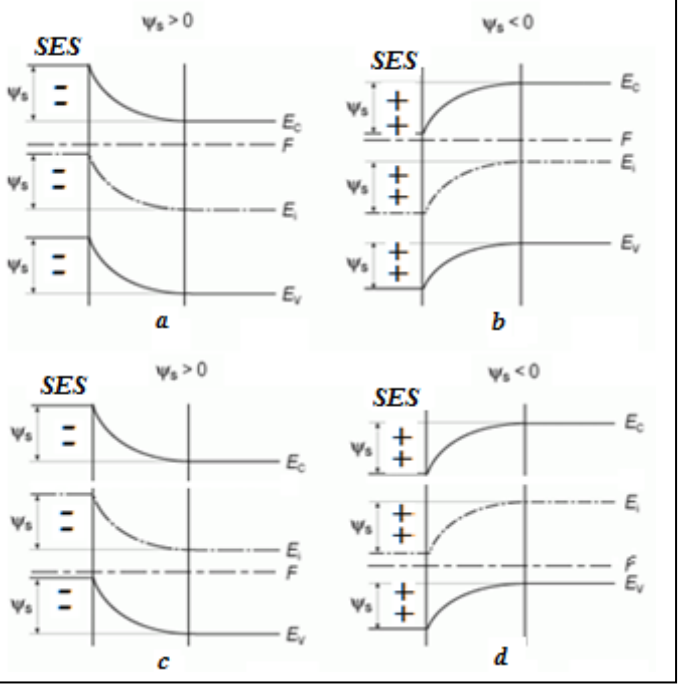



Fig. 2.5.1. Bending of energy bands near the surface of n-type (a, b) and p-type (c, d) semiconductors with donor- (a, c) and acceptor-like (b, d) SESs

If the SES are charged, then, due to the electric neutrality of the semiconductor crystal under its surface (similar to the case of the action of an external field on the ideal surface considered above), there must also be redistribution of free electrons and holes, which leads to accumulation at the surface of those carriers whose charge is opposite to the charge of the SES. As a result, an *SCR* with a bend of energy zones also appears below the surface. This means the occurrence of the *built-in electric field* $E$. The nature of this bending is determined by the sign and the concentration of the SES and the conductivity type of the semiconductor (the concentration of donor or acceptor impurities). Fig. 2.5.1 shows 4 main variants of band bending types at moderate SES concentrations.

Let us first consider SES, which are characterized by the monoenergetic level $E_t$ in the band gap with concentration $N_t$ (per unit area). We will assume that $N_t$ is not too large and the effects of mutual influence of neighboring SES can be neglected, so that the value of $E_t$ does not depend on the concentration or on the degree of filling of SES by electrons.

In thermodynamic equilibrium, the probability of filling SES is subject to Fermi-Dirac statistics:

$$f_{10} = \frac{n_t}{N_t} = \left[1 + g_1 \cdot \exp\left(\frac{E_t - F}{kT}\right)\right]^{-1} \qquad (2.5.1)$$

Here $F$ is the Fermi level, $n_t$ is the concentration of trapped electrons (counting per unit surface), $g_t$ is the "degeneracy factor", which allows us to take into account that a simple acceptor can capture and a simple donor give only one electron (in contrast to the "band" states capable of assuming two electrons). For simple acceptors and donors, $g_t^a = 2$ and $g_t^d = 1/2$, respectively.

Formula (2.5.1) describes the relative number of electrically neutral *SES*. If, for example, the energy level of the acceptor-like SES is much higher than the Fermi level ($E_t - F \geq 3kT$), the concentration of negatively charged acceptors is

$$n_t = N_A^- \cong N_A (g_r)^{-1} \cdot \exp\left(-\frac{E_t - F}{kT}\right) = \frac{N_A}{2} \exp\left(-\frac{E_t - F}{kT}\right) \qquad (2.5.2)$$

In what follows, for convenience, the Fermi-Dirac function (2.5.1) will be used in a more conventional form ($g_t = 1$), but it must be borne in mind that the true position of the energy level is $E_t^{true} = E_1 + kT\ln g$. For simple donor and acceptor SES, the value of $kT\ln g \approx 0{,}7kT$.

The specificity of the surface states in comparison with bulk donors and acceptors is that the positions of the energy levels of the SES with respect to the Fermi level depend on the total surface charge. If the intensity of the electric field created by this charge is significantly less than the intra-atomic field, the $E_t$ position in the forbidden zone of the crystal remains unchanged, and the effect of the SCR field reduces only to a shift in the energy levels of the SES by an amount equal to the bending of the energy bands on the surface: SES levels are "tied" to the energy diagram of the surface. Therefore, it is convenient to introduce the dimensionless energy of the SES $\varepsilon_t = (E_t - E_i)/kT$, which does not depend on the band bending, where $E_i$ is the level of electric neutrality. Using the dimensionless potential $U_s = (F - E_t)/kT$, the function (2.5.1) can be rewritten in the form

$$f_{t0} = \left[1 + \exp(\varepsilon_t - U_s)\right]^{-1}. \tag{2.5.3}$$

Under thermodynamic equilibrium conditions, the charge localized on the SES, with the concentration of $N_t$, and the energy $E_t$, is equal to

$$Q_t^{(a)} = -qN_t^a f_{t0} \tag{2.5.4а}$$

for acceptor SES, and

$$Q_t^{(d)} = +qN_t^a (1 - f_{t0}) \tag{2.5.4б}$$

for donor SES.

From (2.5.3) it is obvious that when the surface potential varies, the charge of the SES can vary, since the filling function $f_{t0}$ depends on $U_s$.

If there are several types of SESs on the surface, the total charge $Q_s$ localized at all SES is equal to the sum of the charges of individual states:

$$Q_s = \sum_i Q_{ti} = q\left[\sum_n N_t^a (1 - f_{t0}) - \sum_m N_{tm}^a f_{t0}\right]. \tag{2.5.5}$$

In the case of a "quasi continuous" energy spectrum of SES, with a large number of different types of SES, the distance between the individual energy levels will be of the order of or less than kT, and the summation in (2.5.5) is replaced by integrating that energy

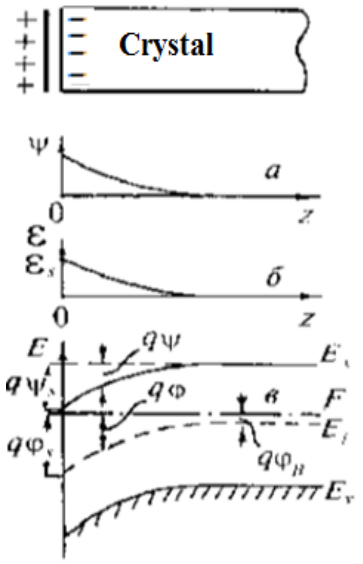
$$Q_s = q\int_{E_v}^{E_c}\left[N_S^{(d)}(E)(1 - f_{t0}) - N_S^{(a)}(E) f_{t0}\right]dE, \tag{2.5.6}$$

where $N_S^{(d)}$ and $N_S^{(a)}$ are the surface densities of donor and acceptor SES (calculated per unit surface and energy), which are usually measured in $eV^{-1}cm^{-2}$.

### 2.5.2. The basic equation of the SCR

As noted above, in the presence of charged SES with a charge density $Q_s$, due to the electric neutrality of a semiconductor crystal in its subsurface region, a bending of the energy bands arises (Figure 2.5.1) with a near-surface SCR. In this section, the main parameters of the near-surface SCR in semiconductors depending on the distance $z$ from the surface will be studied, which do not depend on the specific nature of the appearance of the SCR.

Fig. 2.5.2. The nature of the electrostatic potential dependence versus coordinate (a), the intensity of the electric field (6), and the electron energy at the top of the valence band $E_V$ and the bottom of the conduction band $E_C$ (c) near the surface of the intrinsic semiconductor in a positively charged surface created by acceptor-like SES.




The main regularities in the formation of an SCR can be demonstrated by the example of a homogeneous surface model, where the potential

$$\psi = \frac{1}{q}\int_{\infty}^{z} \varepsilon(z)dz,$$

charge $\rho(z)$, electric field $E(z) = (\rho(z)/2\varepsilon\varepsilon_0)$, concentration of free electrons $n(z)$ and holes $p(z)$, etc. parameters in a semiconducting crystal in the direction of the normal to the surface are functions of only one coordinate $z$ (Figure 2.5.2).

To describe the potential, we use the one-dimensional Poisson equation

$$\frac{d^2\psi}{dz^2} = -\frac{1}{\varepsilon\varepsilon_0}\rho(z), \qquad (2.5.7)$$

where $\psi(z)$ is the electrostatic potential (its value in the crystal bulk is assumed to be zero), $\rho(z)$ is the space charge density, $\varepsilon$ is the permittivity, and $\varepsilon_0$ is the electric constant.

The density of a space charge in a crystal generally consists of charges of free electrons, free holes and charged donor and acceptor impurities, and is described on the basis of the continuity equation

$$\rho(z) = q\left[p(z) - n(z) + N_D^+(z) - N_A^-(z)\right]. \qquad (2.5.8)$$

Here $q$ is the elementary charge magnitude, $N_D^+(z)$ and $N_A^-(z)$ are concentrations of ionized donors and acceptors, respectively.

Principal features of forming an *SCR* can be clarified by simplifying equations (2.5.7) and (2.5.8) as much as possible. To simplify the mathematical side of the problem, let us assume that there is no degeneracy of the free-carrier gas and full ionization of the impurities both in the bulk of the crystal and in the SCR.

The first assumption that there is no degeneracy of the free charge carriers gas allows us to use the Maxwell-Boltzmann statistics. Introducing the *dimensionless surface potential*

$$Y(z) = q\psi(z)/kT \qquad (2.5.9)$$

we get

$$n(z) = n_0 \cdot \exp(q\psi/kT) = n_0 \cdot \exp(Y)$$
$$p(z) = p_0 \cdot \exp(-q\psi/kT) = p_0 \cdot \exp(-Y), \quad (2.5.10)$$

where $n_o$ and $p_o$ are the equilibrium concentrations of electrons and holes in the bulk. The positive magnitudes of $Y$ corresponds to a decrease in the electron energy near the surface, that is, the energy bands are bent downwards (Figure 2.5.2).

As long as the intensity of the electric field in the SCR is much smaller than the intra-atomic field (~$10^8$-$10^9$ V/cm), applying a macroscopic electric field to the crystal causes all energy levels to shift to the same value $q\psi(z) = -kTY(z)$.

The second assumption concerning the total ionization of impurities in the bulk of the crystal and in the SCR (which, as a rule, is valid at room temperature for many semiconductors) allows us to assume that the concentrations of ionized donors and acceptors are equal to their total concentrations (depletion of impurities at room temperatures) $N_D^+(z) = N_D, \quad N_A^-(z) = N_A$.

Taking into account that the crystal bulk is electrically neutral ($\rho = 0$) and, therefore,

$$N_D - N_A = n_0 - p_0, \quad (2.5.11)$$

we have in place of (2.5.8)

$$\rho(z) = q\{[p(z) - p_0] - [n(z) - n_0]\} = q\left[p_0\left(e^{-Y} - 1\right) - n_0\left(e^{Y} - 1\right)\right]. \quad (2.5.12)$$

We characterize the degree of nonintrincity of a semiconductor by a dimensionless parameter

$$\lambda = p_0/n_i = n_i/n_0, \quad (2.5.13)$$

where $n_i$ is the concentration of intrinsic charge carriers. This allows us to reduce the Poisson equation (2.5.7) to the dimensionless form

$$\frac{d^2Y}{dz^2} = L_{Di}^{-2} \cdot \left[\lambda\left(e^{-Y} - 1\right) - \lambda^{-1}\left(e^{Y} - 1\right)\right], \quad (2.5.14)$$

Where

$$L_{Di} = \left(\varepsilon\varepsilon_0 kT / 2q^2 n_i\right)^{1/2}. \quad (2.5.15)$$

If we lower the order of the differential equation (2.5.14) by introducing the function $\zeta = dY/dz$, the solution of equation (2.5.14) can be represented in the form

$$\frac{dY}{dz} = \mp \frac{F(\lambda, Y)}{L_D}, \quad (2.5.16)$$

Where

$$F(\lambda, Y) = \left\{ \frac{2}{\lambda + \lambda^{-1}} \left[ \lambda \left( e^{-Y} - 1 \right) + \lambda^{-1} \left( e^{Y} - 1 \right) + Y \left( \lambda - \lambda^{-1} \right) \right] \right\}^{1/2} . \qquad (2.5.17)$$

In all these equations, the parameter

$$L_D = L_{Di} \left( \frac{2}{\lambda + \lambda^{-1}} \right)^{1/2} = \left[ \frac{\varepsilon \varepsilon_0 kT}{q^2 (n_0 + p_0)} \right]^{1/2} . \qquad (2.5.18)$$

is called the *Debye screening length*. Note that for a proper (undoped) semiconductor, $\lambda = \lambda^{-1} = 1$ and $L_D = L_{Di}$.

The sign of the derivative *dY/dz* in the relation (2.5.16) is positive if the energy bands are curved upwards (Y <0) and negative for Y> 0 (bending the zones downwards).

The first term in (2.5.17) is due to the presence of free holes in the *SCR*, the second one - to free electrons, and the third - to the presence of ionized impurities. For the intrinsic semiconductor ($\lambda = 1$), the last term is absent.

It follows from equation (2.5.16) that the decrease of the field in depth from the surface corresponds to formula

$$\mathcal{E}(z) = -\frac{kT}{q} \left. \frac{dY}{dz} \right|_z = \pm \frac{kT}{q} \frac{F[\lambda, Y(z)]}{L_D} , \qquad (2.5.19)$$

and the field on the surface is

$$\mathcal{E}_S = -\frac{kT}{q} \left. \frac{dY}{dz} \right|_0 = \pm \frac{kT}{q} \frac{F(\lambda, Y_S)}{L_D} . \qquad (2.5.20)$$

In the expressions (2.5.19) and (2.5.20) we have introduced the notation

$$Y_S \equiv Y(0) \text{ и } \mathcal{E}_S \equiv \mathcal{E}(0) . \qquad (2.5.21)$$

Here and in the sequel, the index *s* denotes the physical quantities corresponding to $z = 0$, i.e. related directly to the surface.

From the relations obtained it follows that the total charge of the SCR per unit surface is

$$Q_{sc}(\lambda, Y_s) = -\varepsilon \varepsilon_0 \varepsilon_S = \pm q n_i (\lambda + \lambda^{-1}) L_D F(\lambda, Y_s) \qquad (2.5.22)$$

If the SCR is formed by the accumulation of a charge on the surface, then it follows from the electric neutrality of the entire crystal that the surface charge density is equal to the charge in the *SCR*:

$$Q_S(\lambda, Y_s) = -Q_{sc}(\lambda, Y_s) . \qquad (2.5.23)$$

To characterize the band bending (surface barrier), a dimensionless bulk potential is often introduced for the surface of a semiconductor:

$$u_B = -\ln \lambda . \qquad (2.5.24)$$

It corresponds to the dimensional value of the barrier (the bands bending), equal to

$$\varphi_B = (kT/q) \cdot u_B, \tag{2.5.25}$$

It is easy to show that in the absence of degeneracy

$$u_B = (F - E_i)/kT, \tag{2.5.26}$$

where $F$ is the position of the Fermi level in the impurity semiconductor and $E_i$ is the level of electric neutrality (position $F$ in the intrinsic semiconductor).

The corresponding reduced potentials in the SCR

$$u = q\varphi/kT \tag{2.5.27}$$

and on the surface

$$u_s = q\varphi_s/kT \tag{2.5.28}$$

differ from

$$Y = q\Psi/kT \tag{2.5.29}$$

and

$$Y_S = q\Psi_s/kT \tag{2.5.30}$$

by the magnitude of the bulk potential $u_B$ (see Figure 2.5.2).

It should be pointed out that the results obtained above can not be used for degenerate semiconductors and (or) for incomplete ionization of impurities in the crystal bulk or the SCR. In these cases, it is necessary to use the Fermi-Dirac function when solving the Poisson equation (2.5.7). In this case, the mathematical side of the problem is significantly more complicated. In addition, with incomplete ionization of impurities, it is necessary to know the exact values of the concentrations and positions of the energy levels of all impurities in the bulk of the semiconductor.

Let us analyze a number of important consequences of the equations obtained, in particular, the problems of electrostatic shielding of SES by free carriers and carriers localized at deep levels (in wide-band semiconductors).

***Electrostatic shielding by free charge carriers.*** The near-surface *SCR* "screens" the electrically neutral bulk of the crystal from the external electric field and the field created by SES. From the relation (2.5.16), it is clear that the larger the parameter $L_D$, the smaller the derivative *dY/dz* under other equal conditions and, consequently, the electric field penetrates the crystal to a greater depth.

More precisely, the physical meaning of the screening Debye length can be determined by analyzing the case of small band bends when $|Y| << 1$. Using the expansion of exponentials in (2.5.17) in power series and dividing by quadratic terms, it is easy to show that in these conditions

$$F(\lambda, Y) \cong Y; \tag{2.5.31}$$

Then the solution (2.5.16) takes the form

$$|Y| = |Y_S| \cdot \exp(-z/L_D). \tag{2.5.32}$$

Let us give quantitative estimates of $L_D$ for several typical cases. For extrincic silicon of n-type ($\varepsilon$ = 11,9; $n_o = 1 \cdot 10^{15}$ cm$^{-3}$) at room temperature, the Debye screening length is about $L_D$ = 130 nm. With an increase in the concentration of free charge carriers up to $n_o \approx 5 \cdot 10^{17}$ cm$^{-3}$ (such a concentration of free electrons is common for a bismuth semimetal), calculation by formula (2.5.18) leads to a value $L_D$ = 6 nm. At even higher concentrations of free electrons (for example, $n_o \approx 1 \cdot 10^{22}$ cm$^{-3}$ for metals) $L_D$ < 0.1 nm (i.e. less than the lattice parameter).

The last two numbers should be regarded as purely orientational, since the relation (2.5.18) is obtained on the assumption that there is no degeneracy. For metals and degenerate semiconductors, the thermal energy $kT$ should be replaced by the Fermi energy, and by $\varepsilon$ is understood the permittivity of one lattice, without the electron contribution (usually $\varepsilon$ = 1). Such a calculation yields for the aluminum ($F$ = 11,6 эV, $n_o \approx 1{,}8 \cdot 10^{23}$ cm$^{-3}$) the Debye screening length $L_D \approx$ 0,05 nm.

As the concentration of free charge carriers decreases, $L_D$ rapidly increases. For example, for intrinsic silicon ($n_i = 1{,}5 \cdot 10^{10}$ см$^{-3}$) at $T \approx$ 300 K, the value of $L_D$ = 2500 nm. For intrinsic gallium arsenide with a carrier concentration $n_i = 1{,}5 \cdot 10^6$ cm$^{-3}$ at room temperature, the calculation gives an unreasonably high value - $L_D$ > 2 mm.

***Screening by localized charge carriers.*** In the case of wide-gap semiconductors and dielectrics with low concentrations of free charge carriers $n$ < 1 cm$^{-3}$, the giant $L_D$ values of the order of a meter or more are obtained from the formula (2.5.18). In fact, the Debye screening length, even for good dielectrics, is much smaller. The reason for this discrepancy consists in ignoring the possibility of screening the external field by charge carriers localized at deep levels.

Let us, for example, in a crystal, in addition to fully ionized donors and acceptors, there exist continuously distributed deep energy levels of defects whose density $N_{deep}$ (per unit volume and energy interval $kT$) is constant. In this case

$$\rho(z) \cong q\left\{\left[p(z) - p_0\right] - \left[n(z) - n_0\right] - N_t Y\right\} \qquad (2.5.33)$$

For small band bending | $Y$ | << 1 this expression can be simplified:

$$\rho(z) \cong qY\left(p_0 + n_0 + N_t\right) . \qquad (2.5.34)$$

It is clear that in (2.5.16) - (2.5.18) instead of $(n_o + p_o)$, ($n_o + p_o + N_{deep}$) will appear and, consequently, the Debye screening length in the case under consideration

$$L_D \approx \left[\frac{\varepsilon\varepsilon_0 kT}{q^2\left(n_0 + p_0 + N_{deep}\right)}\right]^{1/2} \quad L_D \approx \left[\frac{\varepsilon\varepsilon_0 kT}{q^2\left(n_0 + p_0 + N_{deep}\right)}\right]^{1/2} . \qquad (2.5.35)$$

Even at a low density of deep defects, $N_{deep}$ = 3,8·10$^{10}$ cm$^{-3}$·eV$^{-1}$ (10$^9$ cm$^{-3}$/$kT$ at T = 300 K), which corresponds to the average distance between deep centers in the SCR of tens of nanometers, the Debye screening length in the dielectric ($\varepsilon$ = 4; $n_o$, $p_o$ << $N_{deep}$) will be only 75 μm.

Obviously, for any distribution of deep centers in energy, only those of them whose energy levels are near the Fermi level will participate in the electrostatic screening, so that by $N_{deep}$ in formula (2.5.35) it is necessary to understand the density of exactly such states in the *SCR*.

In dielectrics with appreciable ionic conductivity, electrostatic screening, at least in part, can be realized due to the accumulation of ion charge in the near-surface region. As a result, the Debye screening length will become even smaller.

### 2.5.3. Spatial characteristics of *SCR* in semiconductors

***Types of SCR in semiconductors.*** Unlike ideal surfaces in an electric field, on the surface of a semiconductor with SES, depending on the direction and magnitude of the bending of the energy bands, there are three types of *SCR*: if the concentration of the majority carriers in the *SCR* is greater than in the crystal bulk, we speak about *enrichment*, if it is lower in the SCR – we may speak about a *depletion* or *inversion*.

For illustration, Fig. 2.5.3 shows the band diagrams of the n-type semiconductor in various cases (different signs and density of SES).

If the charge of the SES is opposite to the sign of the majority carriers at the semiconductor surface, but its concentration is low, so that the Fermi level $F$ does not intersect the electric neutrality level $E_i$ (dashed line in Fig. 2.5.3), the concentration of free carriers is higher than in the bulk of the semiconductor. This situation is shown in Fig. 2.5.3a, where the enrichment of the *SCR* by electrons is observed due to their attraction to acceptor-like SES. If the sign of the SES coincides with the sign of the majority carriers, a depleted region appears in the SCR, in which the concentration of free carriers is less than the concentration in the semiconductor volume (Fig. 2.5.3b).

With an increase in the concentration of SES, which coincide in sign with the majority carriers, the Fermi level can intersect the $E_i$ curve corresponding to the middle of the forbidden band (Fig. 2.5.3c,d). In this case, near the surface of the n-type semiconductor, the Fermi level will be closer to the top of the valence band $E_V$ than to the bottom of the conduction band $E_C$. Because of this, the type of conductivity will change, since the concentration of minority carriers (holes) becomes higher than the concentration of the main carriers (electrons).

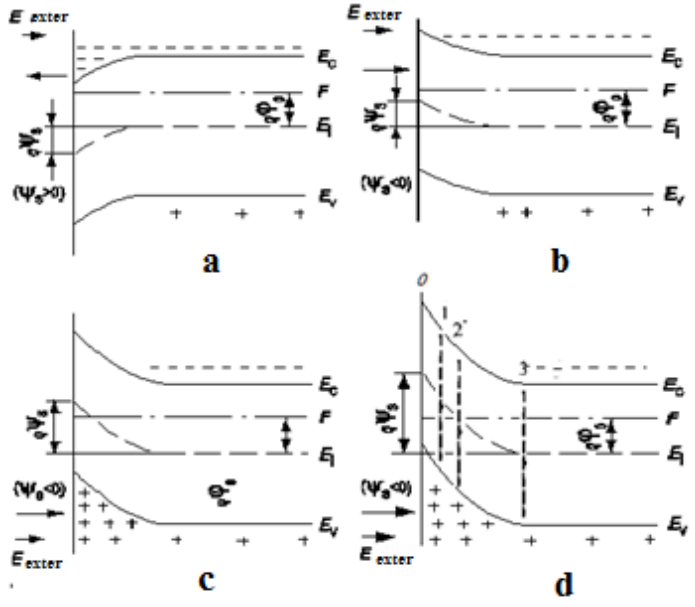



Fig. 2.5.3. Types of SCR in an n-type crystal: enrichment for acceptor-like SES (a); depletion (b), weak inversion (c), and strong inversion (d) for donor-like SES. The digits denote the boundaries of the SCR (3), weak inversion (area 1-2), strong inversion (0-1), and depletion region (2-3).

Just this change in the type of conductivity in Fig. 2.5.3c,d is called the inversion, and the under-surface layer of the *SCR* is called *inversion region*. If the density of the SES is small, a weak inversion can be observed (Figure 2.5.3c), and for a high SES density, a strong inversion (region 0-1 in Figure 2.5.3d). Behind the region of strong inversion in Fig. 2.5.3d in the region 1-2 the region of weak inversion lies, and then in the region 2-3 the *depletion region* lies.

A more precise classification of the types of *SCR* is shown in Table. 2.5.1. In the same place, the orienting values of the band bendings corresponding to the appearance of various types of SCR are given. In particular, the inversion begins since that part of the crystal, where the Fermi level crosses the $E_i$ curve, a strong inversion - where the Fermi energy is at the same distance from $E_i$ as in the bulk, but "symmetrically" with respect to it.

From the basic equation of the SCR (2.5.16) it follows that in the general case the dependence of the potential on the coordinate in the model of a homogeneous surface can be found by integrating

$$\frac{z}{L_D} = \mp \int_{Y_S}^{Y} \frac{dY}{F(\lambda, Y)}. \qquad (2.5.36)$$

This equation has analytical solutions only in particular cases. More precise conditions imposed on the values of the band bendings in different cases follow from the analogy of approximate expressions for the function $F(\lambda,Y)$, which are given in Table. 2.5.1.

Table 2.5.1. Classification of SCR types

| Type of SCR | | Semiconductor of p-type | | Semiconductor of n-type | |
|---|---|---|---|---|---|
| The enrichment | | $p > p_o$ | $Y < 0$ | $N > n_0$ | $Y > 0$ |
| Strong enrichment | | $p >> p_o$ | $Y < -2,5$ | $n >> n_0$ | $Y > 2,5$ |
| The depletion | | $n < p < p_o$ | $0 < Y < \ln\lambda$ | $p < n < n_0$ | $\ln\lambda < Y < 0$ |
| Exhaustion | Strong depletion | $n < p << p_o$ | $2,5 < Y < \mathrm{lu}\lambda$. | $p < n « n_0$ | $\ln\lambda < Y < -2,5$ |
| | Weak inversion | $p < n < p_o$ | $\mathrm{In}\lambda < Y < 21\mathrm{n}\lambda$ | $n < p < n_0$ | $2\ln\lambda < Y < \ln\lambda$ |
| Strong inversion | | $n > p_o$ | $Y > 21\mathrm{n}\lambda$ | $p > n_0$ | $Y < 2 \ln\lambda$ |

As it follows from the table, at enrichment we can neglect in the function $F(\lambda,Y)$ in (2.5.36) the terms, which are associated with minority carriers and ionized impurities, under strong inversion with ionized impurities and majority carriers in the bulk, and finally, in conditions of exhaustion - with all mobile carriers.

***Strong enrichment and strong inversion.*** For definiteness, let us consider an n-type semiconductor with a large concentration of donor-like SES (Figure 2.5.3b). Using the approximation of the function $F(\lambda,Y)$ (see Table 2.5.2) and integrating (2.5.36), we obtain for strong enrichment

$$Y = -2\ln\left(z / \sqrt{2} L_D + e^{-Y_S/2}\right). \qquad (2.5.37)$$

Numerical calculation by this formula for $Y_S = 10$ and $z/L_D = 0.1$ gives $Y = 5$ and, therefore, at a distance of only $0,1L_D$ from the surface, the value of $Y$ decreases by 5 units. Obviously, the strong enrichment layer is characterized by a very small thickness. Since in the strong inversion mode the function $F(\lambda,Y)$ has approximately the same form as in the case of strong enrichment, this conclusion holds true for strong inversion layers.

Table 2.5.2. Function $F(\lambda,Y)$ for various modes in the SCR

| Type of SCR | | Semiconductor of p-type | Semiconductor of n-type |
|---|---|---|---|
| Any | | $\sqrt{2}(e^{-Y}-1+\lambda^{-2}e^{Y}+Y)^{1/2}$ | $\sqrt{2}(e^{Y}-1+\lambda^{2}e^{-Y}-Y)^{1/2}$ |
| $\lvert Y\rvert \ll 1$ | | $Y$ | $Y$ |
| The enrichment | | $\sqrt{2}(e^{-Y}+Y-1)^{1/2}$; $Y<0$ | $\sqrt{2}(e^{Y}-Y-1)^{1/2}$; $Y>0$ |
| Strong enrichment | | $\sqrt{2}\cdot e^{-Y/2}$ ; $Y\le -4$ | $\sqrt{2}\cdot e^{Y/2}$ ; $Y\ge 4$ |
| The depletion | | $\sqrt{2}(e^{-Y}+Y-1)^{1/2}$; $Y>0$ | $\sqrt{2}(e^{Y}-Y-1)^{1/2}$; $Y<0$ |
| Exhaustion | Strong depletion | $\sqrt{2}(Y-1)^{1/2}$ ; $2,5\le Y<\ln\lambda$ | $\sqrt{2}(-Y-1)^{1/2}$ ; $\ln\lambda<Y\le -2,5$ |
| | Weak inversion | $\sqrt{2}(Y-1)^{1/2}$ ; $\ln\lambda<Y<2\ln\lambda+\ln(Y-1)$ | $\sqrt{2}(-Y-1)^{1/2}$ ; $2\ln\lambda-\ln(-Y-1)<Y<\ln\lambda$ |
| Strong inversion | | $\sqrt{2}\lambda^{-1}e^{Y/2}$ ; $Y>2\ln\lambda+\ln(Y-1)$ | $\sqrt{2}\lambda e^{-Y/2}$ ; $Y<2\ln\lambda-\ln(-Y-1)$ |

***Depletion layer.*** Consider a p-type semiconductor. Taking into account the approximation of the function $F(\lambda,Y)$ in Table. 2.5.2, we obtain after integrating (2.5.36) and a small transformation

$$(Y-1)\cong(Y_S-1)(1-z/W_d)^2, \tag{2.5.38}$$

where

$$W_d=L_D\sqrt{2(Y_S-1)}. \tag{2.5.39}$$

Thus, for a depleted layer, a weak, quadratic dependence of the potential on the coordinate is characteristic. In this case, the physical meaning of $W_d$ value corresponds to the width of the depleted layer. The depleted layer forms on the surface a potential barrier for the majority carriers, similarly to the Schottky barrier. The width of this barrier can reach several Debye lengths (for example, $W_d = 4{,}2\ L_D$ for $Y_S = 10$).

Since the band bending in the depletion layer is limited to the maximum value

$$|Y_{max}|=2|\ln\lambda|, \tag{2.5.40}$$

the thickness of the equilibrium depletion layer in a material characterized by the parameter of nonintrincity $\lambda$, can not be greater

$$W_d^{max}=L_D\sqrt{2(2|\ln\lambda|-1)}. \tag{2.5.41}$$

With band bending increase above $|Y_{max}|$, the depletion layer of maximum thickness is "pushed out" into the crystal depth which is close to a layer of strong inversion near the surface.

### 2.5.4. Surface excesses of charge carriers

Surface excesses of electrons $\Gamma_n$ and holes $\Gamma_p$ are defined as their excess amount per unit surface of the crystal, in comparison with the case of plane bands for the semiconductor bulk (when $n = n_o$ and $p = p_o$):

$$\Gamma_n = \int_0^\infty (n(z) - n_o)dz, \tag{2.5.42a}$$

$$\Gamma_p = \int_0^\infty (p(z) - p_o)dz, \tag{2.5.42б}$$

where $n(z)$ and $p(z)$ are the electron and hole concentrations in the SCR, and $n_0$ and $p_0$ are the hole concentrations in the quasineutral bulk. According to (2.5.42), the quantities $\Gamma_p$ and $\Gamma_n$ can be either positive or negative. In accordance with this definition, for $Y_S = 0$, the surface excesses are zero.

Thus, the excess of electrons or holes is an excess of the number of free carriers in the SCR per unit surface area, in comparison with the equilibrium in the neutral bulk. In some literature sources, sometimes these excesses of free carriers $\Gamma_{p,n}$ are called surface concentrations. This is not entirely true, because the surface concentration in its meaning is the number of free charge carriers per unit volume, calculated on the surface of the semiconductor, whereas the excesses $\Gamma_{p,n}$ are the excesses of of free carriers number, integrated over the depth of the SCR and calculated per unit surface area.

Using (2.5.9), we can obtain from (2.5.42) the following expressions for the excesses in terms of the potential $\psi$:

$$\Gamma_n = n_0 \int_0^\infty \left(e^{-\beta\psi} - 1\right)dz = n_0 \int_{\psi_s}^0 \frac{e^{-\beta\psi} - 1}{\frac{d\psi}{dz}} d\psi,$$

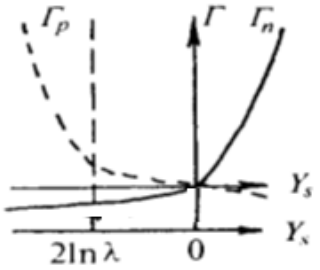
$$\Gamma_p = p_0 \int_0^\infty \left(e^{-\beta\psi} - 1\right)dz = p_0 \int_{\psi_s}^0 \frac{e^{-\beta\psi} - 1}{\frac{d\psi}{dz}} d\psi. \tag{2.5.43a}$$

Using equation (2.5.36), we transform (2.5.42) to the form:

$$\Gamma_n = \int_0^\infty n_0 \left(e^{-Y} - 1\right)(dz/dY)dz = n_0 L_D \int_{\psi_s}^0 \frac{e^{-Y} - 1}{F(\lambda, Y)} dY,$$

$$\Gamma_p = \int_0^\infty p_0 \left(e^{-Y} - 1\right)(dz/dY)dz = p_0 L_D \int_{\psi_s}^0 \frac{e^{-Y} - 1}{F(\lambda, Y)} dY \tag{2.5.43б}$$

Fig. 2.5.4. The nature of the dependence on the surface potential of the excesses of electrons and holes (for an n-type semiconductor)




As before, the upper sign is chosen at $Y > 0$, the lower one - at $Y < 0$. The nature of the surface excess dependences on the surface potential for n-type material is illustrated in Fig. 2.5.4a.

In the modes of strong enrichment (or strong inversion), the excesses of the majority (minority) carriers grow exponentially, in accordance with the expressions $\Gamma_n = \frac{\varepsilon_s \varepsilon_0 kT}{\sqrt{2} q^2 L_D} e^{\frac{\beta\psi}{2}}$ and $\Gamma_p = \frac{\varepsilon_s \varepsilon_0 kT}{\sqrt{2} q^2 L_D} e^{\frac{\beta\psi}{2}}$. For the band bending, giving depletion or inversion, the excesses of the majority carriers tend to the limiting value corresponding to the complete elimination of these carriers from the *SCR*. For minority carriers, the rapid growth of excess begins when a layer of strong inversion occurs, when $|Y_s| \geq 2|\ln\lambda|$.

### 2.5.5. Total charge of the SCR

We express the total charge of the *SCR* per unit surface in terms of the excess quantities for free charge carriers:

$$Q_{sc} = \int_0^\infty \rho(z)dz = q\int_0^\infty \{[p(z) - p_0] - [n(z) - n_0]\}dz = q(\Gamma_p - \Gamma_n) \qquad (2.5.44)$$

Taking into account (2.5.43), we obtain:

$$Q_{sc}(\psi_s) = \pm qn_i L_D \int_0^{\psi_s} \frac{\lambda(e^{-\psi} - 1) - \lambda^{-1}(e^{\psi} - 1)}{F(\lambda, \psi)} d\psi, \qquad (2.5.45а)$$

$$Q_{sc}(Y_s) = \pm qn_i L_D \int_0^{Y_s} \frac{\lambda(e^{-Y} - 1) - \lambda^{-1}(e^{Y} - 1)}{F(\lambda, Y)} dY. \qquad (2.5.45б)$$

We note that, despite the apparent difference between (2.5.45) and the previously presented formula (2.5.22), they are identical, that is easily proved by their differentiation with respect to $Y_S$.

Having the curves of $\Gamma_n(Y_S)$ and $\Gamma_p(Y_s)$ in Fig. 2.5.4, we can construct, using the relation (2.5.45), the dependences of $Q_{sc}(Y_s)$, which are presented below in Fig. 2.5.5a.

### 2.5.6. Differential capacity of SES in a semiconductor

Thus, as shown above, when the crystal electric neutrality condition is fulfilled, the charge of the *SCR* is compensated by the same charge, but opposite in sign, on its surface. Such a system is an analog of a capacitor, one of which is where one plate is the surface of the crystal, and the other is the bulk. The difference from the usual flat capacitor consists in the fact that the charge of such a "capacitor" (SCR + SES) is distributed in space, and the position of the centroid of this charge depends on the surface potential of the crystal (the density of SES).

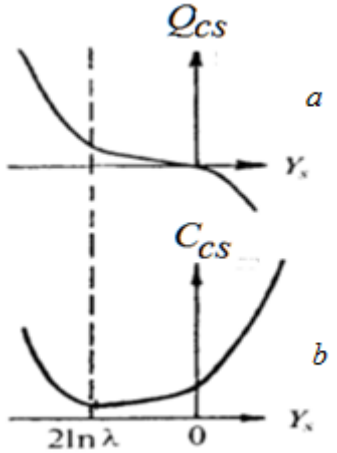
For this reason, the capacitance of the "surface-SCR" capacitor is characterized by a differential capacitance

$$C_{sc} = -\frac{dQ_{sc}}{d\psi_s} = -\frac{q}{kT}\frac{dQ_{sc}}{dY_s}. \qquad (2.5.46)$$

The minus sign in (2.5.46) is introduced so that the value of the $C_{SC}$ is positive. Differentiating (2.5.45) with respect to $Y_S$, we obtain

$$C_{sc}(Y_s) = \mp\frac{\varepsilon\varepsilon_0}{L_D}\left[\frac{\lambda(e^{-Y_s}-1)-\lambda(e^{Y_s}-1)}{(\lambda+\lambda^{-1})F(\lambda,Y_s)}\right]. \qquad (2.5.47)$$

Fig. 2.5.5. The nature of the dependence on the surface potential of the charge of the SCR (a) and the differential capacitance of the SCR (b) for the n-type semiconductor




Recall that the upper sign is chosen for $Y_S > 0$, the lower sign for $Y_S < 0$. The characteristic form of the $C_{SC}(Y_S)$ function is shown in Fig. 2.5.5b.

In the layers of enrichment and strong inversion, the capacity of the *SCR* grows exponentially with increasing of $|Y_S|$: $L_D\sqrt{2}e^{-|Y_s|/2}$ or $L_D\sqrt{2}e^{|u_B|-|Y_s|/2}$, respectively). The capacity of the depletion layer decreases $\left[-(|Y_s|-1)^{-1/2}\right]$ slowly with increasing $|Y_S|$ due to the growth of the the depleted layer thickness. Near the potential corresponding to the onset of a strong inversion ($Y_S = 2ln\lambda$), the $C_{sc}$ value reaches a minimum. Analytical expression for the potential of a minimum of the *SCR* differential capacity can not in general be obtained. For an intrinsic semiconductor, the minimum is observed for plane bands ($Y_S = 0$), and the dependence of $C_{sc}(Y_S)$ is symmetric. The more improper is the material, the greater is the value of the minimum capacity of the *SCR*.

The differential capacitance of the SCR can be expressed in terms of the "effective" distance between the plates of an equivalent flat capacitor

$$C_{sc} = \varepsilon_0\varepsilon/W_{ef}. \qquad (2.5.48)$$

In Table. 2.5.3 approximate formulas are given for calculating the differential capacitance of *SCR*s in different modes, as well as the corresponding values of $W_{ef}$.

Table 2.5.3. Differential capacitance and effective SCR width

| Type of SCR | Differential Capacitance of the SCR $C_{sc}$ | Effective Width of the SCR $W_{ef}$ |
|---|---|---|
| Flat bands | $\varepsilon\varepsilon_0/L_D$ | $L_D$ |
| Exhaustion | $\varepsilon\varepsilon_0/L_D\sqrt{2(Y_s-1)}$ | $L_D\sqrt{2(Y_s-1)}$ |
| Strong enrichment | $\varepsilon\varepsilon_0/L_D\sqrt{2}e^{-\lvert Y_s\rvert/2}$ | $L_D\sqrt{2}e^{-\lvert Y_s\rvert/2}$ |
| Strong inversion | $\varepsilon\varepsilon_0/L_D\sqrt{2}e^{\lvert u_B\rvert-\lvert Y_s\rvert/2}$ | $L_D\sqrt{2}e^{\lvert u_B\rvert-\lvert Y_s\rvert/2}$ |

As follows from Table. 2.5.3, the effective width of the *SCR* for small bending of the energy bands is equal to the Debye screening length. For the depletion layer, $W_{ef}$ is the same as $W_d$ in (2.5.39). The effective width of the layers of strong enrichment or strong inversion can be much less than $L_D$. The capacitance of the SCR $C_{sc}$ for the enrichment region (for $\psi_s < 0$) is due to the

capacity of the free holes $C_{sc} = C_p = \frac{\varepsilon_s \varepsilon_0}{L_D} e^{\frac{\beta Y_s}{2}}$. Estimates show that with strong enrichment $|Y_s|$ = 10 the width of the layer is very small ($W_{ef} = 10^{-2} L_D$).

For the depletion region or weak inversion ($2\varphi_0 > \psi_s > 0$), the *SCR* capacitance $C_{sc}$ is due to the capacity of the ionized acceptors region $C_{sc} = C_B = \frac{\varepsilon_s \varepsilon_0}{W}$. It follows from this relation that the capacitance $C_{sc}$ in the depletion region depends weakly on the surface potential $\psi_s$, decreasing with the growth of the latter. With a strong depletion $|Y_S| = 10|$ the depleted layer has a width $W_{ef} = W_d = 6{,}3L_D$. The minimum capacitance $C_{sc}$ in this case is reached near the threshold value of the surface potential. The capacitance of the *SCR* in the region of depletion and weak inversion is equivalent to the capacitance of a flat capacitor filled with a dielectric with a relative permittivity $\varepsilon_s$ whose plates are separated by a distance $W$ equal to the width of the *SCR*.

For the strong inversion region ($\psi_s > 2\varphi_0$), the SCR capacitance $C_{sc}$ is due to the capacity of free electrons in the inversion layer. For sufficiently large values of the surface potential $\beta(\psi_s - 2\varphi_0) \geq 7$ it will be equal $C_n = \frac{\varepsilon_s \varepsilon_0}{\sqrt{2} L_D} e^{\frac{\beta(\psi_s - 2\varphi_0)}{2}}$.

It follows from the analysis that the capacities of free carriers at enrichment and strong inversion exponentially depend on the surface potential $\psi_s$ and have the same values if the value of the surface potential is counted for inversion from the threshold value $\psi_s = 2\varphi_0$.

As already mentioned in paragraph 2.5.3, under the strong inversion layer there is always a depleted layer of width $W_d^{\max}$. The total charge of the *SCR* in this case consists of the charge of the depletion layer ($Q_d$) and the charge of the strong inversion region ($Q_{in}$): $Q_{sc} = Q_d + Q_{in}$ Differentiating this expression with respect to the surface potential, we find that under equilibrium conditions the total capacitance of the *SCR* in the strong inversion mode is equal to the sum of the capacities of the depleted and inversion layers

$$C_{sc} = -\frac{dQ_d}{d\psi_s} + \frac{dQ_{in}}{d\psi_s} = C_d + C_{in}. \qquad (2.5.49)$$

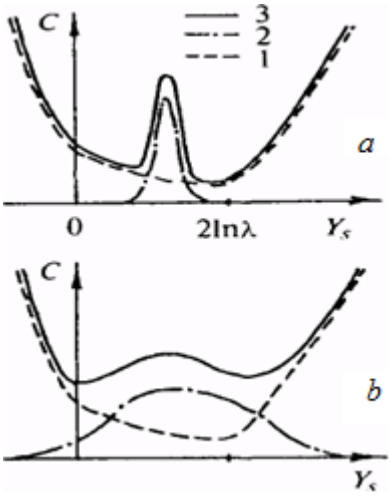
It is obvious that the total charge of the *SES* in (2.5.23) depends on the type of surface states, in particular, on their energy distribution. Therefore, it makes sense to consider the expressions for the differential capacity in the case of different types of energy spectrum of *SES*.

***Discrete energy spectrum of SES.*** For monoenergetic *SES*, the differential capacitance is as follows:

$$C_t = -\frac{dQ_t}{d\psi_s} = -\frac{q}{kT}\frac{dQ_t}{dY} = -\frac{q}{kT}\frac{dQ_t}{du_s} \qquad (2.5.50)$$

As in relation (2.5.46), the minus sign is needed in order that the value of $C_t$ be positive (with

Fig. 2.5.6. The nature of the dependence of the total differential capacities on the surface potential for a discrete (a) and quasi-continuous (b) energy spectrum of *SES*. The figures show the contributions from the *SCR* (1) and *SES* (2).




an increase in the surface potential, the charge of the *SES* always decreases).

Differentiating (2.5.4), with allowance for (2.5.3), we obtain that the differential capacity of acceptor and donor-like monoenergetic *SES* is described by the same formula.

$$C_t = -\frac{q^2}{kT} N_t \frac{df_{t0}}{du_s} = -\frac{q^2}{kT} N_t f_{t0}(1 - f_{t0}) \quad (2.5.51)$$

The function $f_{t0}(1- f_{t0})$ is a fairly narrow peak with a maximum value of 0.25, attained at $F = E_t$, and a width at half-height equal to -3.5$kT$. It is this function that determines the character of the $C_t(Y_s)$ dependence. Calculations show that the capacitance value takes the maximum value

$$C_t^{\max} = (q^2/4kT)N_t \quad (2.5.52)$$

when the Fermi level crosses $E_t$. When $F$ is removed from $E_t$ by an amount of the order of -3$kT$, the differential capacitance of the surface with monoenergetic *SES* becomes negligible. If several monoenergetic *SES* are present on the surface, the total differential capacitance of the *SES* is equal to the sum of the differential capacitances of all states, independently of their type:

$$C_s = -\frac{dQ_s}{d\psi_s} = -\frac{d}{d\psi_s}(\sum_i Q_{ti}) = \sum_i (-\frac{dQ_{ti}}{d\psi_s}) = \sum_i C_{ti} \quad (2.5.53)$$

At a given value of the surface potential, the contribution to the total differential capacitance of the *SES* in (2.5.53) will be given only by those states for which $|E_t - F| \le 2kT$ .

***Quasicontinuous energy spectrum of SES.*** In the case of a quasicontinuous energy spectrum of *SES*, its differential capacitance is determined by integrating over the energy (2.5.51):

$$C_s = -\frac{q^2}{kT} \int_{Ev}^{Ec} Ns(E) \cdot \frac{df_{t0}}{du_{ts}} dE \quad (2.5.54)$$

If the energy spectrum of the *SES* is sufficiently smooth, i.e. within the energy range (3-4) $kT$ near the Fermi level, the *SES* density $N_s^F$ can be regarded as constant, then the relation (2.5.54) is simplified

$$C_s = \frac{q^2}{kT} N_s(F) \int_{E_v}^{E_c} \frac{df_{t0}}{du_{ts}} dE \quad (2.5.55)$$

Since $df = kTdE$, $\frac{df_{t0}}{du_s} = -\frac{df_{t0}}{d\varepsilon}$ and , we obtain:

$$C_s = q^2 N_s(F) \int_{E_v}^{E_c} \frac{df_{t0}}{d\varepsilon} d\varepsilon = q^2 N_s(F) \left[ f_{t0}(E_v) - f_{t0}(E_c) \right]. \qquad (2.5.56)$$

Since in the absence of degeneracy $f_{t0}(E_v) = 1$ и $f_{to}(E_c) = 0$, we finally obtain expression

$$C_s = q^2 N_s(F) \qquad (2.5.57)$$

We emphasize that relation (2.5.57) is valid only under conditions of thermodynamic equilibrium with a sufficiently smooth energy spectrum of the *SES*.

***Total differential capacity of the surface.*** If an external electric field is applied perpendicular to the surface of the crystal, an electric charge is induced in it $\left(Q_{ind} = Q_{sc} + Q_s\right)$, which is distributed between the *SCR* and the *SES*. In this case, the total differential capacity of the crystal surface has the form

$$C = -\frac{dQ_{ind}}{d\psi_s} = -\frac{dQ_{sc}}{d\psi_s} - \frac{dQ_s}{d\psi_s} = C_{sc} + C_S \,. \qquad (2.5.58)$$

This means that an equivalent surface scheme can be represented in the form of parallel-connected capacitances, given by the *SCR* and *SES*. Consequently, the dependence of the total differential capacitance *C* on the surface potential is the result of the addition of the functions $C_{sc}(Y_s)$ and $C_s(Y_s)$. Possible types of such dependencies in the case of the presence of contributions from the *SCR* (1), *SES* (2) to the total capacitance (solid curve 3) are shown in Fig. 2.5.6 for discrete (a), and quasicontinuous (b) energy spectra of SES.

## 2.6. Features of the electronic structure of the semiconductor surface in a nonequilibrium state

### 2.6.1. The concept of quasi-equilibrium

In many cases of practical importance, under external influences, the thermodynamic equilibrium of the electron-hole system in a solid is violated. This happens, for example, when the crystal is illuminated, or when an electric field is applied to it, etc. In this case, the concentrations of charge carriers both in the bulk of the semiconductor and in the *SCR* at the surface can differ substantially from the equilibrium ones. In order to find the dependences of the various characteristics of the *SCR* from the distance to the surface *z*, it is necessary to solve jointly the Poisson equation (2.5.7) and the continuity equations (2.5.8), which in this case should describe the rates of change in the concentrations of free charge carriers due to their generation, recombination, diffusion and drift under external influences. In this case, it is necessary to take into account the dependence of the rate of surface recombination on the potential surface. In general, this problem is very complicated, therefore, as a rule, we should look for simplified ways to solve it.

The consideration of a thermodynamically nonequilibrium situation is greatly simplified in the case of quasi-equilibrium conditions, when one can use the usual statistical relationships for the concentrations of free charge carriers. *Quasiequilibrium* is achieved when *the energy distributions of equilibrium ("thermal") and nonequilibrium carriers (excited by external action) in the allowed zones are identical*. In this case, instead of one equilibrium system Fermi level common to the whole system, the Fermi quasi levels are used - separately for electrons $F_n$ and separately for holes $F_p$.

Immediately after generation (for example, under the action of light), nonequilibrium carriers

can be "hot", i.e. have a kinetic energy significantly higher than the average heat energy ($kT$). However, this excess energy can be transferred to the crystal lattice in a time much shorter than the lifetime $\tau_r$ of the nonequilibrium charge carriers in the allowed bands ($\tau_r \sim 10^{-6}$-$10^{-3}$ s). For example, for semiconductor materials with carrier mobility $\mu = (0.1 \div 1)$ cm$^2$V$^{-1}$s$^{-1}$ and an effective mass $m^*$ close to the free electron mass, the time between two successive collisions of a carrier with a lattice $\tau_l = \mu \cdot m^*/q \approx (10^{-13} \div 10^{-12})$ s. It is easy to estimate that the excess kinetic energy of the hot carrier $\Delta E \approx (0.1 \div 1)$ eV will be transferred to the crystal lattice in the form of phonons after $(10^2 \div 10^3)$ collisions. Consequently, the hot carrier is "cooled" to the lattice temperature in a time $\tau_E \approx (10^2$-$10^3)\,\tau_l \approx (10^{-11} \div 10^{-9})$ c $<< \tau_r$. Therefore, practically all the time of stay of nonequilibrium carriers in the allowed bands, their energy distribution will be similar to the "thermal" ones, and the use of separate Fermi quasi levels for all electrons ($F_n$) and for all holes ($F_p$) is completely justified.

### 2.6.2. The basic equation of an SCR under conditions of quasi equilibrium

When analyzing nonequilibrium processes in the near-surface layers, it is often necessary to encounter generation of charge carriers near the surface. For example, when a semiconductor is illuminated in the self-absorptive band, the electron-hole pairs are generated in the near-surface layer of thickness $h_o = (10 \div 100)$ nm, from which they then diffuse into the crystal by an amount called diffusion

$$L_{dif} = \sqrt{D\tau_r}, \qquad (2.6.1)$$

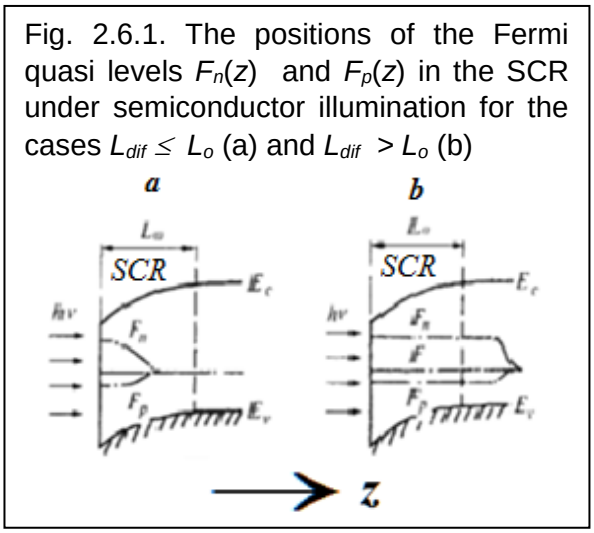


Fig. 2.6.1. The positions of the Fermi quasi levels $F_n(z)$ and $F_p(z)$ in the SCR under semiconductor illumination for the cases $L_{dif} \le L_o$ (a) and $L_{dif} > L_o$ (b)

where $D$ is the ambipolar diffusion coefficient, approximately equal to the diffusion coefficient of the minority charge carriers. Depending on the relationship between $L_{dif}$ and the total $SCR$ width $L_o$, two situations can arise, which are shown in Fig. 3.5:

a) If $L_{dif} < L_o$, the positions of the Fermi quasiparticles $F_n(z)$ and $F_p(z)$ in the forbidden band vary within the $SCR$ and are usually used only for qualitative consideration. Quantitative calculations of the characteristics of the $SCR$ are made on the basis of a joint solution of Poisson's equations and continuity.

b) For $L_{dif} > L_o$ and not very large flows to the surface of the recombining carriers, a quasi-equilibrium is achieved in the $SCR$, as a result of which $F_n(z)$ and $F_p(z)$ are constant throughout the entire $SCR$. In this case, the quantitative calculations of the characteristics of the $SCR$, as will be clear from what follows, are substantially simplified. The quasi equilibrium conditions in the $SCR$ are usually well fullfiled for germanium and silicon.

Under the conditions of quasie quilibrium (for $L_o < z < L_{dif}$), the free electron and hole concentrations in the $SCR$ can be expressed in terms of the Fermi quasilevels by the following relations:

$$n^*(z) = N_C \cdot \exp\left[\frac{F_n - E_c(z)}{kT}\right] = n_0^* \cdot e^{Y^*},$$
$$p^*(z) = N_V \cdot \exp\left[\frac{E_v(z) - F_p}{kT}\right] = p_0^* \cdot e^{-Y^*}, \qquad (2.6.2)$$

where $N_V$ and $N_C$ are the effective density of states in the valence band and the conduction band, respectively, $n_0$* and $p_0$* are the electron and hole concentrations in the volume.

Taking into account the electric neutrality of the semiconductor bulk and assuming, as before, that the impurities are completely ionized, we obtain an expression for the density of the space charge analogous to (2.5.8) under conditions of quasi-equilibrium

$$\rho^* = q\left[\left(p^* - p_0^*\right) - \left(n^* - n_0^*\right)\right]. \tag{2.6.3}$$

After substituting $\rho$* into the Poisson equation (2.5.7), we obtain relations similar to (2.5.10) and (2.5.12), in which all parameters will be marked with "stars". In this case, instead of the equilibrium concentration of eigenstates $n_i$, and the non-intrincity parameter $\lambda$, we get quasi-equilibrium quantities

$$n_i^* = \sqrt{n_0^* p_0^*} \quad \text{и} \quad \lambda^* = p_0^* / n_i^* = n_i^* / n_0^*. \tag{2.6.4}$$

Obviously, the solution of the Poisson equation in the quasi-equilibrium case will have exactly the same form as (2.5.7). The role of the Debye screening length will be played by the quantity

$$L_D^* = \left[\frac{\varepsilon\varepsilon_0 kT}{q^2\left(n_0^* + p_0^*\right)}\right]^{1/2}. \tag{2.6.5}$$

Despite the external similarity of the solutions of the Poisson equation in the quasi-equilibrium and equilibrium cases, the values of all the parameters of the SCR (surface potential, excesses of electrons and holes, charge of the *SCR*, differential capacitance, etc.) will be different in these cases.

### 2.6.3. Levels of injection and quasi-Fermi levels

In order to quantify the amount of deviation of the free-carrier system from thermodynamic equilibrium, injection levels are introduced for the majority and minority carriers. Without concretizing the type of the majority carriers, we determine the injection *levels for electrons* and *holes*

$$\delta_n = \Delta n_0 / n_0; \quad \delta_p = \Delta p_0 / p_0; \quad \Delta n_0 = n_0^* - n_0; \quad \Delta p_0 = p_0^* - p_0. \tag{2.6.6}$$

If in the semiconductor bulk, there is no capture of the nonequilibrium carriers on the traps ("sticking" to deep levels and *SES*), then from the electrical neutrality of the semiconductor bulk it follows that, therefore, the $\Delta n_0 = \Delta p_0$ injection levels for electrons and holes are interrelated

$$\delta_n = \Delta n_0 / n_0 = (\Delta p_0 / p_0)(p_0 / n_0) = \delta_p \lambda^2. \tag{2.6.7}$$

Sometimes is introduced a quantity

$$\delta = \Delta n_0 / n_i = \Delta p_0 / n_i = \sqrt{\delta_n \delta_p}, \tag{2.6.8}$$

which is called the level of injection for majority carriers or simply the level of injection. Taking into account (2.6.6), we can write

$$n_0^* = n_0(1 + \delta_n); \quad p_0^* = p_0(1 + \delta_p). \tag{2.6.9}$$

The values of $\delta_n$ and $\delta_p$ in each particular case can be found, for example, from the change in the bulk conductivity of the crystal when the thermodynamic equilibrium

$$\frac{\sigma_0^* - \sigma_0}{\sigma_0} \equiv \frac{\Delta\sigma_0}{\sigma_0} = \frac{\mu_n \Delta n_0 + \mu_p \Delta p_0}{\mu_n n_0 + \mu_p p_0} = \frac{1 + b}{\lambda^2 + b} \cdot \delta_n = \frac{1 + b}{1 + b\lambda^{-2}} \delta_p, \tag{2.6.10}$$

where $\sigma_o$ and $\sigma_o^*$ are the equilibrium and quasi-equilibrium conductivities in the bulk, $\mu_n$ and $\mu_p$ - the volume mobilities of electrons and holes, and $b = \mu_n/\mu_p$.

Defining $\delta_n$ and $\delta_p$, you can calculate the value of the non-intrincisity parameter

$$\lambda^* = \left(\frac{p_0 + \Delta p_0}{n_0 + \Delta n_0}\right)^{1/2} = \left(\frac{\lambda^2 + \delta_n}{1 + \delta_n}\right)^{1/2} = \lambda\left(\frac{1 + \delta_p}{1 + \delta_n}\right)^{1/2} \tag{2.6.11}$$

The values of $F_n$ and $F_p$ shifts with respect to the equilibrium Fermi level position are directly related to the corresponding injection levels

$$\frac{F_n - F}{kT} \equiv u_n = \ln(\delta_n + 1); \quad \frac{F - F_p}{kT} \equiv u_p = \ln(\delta_p + 1). \tag{2.6.12}$$

When transition from equilibrium to quasi-equilibrium state, the boundaries between the regions of strong depletion and weak inversion, weak and strong inversion are shifted (see Table 2.5.1), since instead of the bulk potential $u_b = -Ln\,\lambda$, we need to use $u_b^* = -Ln\,\lambda^*$. The bulk potential of the crystal decreases by modulo, shifting by half the distance between the Fermi quasi-levels for electrons and holes (the semiconductor becomes more intrinsic, as it was,) $u_B^* - u_B = (u_n - u_p)/2$. The boundary of the strong inversion region under quasi-equilibrium conditions is shifted on the value of the quasi-Fermi level " splitting " $2u_B^* = 2u_B + (u_n - u_p)$. Accordingly, the characteristic minimum point on the dependence of the differential capacitance of the *SCR* on the surface potential (Figure 2.6.1b) will be observed for smaller values of the energy bands bending ($u_n$ - $u_p$).

The injection levels for the majority and minority charge carriers, as well as the shifts of the Fermi quasi-levels, $F_n$ and $F_p$, can differ very much from each other. As an example, let us consider a *p*-type silicon typical for microelectronic engineering with a resistivity of 10 Ohm·cm at 300 K ($p_o = 1{,}5\cdot10^{15}$ cm$^{-3}$, $n_i = 1{,}5\cdot10^{10}$ cm$^{-3}$, $\lambda = 10^5$, $u_b = -Ln\,\lambda = -11{,}5$; $b = \mu_n/\mu_p = 2{,}9$) and assume that under illumination its bulk conductivity has changed by 1% (estimates show that for this it is necessary to bring to the crystal for a distance about 1 meter from light bulb with a power of 100 watts). Using relations (2.6.8) and (2.6.9), we obtain $\delta_p \cong 2{,}6\cdot10^{-3}$; $u_p = \ln(1 + \delta_p) \cong \delta_p$; $\delta_n \cong 2{,}6\cdot10^7$; $u_n \cong \ln\delta_n \cong 17$.

The potential corresponding to the minimum of the differential capacitance of the *SCR* will decrease by 17 *kT/q* (от $2u_b$ = -23 *kT/q* до $u_b^*$ = -6 *kT/q*). It is obvious that when the level of injection the majority charge carriers approaches unity, the semiconductor becomes practically its

intrincic.

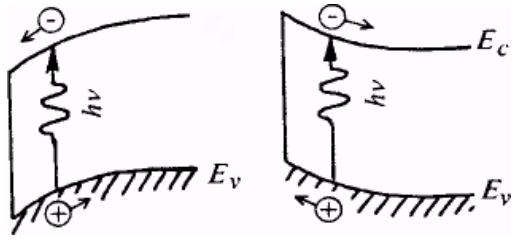
### 2.6.4. Surface photo-EMF with quasi-equilibrium in the SCR

The presence of surface defects that lead to the formation of *SES* and, accordingly, the formation of an *SCR* near the surface of a semiconductor crystal leads to a number of additional consequences. One of them is the appearance of several types of surface photo-*EMF* under illumination. A *surface photo-EMF* is called the additional potential difference that arises between the surface and the bulk of a crystal as a result of a semiconductor illumination.

We indicate three main causes of the of a surface photo-*EMF* appearance.

*1. Separation of the generated electron-hole pairs by the electric field of the SCR.* The built-in electric field in the *SCR* causes separation of the electrons and holes entering the pairs (their movement in different directions), so that carriers of the same sign "cling" to the surface, and the other push off from it (see Figure 2.6.2). The spatial separation of nonequilibrium electrons and holes in all cases leads to a decrease in the magnitude of the band bending. In what follows, we will call this component the photovoltaic *barrier* (or *photoelectric EMF* of the *SCR*).

ig. 2.6.2. Mechanism of the formation of photovoltaic *EMF* under various band bending




If there is no field in the SCR ($Y_S$ = 0), the spatial separation of the non-equilibrium carrier carriers does not occur and the barrier photo-*EMF* does not arise.

*2. Ambipolar diffusion of electron-hole pairs from the generation region.* As a result of the action of light quanta by a thin near-surface layer, electron-hole pairs are generated which, due to the concentration gradient, begin to diffuse deep into the crystal. Moreover, more mobile charge carriers of the pair (usually electrons) are somewhat ahead of the holes, so that a potential difference arises between the surface and the bulk of the crystal, which is called the *Dember EMF*. This component of the photo-*EMF* is in no way connected with the presence of the *SCR*, since it is formed at a distance of the order of the diffusion length from the surface.

*3. Recharge of SES.* The violation of thermodynamic equilibrium as a result of illumination of the semiconductor surface arises not only because of a change of free charge carriers concentrations, but also their concentration on the localized *SES*. Changing the surface charge leads to the appearance of an additional potential difference between the surface of the crystal and the bulk. The magnitude of this component of the photo-*EMF* depends on the initial band bending, the level of injection of nonequilibrium charge carriers, and the parameters of the *SES*. In contrast to the *SCR* and the Dember photoelectric *EMFs*, photovoltage can also arise in the case of flat zones due to charge exchange of the *SES*.

Let us discuss in more detail these components of the photo-emf.

***Photo-emf SCR (barrier photo-emf).*** If we assume that the charge of the SES does not change under illumination of the crystal, then the electric field near the surface remains the same as without illumination, and from (2.5.19) it follows that

$$F(Y_S, \lambda)/ L_D = F(Y^*_S, \lambda)/ L^*_D. \qquad (2.6.13)$$

In this case, the photoelectric power of the SCR is

$$V_S = (\kappa T/q)(Y^*_S - Y_S) = (kT/q) - \upsilon_s. \quad .$$

*For small bending of the energy bands* $F(Y_S, \lambda) \equiv Y_S$ and from (2.6.13) it follows that

$$\upsilon_s \equiv Y'_S - Y_S \equiv -Y_S\,[1-(1+\delta_p)-1/2]. \qquad (2.6.14)$$

In accordance with the qualitative considerations given above, for small initial band bendings, the sign of the photo-*EMF* of the SCR is opposite to the sign of $Y_S$, i.e. at illumination, the bending of the bands decreases. With an increase in the injection level, the photo-*EMF* of the SCR tends to a limiting value equal to the initial bending of the bands.

*In the regime of strong enrichment*, the quantity

$$\upsilon_s \equiv Y^*_S - Y_S \approx ln\,(1+\delta_p). \qquad (2.6.15)$$

The photo- *EMF* of SCR in this regime for $\delta_p < 1$ is very small and does not depend on the initial bending of the bands. The physical meaning of this is that the field in the strong enrichment layer is so large that all nonequilibrium electron-hole pairs are separated in it.

*In the regime of strong inversion*

$$\upsilon_s \equiv -\ln(1+\delta_n) \approx -\ln(1+\lambda^2\delta_p). \qquad (2.6.16)$$

In this case, as in the case of strong enrichment, the photo-*EMF* is independent of $Y_S$, however, it is much larger ($\delta_n >> \delta_p$) for strong inversion at the same illumination level. The reason for this is in the more efficient spatial separation of electron-hole pairs in the regime of strong inversion, since under the inversion layer there is always a wide layer of depletion. The fact that it is possible to achieve a very high level of injection for minority charge carriers seems to be quite realistic to achieve flattening of the bands already at moderate light intensities. In reality, as experiments show, this requires a majority carriers injection level $\delta_p > 10^2$.

Let us pay attention to the interesting feature of the *SCR* photovoltage in the modes of strong enrichment and inversion: its magnitude, regardless of the level of injection, is automatically maintained such that the distance between the edge of the corresponding allowed bands on the surface and the quasi-Fermi level (in particular, $E_c$ and $F_n$ for strong inversion in the p-type material) was constant and equal ($E_c$ - $F$) in equilibrium state, see relations (2.6.15), (2.6.16) and (2.6.12) )

Fig. 2.6.3. Dependence of the Dember EMF (dashed line, right scale) and the resultant surface photo-*EMF* (solid lines, left scale) from the level of injection by the majority carriers for p-silicon (λ = 100). Figures at curves are initial bends $Y_S$.

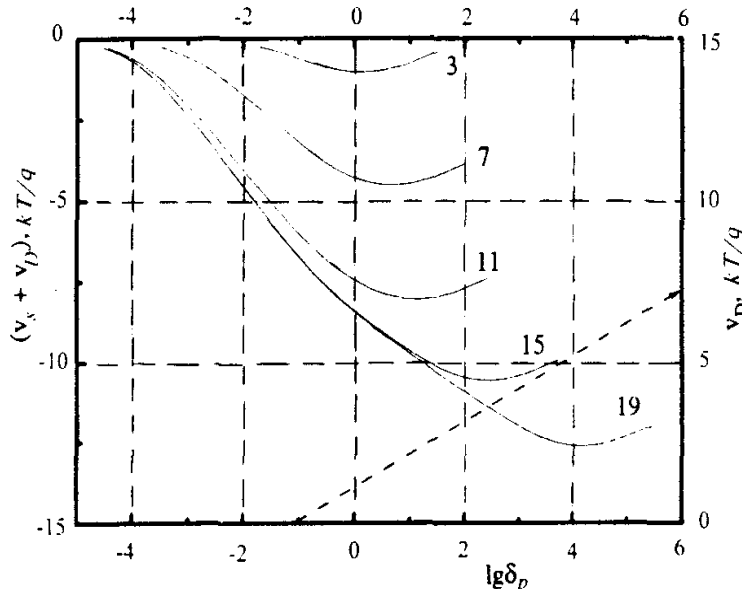



***Photo of Dember***. When a crystal is irradiated with quanta of light with an energy exceeding the width of the forbidden band, the generation of nonequilibrium charge carriers occurs in a thin near-surface layer with tens of nanometers thick, which is usually much less than the thickness of the crystal. Therefore, a diffusion flux of electron-hole pairs is directed from the surface to the bulk. Since the diffusion coefficients for electrons and holes are different, an electric field arises that is directed perpendicular to the surface (along the *z* axis).

According to the theory, the Dember's potential difference between the illuminated and the back surfaces of the crystal is

$$V_D = -\int_{z_1}^{0} \varepsilon_z \, dz = \frac{kT}{q} \frac{\mu_n - \mu_p}{\mu_n + \mu_p} \cdot \ln \frac{\sigma_0^*}{\sigma_1}, \qquad (2.6.17)$$

where $\sigma_0^*$ and $\sigma_i$ are the bulk specific conductivities on the illuminated and back surfaces of the crystal, $z_1$ is the coordinate of the back surface, and

$$\varepsilon_z = \frac{kT}{q} \cdot \left( \frac{\mu_p - \mu_n}{\mu_p + \mu_n} \right) \frac{1}{\sigma} \frac{d\Delta\sigma}{dz}.$$

At $\mu_n > \mu_p$, electrons pass somewhat ahead the holes in the diffusion motion and the potential of the surface increases relative to the bulk value, $V_D > 0$.

If $z_1 > L_{dif}$, then $\sigma_0^* = \sigma_i$ and, consequently,

$$\upsilon_D \equiv \frac{qV_D}{kT} = \left( \frac{b-1}{b+1} \right) \cdot \ln \left[ 1 + \left( \frac{b+1}{1 + b\lambda^{-2}} \right) \delta_p \right]. \qquad (2.6.18)$$

Fig. 2.6.3 shows the dependence of the Dember *EMF* on the level of injection for p-silicon ($\lambda = 10$, $b = 2{,}9$). For $\delta_p < 1$, the value $V_D < kT/q$ and its contribution to the total surface photo-*EMF* in the depletion and strong inversion modes can be neglected. However, at higher injection levels, $V_D$ can be comparable to $V_S$. In particular, for p-type materials, the signs of the Dember *EMF* and the photovoltage of the SCR in the depletion and inversion modes are opposite (in the case of $\mu_n > \mu_p$). Therefore, the value of the resulting surface photo-*EMF* at a certain value of $\delta_p$ passes through a maximum whose position along the ordinate axis is uniquely determined by the initial potential of the surface-see Fig. 2.6.3. One of the methods for determining $Y_S$ is based on this effect.

If necessary, the contribution of the Dember *EMF* to the value of the surface photo-*EMF* can be substantially reduced if measurements are carried out on thin crystals. For $d << L_{dif}$, in relation (2.6.18), $\sigma_1 \equiv \sigma_0^*$ and the Dember potential difference is negligible.

***Photo-EMF of SES recharge.*** If the charge of the SES ($Q_S$) varies with the illumination of the crystal, then, in accordance with the Gauss theorem, the electric field in the vicinity of the surface also changes, and therefore (2.6.13) will no longer be satisfied. Under the condition that the external electric field is constant, the total charge of the semiconductor ($Q_s + Q_{sc}$) remains unchanged. From this, taking into account (2.5.22) and (2.5.23), we obtain the equation

$$Q_S(Y_S) \mp \frac{\varepsilon\varepsilon_0 kT}{q} \frac{F(\lambda, Y_S)}{L_D} = Q_S^*\left(Y_S^*, \delta_p\right) \mp \frac{\varepsilon\varepsilon_0 kT}{q} \frac{F\left(\lambda^*, Y_S^*\right)}{L_D^*}. \qquad (2.6.19)$$

If the surface charge does not change ($Q_S = Q_S^*$), equation (2.6.19) reduces to (2.6.14). In general case, in order to calculate the surface photo-*EMF* with the aim of (2.6.19), it is necessary to have detailed information about the *SES*, which, as a rule, is absent. In principle, this information is contained in the dependencies $\upsilon_S = (Y_S, \delta)$, but its extraction from the experimental data is a complex problem.

We emphasize that the photovoltage of the *SES* can not be separated from the photo-*EMF* of the SCR. As follows from equation (2.6.19), the problem of these contributions in surface photo-*EMF* is self-consistent. In this situation, there is a cardinal difference between the photoelectric

voltage of the *SCR* and the *SES* from the Dember potential.

If the measurement of the photo-*EMF* is used to determine the surface potential of the semiconductor (for example, by saturating of $v_S$ with an increase in the injection level), then one tries to avoid, as far as possible, the complications associated with the recharging of the *SES*. To do this, the crystal is subjected by light pulses of short duration (less than $(10^{-6} \div 10^{-5})$ s) so that the exchange of charges between *SES* and allowed bands does not take place during the time of the pulse. In these conditions, the equation (2.6.14) and all the conclusions derived from it remain valid.

# 3. Thermal Properties of Nanomaterials

In this chapter we describe systems of nonlinear oscillators arranged on low-dimensional lattices, summarize the phenomenology of their transport properties and introduce the features of thermal transport in graphene as an example based on modern theoretical knowledge [1].

## 3.1 Models of heat conduction

Let us consider various models for the thermal conductivity [2].

The simplest microscopic dynamical model for the characterization of heat conduction consists of a chain of N classical point-like particles with mass $m_n$ and position $q_n$, described by the Hamiltonian

$$H = \sum_{n=1}^{N} \left( \frac{p_n^2}{2m_n} + U(q_n) + V(q_{n+1} - q_n) \right). \tag{3.1.1}$$

The potential $V(x)$ accounts for the nearest-neighbour interactions between consecutive particles, while the on-site potential $U(q_n)$ takes into account the possible interaction with an external environment (either a substrate, or some three-dimensional matrix). The corresponding evolution equations are

$$m_n \ddot{q}_n = -U'(q_n) - F(r_n) + F(r_{n-1}),\ n = \overline{1, N}, \tag{3.1.2}$$

where $r_n = q_{n+1} - q_n$, $F(x) = -V'(x)$ and the prime denotes a derivative with respect to the argument. Usually $q_n$ denotes the longitudinal position along the chain, so that

$$L = \sum_{n=1}^{N} (r_n), \tag{3.1.3}$$

represents the total length of the chain. Different kinds of boundary conditions may and will be indeed used in the various cases. For instance, if the particles are confined in a simulation "box" of length *L* with periodic boundary conditions,

$$q_{n+N} = q_n + L. \tag{3.1.4}$$

Alternatively one can adopt a lattice interpretation, in which case, the (discrete) position is $z_n = an$ (where a is lattice spacing), while $q_n$ is a transversal displacement. Thus, the chain length is obviously equal to *Na*.

The Hamiltonian (3.1.1) is generally a constant of motion. In the absence of an on-site potential ($U = 0$), the total momentum is conserved, as well,

$$P = \sum_{n=1}^{N} p_n \equiv \sum_{n=1}^{N} m_n \dot{q}_n. \tag{3.1.5}$$

Since we are interested in heat transport, one can set $P = 0$ (i.e., we assume to work in the center-of-mass reference frame) without loss of generality. As a result, the relevant state variables of microcanonical equilibrium are the specific energy (the energy per particle) $e = H/N$ and the elongation $l = L/N$ (the inverse of the particle density). On a microscopic level, one can introduce three local densities, namely $r_n$, $p_n$ and

$$e_n = \frac{p_n^2}{2m_n} + \frac{1}{2}\big(V(r_n) + V(r_{n-1})\big), \tag{3.1.6}$$

which, in turn, define a set of currents through three (discrete) continuity equations. For instance, the energy current is defined as

$$\dot{e}_n = j_{n-1} - j_n, \tag{3.1.7}$$

$$j_n = \frac{1}{2} a (\dot{q}_{n+1} + \dot{q}_n) F(r_n). \tag{3.1.8}$$

The definition (3.1.8) is related to the general expression, originally derived by Irving and Kirkwood that is valid for every state of matter that, in one dimension, reads

$$j_n = \frac{1}{2}(q_{n+1} - q_n)(\dot{q}_{n+1} + \dot{q}_n) F(r_n) + \dot{q}_n e_n. \tag{3.1.9}$$

In the case of lattice systems, where we assume the limit of small oscillations (compared to the lattice spacing) or in the lattice field interpretation, one can recover formula (3.1.8) setting $q_{n+1} - q_n = a$ in the first term and neglecting the second one. The expression (3.1.9) is useful in the opposite limit of freely colliding particles, where the only relevant interaction is the repulsive part of the potential, that is responsible for elastic collisions. There, the only contribution to the flux

arises from the kinetic term of $e_n$

$$j_n \approx \frac{1}{2} m_n \dot{q}_n^3. \tag{3.1.10}$$

Having set the basic definitions, let us now introduce some specific models. A first relevant example is the harmonic chain, where the potential $V$ is quadratic ($U = 0$). From the point of view of transport properties, we expect this system to behave like a ballistic conductor. The heat flux decomposes into the sum of independent contributions associated to the various eigenmodes. This notwithstanding, this model proves useful, as it allows addressing general questions about the nature of stationary nonequilibrium states. This includes the role of disorder (either in the masses or the spring constants), of boundary conditions, and quantum effects.

In case of anharmonic problem, the most paradigmatic example is the Fermi–Pasta–Ulam (FPU) model

$$V(r_n) = \frac{k_2}{2}(r_n - a)^2 + \frac{k_3}{3}(r_n - a)^3 + \frac{k_4}{4}(r_n - a)^4. \tag{3.1.11}$$

Notice that upon introducing the displacement $u_n = q_n - na$, from the equilibrium position, $r_n$ can be rewritten as $u_{n+1} - u_n + a$, so that the lattice spacing $a$ disappears from the equations.

Another interesting model is the Hard Point Gas (HPG), where the interaction potential is

$$V(y) = \begin{cases} \infty, & y = 0 \\ 0, & y \neq 0 \end{cases}. \tag{3.1.12}$$

The dynamics consist of successive collisions between neighbouring particles

$$v'_n = \frac{m_n - m_{n+1}}{m_n + m_{n+1}} v_n + \frac{2m_{n+1}}{m_n + m_{n+1}} v_{n+1}, \; v'_{n+1} = \frac{2m_n}{m_n + m_{n+1}} v_n - \frac{m_n - m_{n+1}}{m_n + m_{n+1}} v_{n+1}, \tag{3.1.13}$$

where $m_n$ is the mass of the $n$-th particle, $v_n = \dot{q}_n$ and the primed variables denote the values after the collision. For equal masses the model is completely integrable, as the set of initial velocities is conserved during the evolution. In order to avoid this peculiar situation, it is customary to choose alternating values, such as $m_n = m$ for even $n$. This type of dynamical systems are particularly appropriate for numerical computation as they do not require the numerical integration of nonlinear differential equations. In fact, it is sufficient to determine the successive collision times and update the velocities according to (3.1.13). The only errors are those due to machine round-off. Moreover, the simulation can be made very efficient by resorting to fast updating algorithms. In fact, since the collision times depend only on the position and velocities of neighbouring particles, they can be arranged in a heap structure and thereby simulate the dynamics with an event driven algorithm.

Another much studied model involves the Lennard–Jones potential, that in our units reads

$$V(y) = \frac{1}{12}\left(\frac{1}{y^{12}} - \frac{2}{y^6} + 1\right). \tag{3.1.14}$$

For computational purposes, the coupling parameters have been fixed in such a way as to yield the simplest form for the force. With this choice, $V$ has a minimum in $y = 1$ and the resulting dissociation energy is $V_0 = 1/12$. For the sake of convenience, the zero of the potential energy is set in $y = 1$. In one-dimension, the repulsive term ensures that the ordering is preserved (the particles do not cross each other).

In the presence of a substrate potential $U$, the invariance $q_l \to q_l +$ const is broken and the total momentum $P$ is no longer a constant of motion. Accordingly, all branches of the dispersion relation have a gap at zero wavenumber. We refer to them as optical modes. An important subclass is the one in which $V$ is quadratic, which can be regarded as a discretization of the Klein–Gordon field: relevant examples are the Frenkel–Kontorova and "$\phi^4$" models which correspond to $U(y) = 1 - \cos(y)$ and $U(y) = y^2/2 + y^4/4$ respectively.

We will always deal with genuine nonintegrable dynamics. For the FPU model this means working with high enough energies/temperatures to avoid all the difficulties induced by quasi-integrability and the associated slow relaxation to equilibrium. For the diatomic HPG this requires fixing a mass-ratio $r$ not too close to unity.

### 3.2 Theoretical Approaches

Various theoretical approaches to account for the observed phenomenology have been developed and implemented [2]. The methods discussed are

1. fluctuating hydrodynamics approach: here the models are described in terms of the random fields of deviations of the conserved quantities with respect to their stationary values. The role of fluctuations is taken into account by renormalization group or some kind of self-consistent theory;

2. mode-coupling theory: this is closely related to the above, as it amounts to solving (self-consistently) some approximate equations for the correlation functions of the fluctuating random fields;

3. kinetic theory: it is based on the familiar approach to phonon transport by means of the Boltzmann equation;

4. exact solution of specific models: typically in this case the original microscopic Hamiltonian dynamics is replaced by somesuitable stochastic one which can be treated by probabilistic methods.

A sound theoretical basis for the idea that the above described anomalies are generic and universal for all momentum-conserving system was put forward by authors treated the case of a fluctuating *d*-dimensional fluid and applied renormalization group techniques to evaluate the contribution of noisy terms to transport coefficients. The calculation predicts that the thermal conductivity exponent is $\alpha = (2-d)/(2+d)$. From the arguments exposed above, it follows that in 1-D the exponents are

$$\alpha = \delta = \frac{1}{3},\ \beta = \frac{4}{3},\ \zeta = \frac{3}{2}. \tag{3.2.1}$$

According to this approach, any possible additional term in the noisy Navier–Stokes equation yields irrelevant corrections in the renormalization procedure, meaning that the above exponents are model independent, provided the basic conservation laws are respected.

Next we give a flavour of one of the other approaches: the Mode-Coupling Theory (MCT). This type of theories has been traditionally invoked to estimate long-time tails of fluids and to describe the glass transition. In the simplest version, it involves the normalized correlator of the particle displacement, where the discrete wavenumber *k* has been turned to the continuous variable $q$

$$G(q,t) = \frac{\langle U^*(q,t)U(q,0)\rangle}{\langle |U(q)|^2\rangle}. \tag{3.2.2}$$

$G(q,t)$ is akin to the density–density correlator, an observable routinely used in condensed-matter physics. The main idea is to write a set of approximate equations for $G(q,t)$ that must be solved self-consistently. For the problem at hand, the simplest version of the theory amounts to consider the equations

$$\ddot{G}(q,t) + \varepsilon \int_0^t \Gamma(q,t-s)\dot{G}(q,s)\mathrm{d}s + \omega^2(q)G(q,t) = 0, \tag{3.2.3}$$

where the memory kernel $\Gamma(q,t)$ is proportional to $\langle \mathfrak{F}(q,t)\,\mathfrak{F}(q,0)\rangle$, with $\mathfrak{F}(q)$ being the nonlinear part of the fluctuating force between particles. Equation (3.2.3) is derived within the well-known Mori–Zwanzig projection approach. It must be solved with the initial condition0s $G(q,0) = 1$ and $\dot{G}(q,0) = 0$.

The mode-coupling approach basically amounts to replacing the exact memory function Γ with an approximate one, where higher-orders correlators are written in terms of $G(q,t)$. In the generic case, in which $k_3$ is different from zero [see (3.1.11)], the lowest-order mode coupling approximation of the memory kernel turns out to be

$$\Gamma(q,t) = \omega^2(q)\frac{2\pi}{N}\sum_{p+p'-q=0,\pm\pi} G(p,t)G(p',t). \tag{3.2.4}$$

Here $p$ and $p'$ range over the whole Brillouin zone (from $-\pi$ to $\pi$). This yields a closed system of nonlinear integro-differential equations. Both the coupling constant $\varepsilon$ and the frequency $\omega(q)$ are temperature-dependent input parameters, which should be computed independently by

numerical simulations or approximate analytical estimates. For the present purposes it is sufficient to restrict ourselves to considering their bare values, obtained in the harmonic approximation. In the adopted dimensionless units they read $\varepsilon = 3k_3^2 k_B T/2\pi$ and $\omega(q) = 2|\sin(q/2)|$. Of course, the actual renormalized values are needed for a quantitative comparison with specific models.

A more refined theory requires considering the mutual interaction among all the hydrodynamic modes associated with the conservation laws of the system at hand. The resulting calculations are considerably more complicated but they can be worked out.

In case of the interface influence considering the relevant theoretical insight comes from the link with the Kardar-Parisi-Zhang (KPZ) equation. This is a nonlinear stochastic Langevin equation which was originally introduced in the context of surface growth. Let us first consider the fluctuating Burgers equation for the random field $\rho(x,t)$

$$\frac{\partial \rho}{\partial t} = \frac{\lambda}{2}\frac{\partial \rho^2}{\partial x} + D\frac{\partial^2 \rho}{\partial x^2} + \frac{\partial \eta}{\partial x}, \tag{3.2.5}$$

where $\eta(x,t)$ represents a Gaussian white noise. As it is well-known, eq. (3.2.5) can be transformed into the KPZ equation by introducing the function *h* such that $\rho = \partial h/\partial x$

$$\frac{\partial h}{\partial t} = \frac{\lambda}{2}\left(\frac{\partial h}{\partial x}\right)^2 + D\frac{\partial^2 h}{\partial x^2} + \eta. \tag{3.2.6}$$

We argued that the scaling properties of anomalous transport are independent of the microscopic details and correspond to those of the KPZ universality class. One might wonder whether other classes exist and under which conditions they can be observed. A reasonable argument, that can be invoked to delimit the KPZ universality class, is the symmetry of the interaction potential with respect to the equilibrium position. With reference to the MCT, one realizes that the symmetry of the fluctuations implies that the quadratic kernel in (3.2.4) should be replaced by a cubic one, thus yielding different values of the exponents. In the language of KPZ interfaces, whenever the coefficient of the nonlinear term vanishes, the evolution equation reduces to the Edwards-Wilkinson equation that is indeed characterized by different scaling exponents. The argument can be made more precise in the framework of the full hydrodynamic theory. There, different dynamical exponents can arise if the coupling between some modes vanishes.

The FPU model is a natural instance to test this working hypothesis. In fact, systematically larger values of the scaling exponent $\alpha$ have been reported for the FPU-$\beta$ case where the cubic term of the potential is absent. The existence of two universality classes for thermal transport in one-dimensional oscillator systems has been also demonstrated and was further proposed that the criterion for being out of the KPZ class is the condition $\gamma = c_p/c_V = 1$, where $c_p$ and $c_V$ are the specific heat capacities at constant pressure and volume, respectively.

The scenario can be further illustrated by considering a modification of the HPG model, the so-called Hard-Point Chain (HPC), characterized by a square-well potential in the relative distances

$$V(y) = \begin{cases} 0, & 0 < y < a \\ \infty, & y > a \end{cases}. \tag{3.2.7}$$

The infinite barriers at $y = a$ imply an elastic "rebounding" of particles as if they were linked by an inextensible and massless string of fixed length *a*. The string has no effect on the motion, unless it reaches its maximal length, when it exerts a restoring force that tends to rebound the particles one against the other. The potential (3.2.7) introduces the physical distance *a* as a parameter of the model.

As it is well known, the thermodynamics of models like the HPC can be solved exactly and the equation of state is found to be

$$L = N\left(\frac{1}{\beta P} - \frac{a}{\exp(\beta P a)-1}\right) \tag{3.2.8}$$

where $P$ is the pressure of the HPC. Note that, for large values of $a$, the equation of state is the same of an HPC i.e. the one of an ideal gas in 1D. The important point here is that we can choose the parameter $a$ such as $P = 0$. In this particular point, the interaction is symmetric ($^{L}/_{N} = {}^{a}/_{2}$) peculiarity of the HPC model is that energy transfer occurs also at rebounding "collisions" at distance a, this means that besides the contribution defined by (3.1.10) one should include a term $j'_i$ as from (3.1.9). By defining the force between two particles as the momentum difference induced by a collision, $j'_i$ can be written as the kinetic energy variation times the actual distance $a$, i.e. $j'_i = {}^{am_{i(w_i'^2-u_i^2)}}/_{2}$, divided by a suitable time-interval $\Delta t$. In order to get rid of the microscopic fluctuations, it is necessary to consider a sufficiently long $\Delta t$, so as to include a large number of collisions. Since the number of collisions is proportional to the system size, it is only in long systems that fluctuations can be removed without spoiling the slow dynamics of the heat flux. Equilibrium simulations show that for $^{L}/_{N} = {}^{a}/_{2}$ the leading contribution to the heat flux is given by the term $j'_i$ which exhibits a lowfrequency divergence with an exponent $\delta = 0.45$, that is not only definitely larger than 1/3 (the value predicted for the KPZ class), but is also fairly close to the results found for the FPU-$\beta$ model.

In out-of-equilibrium simulations, a compatible exponent $\alpha = 0.4$ has been measured. Those values should be compared with $\alpha = 0.5$, the prediction of mode-coupling theory, thus supporting the conjecture that the case $P = 0$ belongs to a universality class different from KPZ.

Note, that one-dimensional anharmonic chains generically display anomalous transport properties. A prominent exception is the coupled rotors chain described by the equation of motion

$$\dot{q}_n = p_n,\ \dot{p}_n = \sin(q_{n+1} - q_n) - \sin(q_n - q_{n-1}). \tag{3.2.9}$$

The model is sometime referred to as the Hamiltonian version of the XY spin chain. The energy flux is $j_n^e = \langle p_n \sin(q_{n+1} - q_n)\rangle$. As the interaction depends only on the angle differences, angular momentum is conserved and one may expect anomalous transport to occur. Nevertheless, molecular dynamics simulations have convincingly demonstrated normal diffusion.

There are two complementary views to account for this difference. In the general perspective of nonlinear fluctuating hydrodynamics, the chain length defined as $L = \sum_n (q_{n+1} - q_n)$ is not even well defined, because of the phase slips of $\pm 2\pi$, so the corresponding evolution equation breaks down and normal transport is eventually expected. From a dynamical point of view, one can invoke that normal transport sets in due to the spontaneous formation of local excitations, the so called rotobreathers, that behave like scattering centers. Phase slips (jumps over the energy barrier), on their side, may effectively act as localized random kicks, that contribute to scatter the low-frequency modes, thus leading to a finite conductivity. In order to test the validity of this conjecture, one can study the temperature dependence of $k$ for low temperatures $T$, when jumps across barriers become increasingly rare. Numerics indicates that the thermal conductivity behaves $\exp({}^{\eta}/_{T})$ with $\eta \approx 1.2$. The same kind of dependence on $T$ (although with $\eta \approx 2$) is found for the average escape time $\tau$ across the potential barrier: this can be explained by assuming that the phase slips are the results of activation processes. An important extension is the 2D case, i.e. rotors coupled to their neighbors on a square lattice, akin to the celebrated XY-model. As it is well

known, the latter is characterized by the presence of the so called Kosterlitz-Thouless-Berezinskii phase transition at a temperature $T_{KTB}$, between a disordered high-temperature phase and a low-temperature one, where vortices condensate. It is likely that transport properties are qualitatively different in the two phases. Numerical simulations performed on a finite lattice indeed show that they are drastically different in the high temperature and in the low-temperature phases. In particular, thermal conductivity is finite in the former case, while in the latter it does not converge up to lattice sizes of order $10^4$. In the region where vorticity is negligible ($T < 0.5$) the available data suggest a logarithmic divergence with the system size, analogous to the one observed for coupled oscillators. Close to $T_{KTB}$, where a sizable density of bounded vortex pairs are thermally excited, numerical data still suggests a divergence, but the precise law has not be reliably estimated.

Heat conduction in 2D models of anharmonic oscillators coupled through momentum-conserving interactions is expected to exhibit different properties from those of 1D systems. In fact, extension of the arguments predicts a logarithmic divergence of $k$ with the system size $N$ at variance with the power-law predicted for the 1D case. Consideration of this case of great interest for almost-2D materials, like graphene (see 3.3).

There are several numerical evidences in this case in favor of such logarithmic divergence. In [68], a square lattice of oscillators interacting through the FPU-$\beta$ or the Lennard-Jones potentials, was investigated by means of both equilibrium and nonequilibrium simulations. The models are formulated in terms of two-dimensional vector displacements $u_{ij}$ and velocities $\dot{u}_{ij}$, defined on a square lattice containing $N_x \times N_y$ atoms of equal masses $m$ and nearest-neighbor interactions. Periodic and fixed boundary conditions have been adopted in the direction perpendicular ($y$) and parallel ($x$) to the thermal gradient, respectively

The simulations reveal several hallmarks of anomalous behavior: temperature profiles display deviations from the linear shape predicted by Fourier law and the size dependence of the thermal conductivity is well-fitted by a logarithmic law

$$k = A + B\log N_x \tag{3.2.10}$$

with $A$ and $B$ being two unknown constants. A consistent indication comes from the evaluation of the Green-Kubo integrand in the microcanonical ensemble. Indeed, the energy-current autocorrelation is compatible with a decay $1/t$ large times.

Despite these first indications, the numerics turns out to be very difficult, which is not surprising in view of the very weak form of the anomaly, peculiar of the 2D case. As a matter of fact very robust finite-size effects are observed in the calculations for other lattices, which well exemplify the difficulties in observing the true asymptotic behavior with affordable computational resources. Another interesting issue concerns dimensional crossover, namely how the divergence law of the thermal conductivity will change from the 2D class to 1D class as $N_x/N_y$ decreases. This issue has been studied for the two-dimensional FPU lattice.

The harmonic crystal behaves as an ideal conductor, because its dynamics can be decomposed into the superposition of independent channels. This peculiarity can be generalized to the broader context of integrable nonlinear systems. They are mostly one-dimensional models characterized by the presence of mathematical solitons, whose stability is determined by the interplay of dispersion and nonlinearity. This interplay is expressed by the existence of a macroscopic number of conservation laws, constraining the dynamical evolution. Intuitively, the

existence of freely travelling solitons is expected to yield ballistic transport, i.e. an infinite conductivity. From the point of view of the Green-Kubo formula, this ideal conducting behavior is reflected by the existence of a nonzero flux autocorrelation at arbitrarily large times. This, in turn, implies that the finite-size conductivity diverges linearly with the system size.

In this context the most relevant example is the equal-masses Toda chain with periodic boundary conditions, defined, in reduced units, by the Hamiltonian

$$H = \sum_{n=1}^{N} \left( \frac{p_n^2}{2} + \exp(-r_n) \right) \tag{3.2.11}$$

where $r_n = q_{n+1} - q_n$ is the relative position of neighboring particles. The model is completely integrable, since it admits $N$ independent constants of the motion. So, the Mazur-type of inequalities have been recently used as a theoretical basis for the study of thermoelectric coefficients.

Here it was performed the discussion to models where just one quantity, the energy, is exchanged with external reservoirs and transported across the system. In general, however, the dynamics can be characterized by more than one conserved quantity. In such cases, it is natural to expect the emergence of coupled transport phenomena, in the sense of ordinary linear irreversible thermodynamics. Now two models will be discussed: a chain of coupled rotors and the discrete nonlinear Schrödinger equation, where the second conserved quantity is the momentum and the norm (number of particles), respectively.

The evolution of equation defined in (3.2.9) must be augmented to include the exchange of momentum with the external reservoirs,

$$\dot{p}_n = \sin(q_{n+1} - q_n) - \sin(q_n - q_{n-1}) + \delta_{1n}\left(\gamma(F_+ - p_1) + \sqrt{2\gamma T_+}\eta_+\right) + \\ + \delta_{1N}\left(\gamma(F_- - p_N) + \sqrt{2\gamma T_-}\eta_-\right), \tag{3.2.12}$$

where $F_\pm$ and $T_\pm$ denote the torque applied to the chain boundary and the corresponding temperature, respectively; $\gamma$ is the coupling strength with the external baths and $\eta_\pm$ is a Gaussian white noise with unit variance.

Angular momentum is conserved and one can, in fact, define the corresponding flux as

$$j_n^p = \sin(q_{n+1} - q_n). \tag{3.2.13}$$

A chain of rotors is perhaps the simplest model where one can exert a gradient of forces that couples to heat transport, giving rise to nontrivial phenomena, even though the transport itself is normal. For $F_+ = F_-$ all the oscillators rotate with the same frequency $\omega = F$, no matter which force is applied: no momentum flux is generated. In fact, what matters is the difference between the forces applied at the two extrema of the chain. Therefore, from now on we consider the case of zero-average force, i.e. $F_+ = -F_-$. In the presence of such a gradient of forces, the oscillators may rotate with different frequencies and, as a result, a coupling between angular momentum and energy transport may set in. In principle, one could discuss the same setup for general chains of kinetic oscillators, as (linear) momentum is conserved in that context too. However, nothing interesting is expected to arise. For a binding potential, like in the FPU model, the presence of an external force is akin to the introduction of a homogeneous, either positive or negative, pressure all along the chain. In fact, the pressure $P$ is equal to the equilibrium average of the momentum

flux, $P = \langle j^p \rangle$. On the other hand, if the potential is not binding (e.g., the Lennard-Jones chain (3.1.14)) and the applied force is equivalent to a negative pressure, the system would break apart.

In the presence of two fluxes, the linear response theory implies that they must satisfy the equations (angular brackets denote an ensemble, or equivalently, a time average, assuming ergodicity)

$$\langle j^p \rangle = -L_{pp} \frac{\mathrm{d}(\beta\mu)}{\mathrm{d}y} + L_{pe} \frac{\mathrm{d}\beta}{\mathrm{d}y}$$

$$\langle j^e \rangle = -L_{ep} \frac{\mathrm{d}(\beta\mu)}{\mathrm{d}y} + L_{ee} \frac{\mathrm{d}\beta}{\mathrm{d}y}. \tag{3.2.14}$$

where $y = n/N$ is the inverse temperature $1/T$ in units of the Boltzmann constant and $\mu$ is the chemical potential, which, in the case of the coupled rotors, coincides with the average angular frequency $\omega_n = \langle p_n \rangle$. $\mathbf{L}$ is the symmetric, positive definite, $2 \times 2$ Onsager matrix. If $L_{ep} = 0$, the two transport processes are uncoupled.

In the case of the rotor chain, it is important to realize that a correct definition of the kinetic temperature requires subtracting the coherent contribution due to the nonzero angular velocity

$$T_n = \langle (p_n - \omega_n)^2 \rangle. \tag{3.2.15}$$

The effect of coupling between energy and momentum transport could be understood by considering a setup where the two thermal baths operate at the same temperature *T*. Because of the flux of momentum, the temperature profile deviates from the value imposed at the boundaries.

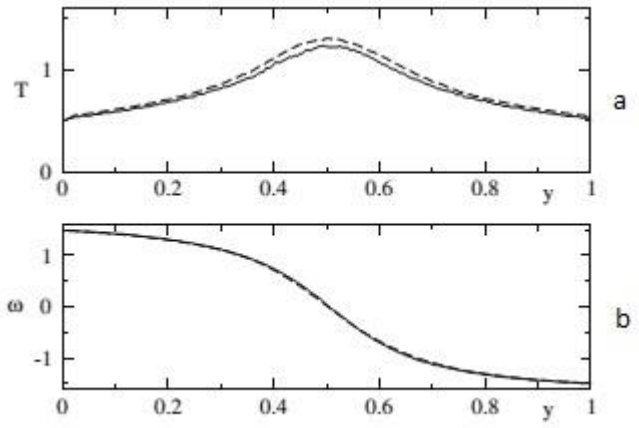



Figure 3.2.1 Stationary profile of the temperature (a) and of the average frequency (b) for $T = 0.5$, $F = 1.5$, $\gamma = 1$. The dashed and solid curves correspond to $N = 100$ and $N = 200$ respectively [2].

In fig. 3.2.1 the results for $T = 0.5$ and $F = 1.5$, and two different system sizes are shown. Notably, the temperature profile displays a peak in the central region, where it reaches a value around 1.2; the average frequency varies nonuniformly across the sample with a steep region in correspondence of the central hot spot. At the same time, the energy flux $j^e$ is zero, so that the anomalous behavior of the temperature profile is entirely due to the coupling with the nonzero momentum flux.

This behavior can be traced back to the existence of a (zero-temperature) boundary-induced transition. In fact, for $T = 0$, there exists a critical torsion $F_c = {}^1\!/\!_\gamma$, such that for $F < F_c$ the ground state is a twisted fully-synchronized state, whereby each element is at rest and is

characterized by a constant phase gradient. Here $T_n = 0$ throughout the whole lattice. For $F > F_c$ the fully synchronized state turns into a chaotic asynchronous dynamics with $\omega_1 = F = -\omega_N$. Remarkably, even though both heat baths operate at zero temperature and the equations are deterministic and dissipative, the temperature in the middle raises to a finite value even in the thermodynamic limit.

The phenomenon can be interpreted as the onset of an interface (the hot region) separating two different phases: the oscillators rotating with a frequency *F* (on the left) from those rotating with a frequency – *F* (on the right). The phenomenon is all the way more interesting in view of the anomalous scaling of the interface width with the system size (it grows as $N^{1/2}$ and its robustness (it is independent of the value of the torsion *F*, provided it is larger than the critical value $F_c$). Accordingly, the interface is neither characterized by a finite width nor it is extensive. A more careful inspection reveals that the $N^{1/2}$ width is due to a spatial Brownian-like behavior of an instantaneously much thinner interface. Nevertheless, even the instantaneous interface extends over a diverging number of sites, of order $N^{1/5}$, thus leaving the anomaly fully in place. Such a state can neither bepredicted within a linear-response type of theory, nor traced back to some underlying equilibrium transition. Even more remarkably, it constitutes an example of a highly inhomogeneous, unusual chaotic regime. Indeed, while the fractal dimension is extensive (i.e. proportional to the number of oscillators) the Kolmogorov-Sinai (KS) entropy is not: it increases only as $N^{1/2}$. The KS entropy measures the diversity of the "ground state" non-equilibrium configurations that are compatible with the given thermal baths. Its lower-than-linear increase with *N* implies that we are not in the presence of a macroscopic degeneracy, as in spin glasses.

The anomaly of the regime is finally reinforced by the scaling behavior of the momentum flux, which scales as $N^{1/5}$. A theoretical explanation of this behavior is still missing. All of these anomalies disappear as soon as the temperature at the boundaries is selected to be strictly larger than zero. In particular, the width of the hot spot suddenly becomes extensive and the scaling of the momentum is normal, when $j^p \sim 1/N$. The nonmonotonous behavior of the temperature is nevertheless a nontrivial consequence of the coupling between heat and momentum transport.

The non-equilibrium transition is not a peculiarity of the rotor model. A similar scenario can be observed also in the Discrete Nonlinear Schrödinger (DNLS) equation, a model with important applications in many domains of physics. In one dimension, the DNLS Hamiltonian is

$$H = \frac{1}{4}\sum_{n=1}^{N}(p_n^2 + q_n^2)^2 + \sum_{n=1}^{N-1}(p_n p_{n+1} + q_n q_{n+1}), \tag{3.2.16}$$

where the sum runs over the *N* sites of the chain. The sign of the quartic term is positive, while the sign of the hopping term is irrelevant, due to the symmetry associated with the canonical (gauge) transformation $z_n \to z_n \exp(i\pi n)$, where $z_n$ denotes the amplitude of the wave function. The equations of motion are

$$i\dot{z}_n = -z_{n+1} - z_{n-1} - 2|z_n|^2 z_n \tag{3.2.17}$$

with $n = \overline{1, N}$ and fixed boundary conditions ($z_0 = z_{n+1} = 0$). The model has two conserved quantities, the energy and the total norm (or total number of particles)

$$A = \sum_{n=1}^{N}(p_n^2 + q_n^2)^2 = \sum_{n=1}^{N}|z_n|^2, \tag{3.2.18}$$

so that it is a natural candidate for the study of coupled transport.

Since the Hamiltonian is not the sum of a kinetic and potential energy, the thermal baths cannot be described by standard Langevin equations. The evolution equation for the first oscillator,

in contact with a thermal bath at temperature $T_+$ and with a chemical potential $\mu_+$ (a similar equation holds for the last particle at site $N$) is

$$\dot{p}_1 = -(p_1^2 + q_1^2)q_1 - q_2 - \gamma\big((p_1^2 + q_1^2)p_1 + p_2 - \mu_+ p_1\big) + \sqrt{2\gamma T} + \xi_1^{/}$$
$$\dot{q}_1 = (p_1^2 + q_1^2)p_1 + p_2 - \gamma\big((p_1^2 + q_1^2)q_1 + q_2 - \mu_+ p_1\big) + \sqrt{2\gamma T} + \xi_1^{//} \qquad (3.2.19)$$

where measures the coupling strength with the thermal bath, while $\xi_1^{/}$ and $\xi_1^{//}$ define two independent white noises with unit variance. It can be easily seen that the deterministic components of the thermostat, are gradient terms. As a result, in the absence of thermal noise, they would drive the system towards a state characterized by a minimal ($H - \mu A$). Notice the nonlinear structure of the dissipation terms in (3.2.19).

An additional problem of the DNLS model is the determination of the temperature, as one cannot rely on the usual kinetic definition (this is again a consequence of the nonseparable Hamiltonian). An operative definition can be, however, given by adopting the microcanonical approach, i.e. by invoking the thermodynamic relationships,

$$\frac{1}{T} = \frac{\partial S}{\partial H},$$
$$\frac{\mu}{T} = \frac{\partial S}{\partial A}, \qquad (3.2.20)$$

where $S$ is the entropy. The partial derivative $\partial S/\partial C_i$, where $i = 1, 2$ with $C_1 = H$ and $C_2 = A$ can be computed by exploiting the fact that $C_i$ is a conserved quantity

$$\frac{\partial S}{\partial C_i} = \langle \frac{W\|\xi\|}{\nabla C_i \cdot \xi} \nabla \cdot \left(\frac{\xi}{\|\xi\| W}\right) \rangle \qquad (3.2.21)$$

where < > stands for the microcanonical average,

$$\xi = \frac{\nabla C_1}{\|\nabla C_1\|} - \frac{(\nabla C_1 \cdot \nabla C_2)\nabla C_2}{\|\nabla C_1\|\|\nabla C_2\|^2},$$
$$W^2 = \sum_{m,n=1}^{2N} \left(\frac{\partial C_1}{\partial x_m}\frac{\partial C_2}{\partial x_m} - \frac{\partial C_1}{\partial x_n}\frac{\partial C_2}{\partial x_m}\right)^2, \qquad (3.2.22)$$

and $x_{2n} = q_n, \quad x_{2n+1} = p_n$. The resulting definitions of $T$ and $\mu$ are nonlocal: numerical simulations, however, show that they give meaningful results even when they are implemented for relatively short subchains. As for the fluxes, they are naturally defined from the continuity equations for energy and norm

$$j_n^e = \dot{q}_n q_{n-1} + \dot{p}_n q_{n-1},$$
$$j_n^p = q_n p_{n-1} - p_n q_{n-1}, \qquad (3.2.23)$$

Notice that for the sake of simplicity we still use the same notations as in the previous setup although here $j_n^p$ denotes the flux of norm/mass rather than momentum. For $T_+ = T_- = 0$, as in the XY model, the control parameter, i.e. the driving force, is given by $\delta\mu = |\mu_+ - \mu_-|/2$. When $\delta\mu$ is larger than a critical value (that here depends on $A$), a bumpy temperature profile spontaneously emerges.

In coupled transport, each conservation law implies the presence of a corresponding thermodynamic variable. In the case of the DNLS equation, there are two of them: the temperature $T$ (equivalently $\beta$) and the chemical potential $\mu$. If the extrema of a given system are attached to

two different points in the $(\mu, T)$-space, a new question arises with respect to the transport of just one variable: the selection of the path in the phase plane. This problem can be solved with the help of the linear transport equations (3.2.14), which can be rewritten as

$$\frac{\mathrm{d}\beta}{\mathrm{d}\mu} = \frac{\langle j^e\rangle \beta L_{pp} - \langle j^p\rangle \beta L_{ep}}{\langle j^e\rangle (L_{pe} - \mu L_{pp}) - \langle j^p\rangle (L_{ee} - \mu L_{ep})}. \tag{3.2.24}$$

The above first order differential equation can be solved once the Onsager matrix is known across the thermodynamics phase-diagram and the ratio of the two fluxes is given. This determines unambiguously the resulting temperature and chemical potential profiles. In the absence of a mutual coupling between the two transport processes (zero off-diagonal elements of the Onsager matrix) such curves would be vertical and horizontal lines in the latter representation. It is remarkable that the solid lines, which correspond to $j^e = 0$, are almost vertical for large $\mu$. This means that in spite of a large temperature difference, the energy flux is very small. This is an indirect but strong evidence that the nondiagonal terms are far from negligible. The condition of a vanishing particle flux $j^p = 0$ defines the Seebeck coefficient $s = -{\mathrm{d}\mu}/{\mathrm{d}T}$.

### 3.3 Thermal transport in graphene

The discovery of unusual heat conduction properties of graphene has led to a surge of theoretical and experimental studies of phonon transport in two-dimensional material systems. The rapidly developing graphene thermal field spans from theoretical physics to practical engineering applications. In this chapter a special attention is given to analysis of the recent theoretical results for the phonon thermal conductivity of graphene, the effects of the acoustic phonon transport in this material [3].

The main experimental technique for investigation of the acoustic phonon transport in a given material system is the measurement of its lattice thermal conductivity. The thermal conductivity is introduced through Fourier's law:

$$\vec{Q} = -k\nabla T, \tag{3.3.1}$$

where $\vec{Q}$ is the heat flux, $\nabla T$ is the temperature gradient and $k = \left(k_{\alpha\beta}\right)$ is the thermal conductivity tensor. In the isotropic medium, thermal conductivity does not depend on the direction of the heat flow and *k* is treated as a constant. The latter is valid for the small temperature variations only. In a wide temperature range, thermal conductivity is a function of temperature ($k = k(T)$). In general, in solid materials heat is carried by phonons and electrons so that $k = k_p + k_e$, where $k_p$ and $k_e$ are the phonon and electron contributions, respectively. In metals or degenerately-doped semiconductors $k_e$ is dominant due to the large density of free carriers. The value of $k_e$ can be determined from the measurement of the electrical conductivity $\sigma$ via the Wiedemann-Franz law:

$$\frac{k_e}{\sigma T} = \frac{\pi^2 k_B^2}{3e^2}, \tag{3.3.2}$$

where $k_B$ is the Boltzmann's constant and e is the charge of an electron. Phonons are usually the main heat carriers in carbon materials. Even in graphite, which has metal-like properties, the heat conduction is dominated by acoustic phonons. This fact is explained by the strong covalent $sp^2$

bonding, resulting in high in-plane phonon group velocities and low crystal lattice anharmonicity for in-plane vibrations.

The phonon thermal conductivity can be written as

$$k_p = \sum_j \int C_j(\omega) v_{xj}(\omega) v_{xj}(\omega) \tau_j(\omega) \mathrm{d}\omega, \tag{3.3.3}$$

where summation is performed over the phonon polarization branches $j$, which include two transverse acoustic branches and one longitudinal acoustic branch, $v_{xj}$ is the projection of the phonon group velocity $\vec{v}_j = {\mathrm{d}\omega_j}/{\mathrm{d}\vec{q}}$ on the x-axis for the $j$-th branch, which, in many solids, can be approximated by the sound velocity, $\tau_j$ is the phonon relaxation time.

$$C_j(\omega) = \frac{\hbar\omega_j \partial N_0(\hbar\omega_j/k_B T)}{\partial T} \tag{3.3.4}$$

is the contribution to heat capacity from the $j$-th branch, where

$$N_0(\hbar\omega_j/k_B T) = \frac{1}{(\exp(\hbar\omega_j/k_B T)-1)} \tag{3.3.5}$$

is the Bose-Einstein phonon equilibrium distribution function. The phonon mean-free path (MFP) $L$ is related to the relaxation time through the expression $L = \tau v$. In the relaxation-time approximation (RTA), various scattering mechanisms, which limit the MFP, usually considered as additive, i.e. $\tau_j^{-1} = \sum_i \tau_{ij}^{-1}$, where $i$ denotes scattering mechanisms. In typical solids, acoustic phonons, which carry the bulk of heat, are scattered by other phonons, lattice defects, impurities, conduction electrons, and interfaces.

In ideal crystals, i.e. crystals without defects or rough boundaries, $L$ is limited by the phonon-phonon scattering due to the crystal lattice anharmonicity. In this case, thermal conductivity is referred to as intrinsic. The anharmonic phonon interactions, which lead to the finite thermal conductivity in three dimensions, can be described by the Umklapp processes. The Umklapp scattering rates depend on the Gruneisen parameter $\gamma$, which determines the degree of the lattice anharmonicity. Thermal conductivity is extrinsic when it is mostly limited by the extrinsic effects such phonon-boundary or phonon-defect scattering.

In nanostructures, the phonon energy spectra are quantized due to the spatial confinement of the acoustic phonons. The quantization of the phonon energy spectra, typically, leads to decreasing phonon group velocity. The modification of the phonon energies, group velocities and density of states, together with phonon scattering from boundaries affect the thermal conductivity of nanostructures. In most of cases, the spatial confinement of acoustic phonons results in a reduction of the phonon thermal conductivity. However, in some cases, the thermal conductivity of nanostructures embedded within the acoustically hard barrier layers can be increased via spatial confinement of acoustic phonons.

The phonon boundary scattering can be evaluated as

$$\frac{1}{\tau_{Bj}} = \frac{v_{xj}}{D}\frac{1-p}{1+p}, \tag{3.3.6}$$

where $D$ is the nanostructure or grain size and $p$ is the specularity parameter defined as a probability of specular scattering at the boundary. The momentum-conserving specular scattering ($p = 1$) does not add to thermal resistance. Only diffuse phonon scattering from rough interfaces ($p \to 0$), which changes the phonon momentum, limits the phonon MFP. The commonly used expression for the phonon specularity is given by

$$p(\lambda) = \exp\left(\frac{-16\pi^2\eta^2}{\lambda^2}\right), \quad (3.3.7)$$

where $\eta$ is the root mean square deviation of the height of the surface from the reference plane and $\lambda$ is the phonon wavelength.

When the phonon-boundary scattering is dominant, the thermal conductivity scales with the nanostructure or grain size $D$ as $k_p \sim C_p v L = C_p v^2 \tau_B = C_p v D$. In nanostructures with $D \ll L$, the thermal conductivity dependence on the physical size of the structure becomes more complicated due to the strong quantization of the phonon energy spectra. The specific heat $C_p$ depends on the phonon density of states, which leads to different $C_p(T)$ dependences in threedimensional, two-dimensional and one-dimensional systems, and reflected in $k(T)$ dependence at low $T$. In bulk at low $T$ $k(T) \sim T^3$, while it is $k(T) \sim T^2$ in 2D systems.

The thermal conductivity $k$ defines how well a given material conducts heat. The thermal diffusivity $\alpha$ defines how fast the material conducts heat. It is given by the expression

$$\alpha = \frac{k}{C_p \rho}, \quad (3.3.8)$$

where $\rho$ is the mass density. Many experimental techniques directly measure thermal diffusivity rather than thermal conductivity.

We now address in more detail some specifics of the acoustic phonon transport in 2D systems. Investigation of the heat conduction in graphene and carbon nanotubes (CNT) raised the issue of ambiguity in the definition of the intrinsic thermal conductivity for 2D and 1D crystal lattices. It was theoretically shown that the intrinsic thermal conductivity limited by the crystal anharmonicity has a finite value in 3D bulk crystals. However, many theoretical models predict that the intrinsic thermal conductivity reveals a logarithmic divergence in strictly 2D systems ($k \sim \ln N$) and the power-law divergence in 1D systems ($k \sim N^\alpha$) with the number of atoms N ($0 < \alpha < 1$). The logarithmic divergence can be removed by introduction of the extrinsic scattering mechanisms such as scattering from defects or coupling to the substrate. Alternatively, one can define the intrinsic thermal conductivity of a 2D crystal for a given size of the crystal.

Graphene is not an ideal 2D crystal, considered in most of the theoretical works, since graphene atoms vibrate in three directions. Nevertheless, the intrinsic graphene thermal conductivity strongly depends on the graphene sheet size due to weak scattering of the low-energy phonons by other phonons in the system. Therefore, the phonon boundary scattering is an important mechanism for phonon relaxation in graphene. Different studies also suggested that an accurate accounting of the higher-order anharmonic processes, i.e. above three-phonon Umklapp scattering, and inclusion of the normal phonon processes into consideration allow one to limit the low-energy phonon MFP. The normal phonon processes do not contribute directly to thermal resistance but affect the phonon mode distribution. However, even these studies found that the graphene sample has to be very large ($> 10$ μm) to obtain the size-independent thermal conductivity.

In Boltzmann transport equation (BTE) approach within relaxation time approximation the thermal conductivity in quasi-2D system are given by:

$$k = \frac{1}{4\pi k_B T^2 h} \sum_s \int_0^{q_{\max}} \left( \left( \hbar\omega_s(q) \frac{\mathrm{d}\omega_s(q)}{\mathrm{d}q} \right)^2 \tau_{\mathrm{tot}}(s,q) \frac{\exp(\hbar\omega_s(q)/k_B T)}{(\exp(\hbar\omega_s(q)/k_B T)-1)^2} q \right) \mathrm{d}q. \quad (3.3.9)$$

Here $\hbar\omega_s(q)$ is the phonon energy, $h$ is the graphene layer thickness, $\tau_{\mathrm{tot}}$ is the total phonon relaxation time, $q$ is the phonon wavenumber, $T$ is the temperature and $k_B$ is the Boltzmann constant.

The specific phonon transport in the quasi-2D system such as graphene can be illustrated with a simple expression for Umklapp-limited thermal conductivity derived in:

$$k_U = \frac{M}{4\pi T h}\sum_{s=\mathrm{TA,LA}} \frac{\omega_{s\max}\bar{v}_s^2}{\gamma_s^2} F(\omega_{s\min}, \omega_{s\max}), \tag{3.3.10}$$

where

$$F(\omega_{s\min}, \omega_{s\max}) = \int_{\hbar\omega_{s\min}/k_B T}^{\hbar\omega_{s\max}/k_B T} \xi \frac{\exp(\xi)}{(\exp(\xi)-1)^2}\mathrm{d}\xi =$$

$$= \left(\ln(\exp(\xi)-1) + \frac{\xi}{1-\exp(\xi)} - \xi\right)\Big|_{\hbar\omega_{s\min}/k_B T}^{\hbar\omega_{s\max}/k_B T}. \tag{3.3.11}$$

In the (3.3.11) $\xi = \hbar\omega/k_B T$ and the upper cut-off frequencies $\omega_{s\max}$ are defined from the actual phonon dispersion in graphene: $\omega_{\mathrm{LAmax}} = 2\pi f_{\mathrm{LAmax}}$, $\omega_{\mathrm{TAmax}} = 2\pi f_{\mathrm{TAmax}}$. The integrand in (3.3.11) can be further simplified near room temperature (RT) when $\hbar\omega_{s\max} > k_B T$ and it can be expressed as

$$F(\omega_{s\min}) \approx -\ln(|\exp(\hbar\omega_{s\min}/k_B T) - 1|) + \frac{\hbar\omega_{s\min}}{k_B T}\frac{\exp(\hbar\omega_{s\min}/k_B T)}{\exp(\hbar\omega_{s\min}/k_B T)-1}. \tag{3.3.12}$$

In (3.3.9) – (3.3.12) the contribution of out-of plane acoustic mode (ZA) phonons to thermal transport has been neglected because of their low group velocity and large Gruneisen parameter $\gamma_{\mathrm{ZA}}$.

There is a clear difference between the heat transport in basal planes of bulk graphite and in single layer graphene. In the former, the heat transport is approximately two-dimensional only up to some lower-bound cut-off frequency $\omega_{\min}$. Below $\omega_{\min}$ there appears to be strong coupling with the cross-plane phonon modes and heat starts to propagate in all directions, which reduces the contributions of these low-energy modes to heat transport along basal planes to negligible values. In bulk graphite, there is a physically reasonable reference point for the on-set of the cross-plane coupling, which is the ZO$'$ phonon branch near 4 THz observed in the spectrum of bulk graphite. The presence of the ZO$'$ branch and corresponding $\omega_{\min} = \omega_{\mathrm{ZO}'}$ under $q = 0$ allows to avoid the logarithmic divergence in the Umklapp-limited thermal conductivity integral (see (3.3.9) – (3.3.12)) and calculate it without considering other scattering mechanisms.

The physics of heat conduction is principally different in graphene where the phonon transport is 2D all the way to zero phonon frequency. Therefore, the lower-bound cut-off frequencies $\omega_{s\min}$ for each *s* are determined from the condition that the phonon MFP cannot exceed the physical size *l* of the flake, i.e.

$$\omega_{s\min} = \frac{\bar{v}_s}{\gamma_s}\sqrt{\frac{M\bar{v}_s\omega_{s\max}}{k_B T l}}. \tag{3.3.13}$$

It should be noted, that using size-independent graphite $\omega_{\min}$ for single-layer graphene (SLG) or few-layer graphene (FLG) is without scientific merit and leads to an erroneous calculation of thermal conductivity.

Equations (3.3.9) – (3.3.12)constitute a simple analytical model for the calculation of the thermal conductivity of the graphene layer, which retains such important features of graphene phonon spectra as different $\bar{v}_s$ and $\gamma_s$ for in-plane longitudinal (LA) and in-plane transverse (TA) branches. The model also reflects the two-dimensional nature of heat transport in graphene all the way down to zero phonon frequency.

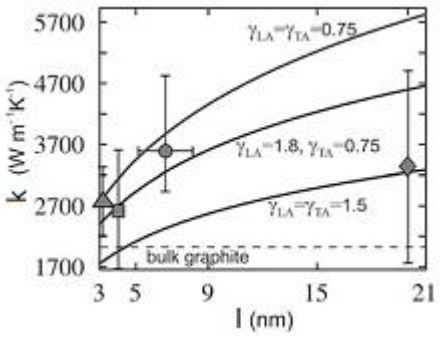



Figure 3.3.1 Calculated room temperature thermal conductivity of graphene as a function of the lateral size for several values of the Gruneisen parameter.

In fig. 3.3.1, the dependence of thermal conductivity of graphene on the dimension of the flake L is presented. The data is presented for the averaged values of the Gruneisen parameters $\gamma_{\mathrm{LA}} = 1.8$ and $\gamma_{\mathrm{TA}} = 0.75$ obtained from ab initio calculations, as well as for several other close sets of $\gamma_{\mathrm{LA,TA}}$ to illustrate the sensitivity of the result to the Gruneisen parameters. For small graphene flakes, the *k* dependence on *l* is rather strong. It weakens for flakes with $l \geq 10$ μm. The calculated values are in good agreement with experimental data for suspended exfoliated and chemical vapor deposition grown (CVD) graphene. The horizontal dashed line indicates the experimental thermal conductivity for bulk graphite, which is exceeded by graphene's thermal conductivity at smaller *l*. Thermal conductivity, presented in fig. 3.3.1, is an intrinsic quantity limited by the three-phonon Umklapp scattering only. But it is determined for a specific graphene flake size since *l* defines the lower-bound (long-wavelength) cut-off frequency in Umklapp scattering through (3.3.13). In experiments, thermal conductivity is also limited by defect scattering. When the size of the flake becomes very large with many polycrystalline grains, the scattering on their boundaries will also lead to phonon relaxation. The latter can be included in this model through adjustment of *l*. The extrinsic phonon scattering mechanisms or high-order phonon–phonon scatterings prevent indefinite growth of thermal conductivity of graphene with *l*.

The simple model described above is based on the Klemens-like expressions for the relaxation time. Therefore, it does not take into account all peculiarities of the 2D three-phonon Umklapp processes in SLG or FLG, which are important for the accurate description of thermal transport. There are two types of the three-phonon Umklapp scattering processes. The first type is the scattering when a phonon with the wave vector $\vec{q}(\omega)$ absorbs another phonon from the heat flux with the wave vector $\vec{q}^{\,/}(\omega^{/})$, i.e. the phonon leaves the state $\vec{q}$. For this type of scattering processes the momentum and energy conservation laws are written as:

$$\vec{q}(\omega) + \vec{q}^{\,/}(\omega^{/}) = \vec{b}_i + \vec{q}^{\,//}(\omega^{//}), \qquad i = \overline{1,3}$$
$$\omega + \omega^{/} = \omega^{//}. \tag{3.3.14}$$

The processes of the second type are those when the phonons $\vec{q}(\omega)$ of the heat flux decay into two phonons with the wave vectors $\vec{q}^{\,/}(\omega^{/})$ and $\vec{q}^{\,//}(\omega^{//})$, i.e. leaves the state $\vec{q}(\omega)$, or, alternatively, two phonons $\vec{q}^{\,/}(\omega^{/})$ and $\vec{q}^{\,//}(\omega^{//})$ merge together, forming a phonon with the wave vector $\vec{q}(\omega)$, which correspond to the phonon coming to the state $\vec{q}(\omega)$. The conservation laws for this type are given by:

$$\vec{q}(\omega) + \vec{b}_i = \vec{q}^{\,/}(\omega^{/}) + \vec{q}^{\,//}(\omega^{//}), \qquad i = \overline{4,6}$$
$$\omega = \omega^{/} + \omega^{//}. \tag{3.3.15}$$

In (3.3.14), (3.3.15) $\vec{b}_i = \overline{\Gamma\Gamma_i}, i = \overline{1,6}$ is one of the vectors of the reciprocal lattice. For each phonon mode $(q_i, s)$, were found all pairs of the phonon modes $\left(\vec{q}^{\,/}, s^{/}\right)$ and $\left(\vec{q}^{\,//}, s^{//}\right)$ such

that the conditions of (3.3.14) and (3.3.15) are met. As a result, in $\left(\vec{q}^{\,/}\right)$-space were constructed the phase diagrams for all allowed three-phonon transitions. Using the long-wave approximation for a matrix element of the three-phonon interaction for the Umklapp scattering rates it could be obtained:

$$\frac{1}{\tau_U^{(I),(II)}(s,\vec{q})} = \frac{\hbar\gamma_s^2(\vec{q})}{3\pi\rho v_s^2(\vec{q})} \times$$

$$\times \sum_{s^{/}s^{//},\vec{b}_i} \iint \omega_s(\vec{q})\omega_{s^{/}}^{/}(\vec{q}^{\,/})\omega_{s^{//}}^{//}(\vec{q}^{\,//}) \left(N_0\left(\omega_{s^{/}}^{/}(\vec{q}^{\,/})\right) \mp N_0\left(\omega_{s^{//}}^{//}(\vec{q}^{\,//})\right) + \frac{1}{2} \mp \frac{1}{2}\right) \times$$

$$\times \delta\left(\omega_s(\vec{q}) \pm \omega_{s^{/}}^{/}(\vec{q}^{\,/}) - \omega_{s^{//}}^{//}(\vec{q}^{\,//})\right) \mathrm{d}q_l^{/} \mathrm{d}q_\perp^{/} \qquad (3.3.16)$$

here $q_l^{/}$ and $q_\perp^{/}$ are the components of the vector $\vec{q}^{\,/}$ parallel or perpendicular to the lines defined by (3.3.14) and (3.3.15) correspondingly, $\gamma_s(\vec{q})$ is the mode-dependent Gruneisen parameter, which is determined for each phonon wave vector and polarization branch, $\rho$ is the surface mass density. In (3.3.16) the upper signs correspond to the processes of the first type while the lower signs correspond to those of the second type. The integrals for $q_l^{/}$ and $q_\perp^{/}$ are taken along and perpendicular to the curve segments, correspondingly, where the conditions of (3.3.14) and (3.3.15) are met.

The main mechanisms of phonon scattering in graphene are phonon-phonon Umklapp (U) scattering, rough edge scattering (boundary (B)) and point-defect (PD) scattering:

$$\frac{1}{\tau_{\mathrm{tot}}(s,\vec{q})} = \frac{1}{\tau_{\mathrm{U}}(s,\vec{q})} + \frac{1}{\tau_{\mathrm{B}}(s,\vec{q})} + \frac{1}{\tau_{\mathrm{PD}}(s,\vec{q})}, \qquad (3.3.17)$$

where

$$\frac{1}{\tau_{\mathrm{U}}} = \frac{1}{\tau_{\mathrm{U}}^{(I)}} + \frac{1}{\tau_{\mathrm{U}}^{(II)}},$$

$$\frac{1}{\tau_{\mathrm{B}}} = \frac{v_s}{l}\frac{1-p}{1+p},$$

$$\frac{1}{\tau_{\mathrm{PD}}} = \frac{S_0\Gamma q_s\omega_s^2}{4v_s}. \qquad (3.3.18)$$

Here $v_s = \mathrm{d}\omega_s/\mathrm{d}q$ is the phonon group velocity, $p$ is the specularity parameter of rough edge scattering, $S$ is the surface per atom and $\Gamma$ is the measure of the strength of the point defect scattering.

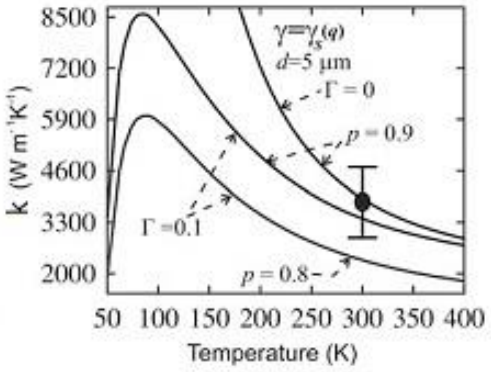



Figure 3.3.2 Calculated thermal conductivity of suspended graphene as a function of the temperature.

The sensitivity of the thermal conductivity, calculated using (3.3.9) and (3.3.14) – (3.3.17), to the value of $p$ and $\Gamma$ is illustrated in fig. 3.3.2. The data is presented for different sizes (widths) of the graphene flakes.

## References

1. S. Lepri (editor): Thermal Transport in Low Dimensions. From Statistical Physics to Nanoscale Heat Transfer, 411 p. Springer (2016).

2. S. Lepri, R. Livi, A. Politi: Heat Transport in Low Dimensions: Introduction and Phenomenology. Thermal Transport in Low Dimensions. From Statistical Physics to Nanoscale Heat Transfer. Chapter 1, pp. 1–38. Springer (2016).

3. D.L. Nika, A.A. Balandin: Thermal Transport in Graphene, Few-Layer Graphene and Graphene Nanoribbons. Thermal Transport in Low Dimensions. From Statistical Physics to Nanoscale Heat Transfer. Chapter 9, pp. 339–364. Springer (2016).

## 4. Arrays of carbon nanotubes: fabrication, properties and applications

The electronic structure, optical and magnetotransport properties of carbon nanotubes attract much attention because they offer great promise as functional elements of nanoelectronics [1, 2]. Single nanotubes possess unique mechanical, electric, and thermal properties. Most interesting are their charge transport properties. Specifically, nanotubes may be used in studies of various quantum phenomena such as single-electron tunneling [3]; observation of the behavior characteristic for Luttinger liquid [4]; ballistic transport [5], etc. On the other hand, single nanotubes were successfully used for the prototypes of field-effect transistors [6], biological [7] and chemical [8] sensors. Unfortunately, the production of sensors and device structures based on single nanotubes is very expensive. Because of this, the formation of the arrays of nanotubes is a very interesting challenge, necessitating further studies of their properties and comparison with the properties of single nanotubes for better understanding of their features and for the creation of the functional nanoelectronic elements and chemical, gas or biological sensors. Samples of different morphology are used as arrays of nanotubes: bundles [9], mats [10], networks [11], films [12]. Due to a high aspect ratio of carbon nanotubes and to the influence of contact barriers between individual nanotubes on the electric properties of the arrays, higher conductivity is expected for arrays of the nanotubes predominantly oriented along one of the preferential direction. At the present time different methods are developed for spatial orientation of individual nanotubes in the macroscopic structures on their basis. Among these methods, the following techniques are of particular interest: orientation of nanotubes in strong magnetic field [13]; forming fibers of the oriented nanotubes by means of electrophoresis [14]; formation of fibers based on the aligned nanotubes using of wet spinning process [15, 16].

Apart from the production of the arrays of carbon nanotubes with different geometries, the techniques have been developed for the formation of various composite materials containing carbon nanotubes as a filling material, whereas the «matrix» is made of polymeric, ceramic or other materials. It should be noted that research of the properties of conductive polymeric composites, offering much promise for the production of antistatic materials, printed-circuit boards, supercapacitors, organic solar cells, biosensors, flexible and transparent displays, and the like, [17], is one of the topical problems in modern materials science. Despite current restrictions on the use of polymeric composites associated with the expensive process that is prone to technological problems for commercial production, conductive polymers attract more and more attention due to new possibilities of their applications in electronics and chemical industry [18], due to their use as heating elements [19] and electromagnetic screens [20-22].

As known, conductivity of the arrays of nanotubes is strongly influenced by contact barriers between individual nanotubes [10]. To determine a role of contact barriers, not only the electric properties of the arrays of nanotubes and their magnetoresistance are studied but also their AC conductance is measured over a wide frequency range enabling one to determine both the active and reactive components of the impedance.

Besides, arrays of carbon nanotubes may be used for the development of THz-range radiation sources and detectors. In principle, the possibility for the development of such devices on the basis of carbon nanotubes has been demonstrated. A potential mechanism to improve conductivity and hence to make feasible the detection of THz waves is heating of the charge carriers as well as activation of the carrier transport over contact barriers between nanotubes in the arrays (at low temperatures). It is well known that a mechanism of heating the charge carriers in semiconductors is also used for the creation of microwave detectors which are used in communication systems and radio locators (radars). However, the feasibility of microwave detectors based on the effect of carrier heating in the arrays of carbon nanotubes is still to be studied.

This chapter presents the formation methods of the arrays of carbon nanotubes (CNTs) and the results obtained in studies of their electric and magnetotransport properties. The studied arrays of CNTs included the samples of different geometry: thin films of single-wall nanotubes; fibers of

single- wall CNTs; coatings of single-wall CNTs deposited onto the surface of $SiO_2$ fibers; CNTs-based composites. This enables one to establish the characteristic patterns for charge transport mechanisms in the arrays of CNTs, on the one hand, and to make recommendations for practical application of the properties under study when designing the functional electronic elements, electromagnetic radiation detectors, and temperature sensors on the basis of the arrays of CNTs.

### 4.1. Carbon nanotube formation techniques.

The formation of single CNTs results from chemical transformations of different carbon-containing materials at high temperatures. Among the most common CNTs synthesis methods, we can name the following: sputtering of graphite in arc discharge; laser ablation of graphite; chemical vapor deposition [23].

The functional nanoelectronic elements; antistatic and reinforcing composite coatings, electromagnetic screens; chemical, gas, and biological sensors based on the nanotubes the output signal of which is dependent on their conductivity are created using the arrays of nanotubes in different morphological forms (bundles, networks, films, fibers) and nanocomposites having the «matrix» of polymeric, ceramic or other materials. Because of this, of great importance are studies of the electric properties of CNT arrays and of the nanocomposite carbon materials containing nanotubes. Besides, it is necessary to establish relations to electrical properties of single nanotubes and a role played by the contact barriers between individual nanotubes.

Thin layers, monolayers, and networks of nanotubes are ordinarily produced by deposition from solutions containing nanotubes onto solid substrates or by the use of the Langmuir-Blodgett (LB) technique. Photomicrographs of thin CNT films formed by the LB technique are given in Fig. 4.1.

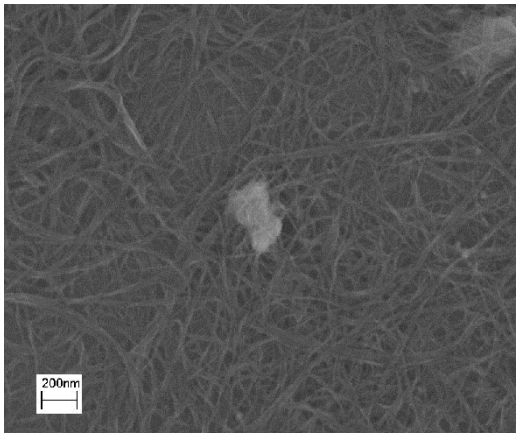



Fig 4.1a. SEM image of the film of CNTs formed using the LB technique.

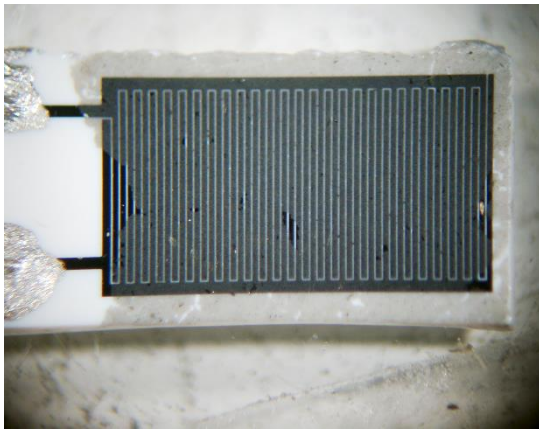


Fig. 4.1б. Image of the substrate surface with an finger-shaped structure of the contacts covered by a layer of the functionalized CNTs.

To form thin layers of CNTs, a drop of the functionalized nanotubes solution in chloroform was deposited on the water surface; then a drop of the dissolved PS-PMMA diblock polymer was applied to attain the surface pressure ~ 9 mN/m [24]. This technique enables one to vary thickness of the formed CNT films.

The fibers based on single-wall carbon nanotubes (SWCNTs) can be fabricated by wet spinning process from the solution containing nanotubes, a surfactant, and water [25]. The wet spinning technique is realized by injection of the CNTs-containing suspension into the rotating tank filled with water and by subsequent thermal treatment of the formed fibers having the diameter from 20 to 100 μm. Fig. 4.2 shows the scanning electron microscopy (SEM) image of

typical SWCNTs fibers. The pristine single-wall nanotubes were obtained by the laser ablation method.

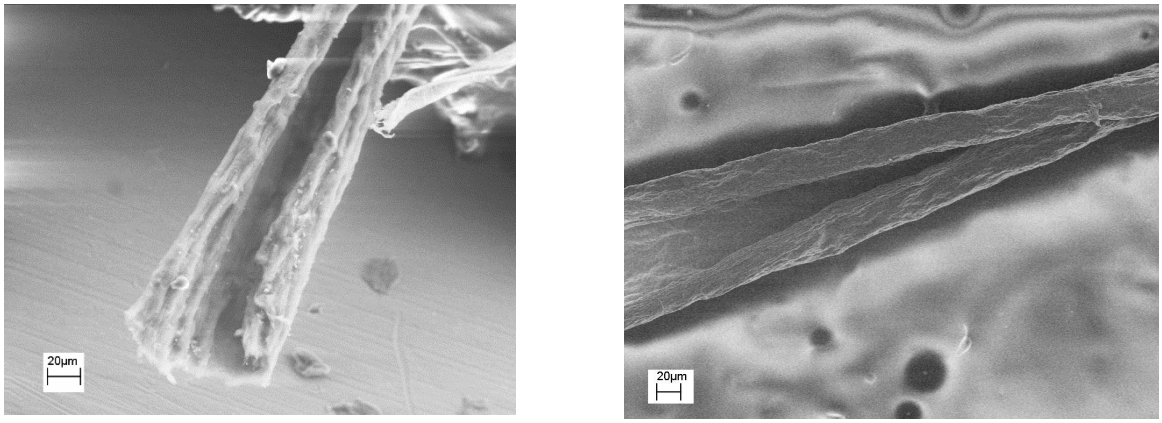



Fig. 4.2 SEM image of typical SWCNTs fibers

The technique used to form coatings of SWCNTs on the surface of $SiO_2$ fibers included the following stages [26]:

1) dissolution of a polyamide coating on the surface of $SiO_2$ fibers in the process of their treatment in toluene for 2 hours;

2) purification of residual polyamide by ultrasonic treatment of the fibers in a saturated solution of KOH/propanol for 20 minutes with subsequent 5-minute ultrasonic treatment in distilled water − this stage was repeated several times;

3) silanization of $SiO_2$ fibers (formation of hydrogen bonds on the surface of the fibers during their treatment in a solution of octadecyl trimethoxysilane in ethanol with the concentration 2 ml/g for 2 hours). Then the fibers were subjected to water dispersion with SWCNTs. After deposition of SWCNTs onto the surface of $SiO_2$ fibers (due to the hydrophobic interaction), they were annealed at 250 ºC.

Cross-section and side views of $SiO_2$ fibers with the deposited arrays of SWCNTs on their surfaces are demonstrated in Fig. 4.3.

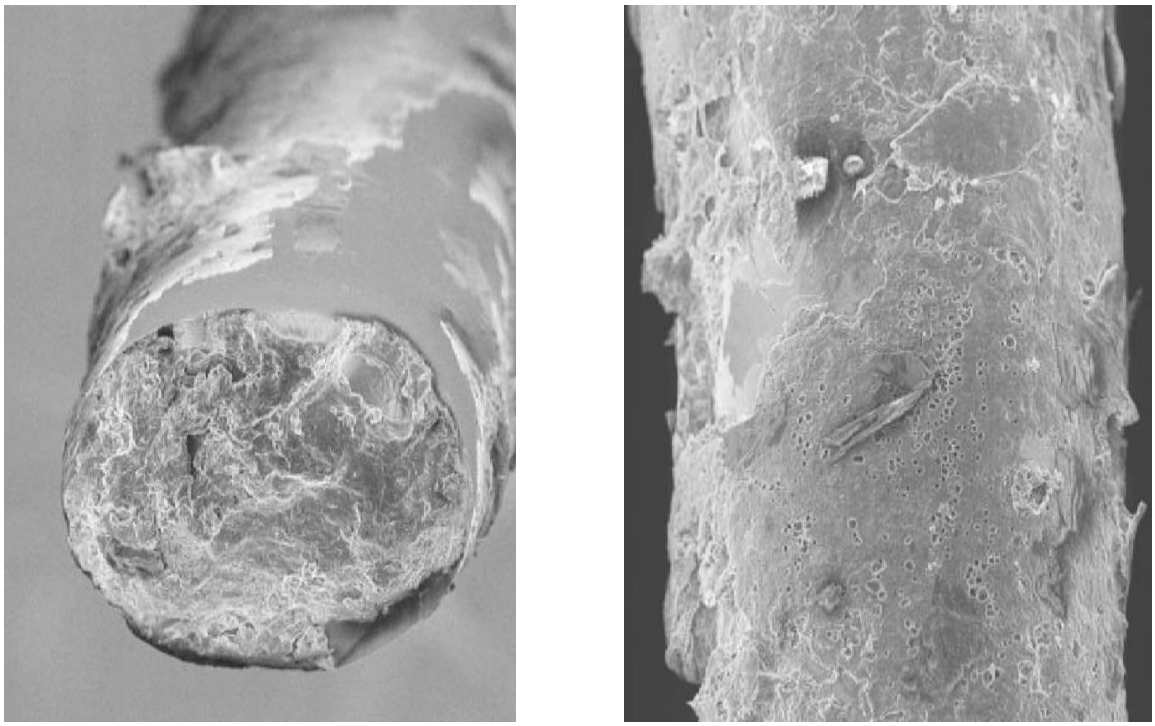


Figure 4.3a. Cross-section view of $SiO_2$-fibers with the arrays of SWCNTs deposited on their surface.

Figure 4.3b. Side view of $SiO_2$-fibers with the arrays of SWCNTs deposited on their surface.

Composites based on single-wall carbon nanotubes (SWCNTs) were produced with the use of Epikote 828 epoxy resin, A1 hardener, and SWCNTs produced by chemical vapor deposition from gas phase. SWCNT have the following parameters: 1-2 nm diameter, 10-20 μm length. The samples were produced as follows [27, 28]: degassing of epoxy resin in the vacuum 1-3 mbar for 12-14 hours and subsequent heating in the oven to the temperature 65 °C. Besides, carbon nanotubes were subjected to ultrasonic dispersion in propanol for 1.5 hours. Then the propanol solution with nanotubes was mixed with epoxy resin and heated in the oven to 130 – 150 °C for evaporation of propanol. Next a mixture of nanotubes and epoxy resin was dispersed in an ultrasonic bath for 1.5 hours. After addition of the curing agent, this mixture was manually stirred for 7 minutes. The obtained mixture was put into special forms for hardening during 20 hours, thermally treated in the oven at a temperature of 80 °C for 4 hours, and taken out of the forms. On the average, composites were 0.7 mm thick. Characterization of the samples was obtained using the scanning electron microscopy (SEM) method. Fig. 4.4 shows the typical SEM images of the composite epoxy resin/SWCNTs.

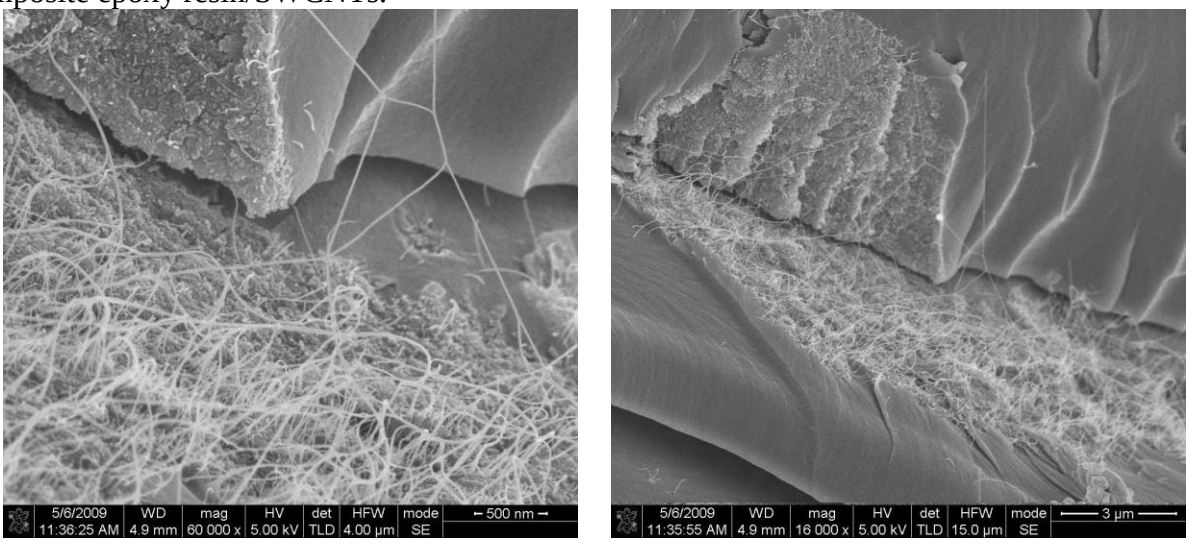



Figure 4.4. SEM images of the polymeric composite with SWCNTs as a filler.

To form thick layers of CNTs, filtration through porous membranes was used.

### 4.2. Electric properties of the arrays of CNTs.

The electric properties of individual CNTs are determined by their structure. Single-wall nanotubes, depending on the chirality, may exhibit semiconductor or metallic properties. According to theoretical calculations of the band structure, a single-wall nanotube with the indices *(n, m)* exhibits metallic properties if $n-m=3i$, where $i$ – integer. The results of theoretical calculations are supported by the experimental data: it has been found that about 2/3 of the synthesized nanotubes possess semiconductor properties and 1/3 – metallic. A width of the band gap $E_g$ for semiconductor nanotubes is dependent on chirality and on the nanotube diameter $d$: $E_g=(4\hbar v_F/3d)$, where $\hbar$ – Planck constant, $v_F$ – Fermi velocity. In this way the band gap $E_g$ of semiconductor nanotubes is inversely to the diameter and may be varied over the range from ~10 meV to ~1 eV.

Multiwall carbon nanotubes (MCWNTs) represent a more complex system: they consist of several single-wall nanotubes, differing in diameters, which are enclosed in one another. Conductivity in multiwall nanotubes is mainly governed by the outer layer due to its larger surface

area relative to the internal layers, most often exhibiting the metallic conductivity. In the majority of experiments devoted to the electric and magnetotransport properties of individual multiwall nanotubes a diffusion character of the charge transport characterized by multiple diffusion processes has been observed.

The electric properties of the arrays of nanotubes are determined not only by the chiral structure of individual nanotubes of which they are composed but also dependent on the quality of contacts between individual nanotubes, contact barriers between them, length of nanotubes, contact geometry and so on. Because of this, different charge transport mechanisms may be developed in the arrays of nanotubes: metallic conductivity; diffusion transport; fluctuation-induced tunneling, etc. Moreover, one can observe different conductivity mechanisms in the same system.

Measuring the temperature dependence of resistance and the dependence of the resistance on magnetic field is a popular and reliable method to study various materials, nanotubes and their arrays including, in an effort to establish the charge transport mechanisms determining conductivity of different systems.

When the volume content of nanotubes in a composite is low, nanotubes have unordered arrangement in the array, structural defects are present, and the like, variable range hopping conductivity is a possible conductivity mechanism in nanocomposites or arrays of nanotubes. The temperature dependence of resistance for this mechanism is described as follows [29]:

$$R(T)=R_0 exp(T_M/T)^{1/n}, \quad (4.1)$$

where $T_M$ – constant depending on the localization radius and on the density of localized sates in a system; $n=1+d$, $d$ – dimension of a system. To find the system's dimension from the experimental data associated with measurements of the temperature dependence for resistance, their analysis may be performed with the use of both linearization in scale $LnR(T^{-1/n})$ (where $n=2,3,4$) and special programs for nonlinear approximation.

When the main contribution into resistance of nanocomposites with carbon nanotubes as a filler or of nanotube arrays is made by the contact barriers between individual nanotubes, a system may exhibit a mechanism of fluctuation-induced tunneling of the charge carriers. The temperature dependence for resistance in the case of this mechanism is described as [30, 31]

$$R(T)=R_0 exp(T_1/T+T_0). \quad (4.2)$$

Here the parameters $T_0$ and $T_1$ are given by the following expressions:

$$T_0 = 16\varepsilon_0 \hbar A V_0^{3/2} / (\pi e^2 k_B (2m_e)^{1/2} w^2), \quad (4.3)$$

$$T_1 = 8\varepsilon_0 A V_0^2 / (e^2 k_B w), \quad (4.4)$$

where $w$ – tunneling barrier width, $A$ – tunneling barrier area, $V_0$ – barrier height.

A model of the fluctuation-induced tunneling was suggested for such disordered heterogeneous systems as, e.g., dielectric-metal composites, granulated metals, disordered semiconductors [30]. According to this model, in such systems the charge transport is mainly determined by transfer of the charge carriers over the distances, large at the atomic scale, which are separated by small tunneling barriers. At the areas with a good conductivity the charge carriers are delocalized, having the ability to move to relatively great in atomic scale distances. Due to thermal motion of the electrons, the potential fluctuations arising close to the tunneling barriers are capable to change the probability of electron tunneling considerably because these barriers are small. We can consider different types of the arrays of carbon nanotubes as heterogeneous disordered systems, where some well-conductive nanotubes are separated by the contact barriers.

Consequently, this model may be used for description of the charge transport mechanisms in the arrays of carbon nanotubes and in carbon nanocomposites with a large volume fraction of nanotubes.

For arrays of metallic carbon nanotubes or for composites containing metallic nanotubes, the temperature dependence of resistance *R(T)* can be described within the scope of a heterogeneous conductivity model, for which different conductivity mechanisms are observed in different temperature intervals:

$$R = \alpha T + R_0 exp(T_1/T+T_0), \quad (4.5)$$

where $\alpha$ - temperature coefficient of resistance characterizing the metallic conductivity, and the second term describes the low-temperature conductivity of a system within the frame of a model for the fluctuation-induced tunneling. This model is based on the assumption that the charge transport process in nanotube arrays and in nanocomposites at high temperatures is determined by the nanotubes with metallic conductivity. However, the presence of electric barriers (in contact regions between different nanotubes and also in the regions of nanotube bending or of structural defects within the nanotubes themselves) results in the development of carrier localization effects at low temperatures. Fig. 4.5 demonstrates the relationship between resistance and temperature for the arrays of single-wall nanotubes; this function is well described within the frame of a heterogeneous conductivity model [32] by formula (2.5). The model is based on the assumption that the charge transport process in arrays of SWCNTs at high temperatures is determined by the nanotubes with metallic conductivity. But, due to the presence of electric barriers, the localization effects of the charge carriers are essential at low temperatures. Consequently, conductivity in the low-temperature region is governed by tunneling of the carriers through the barriers. As found in studies of the thin-layer samples of SWCNTs at intermediate temperatures ($T$ ~ 100 - 185 K), the contribution made by metallic conductivity and by fluctuation-induced tunneling of the carriers into the overall conductivity of a system is comparable in the absolute value. This assumption accounts for the form of the temperature dependence of resistance. In the region of high temperatures ($T$ ~ 185 - 300 K) we can observe that resistance linearly increases with temperature, see the inset in Fig. 4.5a. On the other hand, due to a weak temperature dependence of resistance for the metallic component of a system and to saturation of the dependence with lowering of temperature, in the low-temperature region this dependence is determined by tunneling through the contact barriers according to Sheng's model and is described by the second term in equation (2.5). The dependence *R(T)* for fibers made of SWCNTs is given in Fig. 4.5b. As seen, due to the barrier effects between individual SWCNTs, there is no region with a positive temperature coefficient of resistance $dR/dT>0$ at high temperatures. In the low-temperature region the temperature dependence of resistance was approximated by the Mott law characteristic for variable range hopping conductivity. The dependence *R(T)* is of a similar form for SWCNT layers deposited onto the surface of $SiO_2$ fibers.

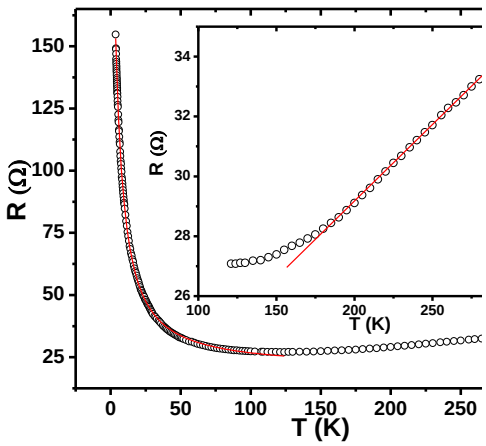



Figure 4.5a. The curve for *R(T)* of SWCNT films. Approximation of the dependence *R(T)* in the region of low temperatures according to (4.2) is represented by the solid line. The inset shows the dependence *R(T)*in the high-temperature region.

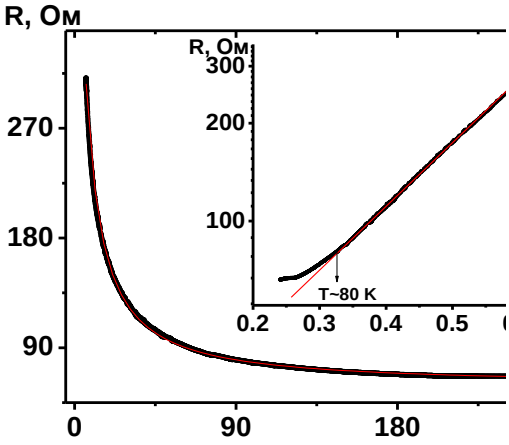



Figure 4.5b. The function *R(T)* of the fibers based on SWCNTs. Approximation of the dependence *R(T)* in the region of low temperatures according to (4.2) is represented by the solid line. The inset shows the dependence *R(T)* at logR-$T^{-1/4}$ scale.

Thus, by measurements of the temperature dependence of resistance of nanotube arrays or of the nanocomposites with carbon nanotubes as a filler we can establish a mechanism of the charge carrier transport in them. Finding the best approximation results of the experimental data, one can determine system's parameters associated with these mechanisms.

### 4.3. Magnetotransport properties of CNT arrays.

Magnetoresistance (MR) of the arrays of carbon nanotubes was measured in a Cryogenics closed-cycle helium cryostat in magnetic fields up to 8 T over the temperature range 2-300 K. Magnetoresistance was measured for the arrays of carbon nanotubes of different geometries: thin layers of SWCNTs; fibers of SWCNTs; layers of SWCNT deposited on the surface of $SiO_2$ fibers; composites comprising epoxy resin/SWCNTs and epoxy resin/MWCNTs. Actually, variations in sign of the magnetoresistance in the region of low temperatures and weak magnetic fields have been revealed for the samples of all types.

For example, only positive magnetoresistance was observed for composites comprising epoxy resin/ SWCNTs (with the weight content of nanotubes 1.5 %) at a temperature of 2 K. With further growth of temperature, a negative peak was observed for the curves of magnetoresistance. At temperatures above 6 K only negative magnetoresistance was observed: beginning from temperatures >10 K, a relative value of the negative magnetoresistance was lowered with increasing temperature.

Positive magnetoresistance, exponentially varying as $R=R_0exp(B^2)$, is characteristic for a mechanism of variable range hopping conductivity when an increase of resistance in a magnetic field is associated with the shrinkage of electron wave functions, leading to lower tunneling probability of the carriers between the localized states. However, positive magnetoresistance in the low-temperature region was observed even for the samples with a sufficiently high conductivity, where an analysis of their current-voltage characteristics and *R(T)* functions revealed no variable range hopping conductivity mechanism. Besides, we can hardly explain the nonmonotonic behavior of relative values of the negative magnetoresistance within the frame of the Sivan model [33] that includes the quantum-interference effects in regime of strong localization of the carriers. One of the ways to explain a negative magnetoresistance is consideration of the

contribution made by weak localization of the carriers [34] into the total magnetoresistance. We can assume that the main contribution into magnetoresistance of the arrays of carbon nanotubes is made by the influence of a magnetic field on the charge-carrier scattering mechanisms in single nanotubes. Positive magnetoresistance was observed in the region of temperatures and magnetic fields associated with electron-electron scattering as a principal scattering mechanism of the carriers [35]. With a growth of temperature, the arrays of nanotubes revealed only negative magnetoresistance due to the predominant electron-phonon scattering of the carriers [35].

A minimum on the magnetoresistance curves was most clearly seen for thin layers of SWCNTs deposited from the suspension (containing nanotubes 1-3 μm long in the average) onto the substrates with strip contact structures. The curves for magnetoresistance (defined as variations in resistance divided by the resistance of zero field) for these samples were measured at different temperatures. The curves for a relative change of conductivity in a magnetic field, reconstructed on their basis, are demonstrated in Figs. 4.6a and 4.6b, respectively.

As seen, negative magnetoresistance is observed for these structures. But, with an increase of magnetic field in the low-temperature region, the magnetoresistance changes the sign from negative to positive one. With an increase in temperature, we can observe only negative magnetoresistance, its absolute value decreasing as $T$ increases. Analyzing the form of the curve for resistance as a function of magnetic field, we can define the principal mechanisms contributing to the resistance [35]. For high magnetic fields and low temperatures, magnetoresistance is positive. This points to the fact that electron-electron interactions represent the principal scattering mechanism in the temperature interval 2-15 K [35].

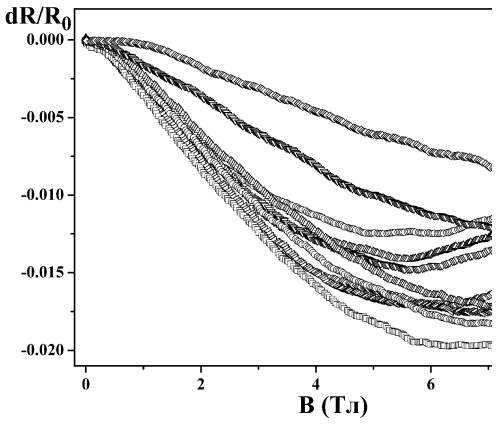



Figure 4.6a. Relative changes in the relationship between resistance *(R(B)-R(0)/R(0))* and magnetic field for layers of SWCNTs 1-3 μm in length (formed by deposition of the nanotubes containing suspension) measured at different temperatures.

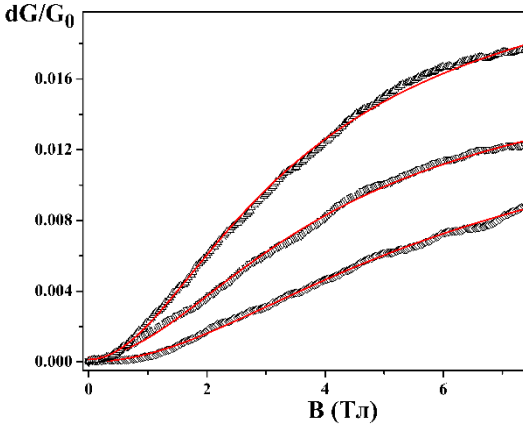



Fig. 4.6b. Relationships between *G(B)-G(0)/G(0)* and magnetic field constructed from the dependences of *R(B)-R(0)/R(0)* on magnetic field for layers of SWCNTs 1-3 μm in length, which were measured at different temperatures. Solid lines represent approximation of the relationship between *G(B)-G(0)/G(0)* and magnetic field by the law (4.6) characteristic for one-dimensional weak localization.

For low magnetic fields and high temperatures, the negative magnetoresistance is observed. This points to the fact that in the temperature interval 15-80 K the principal scattering mechanism is associated with weak localization [35]. Besides, we have found that in the region of high temperatures, where only negative magnetoresistance was observed, the curves for magnetoconductivity can be well approximated by the law characteristic for one-dimensional weak localization [36]:

$$G(B) = G_0 - \frac{2e^2}{\pi\hbar L}\left(\frac{1}{L^2_{\varphi}} + \frac{(eBw)^2}{3\hbar^2}\right)^{-1/2} . \quad (4.6)$$

Here $G_0$ – zero magnetic-field conductivity; $\hbar$ – Planck constant; $e$ – electron charge; $L$ – intercontact distance; $w$ – diameter of a one-dimensional system; $L_\varphi$ – phase coherence length of the carriers.

This supports our assumption that the principal contribution into magnetoresistance of the arrays of CNTs is made by the magnetic field effects on scattering mechanisms of the charge carriers in single nanotubes.

In this way the magnetoresistance of the arrays of carbon nanotubes with different geometries has been studied in magnetic fields up to 8 T over the temperature range 2-300 K. It has been found that the principal contribution into magnetoresistance of the arrays of carbon nanotubes is made by the magnetic field effects on scattering mechanisms of the carriers in single nanotubes. Positive magnetoresistance was observed in the region of temperatures and magnetic fields associated with electron-electron scattering as the principal mechanism of scattering.

As temperature increased, only negative magnetoresistance was observed due to the predominant electron-phonon scattering mechanism of the carriers.

**4.4. Devices and sensors based on CNTs.** It is well known that single carbon nanotubes are characterized by unique mechanical and electric properties, high thermal conductivity, and chemical stability. Because of this, carbon nanotubes offer much promise in different fields of applications [37-39]. In the last years considerable advances have been achieved in the development of synthesis and purification methods for nanotubes, making it possible to produce nanotubes of good quality in large quantities and to lower prices per product unit.

It should be noted that practical applications of carbon nanotubes are based on the properties of their arrays. This is due to the fact that the production of devices based on the properties of single nanotubes is expensive and unreasonable. One of the possible applications of single nanotubes is the creation of the probes for atomic-force and electron microscopy.

Presently, devices and materials on the basis of carbon nanotubes are serially produced. First of all, nanotubes are used in composite materials, e.g., for the production of lithium-ion batteries or electrostatic screens, where multiwall nanotubes are used as the agents contributing to the improved properties of the composites.

Owing to high strength of carbon nanotubes, they may be used for the production of reinforcement composites both for daily living needs and for aerospace industry.

By the present time, the prototypes of flat-panel displays with the use of the arrays of vertically oriented nanotubes have been designed. Unfortunately, such displays are not produced commercially.

Single-wall nanotubes have a variety of potential applications in optoelectronics and photonics. It has been demonstrated that nanotubes may be used as photodetectors for transformation of optical radiation to electric signals. Moreover, the operation of a detector based on individual carbon nanotubes has been demonstrated [40]. A single-wall carbon nanotube (SWCNT) may be considered as a one-dimensional wire, the electron energy spectrum of which represents a series of the peaks (so-called Van Hove singularities) with the distances between the energy levels coming to 0.1-2 eV [41]. The electron transitions between pairs of the Van Hove singularities underlie the operation of optical detectors, based on single carbon nanotubes, over a wide wavelength range.

The development of detectors of electromagnetic radiation in the visible or ultraviolet wavelength range for large-scale serial production, with the use of the nanotube arrays rather than individual nanotubes, necessitates solution of a rather complex problem – separation of nanotubes on their chirality. When using the arrays of nanotubes with identical chiral properties and hence with a similar band structure, one can produce the detectors demonstrating higher sensitivity with the selectivity comparable to that of the detectors on the basis of individual nanotubes.

Films and bundles of nanotubes are also suitable for the photocurrent generation but, in difference from single nanotubes, their detection mechanism is thermal (bolometric) in character [40]. Under the effect of radiation, the energy is imparted to the whole array of SWCNT and this leads to heating of the sample and to variations in its resistance. As a bolometric signal may be observed in a significantly higher wavelength region of electromagnetic radiation with a characteristic relaxation time of several milliseconds, the arrays of carbon nanotubes have potentialities for application as active components of detectors for terahertz-range radiation [42, 43].

However, the signals with a low relaxation time observed by the authors in their structures are nonthermal in nature [24]. Besides, it has been found that a signal is dependent on the ordering of nanotubes in the arrays. It seems expedient to continue this direction of our studies. Because in the arrays created from nanotubes of the same length the antenna effects may appear when the nanotube lengths are coincident with the terahertz-range radiation wavelength (the frequency 1 THz corresponds to the wavelength on the order of 0.3 mm), one can improve selectivity of detectors for terahertz radiation by choosing nanotube with the same length.

Carbon nanotubes offer much promise for the development of different sensors. It has been found that carbon nanotubes are sensitive to chemical, biological, and gas media and can be used in the production of gas, chemical, and biosensors [44-47]. In this case there is no need to separate nanotubes in the arrays according to their chiral properties as their operation is based on registration of variations in conductivity of the whole array. It has been found that a conductivity of the array of nanotubes is strongly dependent on the quality of contacts between individual nanotubes and not only on the electric properties of the individual nanotubes. A great aspect ratio – *surface area/volume* – characteristic for the structures based on carbon nanotubes makes it possible to increase sensitivity of the sensors produced with their use.

Owing to their good conductivity, sensitivity to the external effects, relatively simple fabrication technology, the arrays of carbon nanotubes with different geometries (fibers of single-wall nanotubes; layers of single-wall nanotubes deposited onto the surface of $SiO_2$ fibers; thick layers of multiwall carbon nanotubes) hold the greatest promise for the development of various sensors to be used in biological, chemical, and gas media.

One of the fields for application of the arrays of carbon nanotubes is electronics. According to the predictions, serial production of integrated circuits with nanotubes as active elements has number of problems. The main difficulty is separation of individual nanotubes according to their chirality. At the same time, high thermal conductivity (nearly by an order of magnitude higher than that of copper) and good electric conductivity of nanotubes enables the formation on their basis of interconnections and heat sinks or field-effect transistor gates for integrated microcircuits even in the near future. As compared to the standard copper conducting paths, interconnections on the basis of carbon nanotubes are more preferable due to a considerably increased role of the surface scattering with the proportional miniaturization of integrated circuits.

**References:**


1. Dresselhaus M.S., Dresselhaus G., Avouris Ph. Carbon nanotubes. – Springer, Berlin, Germany, 2001.
2. Avouris Ph. et al. // Appl. Surf. Sci. – 1999. V. 141. – P. 201.
3. Bockrath M., Cobden D. H., McEuen P. L., Chopra N. G., Zettl A., Thess A. and Smalley R. E. // Science –1997. –V. 275. – P. 1922.
4. Bockrath M., Cobden D. H., Lu J., Rinzler A. G., Smalley R. E., Balents L. and McEuen P. L. // Nature. –1999. – V. 397. –P. 598.
5. Frank S., Poncharal P., Wang Z. L. and DeHeer W. // Science. –1998. –V. 280, –P. 1744.
6. Martel R. et al. // Appl. Phys. Lett. –1998. –V. 73. – P. 2447.
7. Guo Z. et al. // Advanced Materials –1998. –V. 10. – P. 701.
8. Kong J. et al. //Science. –2000. –V. 287. – P. 622.

9. Krstic V., Roth S. and Burghard M. // Phys. Rev. B. –2000. –V. 62. – P. R16352.
10. Kaiser A. B., Dusberg G. and Roth S. // Phys. Rev. B. –1998. –V. 57. – P. 1418.
11. Kim G.T., Jhang S.H., Park J.G., Park Y.W. and Roth S. // Synthetic Metals.–2001. –V. 117. – P. 123.
12. Baumgartner G., Carrard M., Zuppiroli L., Bacsa W., deHeer W. A. and Forro L. // Phys. Rev. B. –1997. –V. 55. – P. 6704.
13. Smith B.W., Benes Z., Luzzi D.E., Fischer J.E., Walters D.A., Casavant M.J., Schmidt J. and Smalley R.E. // Appl. Phys. Lett. – 2000. – V. 77. – P. 663.
14. Gommans H.H., Alldredge J.W., Tashiro H., Park J., Magnuson J. and Rinzler A.G. // J. Appl. Phys. –2000. –V. 88. –P. 2509.
15. Launois P. and Poulin P. // Encyclopedia of Nanoscience and Nanotechnology – 2004. – V. 4. - P. 1.
16. Vigolo B., Penicaud A., Coulon C., Sauder C., Pailler R., Journet C., Bernier P. and Poulin P. // Science. –2000. –V. 290. –P. 1331.
17. György Inzelt. Conducting polymers. A new era in electrochemistry, Series: Monographs in Electrochemistry, Editor: Fritz Scholz, (Springer, 2008), XII, 282 p., ISBN: 978-3-540-75929-4.
18. Dantal B., Saigal A., and Zimmerman M. // Experimental Analysis of Nano and Engineering Materials and Structures. – 2007. – Vol. 1T3. – P. 37.
19. Boysen D.A., Chisholm C.R.I., Haile S.M., and Narayanan S.R. // J. Electrochem. Soc. – 2000. – Vol. 147. – P. 3610.
20. Subramoney S. // Adv. Mater. – 1998. – Vol. 10. – P. 1157.
21. Hu Ya., Shenderova O.A., Hu Z., Padgett C.W., and Brenner D.W.// Rep. Prog. Phys. – 2006. – Vol. 69. – P.1847.
22. Reich S., Thomsen C., and Maultzsch J., *Carbon Nanotubes: Basic Concepts and Physical Properties* (Wiley-VCH, Berlin, 2004), 224 p.
23. А.В. Елецкий. Углеродные нанотрубки и их эмиссионные свойства. УФН. – 2002. – Т.172, №4. – с. 401-438.
24. Ksenevich V., Galibert J., Samuilov V. Charge transport in carbon nanotubes films and fibers // In: Carbon Nanotubes, Edited by J. M. Marulanda, p.p. 123-145, INTECH, 2010, ISBN 978-953-307-054-4.
25. Kozlov M.E., Capps R.C., Sampson W.M., Ebron V.H., Ferraris J.P. and Baughman R.H. Spinning Solid and Hollow Polymer-Free Carbon Nanotube Fibers // Adv. Materials. – 2005. –V. 17. –P. 614.
26. Q. Lu, V. Samuilov, V. Ksenevich, T. Dauzhenka, R.S. Helburn. Temperature Dependent Resistance and Magnetoresistance of Single Wall Carbon Nanotubes Mounted on Silica Fiber Surfaces // In: Interfaces and Interphases in Analytical Chemistry, Edited by Robin Helburn, Mark F. Vitha, p.p. 185-196, ACS Symposium Series, Vol. 1062, 2011, ISBN13: 9780841226043, eISBN: 9780841226050, Chapter DOI: 10.1021/bk-2011-1062.ch008.
27. Bellucci S., Balasubramanian C., De Belli G., Micciulla F., and Rinaldi G.// Macromolecular Symposia. – 2008. – Vol. 263. – P.21.
28. Bellucci S., Balasubramanian C., Micciulla F., and Rinaldi G.// Journal of Experimental Nanoscience. – 2007. – Vol. 2.– P. 193.
29. B.I.Shklovskii, A.L.Efros, *Electronic Properties of Doped Semiconductors*, Springer-Verlag, Berlin 1984.
30. P.Sheng // Phys. Rev. B. – 1980. – Vol.21. – P. 2180.
31. G.T. Kim, S.H. Jhang, J.G. Park, Y.W. Park, and S. Roth // Synthetic Metals. – 2001. – Vol. 117. – P.123.
32. Kaiser A.B., Dusberg G. and Roth S. Heterogeneous model for conduction in carbon nanotubes // Phys.Rev.B . –1998. – V. 57, No. 3. –P.1418-1421.
33. Sivan U., Entin-Wohlman O. and Imry Y. // Phys. Rev. Lett. – 1980. – V. 60. – P. 1566.

34. Lee P. and Ramakrishnan T.V. Disordered electronic systems // Rev.Mod.Phys. – 1985. – Vol. 57, No. 2. – P.P. 287-337.
35. Choudhury P.K., Jaiswal M., and Menon R. Magnetoconductance in single-wall carbon nanotubes: Electron-electron interaction and weak localization contributions // Phys. Rev. B. – 2007. – V. 76. – P. 235432.
36. Strunk C., Stojetz B., Roche S. Quantum interference in multiwall carbon nanotubes // Semicond. Sci. Technol. – 2006. – Vol. 21, № 11. – P.P. S38-S45.
37. Dai H. Carbon nanotubes : opportunities and challenges // Surface Science. – 2002. – Vol. 500, No. 1-3. – P.P. 218–241.
38. Robertson J. Realistic applications of CNTs // Materials Today. – 2004. – Vol. 7, No. 10. – P.P. 46-52.
39. Robertson J. Growth of nanotubes for electronics // Materials Today. – 2007. – Vol. 10, No. 1-2. – P.P. 36–43.
40. P. Avouris et al. // Nature Photon. – 2008. – Vol. 2. – P.P. 341.
41. A. Mohite et al. // Appl. Phys. Lett. – 2005. – Vol. 86. – P. 061114.
42. K. Fu et al. // Appl. Phys. Lett. – 2008. – Vol. 92. – P. 033105.
43. M. Itkis et al. // Science. – 2006. – Vol. 312. – P. 413.
44. P.G. Collins, K. Bradley, M Ishigami, A. Zettl. Extreme Oxygen Sensitivity of Electronic Properties of Carbon Nanotubes // Science. – 2000. – Vol. 287. – P.P.1801-1804.
45. O.K. Varghese, P.D. Kichambre, D. Gong, K.G. Ong, E.C. Dickey, C.A. Grimes. Gas sensing characteristics of multi-wall carbon nanotubes // Sensors and Actuators. – 2001. – Vol. B81. – P.P. 32-41.
46. 11. J. Kong, N.R. Franklin, C. Zhou, M. Chaplin, S.Peng, K.Cho, H.Dai. Nanotube molecular Wires as Chemical Sensors // Science. – 2000. – Vol. 287. – P.P. 622-625.
47. Z. Guo, P.J. Sadler, S.C.Tsang. Immobilization and Visialization of DNA and Proteins on Carbon Nanotubes // Advanced Materials. – 1998. – Vol. 10. – P.P. 701-703.

## 5. Conductive polymers

Conductive polymers represent a new class of materials coming into the use quite recently and swiftly developing at the present time. The first conductive polymer starting a new trend of investigations was polyacetylene [$(CH)_n$]. The 2000 Nobel Prize winner in chemistry Hideki Shirakawa synthesized it from acetylene in the form of solid silvery films in 1974 at the Tokyo Institute of Technology. In 1977 he, simultaneously with the USA researchers, found that, due to partial oxidation of polyacetylene by molecular iodine or other reagents, its conductivity is increased by a factor of $10^9$.

Such materials are highly promising for microelectronics and photonics owing to their relatively low cost, simple processing techniques, mechanical flexibility as compared to inorganic layers. Currently, more than 90% of photoreceptors in copy machines and laser printers are produced of molecularly-doped polymers and low-molecular organic compounds. The major electronic companies (Sony, LG, Panasonic, Samsung, etc.) are manufacturing displays on the basis of organic electroluminescent materials. In laboratories all over the world the researchers design and test new types of light-emitting devices, organic photodetectors, photovoltaic units, electrooptical modulators, field transistors, photorefractive materials.

The activities in this field are stimulated by advances in synthesis of the functional materials based on polymers with the desired and reproducible properties. The use of special synthesis methods with subsequent modification enables one to vary conductivity of polymers over a wide range, beginning from the conductivities characteristic for dielectrics (below $10^{-10}$ S/cm) to the metallic conductivity (higher than 1 S/cm). These possibilities contribute to more and more extensive applications of the polymers in various fields.

### 5.1. Structure, properties, and classification

Polymeric materials are characterized by low specific weight, they are easily processes and chemically inert. The majority of polymers are insulators. Because of this, they are widely used as dielectrics and insulating materials. A great interest to electrically conductive polymers in the last few years is associated with their potentialities in electronics as materials for functional elements. Of particular importance is a character of their interatomic bonds in a molecule which influence their physical and chemical properties. Carbon is the principal element in the matrix of the majority of polymeric materials. Its atom has the atomic number 6; the electronic configuration in the ground state is $1s^2 2s^2 2p^2$. Being aggregated in molecules, the atomic orbitals, due to redistribution of the electron density with respect to the nucleus, form strictly directional hybrid orbitals. When a chemical bond is formed, atoms are overlapped by the orbitals in the direction, where the electron density is maximal. Uniqueness of carbon is due to its ability to form the structures comprising the atoms which are in one of the three states with hybridization of atomic orbitals: $sp^3$, $sp^2$ or sp (Fig. 5.1).

In a wide sense, polymeric materials are compounds constituting of a large number of atoms or atomic groups which are connected by the valence bonds. They consist of the repeated groups (monomer units). Molecular mass of polymers varies from several thousands to millions of monomer units.

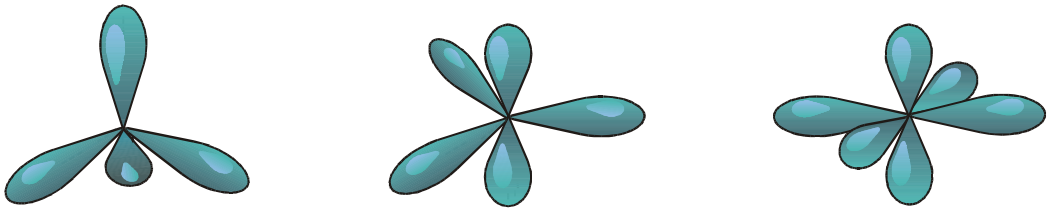


*Figure 5.1 – Hybridization of carbon atomic orbitals*

In general, in polymers two bond types are realized: σ-bonds and π-bonds. σ-bond and π-bond may form a double bond between the adjacent atoms in a molecule. Regular alternation of single and double bonds leads to the formation of conjugate systems in polymers. σ-orbitals are symmetrical with respect to the bond axes, forming the localized bonds: C–C and C–H. The electrons filling these orbitals are concentrated mainly along the line that connects two carbon nuclei and are localized. On the contrary, the electrons of π-orbitals are not involved in any bonds, being delocalized over all the carbon atoms so that their free motion in a molecule is possible. Sometimes they are called the mobile or unsaturated electrons. Owing to the presence of mobile π-electrons, electric effects are easily transferred from one part of the molecule to another. Besides, π-electrons are responsible for transfer of chemical and physical effects along the whole length of large molecules in organic systems.

Polymers are characterized by the submicroscopic structure. Such structures are understood as a totality of various isolated formations (differing in their shape, size, and internal constitution) composed of many chain macromolecules: amorphous (glassy) polymer with randomly packed macromolecules; folded polymer chains forming the lamellae which represent analogs of two-dimensional molecular crystals; crystalline polymer; microvolumes of amorphous crystalline polymers. Actually, polymers represent a mixture of the crystalline and amorphous phases.

Structural inhomogeneity of polymeric materials is associated with differences in thermodynamic behavior of macromolecules and of the ordinary molecules. Macromolecular systems can exist simultaneously in several crystalline or supermolecular forms. This phenomenon is known as polymorphism.

A characteristic feature of polymers is combination of the properties intrinsic in solids and in liquids, i.e., these materials posses the strength of solids being capable of reversible deformations due to their plasticity. Polymers exhibit a high anisotropy of their mechanical properties that allows for the formation of different films and fibers.

The boiling temperature of polymers is higher than the dissociation temperature and hence polymers can be in the condensed phase only and can not be in the gaseous state. One variety of the solid state is a glassy state in which molecules loose their ability for motion. High-elasticity state is another variety. Polymers in this state have the ability for reversible deformations, being in a particular intermediate state between the liquid and glassy state. Deformational motion involves only small segments of chain macromolecules.

The properties of polymers are greatly dependent on a character of the above-mentioned microscopic supermolecular structures. Changes of the supermolecular polymeric structure – structural transition – result in variations of the physical and chemical properties influencing the reactions which proceed within the polymeric matrix. To illustrate, melting may be considered as a structural transition in the process of which an ordered structure becomes less ordered or disordered.

The properties of polymers, specifically their mechanical properties, are very important when polymers are used as functional materials. The quality of products manufactured on the basis of polymers is greatly dependent on such mechanical parameters as hardness, elasticity or tensile, compression, and bending strength.

Of great importance are such physicochemical properties of polymers as hygroscopicity, moisture resistance, wettability. For example, moisture penetrating the pores of a material leads to lowering on conductivity, limiting the use of polymers in conditions of high humidity.

Sometimes the quality of polymeric materials is affected only after prolonged external effects associated with high variations in the ambient pressure or composition (changes in the concentration of oxygen or different chemical reagent, and so on). This phenomenon known as aging must be taken into account for long-term usage of the products made of polymers.

According to their electric properties, polymers are subdivided into dielectrics, semiconductors, electrically conductive materials. This division is very conventional. It is accepted that conductivity of dielectrics is below $10^{-10}$ S/cm. Polymers with high conductivity are attributed to the class of semiconductors, whereas at conductivities above 1 S/cm they may be considered conductive.

Semiconductor and conductive polymers are produced by the following methods:

- introduction (filling) of conductive components into the polymeric matrix;
- formation of conjugated double bonds in the process of polymerization;
- creation of polymeric charge-transfer complexes (CTC);
- high-temperature processing of polymers in the vacuum (pyrolysis);
- ion implantation of polymers.

In semiconductive polymers the mobility is usually very low coming to $10^{-3}$-$10^{-1}$ $cm^2(V \cdot s)^{-1}$ and it is impossible to use the Hall effect for its measurement. We should except pyrolized polymers with the mobility 1-100 $cm^2(V \cdot s)^{-1}$ and filled polymers, where the mobility is greatly determined by the filling material.

## 5.2. Fundamentals of polymer electrophysics

According to their ability to move under the effect of an electric field, charges in polymers may be subdivided into two types. Charges of the first type (free charges) subjected to the electric field effect can move in an object within macrodistances and be discharged at the electrodes creating the conduction current. They may be represented by electrons or ions and are mostly characteristic for semiconductive and conductive materials. Charges of the second type (bound charges) are strongly coupled with structural units of the material (atom, molecule, cluster and the like). Due to the effect of an electric field, these charges can move only within the limits of their structural unit. The process of the finite charge displacement in a stationary electric field is referred to as the dielectric polarization. This phenomenon is characteristic for low-conductivity dielectric materials. It should be noted that in reality a polymer always contains both types of charges because, even in dielectrics with an ideal crystalline structure which, according to the band theory of solids, should not conduct the current, conductivity is nonzero due to the presence of impurities and structural defects – sources of mobile charge carriers.

The term “static electrification” denotes the generation and separation of positive and negative charges as a results of contact or interaction of the material surface with other material or medium: surface friction of solids; contact between the solid surface and liquid; passing of gasses (ionized gasses including) over the solid surface; interaction between the solid surface and beams of charged particles, etc.

A mechanism of contact electrification is most common and thoroughly studied. Electrification of dielectrics in contact with each other may include two effects: kinetic and equilibrium. For example, in the case of surface friction between two objects of the same material, when one of the objects is motionless and the other is sliding over the first one, we can often observe reciprocal electrification. The kinetic effect resulting in electrification is caused by higher heating of the stationary object compared to the sliding one. The equilibrium effect is simply caused by contact between different materials. As demonstrated by experiments in the vacuum, electrification of polymers, when the contacting surfaces are metal and polymer or polymer and other polymer, is due to transfer of the charge by electrons.

Besides, a static charge may arise on the surface of dielectric polymers under the effect of a beam of charged particles when electrification is associated with ionization of the surface dielectric layer by the penetrating particles.

Electrification of a polymeric material is greatly influenced by its conductivity. If the conducting surfaces are conductive enough, a greater part of the electrifying charge is neutralized on its separation due to the inverse electron flow. The structure of a polymer is of crucial importance for its ability to accept a positive or a negative charge of static electricity. This feature

was used to make a list of the so-called triboelectric series for polymeric materials, where polymers are positioned in a descending order according to the probability of accepting a positive charge or in an ascending order with respect to a negative charge. This series is if great practical importance. At the same time, we should take into consideration that a sign of the static charge is variable with varying environmental conditions. It has been sown that a static charge may be compensated by the adequate selection of compositions for the materials positioned at different ends of this series. Moreover, the position of a dielectric in the triboelectric series is related to its dielectric permeability: friction of two dielectrics against each other leads to positive charging of the dielectric having a higher dielectric permeability; a quantity of electricity generated on friction of the bodies against each other is proportional to the difference in their dielectric permeabilities.

Electrification of polymeric dielectrics involves both the charge generation and its decay (relaxation). Relaxation may be due to internal factors, e.g., migration and recombination of charges, or external factors such as neutralization of the charge by atmospheric ions.

In general, in polymeric dielectrics free carriers are represented by the ions generated as a result of molecular dissociation and the electrons injected from the contacts or thermally excited from impurity levels. In semiconductive and conductive high molecular compounds the principal charge carriers are electrons. Describing the carrier motion, researchers highlight two limiting cases or mechanisms. In the first case motion of the carriers is considered as propagation of a delocalized plane wave in the conduction band; in the other case the carrier is strongly localized and moves by jumping or hopping, each jump being associated with scattering.

Though electron conduction is characteristic, first of all, for polymeric semiconductors and conductive materials, it may be observed for dielectric polymers too.

The following models may be used to describe a motion mechanism of electrons and holes in a polymer: activation-free band model, hopping model, and activation hopping model. In some cases we should take into account emission of the electrons from electrodes and conductivity due to impurities and defects.

When conductivity is considered within the scope of a band model, the notions and inferences of a band theory for solids, developed for low-molecular crystals, are used. The theory may be applied to polymeric solids because the crystalline formations in polymers are characterized by a high degree of ordering and by the long-range order existent along a macromolecule. However, a width of the forbidden band in some polymers is great, e.g. for polytetrafluoroethylene $E_g$=10.07 eV, and transport of the electron from the valence to the conduction band should result in destruction of the macromolecule due to breaking of the σ-bonds. If in a polymer there is a chain of conjugated double bonds along macromolecules, migration of electrons from one π-band to another is not accompanied by breaking of the macromolecules and a width of the forbidden band is not great. To illustrate, let us consider the organic polymer polyacetylene representing the simplest linearly-conjugated polymer with the structural formula $(–CH=CH–)_n$. The polymer consists of the chaotically oriented fibers several hundred angström in diameter. Being doped with donor (Li, $NH_3$) and acceptor ($I_2$, $AsF_5$) impurities, polyacetylene acquires the properties of a *p*- or *n*-type semiconductor, respectively, its conductivity varying by 12 orders of magnitude from $10^{-9}$ to $10^3$ $\Omega^{-1}\cdot cm^{-1}$.

High electron mobility along the conjugate chain is the necessary but not sufficient condition for a considerable polymeric conductivity. Additionally, the transport of carriers from one macromolecule to another is required. These intramolecular transitions may be realized by activation surmounting of the potential barriers between the molecules due to the activation, hopping or barrier mechanism.

## 5.3. Synthesis of semiconductive or conductive polymers

Filled conductive polymers represent a broad class of the materials characterized by the electric conduction due to the introduction of conductive components into the polymer structure. According to their conductivity, these polymers are subdivided into resistive ($10^{-4}$-$10^{8}$ Ω·m) and contact ($10^{-7}$-$10^{-4}$ Ω·m) groups. Filled polymers are heterogeneous systems composed of conductive components, binders, and fillers. The initial polymer plays a role of the binding material that determines the physical and chemical properties of the resultant polymer.

The following powder materials are commonly used in filled polymers (FP) as conductive components: carbon, graphite, fine-dispersed particles of silver and other metals. Of critical importance for conduction is the surface of particles of a conductive component. Binding components of conductive FP bound and fix the structure of a conductive component, i.e. provide for dispersion of the conductive component, formation of conductive chains, and regulation of a width of dielectric interlayers between the working particles, retaining electrophysical properties of the material. Fillers are introduced into conductive FP to improve the mechanical characteristics and to attain better heat and moisture resistance. Often the use of fillers is aimed at lowering the cost of a particular material because fillers are usually less expensive than polymers.

The principal electric characteristics of filled polymers in general are close to those of conductive components. Consequently, in the first approximation we can consider filled polymers to be a system of conductive particles the total contact resistance between which governs resistance of the material as a whole.

A double bond in polymers is formed between the adjacent atoms in a molecule by means of σ- and π-electrons. The double bonds exhibit special properties when they are positioned in a molecular chain alternately with single bonds. In this case we can speak about the conjugation effect because the "clouds" of π-electrons for all the atoms forming double bonds at a particular section of a molecular chain are positioned in one plane and overlap. In this case π-electrons are no longer localized and belong to the whole conjugate system. Such systems are characterized by higher stability than the corresponding nonconjugate systems. As the conjugate chain length is growing, conductivity of the polymer as a whole is increased. It has been found that in semiconductive polymers with a system of conjugate bonds the electronic conduction is realized.

Polymers with a system of conjugated bonds may be synthesized according to the following principles: creation of a conjugated bond in the available nonconjugate molecule of the polymer by thermal, chemical or radiative detachment of side groups; production of the conjugate macromolecule by polymerization or polycondensation of the monomers containing multiple bonds or aromatic nuclei; chelation of polymers by metals; formation of polymer complexes with halogens.

It is assumed that for polymers with a system of conjugated bonds within the conjugate chain the macromolecular π-electrons are moving as in a common potential well with a periodic potential determined by the chain structure. In the first approximation we can study such electron motion using a model of the electrical conductivity band theory. However, apart from motion of electrons along the conjugate chain, the carriers should be able to migrate from one macromolecule to another, i.e. from a conjugate system to the other. These intermolecular transitions are possible due to activation surmounting of the potential barriers between the molecules.

Charge-transfer complexes (CTC) are donor-acceptor systems, where one molecule has the properties of a donor, whereas the other – of an acceptor. Acceptors and donors are the formations with completely packed molecular orbitals. But, when one of the molecules has a considerable electron affinity and the other is characterized by a fairly low ionization potential, there is the probability for polarization of donor and acceptor molecules. This means that one or several π-electrons of a donor molecule will belong, for a very short time, to the acceptor molecule with retention of the spin orientation. In other words, the two structural subunits are exchanging their electrons: $A+D \Leftrightarrow A^{-}+D^{+}$. The electron remaining in the molecular orbit of a donor and the electron for some time residing in the orbit of an acceptor form some kind of a homopolar bond between the molecules. Precisely this additional bond is responsible for the energy gain when the complex is formed.

One of the ways to form CTC consists in mixing the solutions of both components in the required proportion and subsequent evaporation of the solvent or precipitation of the formed complex by cooling of the solution. If the components are insoluble, one can prepare their suspensions and mix them.

The transport processes of charge carriers vary depending on the CTC type, molecular packing, and extent of overlapping between the adjacent chains. Their mechanisms may be both electronic and ionic.

Pyrolysis – decomposition of an organic compound at high temperature (several hundred degrees). On the one hand, pyrolysis enables one to identify the structure of high-molecular compounds and, on the other hand, provides for synthesis of novel materials with the desired properties, specifically pyrolized polymers (PP).

Polymers are characterized by minor thermal stability. Depending on the polymer nature, thermal destruction is observed even at 250-500 °C. In the presence of oxygen, thermolysis of polymers, destructive in character, is the case at lower temperatures.

Polymer destruction due to the effect of heating may result in the formation of low-molecular products (monomers and products of major destruction) and in the origination of transverse bonds between the molecules. These processes, especial the latter, lead to cross-linking of a polymer with the formation in a polymeric chain of double or, when transformations are more profound, of conjugated –C=C–C=C– bonds. The presence of conjugate systems is a prerequisite for increase of polymer conductivity.

The process of polymer pyrolysis with the formation of a solid residue may be considered as the formation, at first, of the precarbon and, finally, carbon structures having different degrees of metamorphism which contribute to the conductive properties of a material, similar to the filled polymers.

As a result of the formation of the graphite-like networks in PP, some isolated sites have a high degree of conjugation and are divided by nonconjugate interlayers. Within every of such sites (clusters) the π-electrons strongly interacting with each other to form a common electronic system. As this takes place, electron interactions between the conjugation sites are hindered. Heating at temperatures above 500 ºC leads to a gradual increase in size of the conjugated sites and to higher structural ordering of the carbonized phase that step-by-step evolves to turbostratic carbon. Note that the formation of polyaromatic networks makes it possible to produce a long sequence of conjugated double bonds or a continuous system of π-electrons with a certain concentration of the macromolecules distributed over the rigidly bound structure. As temperature is growing, a number of the ensembles of microparticles (fractal clusters) and their sizes are increased. In some cases carbon atoms form a hexagonal lattice with the unit cell parameters $c$=0.7474 nm and $a$=0.255 nm. The indicated structural transformations are associated with cardinal changes of conductivity (percolation transition if the charge transport between the discrete π-bound carbon clusters is possible).

The effect of ion implantation on the structure and composition of a polymer is a complex of physical, chemical processes and phenomena including the interaction between the impinging ions and the polyatomic target as well as radiative and pyrolytic processes of the polymeric matrix transformations. Characteristics of these phenomena depend on the energy transferred to the polymer during implantation, on the composition and structure of the polymer itself, on its interactions with the ambient before and after ion implantation, etc. The existence of two principally different mechanisms of the energy transfer from the implanted ions to the polymer (that is realized by means of electron or nuclear deceleration) is of crucial importance for the ion-induced processes proceeding in polymeric materials.

Nuclear collisions are associated with transport of a considerable energy and momentum; actually, for collision of an ion and an atom approximately equal in mass all the energy is transferred to the atom. The interacting atom is knocked out of the equilibrium position in a polymeric chain, migrates within the material volume, and interacts with other atoms until the whole energy acquired due to the first collision will be dissipated. As a result of this process,

during the period $10^{-8}$-$10^{-7}$ s, in some region of the material collision cascades occur with the sizes on the order of 10-100 nm. Such a region is characterized by a high density of the defects and is located at the depth corresponding to the finite range of the ion path. Approximately in 100 ps after the ion trip, the formed defects recombine partially or relax to be localized within the limits of the ion track. Nuclear collisions and the associated cascade phenomena may lead to desorption or sputtering of atoms or, possibly, of molecular fragments from the surface of a sample. This phenomenon becomes significant at very high ($>5\times10^{16}$cm$^{-2}$) implantation doses for polymers and with the use of heavy ions.

At high implantation energies, the energy is transferred predominantly to electrons of the target atoms with excitation of the latter; ionization and secondary processes (e.g. breaking of chemical bonds) take place. Electron interactions are most essential at the initial part of the ion path range. The energy transfer from the point of interaction results in excitation of the electrons in adjacent parts of the polymeric chain. The effective length of the energy transfer is dependent on the polymeric molecular structure; existence of the delocalized π-electrons, quantity of conjugate systems and conjugated bonds. The secondary electrons initiate further excitation and ionization of a material, and also the development of cascades with the characteristic size coming to a few tens or hundreds of nm. As a density of the energy released within the cascade volume is lower than in the track nucleus, the processes proceeding in the cascade are considered to be pure radiative, thermal processes are disregarded.

Because of the processes of nuclear and electron deceleration, in a polymer, along the ion path range, a region of the structural disturbances is formed. A core of the region, characterized by a high defect density and a lower density of the material as compared to the unmodified polymer, is known as a latent track (Fig. 5.2). The track radius is depending on the implantation conditions; usually it is from a few nm for light ions to several tens of nm – for heavier ions. The track length is growing with the implantation energy. When a certain dose is attained, the surface of a sample is completely filled with tracks and further implantation is realized into the polymer subjected to the primary modification. It is expedient to define two implantation regimes: single-track implantation, when the tracks formed in a polymeric matrix are not interacting with each other, and implantation with overlapping tracks.

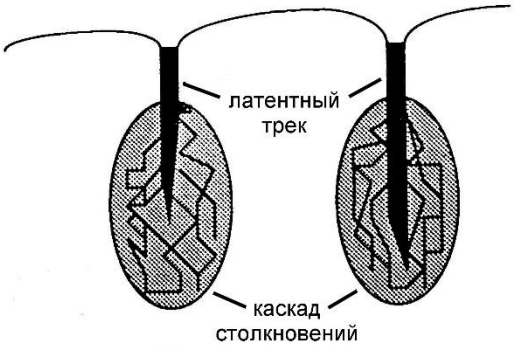



Figure 5.2 – Formation of latent tracks and collision cascades in the case of ion deceleration in a polymer : latent track, collision cascade

In description of the thermal component in the polymer modification process at conditions of ion implantation we should take into account high local heating close to the ion track rather than heating in the bulk of the irradiated target, the consequences of which are similar to the ordinary vacuum pyrolysis. To exclude pyrolytic degradation of a polymer in the process of ion irradiation, it is required to maintain a low level of the ion current and to cool the substrate. With nuclear deceleration of ions, the atoms of a target may be activated to the state with high-amplitude oscillations about the equilibrium position. The kinetic energy formed in this way is transferred to the lattice, whereas an increase in the local energy may be interpreted as a heat transfer with

growing temperature and energy dissipation. Based on the calculations, it has been found that a value of the local temperature is about ~$10^4$ K for the ions implanted into different polymers at the energy ranging 100 keV – 1 MeV. Lowering of temperature to that of the target set for implantation happens in time of $10^{-9}$–$10^{-10}$ s. This estimate is very approximate but it gives some knowledge about a character of thermal effects involved in ion bombardment of polymers.

Implantation of polymers leads to their degradation due to thermal processes and modification of chemical bonds. Considering chemical transformations associated with ion implantation, we should take account of the type of interactions (electronic or nuclear) for which the energy transfer from the implanted ions to the polymeric matrix is realized. Nuclear collisions are accompanied by direct breaking of the chemical bonds, whereas electronic interactions lead to excitation of the macromolecular fragments. Considering weak bonds in polymeric macromolecules and the possibility for long-distance migration of the excited state during its lifetime (~$10^{-12}$ s), we can conclude that electron deceleration also may be responsible for breaking of the chemical bonds. Obviously, in the case of a nuclear nature of the energy transfer, breaking of chemical bonds is mass, predominantly statistical in character. But destruction of extent electron systems with a high level of delocalization in conditions of electron deceleration is selective in character.

When in the process of polymer implantation a drastic increase of temperature takes place, considering low thermal conductivities and degradation of the chemical bonds, we should observe the phenomenon of radiolysis accompanied by egress of different easily volatile gas products from the radiation-damaged layer depending on the polymer type: $C_2H_2$, $CH_4$, $H_2$, CO, $CO_2$ and so on. In the process release of $H_2$ leads to dehydration with the formation of free radicals, CO – to amidization, $CO_2$ – to dissociation of carboxyl groups.

It is important that the processes proceeding in a polymeric matrix on ion implantation represent a single process of raiothermolysis but not isolated thermal (despite high temperature peaks in the track) or pure radiative processes.

Degradation and splitting of low-molecular hydrocarbons on ion implantation of polymers result in gradual enrichment of the implanted layer with carbon, i.e. carbonization that is one of the secondary effects in the process of implantation. When heavy ions are implanted, ion-induced carbonization of polymers is finished at lower doses compared to implantation of light ions.

In this way the radiative-pyrolytic transformations of the polymeric matrix proceeding in the track of impinging ions lead to the formation of carbon-enriched clusters. Sizes of these clusters are growing with the implantation dose. The clusters begin to overlap at the threshold implantation dose. Overlapping may be associated with the formation of «carbon coagulants» and, at high implantation doses, - with the formation of a continuous modified layer enriched in carbon and characterized by numerous conjugated bonds.

One of the important aftereffects of the ion-induced polymer surface modification is a significant increase of its microhardness owing to strengthening of the implanted layer by carbon clusters formed on ion bombardment. Hardness may be as high as 7—8 Mohs units approaching the hardness of diamond-like films. As a rule, improvement of harness of polymeric surfaces due to ion implantation results in a considerable increase of their abrasion resistance. It should be noted that resistance is an extremal function of the implantation dose as fragility of the implanted layer is growing with an increase in its hardness.

The ion-implanted polymer is characterized by the presence of micropores in the near-surface layer and by deep modification of the surface, cardinally influencing its chemical activity and adsorptive properties. To illustrate, effective implantation of molecular iodine, of some metals and salts into the modified polymers has been demonstrated. Adsorptive capacity is determined by the energy loss density on implantation; free radicals generated in the process of implantation and retained for a long time contribute considerably to adhesive properties of the surface of ion-modified polymers.

The formation of surface pores, variations in the polar component of free surface energy, and destruction or formation of the specific adsorption centers due to ion bombardment may

contribute to adsorption of living cells on implanted polymers, offering the possibility to control biocompatibility of polymeric materials.

Implantation of optically transparent polymeric films leads to induced coloring that is observed even at the dose $1\times10^{13}$ $cm^{-2}$. With the increased dose, color of a sample is varying from pale yellow to dark brown. At implantation doses $>5\times10^{16}$ $cm^{-2}$ the polymeric surface acquires the characteristic metallic luster.

The development of conductive properties is one of the most important consequences resultant from modification of the surface layer of a polymer on ion implantation. Resistivity of the typical insulators such as polyethylene, polyamide-6, polyethylene terephthalate, polyimide and the like is on the order of $10^{12}$—$10^{18}$ Ω·cm. In the process of ion implantation the surface layer of a polymer is enriched with carbon. Excessive carbon is clustering with the formation of polynucleated aromatic structures. Note that, because a hybridization character of carbon is changing from $sp^3$ to $sp^2$, each atom of the carbon cluster is associated with one unpaired electron having the ability to participate in π-bonding. The carbon-enriched «clusters» formed within the limits of ion tracks, their growth and overlapping at a certain dose lead to the carrier transfer within the implanted layer. Depending on the material and implantation conditions, conductivity of the layer may be as high as 1-10 S/cm.

Electric conduction of the implanted layer is very sensitive to the structural features of the carbonized phase: size of carbon clusters, existence of heteroatoms, character of the cluster aggregation, etc. Because of this, resistivities of different polymers implanted at identical conditions can differ by several orders of magnitude. Conductivity of polymers, apart from the dose, depends on other implantation conditions: ion energy, ionic current, thickness and temperature of the polymeric target. The literature data concerning variations in conductivity on implantation for one and the same polymer differ considerably. At the same time, the general pattern is an increase in conductivity of polymers with the increased ion energy and relative density of electron deceleration mechanism. As origination of the conductive phase in the implanted layer is due to the formation of carbon clusters, the electric conduction phenomenon is threshold in character, whereas the «insulator – conductor» transition itself can be considered percolational.

Of particular interest for polymeric ion implantation is introduction of ions of the iron-group transition metals enabling the formation of both conductive and magnetically-ordered layers in polymers.

## 5.4. Applications of polymers in electronics and eletro engineering

Polymeric materials have been always associated with revolutionary advances in science and technology. Polymers have been successfully used instead of the standard insulating materials for a long time. A great variety of polymeric compositions and the possibility for the production of materials with a wide range of physical and chemical properties on their basis enables their application as structural materials in microelectronics and radio engineering. But an interest to polymeric materials is ever growing. Currently the researchers are still engaged in search for modified polymers with the conductive properties; compositions exhibiting nonlinear characteristics under the field effects are designed.

Based on the filled polymers, a great series of variable resistors, where conductive polymer is applied to dielectric in the form of a film, is manufactured: one contact of a high-conductivity material is formed at one of the film edges; the second contact is movable over the film surface and allows for adjustment of resistance. Besides, fixed resistors used for operation in direct, alternating, and pulse current circuits and high-voltage resistors are manufactured. Resistance heaters designed on the basis of the plates with conductive polymeric coatings are characterized by the dissipation power up to 1 kW. Filled conductive polymers are used in resistive elements applied as temperature gages, i.e. they are sensitive to irregular excess of temperature. Such elements are essential for specific processes, e.g. pasteurization of liquids in food industry. Filled

polymers are widely used in different devices for static charge removal in electro engineering: plates, films, rollers, V-belts, conveyor belts, etc.

We should isolate a group of composite semiconductive polymers filled with fullerenes. Owing to the use of these materials with Au or In/SnO layers, the heterojunctions exhibiting high photosensitivity (below 100 $pA/cm^2$ for dark current) have been realized, offering much promise for the development of new-type photodetectors and photoelectric converters.

One of the currently central trends in usage of conductive polymers is photopolymerization applied in photolithography. We could hardly find any serial production of integrated microcircuits on a silicon platform without the use of mask photoresists.

Polymers with a system of conjugated bonds are applicable for the creation of p-n-junction electronic devices. However, more common are hybrid elements such as MOS (metal-oxide-semiconductor) diodes and field MOS transistors, light-emitting diodes, etc. They are at the base of polymeric resistors, thermal resistors, temperature sensors, varistors, and the like. Polymers with a system of conjugated bonds offer much promise for the production of miniature condensers and inductive coils. The possibility to use semiconductive polymers (owing to their photoconductivity) for photoelectric generators or solar cells has been demonstrated. Polymers with a system of conjugated bonds are very promising for manufacturing of electrochemical batteries.

Charge-transfer complexes have found an extensive use in the development of electroluminescent devices with a molecularly-doped polymeric layer.

Among the applications of conductive pyrolized and implanted polymers, there is manufacturing of inexpensive film resistors on their basis. To illustrate, resistance of polyimide due to implantation varies within the interval $10^{-2}$–$10^{16}$ Ω·cm. As variations in resistance are conditioned by the structural changes of a polymer, the resultant resistive elements are highly stable.

At high implantation doses ($>5\times10^{16}$ $cm^{-2}$) the electrochemical behavior of ion-implanted polymers is fairly similar to that of carbon electrodes. The same effect is characteristic for pyrolized polymers as well. This in turn makes it possible to use the electroplating methods for metallization at the sites of the planar conductive structures formed on the surface of modified polymers. Besides, there are prospects for integration of electronic devices produced using implantation or pyrolysis and of electrochemical sensor elements. In the case of ion implantation the carbonized polymeric films with numerous nanopores piercing them offer much promise as a material for current sources operating by means of ion accumulation within the electrode volume.

As regards pyrolized polymers, variations in their voltage-current characteristics depending on the thermal treatment conditions have been revealed. As pyrolysis temperature is growing and the treatment period is increased, an increase in the threshold voltage makes it possible to use such a material as a switch. Under the effect of alternating current the material reveals a certain memory (regions of high, low or negative resistance on the voltage-current characteristics), i.e. reversibility and high reproducibility of the parameters.

Using special implantation conditions, one can form in a polymer a hidden carbonized conductive layer. Spatial characteristics of the formed structure enable the control of resistance of this hidden layer by application of an external electric field, offering possibilities for designing of the transistor-type electronic switches. A depth of the carbonized layer and its thickness is dependent on the implantation energy and on the type of implanted ions.

It is known that, for the majority of implanted polymers, resistance is growing exponentially with lowering of temperature in line with a negative temperature factor of electric resistance. This effect may be used to create temperature sensors on the basis of implanted polymers. This possibility is due to their high impedance and significant resistance to radiation damage – conductivity of implanted polymers used to this end is invariable in the process of the subsequent radiative action. Besides, owing to the formation of sensor structures on dielectric surfaces, one can avoid short circuits and designing of complex-configuration planar temperature sensors is possible. One of the trends in polymer electronics is printed electronics enabling one to

produce thin and flexible devices, e.g. with the use of inexpensive roll-to-roll process. Electronic devices and their components can be printed by an ordinary printer when, instead of ink, one uses various polymeric materials. Because of this, there is a possibility for the production of a series of the components for microcircuits (from transistors, storage blocks, and batteries to processors, sensors and displays) by printing light-emitting polymers on plastic, metallic or paper films. Among the products, we can name the following: «intelligent» packing, OLED (organic light-emitting diode) illumination, inexpensive electronic markers for RFID (radio frequency identification), rolled displays, flexible solar cells, disposable diagnostic devices, toys, flexible touch screens, printed batteries, transistors, and memory devices.

Organic electronics has undeniable advantages over inorganic electronics that approaches to its physical limits. Among the advantages, we should name the following: low cost, economic and ecologically safe production; such universal properties as low weight, flexibility, transparency, long shelf life, reliability, possibility of printing various electronic components with the use of the same facilities. Moreover, there is no need to use expensive equipment for vacuum processes that lowers a cost of the equipment by a factor of ten as compared to the microelectronic production. Any idea of circuit engineering is realized rapidly and at low expense due to the use of computers and subsequent printing on the substrate. The cost of such production compared with manufacturing of silicon electronics is lower by the order of magnitude.

**References**


The Nobel Prize in Chemistry, 2000: Conductive polymers //www.nobelprize.org.

Shirakawa H. The discovery of polyacetylene films. The dawning of a new era of conducting polymers.// Synthetic Metals. -2002. -V.125, N.l. -Pp.3-10.

MacDiarmid A.G. Synthetic metals: a novel role for organic polymers.// Synthetic Metals. 2002. - V.125, N.l. - Pp.11-22.

Heeger A.J. Semiconducting and metallic polymers: the fourth generation of polymeric materials.// Synthetic Metals. 2002. - V.125, N.l. - Pp.23-42.

Brabec C.J., Sariciftci N.S. Hummelen J.C. Plastic solar cells // Advanced Functional Materials 2001. - V.ll, N.l. - Pp.15-26.

Оджаев В.Б., Попок В.Н., Азарко И.И.. Физика электропроводящих полимеров. – Мн.: Белгосуниверситет, 2000. – 82 с.

Correlation of Electronic and Magnetic Properties of Thin Polymer Layers with Cobalt Nanoparticles / A.Kharchenko, M. Lukashevich, V.Popok, R.Khaibullin, V.Valeev, V.Bazarov, O.Petracic, A.Wieck and V.Odzhaev //Particle & Particle Systems Characterization. – 2013. – Vol. 30 – Issue 2. – P. 180–184.

Modification of the subsurface layers of polyimide films upon boron-ion implantation Kharchenko A.A., Brinkevich D.I., Brinkevich S.D., Lukashevich M.G., Odzhaev V.B. //Journal of Surface Investigation. X-ray, synchrotron and neutron techniques// - 2015, - №1, - P.87-91.

Popok V., Karpovich I., Odzhaev V., Sviridov D. Structure Evolution of Implanted Polymers: Buried Conductive Layers Formation // Nucl. Instr. and Meth. B. - 1999. - V.148. - P. 1106-1110.

M.G. Lukashevich, V.N. Popok, V.S. Volobuev, A.A. Melnikov, R.I.Khaibullin, V.V. Bazarov, A. Wieck and V.B. Odzhaev. Magnetoresistive effect in PET films with iron nanoparticles synthesised by ion implantation. //The Open Applied Physics Journal. 2010, - Vol 3,- P.1-5.


## 6. Electrical conductive composite materials

### 6.1. Statement of the problem

To specify the subsequent presentation of the material concerning the electrical properties of composite materials, we introduce several definitions.

*The phase* is a part of the medium (system) confined by the interface and characterized in the absence of external forces by the constant physical and chemical properties at all points.

*Homogeneous* (or uniform) *medium* (system) is a system, where all constituent parts form a single phase (there are no interfaces).

*Heterogeneous* (or nonuniform) *medium* (system) is a system whose constituent parts form several phases separated from one another by interfaces. For example, a mixture of some different crystalline substances, a saturated aqueous solution of salt with its sediments, a mixture of liquids with limited solubility in each other, eutectic alloys, etc.

*Binary* heterogeneous medium (system) is a heterogeneous system consisting of two phases.

*Cluster* is a collection of identical objects connected to each other in any, but in a predetermined way.

In the case of heterogeneous media, the terms disperse phase and dispersion medium are often used. The *disperse phase* is understood as a "fragmented" or "crushed" component, and under the *dispersion medium*, the medium in which the dispersed phase is located. It is assumed that the dispersion medium forms a continuous volume in the system (that is, from any randomly chosen point of the dispersion medium, it is possible to leave the heterogeneous system only within the dispersion medium), but the disperse phase does not. It should be borne in mind that in for 3D (bulk) systems both the dispersed phase and the dispersion medium can form continuous volumes.

In the case of composite materials, the disperse phase is the *filler*, and the dispersion medium is the *matrix*.

*Macroscopically inhomogeneous media* (systems) are understood to mean such media, the characteristic size of the inhomogeneity of which is much larger than any characteristic macroscopic lengths (correlation parameters) that determine physical phenomena. For example, when describing the electrical conductivity, the size of the inhomogeneities should be much larger than the characteristic free path lengths of free charge carriers. The concept of the phase is attributed to physically small volumes so that within the limits of these volumes "macroscopic values" of physical quantities make sense.

In the case of a homogeneous (homogeneous) isotropic medium, the vectors of the electric current density $\boldsymbol{j}$ and the electric field intensity $\boldsymbol{E}$ are related to each other by the expression

$$\boldsymbol{J} = \sigma\boldsymbol{E}. \tag{6.1}$$

In irregular (disordered) heterogeneous systems, the $\boldsymbol{j}$ and $\boldsymbol{E}$ vectors, as well as the electrical conductivity and permittivity tensors, are random functions of the radius vector $\boldsymbol{r}$. In this case, the medium can be characterized by local execution of Ohm law, which connects the electric current density $\boldsymbol{j(r)}$ with the electric field intensity $\boldsymbol{E(r)}$ at each point of the medium with coordinates $\boldsymbol{r}$:

$$\boldsymbol{J(r)} = \sigma\boldsymbol{(r)E(r)}. \tag{6.2}$$

The inhomogeneity of the medium is then given by the dependence of the conductivity on the coordinate $\boldsymbol{r}$. A macroscopic inhomogeneity can be both continuous, if $\boldsymbol{r}$ is a continuous function, and discrete. In the latter case, we speak about 2-, 3-phase and so on media (systems).

Owing to the small size of the inhomogeneities and their random distribution over the volume $V$ of the of the heterogeneous sample, it is possible to identify in it some volumes of $l^3$ (so-called *representative volumes*) whose properties are the same and correspond to the properties of the entire sample as a whole. In other words, for scales of larger than $l$, the heterogeneous system can be considered macroscopically homogeneous and characterized by a set of effective parameters, including electrical conductivity and dielectric permittivity.

We shall assume that the electrical conductivity of the phases constituting the heterogeneous system is a scalar quantity, and the choice of the coordinate system makes it possible in all cases to use instead of the $\boldsymbol{j}$ and $\boldsymbol{E}$ vectors their moduli.

The effective electrical conductivity, $\boldsymbol{\sigma}_{\text{eff}}$, is the main characteristics which describes the charge transfer in a randomly inhomogeneous heterogeneous system. It is defined as the coefficient connecting the averaged by $V$ values of the local electric current density $j = j(\mathbf{r})$ and the local field intensities $E = E(\mathbf{r})$:

$$<\boldsymbol{J}> = \boldsymbol{\sigma}<\boldsymbol{E}>, \tag{6.3}$$

where

$$< J > = \frac{1}{V} J(\boldsymbol{r}) dV, < E > = \frac{1}{V} E(\boldsymbol{r}) dV.$$

The effective electrical conductivity of heterogeneous systems is primarily determined by the electrical conductivity of the components of $i$ and their volume content: $p_i = V_i/V$, where $V_i$ is the volume of the $i$-th component; $V$ is the total volume of the heterogeneous system ($\sum_i p_i$=1). The binary heterogeneous system consisting of two phases is described by one parameter $p = p_1$, since $p_2 = 1 - p$. A two-phase heterogeneous system is considered highly non-homogeneous if the electrical conductivity $\sigma_2$ of the dispersion medium (matrix) is much smaller (or much larger) than the conductivity of the dispersed phase $\sigma_1$. As $p$ increases from 0 to 1, the effective electrical conductivity of the heterogeneous system increases from the minimum electrical conductivity value $\sigma_2$ (at $p = 0$) to values close to the maximum electrical conductivity $\sigma_1$ (for $p$ =1).

The change in $\sigma$ for mixtures of the dielectric-metal type in the region $0 < p < 1$ can be tens of orders. When the dispersed phase is randomly distributed in the dispersion medium (the filler in the matrix), the growth of the effective electrical conductivity of the heterogeneous system occurs very rapidly (unevenly). The greatest increase in $\sigma$ is observed in the vicinity of a certain (critical) concentration $p = p_c$, which is determined by the percolation theory *as the percolation threshold*. For $p = p_c$, a *continuous* ("infinite") *cluster* (*IC*) is formed from a conductive component (phase) in a heterogeneous medium, piercing the sample from the electrode to the electrode, wherever they are located. In this case, the heterogeneous medium actually undergoes a transition in which its electrical conductivity changes cardinally (by many orders of magnitude). Such a transition is commonly called the *percolation transition* dielectric-metal.

It is assumed that if in a heterogeneous system the characteristic size $d$ of the inhomogeneity (for example, the particles of the dispersed phase or the interlayer of the matrix between them) is much larger than the characteristic microscopic mean free paths of the charge carriers, then the Ohm law holds locally: $j(\mathbf{r}) = \sigma(\mathbf{r})E$, where $\sigma(\mathbf{r})$ is the value of electrical conductivity at a point with a radius vector $\mathbf{r}$ (local electrical conductivity). For a heterogeneous system consisting of non-interacting particles of a dispersed phase and a dispersion medium, $\sigma(\mathbf{r})$ randomly takes on the values: $\sigma_1$ (electrical conductivity of the dispersed phase) and $\sigma_2$ (electrical conductivity of the dispersion medium). In real heterogeneous systems, transition layers are always present between the phase particles of filler, which are characterized by their local conductivity values, and therefore $\sigma(\mathbf{r})$ can vary continuously over a wide range.

The effective electrical conductivity of a heterogeneous system does not coincide with the average electrical conductivity in the volume. The establishment of its connection with the values of specific electrical conductivities and volume concentrations of phases is one of the important problems in the physics of heterogeneous systems. The analytical solution of the problem is carried out within the framework of the theories of effective medium and percolation (percolation).

The most common approach to the study of transport processes (including charge transfer) in disordered and random inhomogeneous heterogeneous systems based on the use of *percolation theory*. It describes the phenomena and effects caused by the connectivity of a very large (macroscopic) number of objects (elements that make up clusters), provided that the connection of each object with its neighbors is random, but is specified in a certain way. In other words, the percolation theory has to be related with linked objects in the disordered and heterogeneous media.

As noted in [1], percolation is a convenient model to describe a broad class of phenomena, which are called critical. On the other hand, the task turned out to be very interesting from the point of view of pure mathematics, as well as physics, chemistry and engineering. For example, in terms of mathematics, the percolation theory should be attributed to the probability theory on graphs. From the point of view of physics, percolation is the geometric phase transition. From the point of view of engineering, the TP is intended to solve various applied problems.

Different tasks of percolation theory are combined by the same geometry of related objects near the percolation threshold. Universal large-scale geometry determines the universal dependencies of physical quantities on the structure of large clusters as a set of interconnected objects. Objects can be linked both directly with each other, and through other single-type objects.

## 6.2. Elements of the theory of percolation

Historically, percolation theory has been traced back to the works of Flory and Stockmayer, published in 1941-1943, which examined the process of gel formation during polymerization. However, usually the beginning of percolation theory is associated with the publication in 1957 of the work of Broadbent and Hammersley [2], concerning a new class of mathematical problems for studying the flow of a fluid through a random labyrinth, in which the authors introduced the term "percolation theory" and examined the process from a mathematical point of view. The term percolation was used to contrast diffusion. If in the case of diffusion we are talking about random displacements of particles in a homogeneous medium, in the case of percolation, we are talking about regular motion (for example, the flow of a fluid or an electric current) in a random environment.

Since percolation is performed by specialists from different fields of knowledge, two approaches are used for its description [1]. Mathematicians apply a formal approach, based on representations of probability theory and graph theory, rigorous proofs and formal notation. Applied specialists use a more visual (and simple) approach without rigorous evidence and with references to analogy, intuition and common sense. We believe that for the initial acquaintance with TP it is reasonable to use the second approach.

Percolation theory develops and investigats the simplest models of disordered systems. The groups of percolation theory models, that have common features in the methodology of their construction and are used in this formulation, are called the problems of percolation theory. When setting the percolation theory tasks, it is mandatory to determine: 1) the element from which the connected structures (clusters) are composed; 2) the way of the elements arrangement in space; 3) the way in which the connection between the elements is carried out.

Depending on the elements of the system (or clusters) in the percolation theory, the following tasks are distinguished: node tasks, problems of interconnections (bonds), sphere problems, continual tasks. In interconnections problems, the edges of the lattices are the elements of the system (very often the term "connection" is used instead of the term "link" in the sense that the edges of the lattice connect the nodes of lattices, although they actually operate with the edges of the lattices in the bond problem), whereas in continual tasks they are the lattice nodes. The fragments of the lattices for the problem of bonds and nodes are shown in Figs 6.1*a* and *b*. In the spheres problem, clusters are formed from spheres of certain radius *a* (see Fig. 6.1*c*), and in continual problems - from points of space. The most common problems of percolation theory are lattice problems: *the nodes problem* and *the bonds problem* [3-5]. Lattice models are primarily of interest from a theoretical point of view, since a number of rigorous statements and relations have been proved for them.

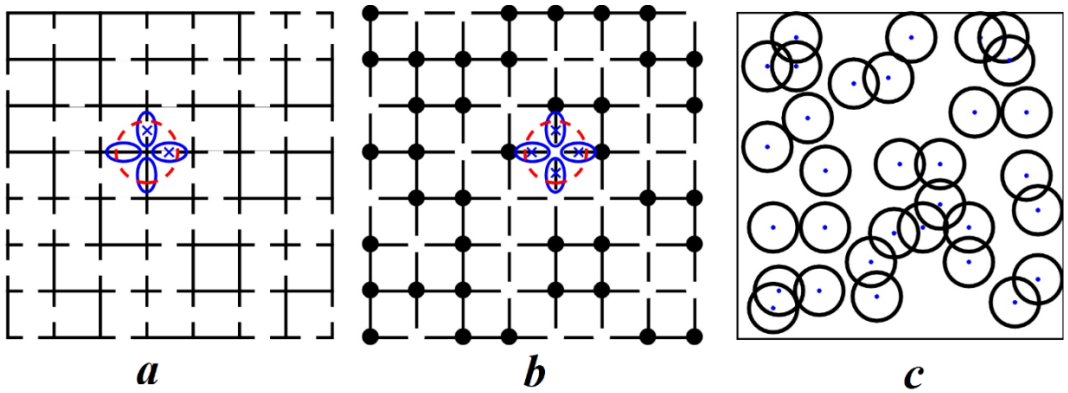



Figure 6.1. Fragments of lattices for the problems of bonds (a), nodes (b), and an example of statement of the problem of circles (c)

The method of arranging the elements of the system in space allocates lattice problems of percolation theory. and problems at random nodes. Lattice problems, as follows from the name, are formulated on the example of lattices having a translational symmetry, located in two-, three- and n-dimensional space.

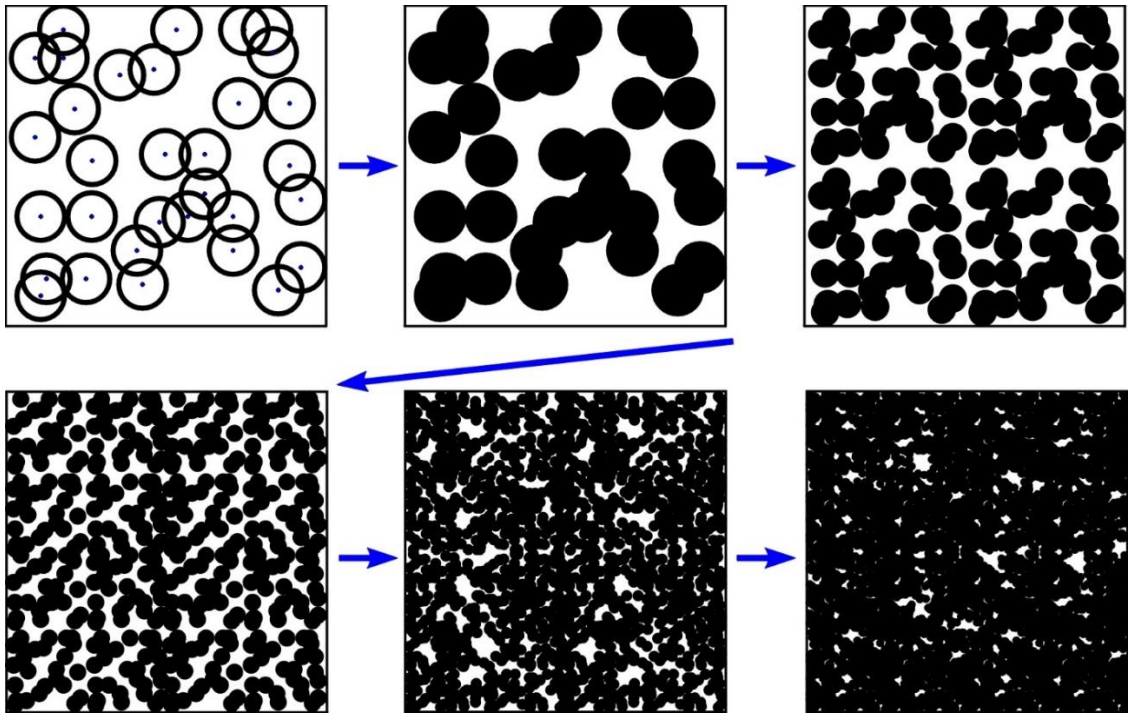


Figure 6.2. Illustration of the transition from the circle problem to the continual problem

The problems of circles (spheres) are formulated on the example of combining randomly located regions of space confined by circles (spheres). Thus, in lattice problems of percolation theory, the elements are arranged in an orderly way, in random-node problems - in an arbitrary way. Continual problems, as well as the sphere problems, are formulated for structureless regions of space and can be considered as the limiting case of the sphere problem (see Figure 6.2).

The order (law) of the location of the nearest elements in space and/or the nature (mode) of their interaction, which determine the relationship between the elements constituting the percolation system, is specified during the formation of the model. Interconnection for each particular model can be formulated differently, even if the elements of the system and the way they are arranged in space are coincided. For example, in some node problems, the elements of

not only the first coordination sphere, but also the subsequent ones are taken into account. This can be useful, for example, in the analysis of ferromagnet-paramagnet transitions.

One of the key concepts of percolation theory is the concept of a cluster. Therefore, we will once again concentrate on the definition: in the general case, a *cluster* is understood to mean a set of similar objects connected with each other, provided that the relationship between objects is uniquely determined. If the cluster permeates the entire volume of the investigated macroscopic system (sample), then it is considered an *"infinite cluster"* (*IC*). In Fig. 6.3 several methods are shown for determining the bonds between the objects in the circle problem. Clusters are indicated by dashed lines. Circles can be considered bound (forming a cluster) either when they are at least touched, or when they are stopped (in this case, one point of contact is no longer sufficient), or when the overlapping area of two circles covers the center.

As a rule, when specifying the bond (connection), it is considered that if two elements are associated with the third, but do not have a direct connection with each other, they are still connected (*transitivity rule*). For example, in Fig. 6.3, two clusters on a square lattice are separated, which are formed under the condition of defining of horizontal or vertical connection, but not diagonal. One of the clusters consists of two nodes, directly linked to each other. The second consists of the four nodes. Moreover, the extreme left and right nodes are "connected" with each other through bonds with other nodes, and all together they form a cluster.

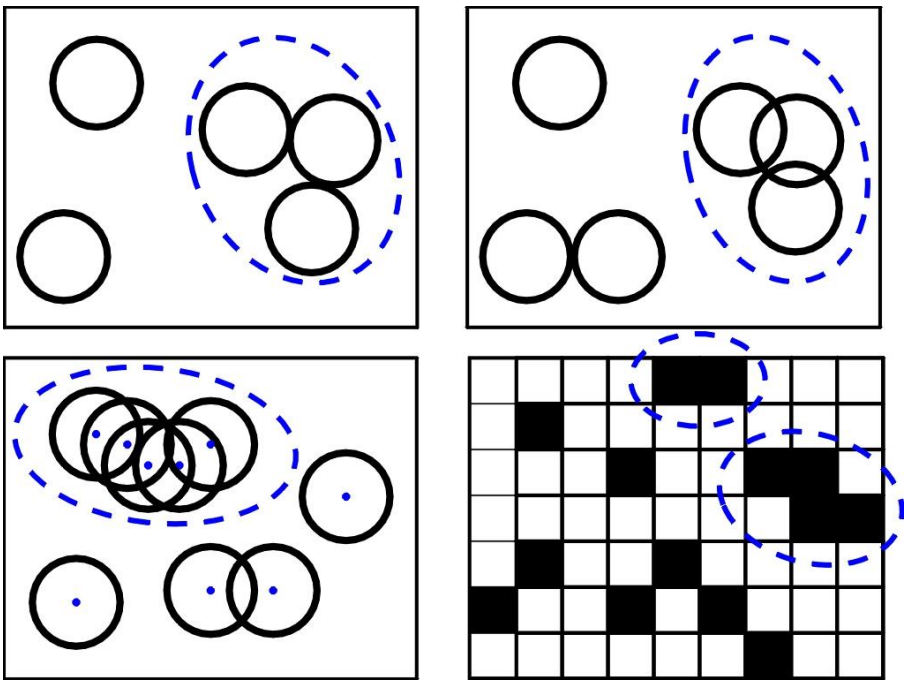
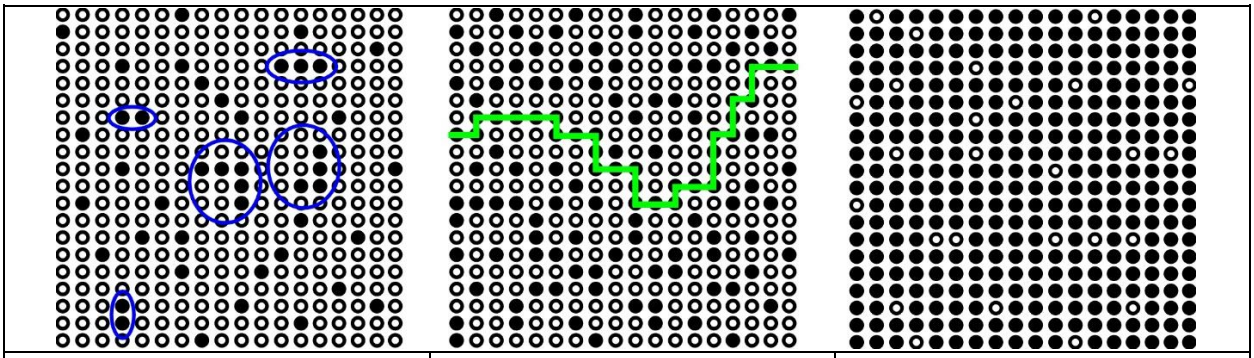


Figure 6.3. Example of assignment of bonds in a percolation system

The main issues that are considered in the framework of tasks of percolation theory are the average number of elements in the cluster; distribution of clusters according to the number of elements; the appearance of an infinite cluster and proportion of the elements included in it.

Consider the formulation of the nodes problem on a two-dimensional square lattice. The filled (occupied) node will be painted black, not filled - in white. Two filled nodes (black) will be considered connected with each other if they are located in the nearest positions of the lattice (vertically or horizontally, but not diagonally). For a squared two-dimensional lattice, this definition corresponds to the first coordination sphere. As indicated above, we assume that if two filled nodes are linked to the third, but do not have a direct connection with each other, they are still connected. For clarity, we can assume that we are talking about electrically conducting and dielectric balls.

Electrically conductive balls are located in the places of the occupied (black) nodes, whereas the unoccupied ones are dielectric. Accordingly, the electric current can flow only between the nearest conducting nodes (the balls "contact" with each other). The ratio of the number of occupied lattice sites to the total number of nodes is denoted by $p$ (concentration or otherwise - the relative fraction of occupied nodes, i.e., $p < 1$). One of the main questions that the percolation theory tries to answer is: for what fraction of $p_c$ of filled squares does a chain of black squares (cluster) appear that connects the top and bottom sides of the net?



*a* | *b* | *c*

Figure 6.4. "Evolution" of the percolation system as the percentage of occupied (black) nodes in the squared lattice increases: *a* - $p < p_c$; *b* - $p \approx p_c$; c - $p > p_c$,

In the case of low concentrations $p$, the conducting (black) nodes are either isolated, or (as shown in Fig. 6.4 *a*) form small clusters of several nodes. At low values of p, there is no linked route through the black nodes. In this case, the net consists of a mixture of black (conductive) and white (non-conductive) assemblies and is a dielectric, since there is not one path (continuous cluster) between the opposite ends of the net, over which an electric current can flow.

On the contrary, for large values of $p$, between the opposite ends of the lattice, there are many conducting routes along black nodes, and a mixture of black and white nodes is a conductor (Figure 6.4*c*).

To characterize the entire system as a whole, it is convenient to speak not about one particular initial node, but about the probability that an arbitrary initial node is associated with an infinite number of nodes. This allows us to introduce the so-called density function of an infinite cluster $P(p)$, which in an infinite lattice depends only on $p$, and does not depend on a particular implementation, i.e. from the location of the occupied nodes on the net.

Since for small $p$ values the quantity of $P(p) = 0$, i.e. the occupied nodes (black nodes) are separated from each other by free (white) nodes, and connection between any randomly chosen nodes is practically eliminated (see, Fig. 6.4*a*). For $p$ close to unity, the probability $P$ is also close to unity $P(p) = 1$ (see Fig. 6.4*c*). The most important is the concept of the above mentioned *percolation threshold* (critical probability) introduced by Broadband and Hammersley. In strict analysis, the percolation threshold $p_c$ is the upper limit of those values of $p$ for which $P(p) = 0$. It is generally accepted [3-5] that beyond the percolation threshold the function $P(p)$ increases continuously to unity by law

$$P(x) \sim (p - p_c)^{\beta} \qquad (6.4)$$

where β is the critical index of the "density" of an infinite cluster.

Thus, the behavior of $P(p)$ with increasing $p$ resembles the behavior of the order parameter for a second-order phase transition with decreasing temperature. For example, this is characteristic for the behavior of the spontaneous magnetization in a ferromagnet or the polarization in a ferroelectric. Keeping in mind the threshold behavior of *P(p)*, it is often said that there is no

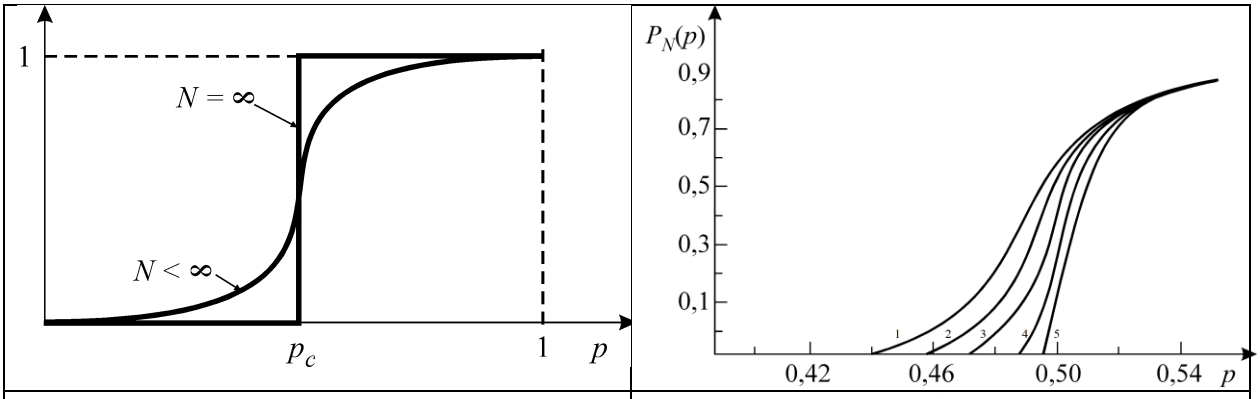
percolation at $p < p_c$, and at $p = p_c$ the flow occurs (see Fig. 6.4b, on which a cluster is separated which " penetrates through" the system from left to right).

The existence of a definite value of the percolation threshold $p_c$, as well as of the function $P(p)$ itself, is connected with the fact that an infinite lattice is considered for which all random realizations of the system with a given value of $p$ from the point of view of percolation are equivalent. If the lattice has finite dimensions, then we can talk about the probability of $P_N(p)$ that this node is connected with at least $N$ nodes, where $N$ is a large but finite number.

Thus, the probability $P(p)$ is obtained from $P_N(p)$ by the passage to the limit [3-5] for $N \to \infty$:

$$P(p) = \lim_{N\to\infty} P_N(p) \tag{6.5}$$

Obviously, for a net of finite dimensions, chains of connected nodes that penetrate the system through (Fig. 6.4b) can occur at different concentrations $p$. However, if the number of elements $N$ (the size of the net) tends to infinity, then the value of $p_c$ takes a definite value [3-5]. Just such a critical concentration is called the *percolation threshold* $p_c$. Fig. 6.5 shows the $P_N(p)$ functions obtained by numerical simulation in [6]. It can be seen that the function $P(p) = 0$ for $p < p_c$ and monotonically increases for $p > p_c$.




*a* *b*

Fig. 6.5. (*a*) Probability of percolation $P$ as a function of filled nodes ratio $p$. A smooth curve corresponds to a lattice of finite size, whereas a step-like curve conforms with an infinitely large lattice. (*b*) The $P_N(p)$ function for the bonds problem on a two-dimensional squared lattice: *1* - for $N = 667$; *2* - 1000; *3* - 2000; *4* - 6000; *5* - $N \to \infty$.

A similar regularity is observed for the bonds problem. In this problem, the elements constituting the percolation system are the edges of the lattices (often referred to simply as bonds). The bond (lattice edge) can be either whole or torn (break off), $p$ is the relative fraction of entire edges. Two whole edges are considered to be connected if they join at a node. Visually, illustrating the problem of the connections, there can be wetting of the lattice edges by the liquid (the liquid can be moved "to wet the lattice edges" along the lattice edges connected in the node). In an infinite system for a small value of $p$, the probability that, when a drop of liquid is placed at randomly selected edge leads to the wetting of an infinite number of edges, it is zero whereas for $p = p_c$ it is nonzero. Just this value $p_c$ is the percolation threshold.

With the help of the bond problem, the process of electric current flowing through the lattice is well described, when instead of non-broken bonds we have the same electrical resistances, and the dangling bonds correspond to an infinite electrical resistance (insulator). The effective electrical conductivity $\sigma$ of such a medium differs from zero only at $p > p_c$.

In the general case, in a given lattice the connection (bond) has more neighbors than one node. For example, in a cubic lattice, a bond is linked with six neighbouring connections, while every node has only four nearby nodes. Thus, large clusters of bonds can be formed with greater efficiency than large clusters of nodes, and in this lattice the percolation threshold for the bonds will have a lower value than the percolation threshold for the nodes (see, Table 6.1). Thus, the value of $p_c$ is individual for the tasks of bonds and nodes. This also depends on the type of lattice and the dimensionality of space (the definition of percolation problems of nodes and bonds for three-dimensional space can easily be generalized to lattices of arbitrary dimension $d$).

In Table 6.1 the values of the percolation threshold for nets in the problems of nodes and bonds are presented. In some cases, the exact value of the percolation threshold can be analytically determined using the renormalization group method. If this is not possible, then numerical calculations, for example, using the Monte Carlo method are performed.

Table 6.1 - Percolation thresholds for nodes and bonds problems for various nets [3-5]

| Net type | Conductivity by: | |
|---|---|---|
| | nodes | bonds |
| Triangular | 0.5 | $2\sin(\pi/18)$ |
| Quadratic | 0.5927460 | 0.5 |
| Hexagonal | 0.6962 | $1-2\sin(\pi/18)$ |
| Simple cubic | 0.31161 | 0.248814 |
| Face-centered | 0.198 | 0.119 |
| Body-centered | 0.245 | 0.1803 |

We note once again that in the language of clusters the dynamics of the occurrence of a flow in a mixture of black (conducting) and white (non-conducting) nodes with increasing $p$ looks as follows. For small $p$ all clusters are small. However, as we approach the percolation threshold, individual small clusters join and the average cluster size increases. At the point $p = p_c$, the *IC* originates for the first time from conducting black squares. It resembles a random net and penetrates through the entire space. In the "pores" of the infinite cluster, finite and isolated small clusters are located. If we further increase $p$, then the density of the infinite cluster increases, because more and more nodes become part of an infinite cluster. Accordingly, the average size of finite clusters that do not belong to the *IC* is reduced. For $p = 1$, obviously, all the nodes belong to the *IC*.

The concept of the percolation threshold can also be introduced in another way. Consider a lattice of finite size, for example, a cubic lattice, on each side of which there are $N$ nodes. First let all the lattice sites be white. We will then recolour randomly chosen nodes in black, gradually increasing the proportion of repainted nodes $p$. For a certain value of $p = p_c$, a path will appear for the first time along the connected black nodes from the left side of the cube to the right. We call $p_{cN}$ the percolation threshold for a finite lattice. If you repeat this procedure, the black nodes will be in other positions and the percolation threshold will also be, generally speaking, different. Thus, $p_{cN}$ is a random variable. Many times repeating the described procedure, we can calculate the average value of $\langle p_{cN} \rangle$. In Fig. 6.6 distributions of $p_c$, calculated for systems with differing $N$, are schematically depicted. The percolation threshold corresponding to an infinite system is obviously determined by the expression

$$p_c = \lim_{N\to\infty} \langle p_{cN} \rangle \cdot \tag{6.6}$$

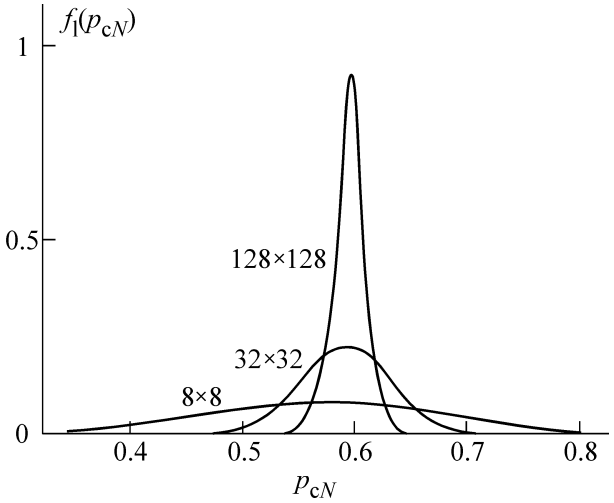



Figure 6.6. The function of distribution for the percolation threshold $p_{cN}$, as a random variable, calculated for the nodes problem on a square lattice by the Monte Carlo method. The dimensions of the nets are indicated in the figure [3].

The closest to the problem of the electrical conductivity of heterogeneous metal-insulator systems is the continual percolation problem, where the positions of the two components of a random mixture are not confined to discrete nodes of a regular lattice. An example is a sheet of conductive material with round, randomly distributed holes. The parameter determining the electrical conductivity of such system is considered to be the part $p = \frac{S_e}{S}$ of the remaining conductive material, where $S$ is the area of the initial sheet, and $S_e$ is the area of its remaining part after the perforation.

This example can be easily generalized to a three-dimensional space where spherical voids are randomly distributed in a cube. Because of its external similarity with Swiss cheese, this model is called the "Swiss cheese" model. Similar models, where the size of spheres can also vary, are used to describe porous materials. The mathematical formulation [3-5] of the continual problem requires the introduction of a random function $V(r)$, where $r$ is the radius of the random point vector. Proceeding from the definition of the distribution function $f(V)$, where $V$ is the probability of finding the value of the function $V(r)$ inside a certain interval $V+dV$ at a randomly chosen point $r$. The function $f(V)$ is characterized by a Gaussian distribution with dispersion $\delta$. For clarity, it is assumed that the part of the space where $V(r) < V'$ is covered with white paint, and the rest - black. The condition of percolation in this case is the critical value $V' = V_c$ (it is sometimes called the percolation potential) at which a white cluster appears in the system, penetrating through the entire system. The critical fraction of the white space $x_c$ is analogous to the percolation through filled nodes in the lattice problem of percolation theory and is expressed in terms of the distribution function

$$x_c = \int_{-\infty}^{V_c} f(V)dV \,. \tag{6.7}$$

On the plane percolation along the white and black can not be simultaneously, and therefore proceeding from the symmetry of the problem, $p_c = 0.5$. In the 3D case, there is no such restriction, and the percolation threshold has $p_c \approx 0.16$ - a much smaller value.

Let us consider the question of what a percolation cluster is and what models exist for its description. The percolation cluster is a fractal formation, in which some substructures can be separate out. *The skeleton of a cluster* is the association of all shortest paths from a given node to nodes at a given distance. *Dead ends* are parts of a cluster connected to the core by a single node (bond). They form a large part of the cluster, but they do not participate in the conductivity. *Red links* are single bonds, at the destruction of which the percolation cluster ceases to conduct current. *Loops* belonging to the *IC* duplicate the shortest routes between the nodes of the *IC* skeleton. *The shell* or *external perimeter* consists of those cluster nodes that come into contact with empty nodes and are connected to infinity by means of empty nodes. *The full perimeter* also includes voids within the cluster. Figure 6.7 shows schematically the structure of an *IC* fragment. Above the percolation threshold, the mass of a cluster (or the number of its nodes) inside a sphere of radius $r$ depends on $r$ as $M(r) \propto r^d$ , and below the percolation threshold as .

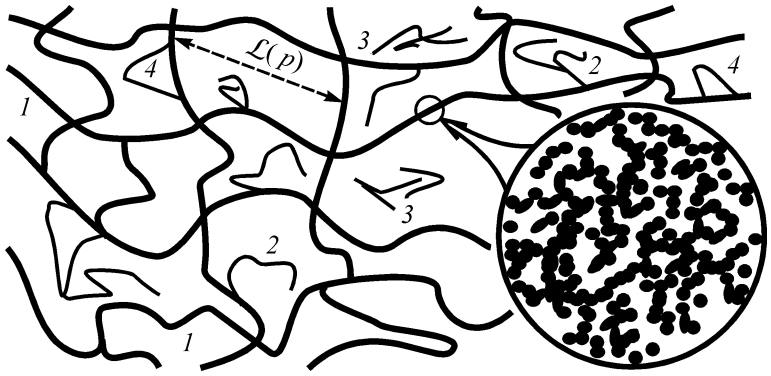



Figure 6.7. Schematic structure of an infinite cluster: 1 - skeleton, 2 - dead ends, 3 - isolated clusters, 4 - loops shunted by sections of an infinite cluster and therefore practically not involved in charge transfer [7]

An important characteristic of percolation theory, which describes percolation systems and an infinite cluster, is the *correlation length*, which also depends in a power manner on $p$ [3-5]:

$$L(p) \approx \left| p - p_c \right|^{-\nu} \tag{6.8}$$

where $\nu$ is the critical index of the correlation length.

Below the percolation threshold ($p < p_c$), the correlation length $L$ is of the same order of magnitude as the average cluster size. Since the number of nodes of a critical cluster approaches infinity as it approaches the percolation threshold, the correlation radius $L$ should also increase.

For $p > p_c$, the small probability of occurrence of finite clusters with a size larger than $L(p)$ means simultaneously a low probability of the pores in an infinite cluster of this size. Therefore, for $p > p_c$, the correlation radius can be interpreted as the characteristic pore size for the net of an infinite cluster. We can say that the correlation radius gives a presentation of the average distance between the nodes of net for an infinite cluster.

At distances less than the correlation length, an infinite cluster has a fractal structure, and at distances greater than $L$ there is a homogeneous structure consisting of cells with the characteristic size $L$. There is strict proof that there is only one infinite cluster above the percolation threshold. The transition from a fractal to a homogeneous state is commonly called a *crossover*. Figure 6.8 illustrates the characteristic behavior of the percolation system before and after the percolation threshold [3-5].

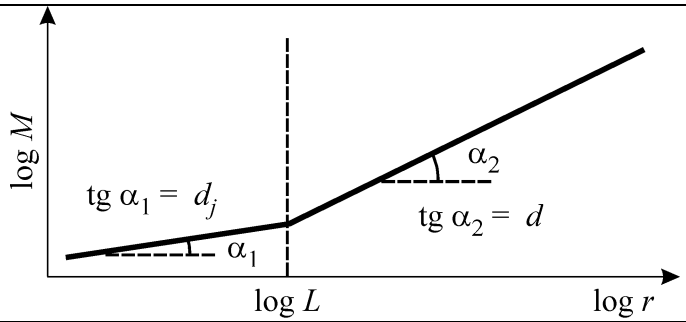



Fig. 6.8. The "cluster mass" dependence (in fact, the number of its nodes) inside a sphere of radius *r* from the radius value. Schematic representation of the crossover. Changes in the slope at the "transition point" $r = L$. The figure also illustrates the method for determining the critical indices of the "cluster mass" $d$ and $d_j$.

At the moment, there are several models describing the structure of the percolation cluster. *The Skal-Shklovsky-de Gennes* model [3] describes the structure of a percolation cluster by neglecting the dead ends. The model was proposed to describe such effects as the Hall effect, thermoEMF, etc. In this model, the network is assumed to be contained from single-stranded wire-like conductors, i.e. the most connections are not duplicated by loops that are not effective in terms of electrical conductivity. Here clusters of the conducting phase particles are characterized by a correlation radius (or correlation length) $L \propto (p - p_c)^{-\varphi}$ is the average distance between particles belonging to the same cluster. These individual clusters are connected in nodes located at a distance of the so-called length of the macrobond $L^* \propto (p - p_c)^{-\nu}$. To calculate the electrical conductivity of a heterogeneous system in this model, the disordered skeleton network of the *IC* is replaced by an ideal cubic lattice with a period equal to the correlation radius. The model has one important advantage: it is an exact measure of the density of the conducting routes in a typical section of the sample.

The main drawback of the Skal-Shklovsky-de Gennes model is the contradiction that arises in the 2D case, when the distance between the nodes of the "net" of an infinite cluster changes faster than the length of single-stranded channels along which charge transfer takes place. To eliminate this contradiction, models were developed that took into account more complex structural elements, for example, the drop model of a randomly inhomogeneous heterogeneous system (Fig. 6.9). In this model, it is assumed that at scales smaller than the correlation radius *L*, the structure of the skeleton of an infinite cluster does not reduce to a single-stranded macrobonds [3-5]. In the arising infinite cluster there are fragments that are connected to each other by single bonds (macroconnections). The system itself can have self-similarity. A drop having a size *b*

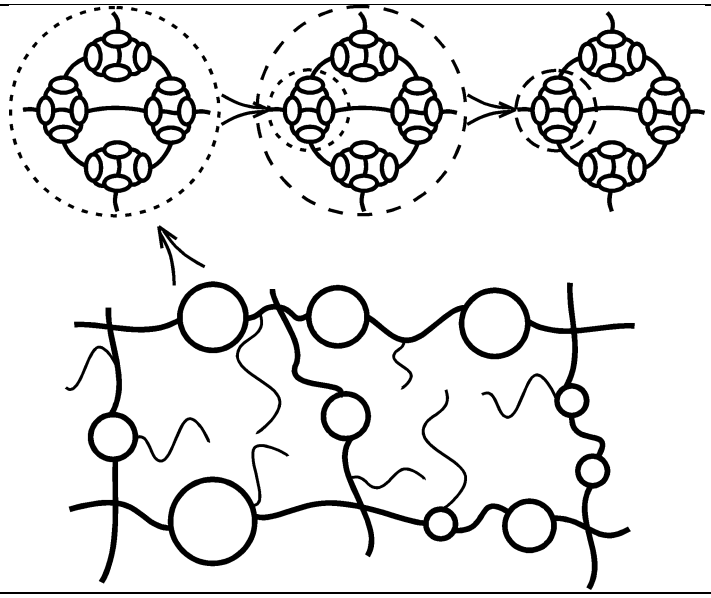


Figure 6.9. Percolation cluster in the model of drops and bonds. Only a small part of the dead ends (thin lines) are shown. Drops are represented in the form of circles. The distance between the droplets and their diameters is of the order of the correlation length.

consists of droplets of size $b/2$, connected by solid macroconnections of length $b^{1/\nu}$. In turn, on a scale of $2b$, droplets of size b are connected by macro-bonds of length $(2b)^{1/\nu}$ into a new drop of size $2b$.

## 6.3. Application of the theory of percolation to describe the electrical conductivity of heterogeneous systems

Simple representations of the percolation theory proved to be useful in explaining the properties of a large number of disordered systems: porous and amorphous materials (including thin films), disordered ionic conductors, branched polymers, galactic structures. For this part of the Chapter, the most interesting is the use of percolation theory to describe the electrical conductivity of heterogeneous systems.

As already mentioned, the bond problem is one of the most studied problems in percolation theory. If we assume an unbroken connection as a conductor, and a dangling - as dielectric with conductivity $\sigma_D = 0$, then it is obvious that for $p < p_c$ the effective conductivity $\sigma(p) = 0$. For $p > p_c$, the conductivity increases according to the law [3-5]:

$$\sigma(p) = \sigma_M (p - p_c)^t, \qquad (6.9)$$

where $\sigma_M$ is the lattice conductivity when all bonds are present ($p = 1$). Here, the numerical coefficients that must be present in expression (6.9) and in subsequent expressions are omitted. It is usually assumed that the critical index $t$ does not depend on the type of the percolation problem, but depends on the dimensionality of the space (the universality hypothesis). This means, for example, that the index $t$ is the same for a flat lattice and for a flat model of "Swiss cheese". For the case of a 3D space $t_3 = 1.6$, and for the 2D case $t_2 = 1.3$.

However, the assumed simplification in which $\sigma_D = 0$ (in other words, the participation of the dielectric phase in the charge transfer is excluded) is not always true for real heterogeneous

systems, and therefore a more complicated case is often analyzed: when, as before, unbroken bonds are conductive with high values of $\sigma_M$, and previously the dangling bonds are considered to be low conductive $\sigma_D$. Sometimes such bonds are called substituted. Let the conductivity of the lattice, in which all the conducting ("metalic") bonds are replaced is $\sigma_D \neq 0$, and $h \equiv \sigma_D / \sigma_M << 1$. Obviously, $\sigma(p)$ is a differentiable function of $p$ for small but non-zero values of $h$. Thus, $h$ plays the same role as the magnetic field in the ferromagnetic theory of phase transitions. Near the percolation threshold, conductivity obeys a power law:

$$\sigma(p_c) = \sigma_M \left( \frac{\sigma_D}{\sigma_M} \right)^s = \sigma_M h^s . \tag{6.10}$$

Thus, the critical index $s$ is introduced. In the 2D case $s = 1/2$, there is no exact data for the 3D case. An analysis within the framework of the theory of an effective medium gives $s = 1/2$. However, the approximations used in the theory of an effective medium are not always valid in the vicinity of the percolation threshold.

If $\sigma(p_c) >> \sigma_D$, then the effective conductivity $\sigma(p)$ should increase with increasing $p$ even below the percolation threshold $p < p_c$. Conductivity at $p < p_c$ has the form [3-5]:

$$\sigma(p) = \sigma_D (p_c - p)^{-q}; \; q > 0, \tag{6.11}$$

where $q$ is another critical index. For the 3D case $q_3 \approx 1$.

A smooth transition from (6.9) to (6.11) occurs in a certain small interval $\Delta$ near the point $p_c$ (see Fig. 6.10), which is called the "smearing region" [8]. In this interval, in the neighborhood of the percolation path, the use of expression (6.10) is legitimate. This means that (6.11) is valid for ($p_c$—$p$) >> $\Delta$ (i.e., below the percolation threshold). Equation (6.9) works under the condition ($p$—$p_c$) >> $\Delta$, (that is, beyond the percolation threshold) $\sigma(p_c) << \sigma(p) << \sigma_M$ ).

Between the three critical indices there is a relationship [3-5]:

$$q = t\left( \frac{1}{s} - 1 \right). \tag{6.12}$$

For the 2D case, $s_2 = 1/2$, and it follows from (6.12) that $q_2 = t_2$. In the 3D case $q_3 \approx 1$, and from (6.12) we obtain that $s_3 \approx 0.62$.

Above we discussed the problem of effective DC conductivity of a two-component system. It is called a ($\sigma_M$, $\sigma_D$) problem. The effective conductivity in this case depends on two variables: $h = \sigma_D / \sigma_M$ and $\Delta = p - p_c$, and can be written in the form

$$\bar{\sigma} \equiv \frac{\sigma}{\sigma_M} = h^s \varphi\left( \frac{\tau}{h^m} \right), \tag{6.13}$$

where $\varphi$ is a function that depends on the ratio of the conductivities of the dielectric and metallic bonds (phases), as well as on the dimensionless the so-called "proximity to the percolation threshold" ($\tau = (p - p_c)/p_c = \Delta/p_c$) [3-5]. One can further complicate the problem and assume that

the conductivity of the dielectric component is a complex quantity and depends on the frequency ω. In this case, the real and imaginary components of electrical conductivity will also obey the power laws.

Accordingly, the effective dielectric constant $\varepsilon$ near the percolation threshold varies according to a power law: its growth is observed as it approaches $p_c$. Confirmations of such a phenomenon can be found in well-known works that generalize theoretical and experimental studies of heterogeneous systems [3-5]. In particular, it was shown that the static dielectric constant near the percolation threshold has the form:

$$\varepsilon(0,x) = \frac{\varepsilon_0}{|p - p_c|^q}. \qquad (6.14)$$

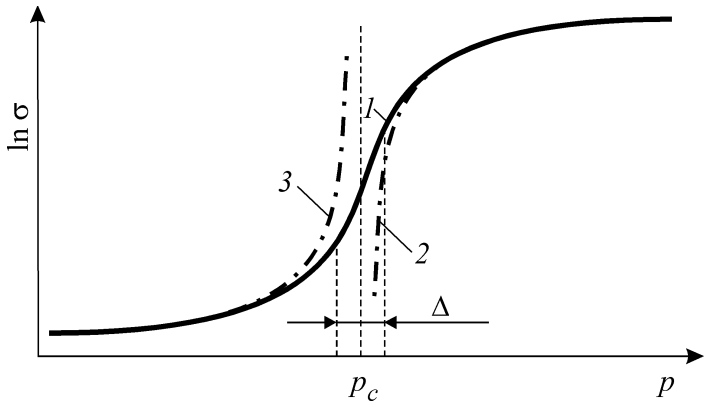



Figure 6.10. General view of the dependence of the effective conductivity of a heterogeneous metal-insulator system: *1* - $h \neq 0$; *2* - $h \to 0$ and $\sigma_D \to 0$, $\sigma_M \neq \infty$; *3* - $h \to 0$ and $\sigma_D \neq 0$, $\sigma_M \to \infty$.

Within the framework of the approximation of the effective medium, $q = 1$ regardless of the dimensionality of the space. In the 2D case $q_2 = t_2$. Then $\varepsilon \propto |p - p_c|^{-t_2}$. In the 3D case $\varepsilon \propto |p - p_c|^{-q_3}$ and $q_3 \approx 1$.

### 6.4. Electrically conductive composite materials

Of all the heterogeneous materials, electrically conductive composite materials (ECCM) have found the most widespread use in microelectronics and electrical engineering, which will be considered further. Most often they consist of two phases - a dielectric (binder) and an electrically conductive filler (particles), that is, they belong to strongly heterogeneous (in the sense of electrical conductivity) binary systems. In this section we consider ECCM, where the filler consists of conducting particles with micron and larger dimensions. In this case, the equivalent circuit of the such ECCM (from the point of view of its behavior on alternating current) is a net of resistors and capacitors.

By the nature of the distribution of the filler in the ECCM can be summarily separated into several large groups: ordered, statistically random (disordered), and structured ECCM. Ordered composite materials include materials whose filler particles are in more or less regular "lattice sites". In statistical ECCM, the filler particles are randomly distributed in matrix (binder). In structured ECCM the filler (particles) form 1D (chain-like), 2D (flat) or 3D (wireframe) structures. A special class is made up of functional gradient materials. Their principal difference from other ECCM is the change in properties (for example, effective electrical conductivity) along the chosen direction inside the body due to the uneven distribution of the filler.

Under the same external conditions, the electrical conductivity of disordered ECCM depends on the type of filler and its concentration, on the shape of the filler particles, the physical/chemical state of the matrix-filler interface, the type of matrix, and also on the method and technological manufacturing parameters. Because of the huge number of types and production technologies of ECCMs, the analysis of their electrical conductivity depending on the above parameters is a rather difficult task. To facilitate its solution, there are usually several groups of primary characteristics of ECCM:

1) the shape and size of the filler particles, its content in the matrix, the intrinsic conductivity of the filler;

2) the distribution and orientation of the filler particles in the matrix bulk;

3) contact resistance between the filler particles.

These characteristics have a different effect on the electrical conductivity of disordered ECCM. Just as for other heterogeneous systems, the volume fraction (concentration) $p$ of the electrically conductive filler (metal-like) is the most significant parameter affecting the electrical conductivity: $p = V_M/V$, where $V_M$ is the total volume of the filler particles, $V$ – the total volume of the composite (heterogeneous system). In the future, for brevity, the volume fraction $p$ will often be called concentration. The $p$ introduced in this section is close in meaning to the volume fraction of the phases in the continuum problems of percolation theory. When $p$ varies from 0 to 1, the effective electrical conductivity of the ECCM increases from the electrical conductivity of the dielectric matrix to values close to the electrical conductivity of the filler, which can be up to 10 orders of magnitude or even more. The other characteristics of ECCM proves differently depending on the concentration of the filler. Therefore, high-filled ($p >> p_c$) and low-filled ($p << p_c$) composite materials are considered as a rule. The natural boundary between them is the experimentally determined percolation threshold $p_c$.

With a small content of an electrically conductive filler (at $p << p_c$), its particles (in the general case of an arbitrary shape) can be placed in the insulating matrix in a large number of ways. Therefore the electrical conductivity of the low-filled ECCM is determined mainly by the distribution and orientation of the filler particles. In high-filled ($p >> p_c$) ECCM conductivity approaches by its values to the electrical conductivity of the filler and is determined to a greater degree by the contact resistance between the filler particles.

The main ECCMs, which are used in electronics are ceramic and polymer composites, are differ by both manufacturing technology and structure.

Ceramic ECCM are most often obtained by sintering compacted powder mixtures from dielectrics (oxides, nitrides and metallic carbides) and a conducting filler particles of various kinds (for example, metals, semiconductor elements - silicon and germanium, semiconductor chemical compounds, etc.). Ceramic itself is a material produced under the influence of high temperature from inorganic substances and their mixtures with fillers, which determine additional functional properties, including electrical conductivity. The scheme of the approximate technological process for manufacturing of ceramics is shown in Fig. 6.11. Ceramics is a complex multiphase system. In its composition, crystalline, vitreous and gas (in closed and sometimes open pores) are distinguished. The crystalline phase, both in content and in properties that it provides the material, is the main constituent of the ceramics. The vitreous phase is a glass interlayer as a rule, which ties together the grains of the crystalline phase. The amount of vitreous phase depends on the features of the manufacturing technology (sintering temperature and duration, composition) and determines the mechanical properties of the ceramics (ductility, toughness, etc.). Depending on the type of ceramic, the content of the vitreous phase varies in the range 1-40 %. The gas phase in ceramics is caused by the peculiarities of the technological process and is often undesirable.

The following technologies are widely used in the production of polymer ECCM: powder mixing; cold extrusion; mixing in a melt or polymer solution; polymerization filling; chemical and electrochemical filling of porous polymers; nano-hybrid technologies.

Powder technology for the creation of electrically conductive polimer-based ECCMs is to mix powders of a conductive filler and a dielectric polymer matrix, and then short-term pressing the mixture at a temperature higher than the temperature of the transition of the polymer to the viscous state. This allows one to achieve a nearly random distribution of conductive particles. This is due to the fact that the matrix and the filler in the form of powders practically do not interact with each other, and the pressing time is short. The dependence of the electrical conductivity of such composites on the filler concentration is satisfactorily described by the percolation theory. The values of the percolation threshold for such ECCM lie as a rule in the range of $p_c \sim 0.15\text{-}0.20$.

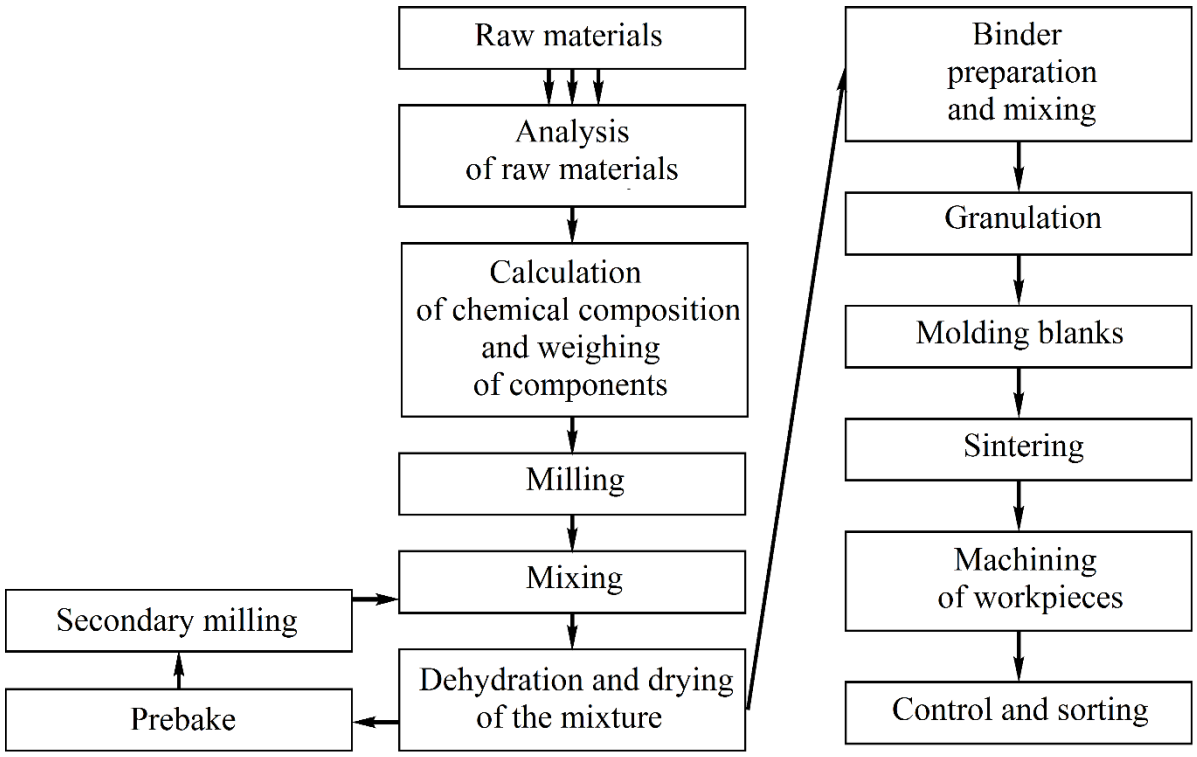



Fig. 6.11 The example of block diagram of technological process for ceramics manufacturing

If the particle size in the matrix is several times larger than the particle size of the filler, the filler is located on the surface of the polymer granules. Pressing thus only deforms the polymer particles, without changing the distribution of the filler, and in fact, structured (wireframe) ECCMs are formed. The percolation threshold for such ECCMs is less than for composites with chaotic filler distribution.

In the cold extrusion method, a mixture of filler and matrix particles is passed through the extruder at a temperature of 5-10 °C below the temperature of the polymer transition into the viscous state. The electrically conducting properties ECCM produced by the cold extrusion method are similar to the properties of ECCMs created by powder technology.

The properties of ECCM produced by mixing components in the melt or solution of polymers strongly depend on the ratio of the energies of the intermolecular filler-filler and filler-polymer interaction. If the energy of the the filler-filler intermolecular interaction is close in value to the interaction energy of the filler-polymer, then it is possible to obtain composites with statistical (random) distribution of the conducting particles in the matrix. The electrical conductivity of such

ECCMs is well described by the percolation theory. If the interaction energy of the filler-filler is lower than that of the filler-matrix interaction, strong bonds of the polymer with the filler are formed, as a result of which each conductive particle is more or less covered with an insulating polymer layer and the percolation threshold increases as compared to the predicted by the percolation theory. If the energy of interaction between the particles of the filler is larger, then it is possible to form chain-like and wireframe-like structures of the filler. To obtain a conducting ECCM in this case, a smaller volume of filler is required, and the percolation thresholds decrease to values of ~ 0.01.

The polymerization filling method consists in the polymerization of the monomer on the surface of the fillers. Therefore, in the polymerization process, the filler particles are often coated with a polymeric layer. By increasing the thickness of this layer, it is possible to shift the percolation threshold toward values substantially exceeding those predicted by the percolation theory. In addition, polymerization filling allows to obtain ECCM with high homogeneity of the particles arrangement in the matrix bulk, and such composites are closest to matrix-like systems.

Features of technological regimes of manufacturing products from ECCM also affect the value of electrical conductivity. Such influence is usually explained by the redistribution of the filler in the bulk of the matrix. For example, if the energy of the polymer-filler interaction is lower than the filler-filler interaction energy, then during the long pressing of the product, chains of filler particles are formed and through-going structures penetrating the sample are formed. The time of formation of such through conducting channels is called the "percolation time". It depends on the volume concentration of the filler, the temperature, the viscosity of the polymer, and the pressure. The pressing time can change by orders of magnitude the electrical conductivity of ECCM with the same filler concentration.

At the study of composite materials, in the first approximation, it is assumed that an infinite current-conducting cluster arises only at filler concentrations larger than the percolation threshold and it consists of the same objects - the particles of the electrically conducting phase. Such a structure of ECCM is answered by the simplest model objects of percolation theory, for example, lattices of resistors with definite and constant resistance, or coupled regions of space with a constant value of the specific electrical conductivity.

The participation of the dielectric phase in DC charge transfer is usually assumed only for composite materials up to the percolation threshold. This corresponds to the basic assumption of percolation theory, according to which, beyond the percolation threshold, there must necessarily exist an infinite cluster of particles of the electrically conducting phase, connected exclusively to each other by similar particles. Sections of the current-carrying skeleton *IC* shunts any dielectric strata.

Note that in describing the electrical conductivity of disordered composite materials, it is preferable to speak not of an infinite cluster of particles of any phase, but about a *conductive cluster*. For $p > p_c$ values (i.e. above the percolation threshold), a conductive cluster is often also referred to as a percolation cluster, implying, firstly, its limited dimensions and, secondly, the possibility of binding the objects (particles) of the conducting phase with thin layers of the dielectric, having local resistance, comparable to the resistance of particles in clusters (in particular, on alternating current).

In real composite materials, because of fundamental physical limitations (for example, wettability and surface tension), often only one of the phases forms a continuous bulk. In addition, for compositional binary materials based on base metals, the presence of additional dielectric component of a heterogeneous system - oxide layers on metallic particles - is possible. In this case, because the composite is not biphasic, it is more rigorous to speak not of the percolation threshold, but of the "touching threshold" of the filler particles in the conductive cluster.

Thus, according to the composition of the conductive cluster, it is possible to distinguish two types of composite materials: ECCM, in which, below the percolation threshold ($p < p_c$) at a constant current, charges are transferred only over the particles of the electrically conducting phase, and ECCM, in which, $p > p_c$, carrier transport is supplied also by strata of dielectric matrix.

The desire to take into account both the resistance of dielectric layers and the nonuniform distribution of currents along the skeleton of an infinite conducting cluster due to the difference in the local resistances of sections of the *IC* was reflected in the creation of a *weak link model* (see, its schematic representation in fig. 6.12). Morozovsky, Snarsky and others [9, 10] have developed several hierarchical stages in the weak link model. The authors of the model consider the volume $L^d$, where $L$ is the correlation length, and $d$ is the dimensionality of the space. In the volume $L^d$ there is a definite (depending on the degree of hierarchy) the number of conditionally allocated objects: the bridges of the conducting phase and/or the dielectric phase strata. It is believed that the main voltage drop falls on them. It is assumed that the volume $L^d$ in question is sufficient (representative) to describe the electrically conductive properties of the heterogeneous system as a whole. The bridges and interlayers, from which form the main resistance of the ECCM, have in the overwhelming majority the same resistance and are scattered chaotically by the system.

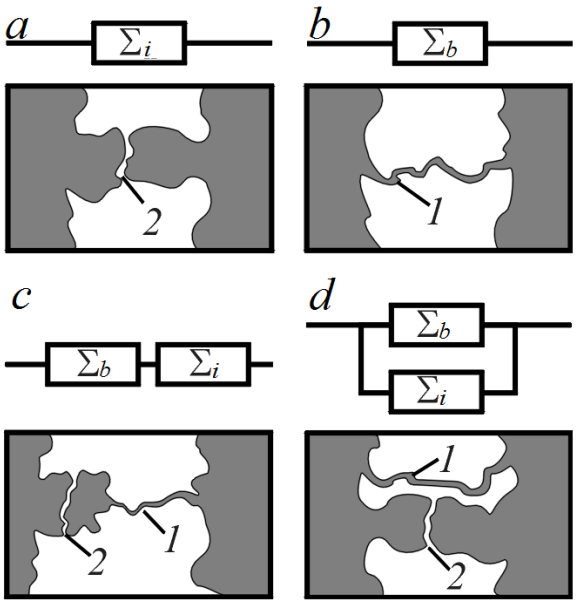



Figure 6.12. Structure of the correlation volumes $L^d$ in the hierarchical weak link model [7]: *a*, *b* - the first stage of the hierarchy; c, в - the second stage; *a*, *c* - below the percolation threshold; *b*, *d* - beyond the percolation threshold; *1* - the bridge; *2* – interlayer (stratum); $\Sigma_b$ is electrical conductivity of bridge; $\Sigma_i$ -the conductivity of the interlayer. The conductive component is shown in gray.

Below the percolation threshold ($p << p_c$, Fig. 6.12a), the first stage of the hierarchy of the weak link model is an interlayer connecting the particles of the electrically conducting phase. The voltage drop on the particles in this cluster is neglected. Beyond the percolation threshold ($p >> p_c$, Fig. 6.12b), the first stage is a bridge of the electrically conducting phase, connecting two segments of an infinite cluster, the voltage drop on which is neglected. Thus, in the weak link model, at the first hierarchical level, the heterogeneous structure of the *IC* is taken into account. It is actually made up of objects of two types: bridges and the rest volume of the electrically conducting phase.

The second stage of the hierarchy takes into account the finiteness of the ratio of the electrical conductivities of the "dielectric" (non-conducting) and "metallic" (conducting) phases. According to the model, this means that in the case of $p << p_c$ (Fig. 6.12c), the bridge is "connected" in series with the interlayer. At $p >> p_c$ (Fig. 6.12d) a stratum layer is included parallel to the bridge. This takes into account not only the current in the conducting phase, but also the current in the dielectric phase. Therefore, the participation in the *IC* of another type of objects - the dielectric stratum, which better reflects the structure of real heterogeneous systems and the current transfer in them - is considered. However, according to the model, the main role in charge transfer in heterogeneous systems beyond the percolation threshold is still played by the regions consisting of the filler, i.e. bridges. The interlayer of the dielectric makes only corrections to the value of the effective electrical conductivity of the ECCM.

## 6.5. Effect of matrix interlayers in a conducting cluster on the impedance of composites

When studying the electrical conductivity of composites on alternating current (AC), in addition to the resistance of the matrix interlayers, it is also necessary to take into account the capacitive resistance of the "capacitors" formed by the particles (or cluster sections) of the electrically conducting phase (playing the role of metal probes) and the interlayers of the dielectric matrix between them.

For this case, a simplified schematic representation of the composite up to the percolation threshold is shown in Fig. 6.13a. The resistive-capacitive coupling between interlayers and particles of the electrically conducting phase or their clusters is symbolically represented by parallel-series RC-circuits. Since an infinite cluster is not formed in composites at $p < p_c$, the equivalent circuit of ECCM (Fig. 6.13b) should contain two series-connected resistors. The first corresponds to the resistance $R_c$ of particles or clusters of the electrically conducting phase. The second resistor takes into account the resistance $R_d$ of the interlayers of the dielectric matrix. The resistor $R_d$ is shunted by a capacitor $C_d$ whose capacitance corresponds to the capacity of the interlayer layers between the particles (or between clusters) of the electrically conducting phase.

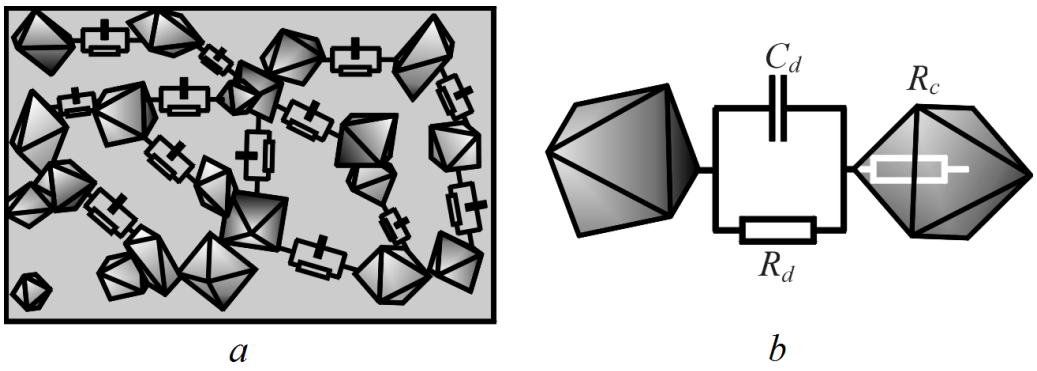



Fig.6.13 Schematic representation of the composite up to the percolation threshold (*a*) and its substitution equivalent circuit (*b*) [7].

To create equivalent nets and extract $R_c$ , $R_d$ and $C_d$ we should make measurements of the electrical conductivity of a composite in AC regime in the form of the real and imaginary components of the impedance $Z(\omega) = Z'(\omega) + iZ''(\omega)$. Then, we represent these as the function $Z'$ *vs* $Z''$ (hodograph) in Figure 6.14a, which will have the form of a semicircle that intersects the abscissa at two points. At high frequencies $\omega << 1/R_dC_d$ (the left intersection point of the hodograph with the $Z'$ axis), the real part of the impedance $Z' \approx R_c$. Thus, in order to determine the resistance of particles and clusters of a dispersed (conducting) phase in a composite material at concentrations $p < p_c$, it is necessary to perform measurements at the highest possible frequencies. At low frequencies $\omega << 1/R_dC_d$ (the right intersection of the hodograph with the $Z'$ axis), the measured impedance values approximately correspond to the total resistance of the dielectric matrix interlayers and the clusters of the conducting phase: $Z' \approx R_c + R_d$. Therefore, when the condition $R_d >> R_c$ is satisfied, that is, when the resistance of the matrix interlayers is much greater than the dispersion phase resistance, low-frequency impedance measurements allow one to estimate directly the resistance $R_d$. The condition $\omega \approx 1/R_dC_d$, is established during the analysis of the experimental data, as the frequency at which the dependence $Z'(\omega)$ reaches its maximum value. Fig. 6.14b shows the dependence of the real part of the impedance on the frequency $Z'(\omega)$ for the composite up to the percolation threshold. In accordance with the above arguments, we can separate out on the $Z'(\omega)$ dependence 3 specific regions: a low-frequency and high-frequency plateau, as well as a region with dispersion (frequency-dependent) decay between them.

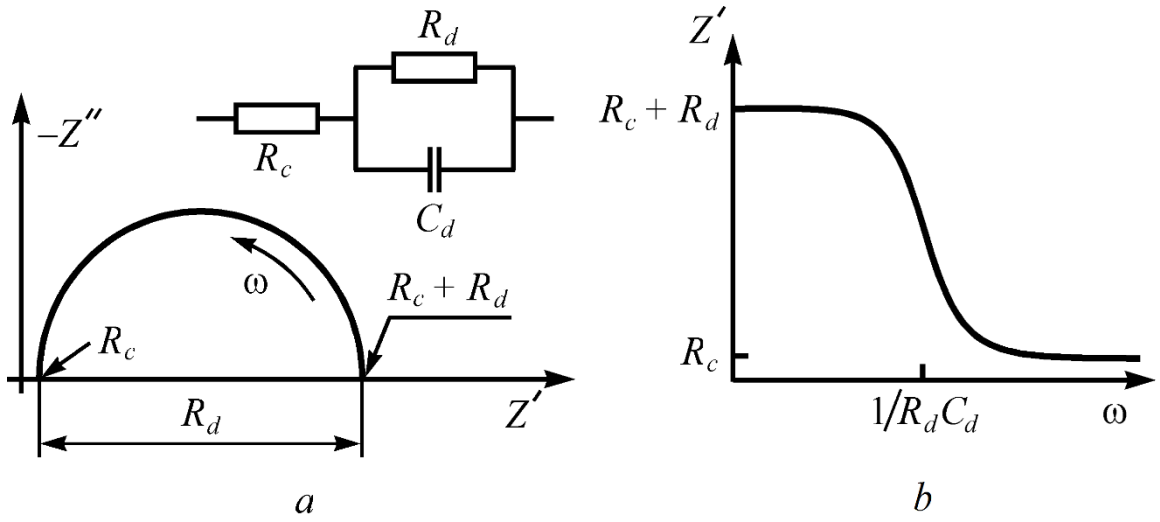



Figure 6.14. The hodograph of the impedance of the composite before the percolation threshold (*a*) and the dependence of the real part of $Z'$ impedance on frequency (*b*) [7]. An equivalent circuit is shown in inset.

The participation of the dielectric matrix interlayers in the currier transport in composites below the percolation threshold is also manifested in the fact that the dependences of the real part of the complex conductivity measured at different frequencies on temperature $T$ differ from each other. Fig. 6.15a shows the examples of curves $G(T) = Z'(T)$ at different frequencies for a sample of a ceramic silicon/silica composite with a mass concentration of silicon particles $C_m$(Si) = 17 wt.% [11]. The values of $G = Z'$ are normalized to the value of the real part of the complex conductivity at $T$ = 318 K. It follows from Fig. 6.15a, with increasing temperature, the real part of

the conductivity *G* at low and intermediate frequencies increases more sharply. This feature is also clearly visible on the impedance hodographs.

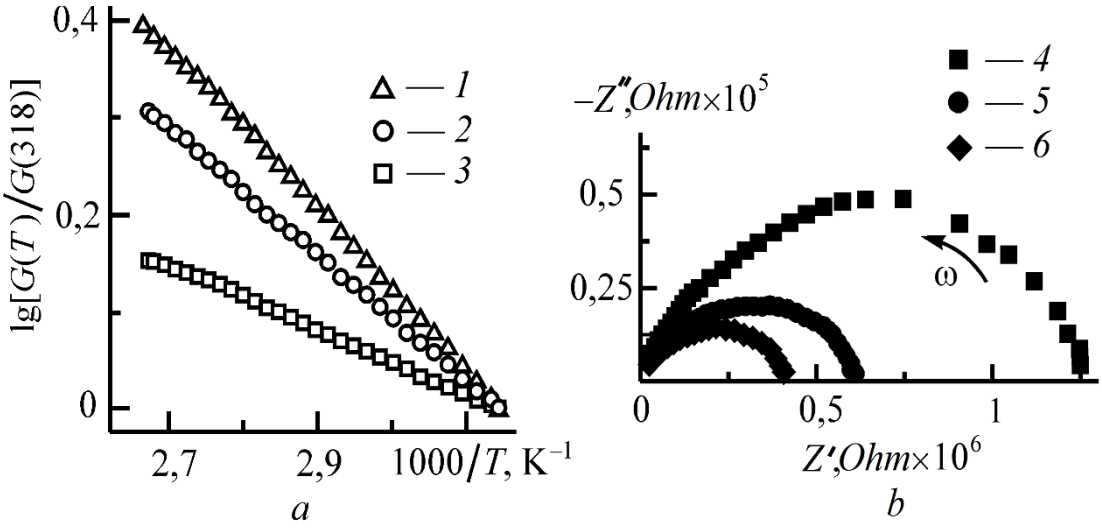



Fig. 6.15. Dependences of the real part of conductivity *Z'(T)* on temperature (a) of a ceramic composite Si/$SiO_2$ sample below the percolation threshold measured at frequencies of 100 Hz (*1*); 10 kHz (*2*); 1 MHz (*3*), and impedance hodographs (*b*) at temperatures of 25 ° C (*4*); 75 ° C (*5*); 100 ° C (*6*) [11].

As is seen in Fig. 6.15*b*, are plotted based on the *Z*(□) measurements at three temperatures: 25 ° C (dependence *4*), 75 ° C (*5*) and 100 ° C (*6*).

The strongest changes of *Z'(T)* are observed in the low-frequency regions of hodographs corresponding to the resistance of silicon dioxide interlayers between the silicon particles. At high frequency regions of hodographs, where the resistance of Si particles or clusters is manifested, the *Z'(T)* values differ little from each other. It should be noted that as the frequency of the alternating current increases, the activation energy of the effective electrical conductivity of the composite samples decreases (from 0.15 eV for $f$ = 100 Hz to 0.06 eV for $f$ = 1 MHz). This is due to the increase in the overall electrical conductivity of the composite from clusters of silicon particles at high frequencies, and the fact that the temperature dependence of the electrical conductivity of the Si/$SiO_2$ samples beyond the percolation threshold (with $C_m$(Si) = 19 wt. %) is consistent with the temperature dependences of the electrical conductivity for single-crystal silicon.

The structure of the composite beyond the percolation threshold is schematically shown in Fig. 6.16*a*. The figure reflects the fact that when the concentration of the electrically conductive phase exceeds the threshold value $p_c$, a conductive percolation cluster is necessarily present in the ECCM. The cluster consists exclusively (in contrast to the case with $p < p_c$) of the electrically conducting particles and permeates the sample from contact to contact. The resistance of dielectric layers separating the sections of the cluster of the conducting phase, as in the case of composites

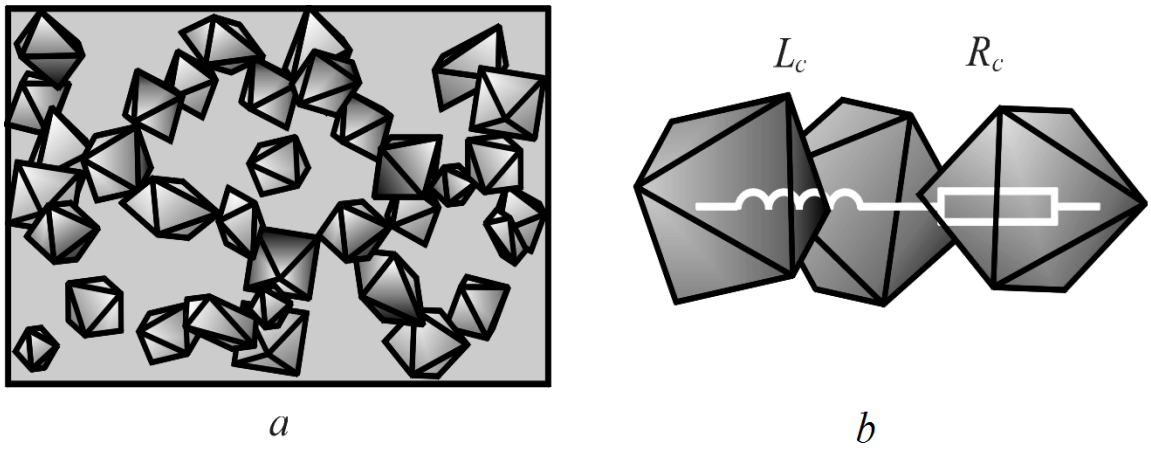



Fig. 6.16 Schematic representation of the composite beyond the percolation threshold (*a*) and its substitution equivalent circuit (*b*) [7].

below the percolation threshold ($p < p_C$), is much greater than the resistance of the clusters themselves. However, for any interlayer of the dielectric matrix in the composite beyond the percolation threshold ($p > p_C$), it is possible to find a shunting path through the particles of the electrically conducting phase.

Therefore, the preferential current transfer is carried out by a percolation cluster. Thus, beyond the percolation threshold, the equivalent circuit of the composite can be represented as a series-connected resistor $R_c$ of the electrically conducting percolation cluster and the inductance $L_c$ (Fig. 6.16b). In this case, the hodograph of the composites impedance will have the form of a ray emanating from the point $R_c$ on the axis parallel to the ordinate axis $Z''$ (Fig. 6.17). It should be noted that in this case the inductive component will be dominated in the reactive resistance. The experimental values of the inductance of the composite sample in this case are smaller than the true inductance of the skeleton of the infinite cluster, since in the measurement process the contribution of the capacitive component of the reactive resistance from the sections of the *IC* separated by the interlayer of the matrix also gives contribution.

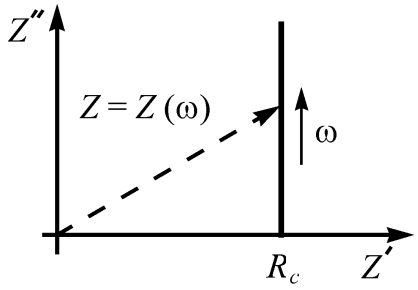



Figure 6.17. Hodograph of the composite impedance beyond the percolation threshold [7]

In real ECCM, a typical example of which can serve composites with polymer matrix, interlayers of the dielectric matrix participate in charge transfer even when $p > p_c$. A feature of polymer ECCM is the presence, at equal filler concentrations, of the electrical conductivity of the composite from the manufacturing process. The effect of the preparation technology on the electrical conductivity of ECCM is primarily due to the processes of interfacial interaction between the filler particles and the polymer matrix in the stage of components mixing (or at the stage of chemical production of the conducting phase). As a result of this interaction, firstly, there is a

redistribution of the filler particles along the matrix bulk, and the geometry of the percolation cluster can no longer be considered to be accidental. Secondly, interphase layers are formed, the properties of which differ from the properties of both the filler and matrix.

The percolation thresholds in polymeric composites prepared by different technologies lie in wide ranges of concentrations from $p_c \approx 0.05$ to $p_c \approx 0.5$. They are recorded from the bend of the $\sigma(p)$ dependence (or by the maximum of its derivative $d\sigma(p)/dp$). Fig. 6.18a shows the experimental dependences of the effective electric conductivity of composites for two typical preparation technologies [12]. It can be seen that with comparable filler concentrations, the electrical conductivity of composites can differ by several orders of magnitude. Fig. 6.18b illustrates an effective and intuitive way to determine the experimental values of the percolation threshold with respect to the maximum of the derivative $d\sigma(p)/dp$. It should be borne in mind that it is possible to construct the derivative of the $\sigma(p)$ dependence correctly only in the case of a large number of experimental points, which is a considerable difficulty for composite materials.

In the case of ceramic composite materials, interfacial interaction processes also play a significant role in the formation of the structure. First of all, we are talking about the technological stages of molding and sintering. In addition, pores, especially open pores, can influence the effective conductivity of ceramic composite materials. In open pores, adsorption of moisture is observed, which leads to an increase in both the active (due to the transfer of ions) and the reactive components of electrical conductivity. The effect of adsorbed moisture on the electrical conductivity of porous heterogeneous systems is particularly clearly seen in the case of porous glasses and xerogels [13].

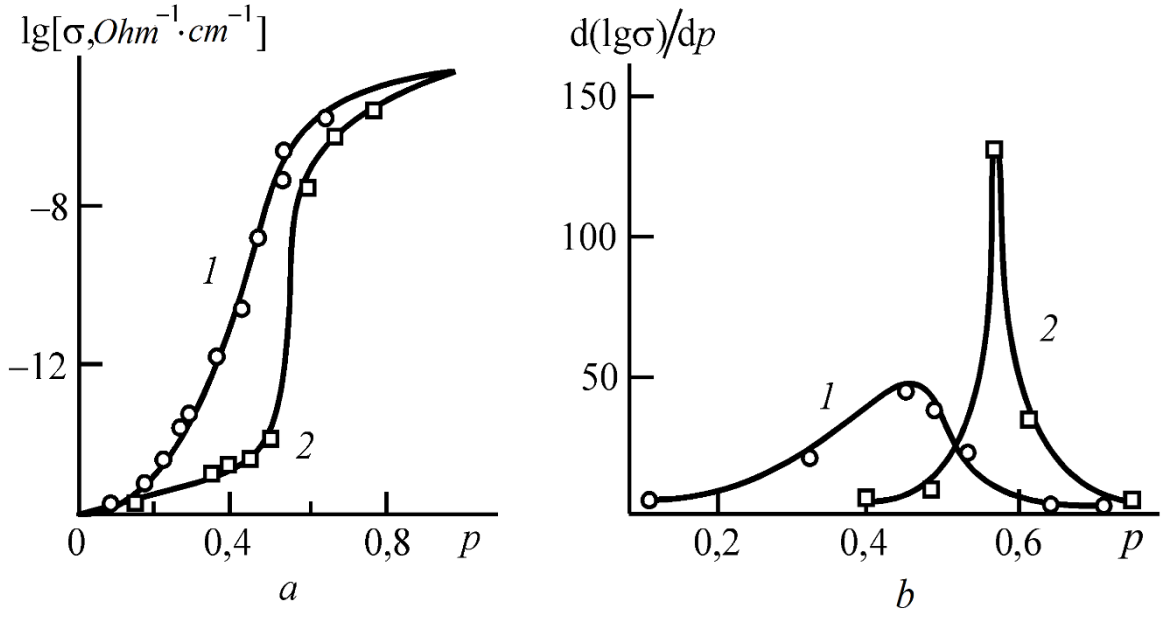



Fig. 6.18 Dependence of the effective electrical conductivity (*a*) and its derivative (*b*) on the filler concentration for polypropylene/aluminum composites obtained by the polymerization filling method (*1*) and the rolling of the mechanical mixture of components (*2*) [12]

Thus, the electrical conductivity of polymer and ceramic composites depends on many factors. The contribution of these factors to the overall course of the $\sigma(p)$ dependence can only be separated

by additional methods of investigation, for example, optical and tunneling microscopy. However, there is a fairly large number of works, the results of which indicate the prevailing effect of dielectric layers and interphase layers in the formation of electrical properties of composites. First of all, this is manifested by the shift, under otherwise equal conditions, of the experimental values of the percolation threshold toward larger (in comparison with the predicted by the theory) concentrations of the conducting component. In addition, the presence of thin insulating layers around the particles of the dispersed phase leads to a significant difference in the behavior of the curves of real ECCM from the dependences predicted by the models of percolation theory. In composite materials, the so-called "fuzzy" dielectric-metal transition is observed. On the graphs of $\sigma(p)$, this is manifested in the fact that the growth of effective electrical conductivity is observed in a wide range of $p$ (usually from $p \approx 0.1$ to $p \approx 0.6$).

Fig. 6.19 schematically shows the dependences of $\sigma(p)$ for typical cases of a dielectric-metal transition in an idealized composite and in a composite for which interphase layers play an important role. In this connection, consider the behavior of a typical representative of conductive composites - a carbon black/polyethylene composite. Particles of carbon black have a high adsorption capacity. It is known that they can enter into chemical reactions with the matrix material. As a result, each particle is surrounded by a dielectric layer of polyethylene. At the same time, due to the presence of a residual non-polymeric adsorbate (for example, oxygen) and functional groups arising from chemical reactions with carbon black, the electrical properties of the interlayers may differ significantly from those of the rest of the matrix. Therefore, even for $p > p_C$, the percolation cluster in the carbon black/polyethylene composite is non-uniform [7].

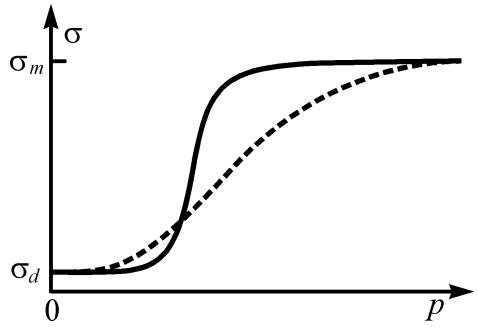



Fig. 6.19. Dependence of the effective electrical conductivity on the volume fraction of the electrically conductive disperse phase. The solid line is an idealized dependence; dotted line - dependence for the "smeared" transition [7]

Dielectric strata of polyethylene are those sections of the PC, which, first, determine its overall resistance; secondly, lead to a nonlinearity of the current-voltage characteristics of the polymer composite at $p > p_c$. A layer of polyethylene that bonds two carbon black particles to each other can be represented as a parallel resistor (with resistance $R_d$) and a capacitor (with a capacitance $C_d$). Note that the percolation cluster will necessarily include carbon black particles, directly related to each other. They constitute carbon black agglomerates, to avoid formation of which is practically impossible with existing manufacturing technologies. Therefore, the carbon black particles and the interlayers of the surrounding them matrix will participate in the current transfer. The interlayers that enter the conductive percolation cluster have a minimal (in relation to other

parts of the matrix) local resistance, the value of which is comparable to the resistance of carbon black agglomerates. This condition can be satisfied for interlayers having a minimum thickness, the greatest number of defect and impurity centers providing hopping electrical conductivity.

Proceeding from such a percolation cluster model, an equivalent substitution circuit similar to that shown in Figure 6.13 can be used to describe the electrical conductivity of the composites at AC regime beyond the percolation threshold. Resistance of $R_c$ in this case will correspond to the resistance of carbon black particles and their agglomerates, and $R_d$ and $C_d$ are the resistance and capacitance of interlayers of the polymer matrix, or rather, those layers that are "adsorbed and modified" by carbon black particles. As a result, the hodograph of the samples of the carbon black/polyethylene composite beyond the percolation threshold has the form similar to the hodograph shown in Fig. 6.14.

At high frequencies, the current routes will include RC chains built of filler particles surrounded by polymer interlayers. If such "links" of the percolation cluster have high effective capacitance values, they close the gaps between the "dead ends" shunting the long loops. This leads to a "straightening" of the conducting paths and a significant change in the topology of the infinite cluster. As the filler concentration increases, the dispersion section on the dependence of the real part of the electrical conductivity $G$ on the frequency $\omega = 2\pi f$ is shifted towards higher frequencies. The difference between the high-frequency and low-frequency plateaus is also less noticeable. A set of dependencies $G(\omega)$ for composites beyond the percolation threshold at different filler concentrations $p$ is shown in Figure 6.20. The same figure shows two dependences of the electrical conductivity $G$ on the concentration $p$ of the conducting phase: for the low-frequency region, $G_l\,(p)$, for the high-frequency region, $G_h(p)$. In its most general terms, the experimental dependence of $G_h(p)$ for high frequencies differs from the analogous dependence of $G_l\,(p)$ for low frequencies by a shorter interval of variation $\Delta G$ and "smoother", i.e., all other conditions being equal, in the vicinity of the percolation threshold $\mathrm{d}G_h(p)/\mathrm{d}p < \mathrm{d}G_l\,(p)/\mathrm{d}p$.

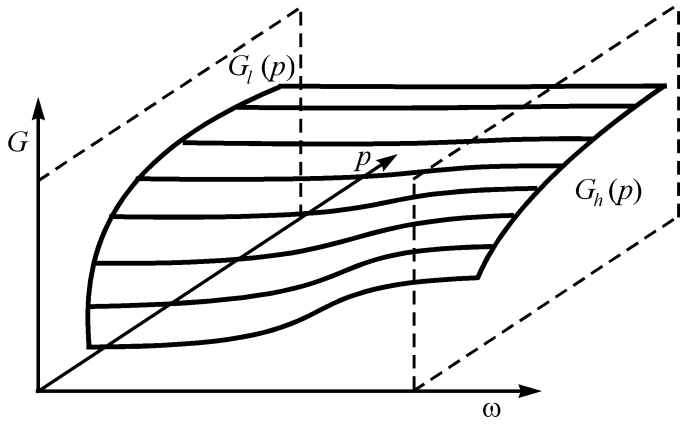



Fig. 6.20 The family of frequency dependences of the real part of admittance $G$ of composites beyond the percolation threshold with different filler concentration $p$ [7]

The physical-chemical processes occurring at the interface lead to the the increase of local polymer interlayers electrical conductivity with increasing filler concentration. This causes a partial decrease in the smearing of the dielectric-metal transition on the side of concentrations larger than the threshold, decreases the hodograph radius, increases the exponent of the power-law dependence of the actual part of the electrical conductivity on frequency. The most probable reason

for the increase in the local electrical conductivity of the dielectric interlayers is the decrease in the effective hopping length of charge carriers in the polymer interlayers of the percolation cluster.

## 6.6. Models of an effective medium for describing the dielectric properties of nonconductive (porous) composite materials

One of the main models used to calculate the properties of composite media is the effective medium theory, which uses the principle of self-consistency: when calculating the field inside a microscopic element of a composite material, it is considered that it is surrounded by an "effective medium" whose permittivity is identical to the desired one. It is assumed that the average field due to this inclusion is not changed. Averaging over the entire sample the field calculated under such assumptions, it is equated to a given macroscopic field and equations are obtained for determining the effective permittivity.

Depending on the morphology of the environment, two different design schemes are used. The first of these is a symmetrical model which is used for mixtures with two continuous phases, i.e. when there is no explicit division of the components into a matrix and a filler. In this case, the microscopic element over which the averaging is carried out is the space portion filled with the material of either the first or second component of the composite. As a result of these assumptions, for the dielectric constant ε of the mixture, the Bruggeman formula is obtained:

$$(1-c)\frac{\varepsilon-\varepsilon_2}{\varepsilon+2\varepsilon_2}=c\frac{\varepsilon-\varepsilon_1}{\varepsilon+2\varepsilon_1}, \tag{6.15}$$

where $\varepsilon_1$ and $\varepsilon_2$ are the complex dielectric permittivities of components 1 and 2, $c$ is the volume fraction of component 2. If we consider a liquid-filled porous medium and component 2 is a liquid, then $c = \phi$, where $\phi$ is the specific porosity.

The second calculation scheme using the effective medium model is an asymmetric approach applicable to mixtures in which one component can be called a matrix and the second one is a filler. In this case, the microscopic element on which the averaging is performed is two-phase and represents a core of the filler material surrounded by a layer of matrix material. The asymmetric model of the effective medium, depending on the averaging method for finding the permittivity, yields the following equation

$$\varepsilon=\varepsilon_1+3c\varepsilon\left(\frac{\varepsilon_2-\varepsilon_1}{\varepsilon_2+2\varepsilon}\right), \tag{6.16}$$

or the so-called Maxwell-Garnett equation

$$\varepsilon=\varepsilon_1\frac{2\varepsilon_1+\varepsilon_2-2c(\varepsilon_1-\varepsilon_2)}{2\varepsilon_1+\varepsilon_2-c(\varepsilon_1-\varepsilon_2)}. \tag{6.17}$$

The following expression was obtained in [14,15] in the framework of the differential model of an effective medium:

$$\frac{\varepsilon-\varepsilon_1}{\varepsilon_2-\varepsilon_1}\left(\frac{\varepsilon_2}{\varepsilon}\right)^{1/3}=c, \tag{6.18}$$

which coincides with the results of [16], starting from the self-similar model of a porous medium. There are various modifications of this expression for cases of non-spherical pores with the same orientation and isotropic distribution of orientations, which do not generally change the qualitative picture of the dependence of the effective dielectric constant on the composition of the heterogeneous system.

In a number of works [16, 17] the dielectric properties of porous media are described using the Luyengi equation

$$\varepsilon^{1/3} = (1 - c)\varepsilon_1^{1/3} + c\varepsilon_2^{1/3}, \tag{6.19}$$

as well as the Lichteneker equation

$$\varepsilon = \exp[(1 - c)\ln(\varepsilon_1) + c\ln(\varepsilon_2)], \tag{6.20}$$

or Brown equation

$$\sqrt{\varepsilon} = (1 - c)\sqrt{\varepsilon_1} + c\sqrt{\varepsilon_2}, \tag{6.21}$$

which are a particular case of equation

$$\varepsilon_k = (1 - c)\varepsilon_1^k + c\varepsilon_2^k. \tag{6.22}$$

Which model best describes the properties of porous media, is the subject of discussion. For example, in [16], where the dielectric properties of polypropylene filled with aromatic liquids are described, it is shown that the values obtained with the help of the Lichteneker equation are the closest to the experimental results. Some authors, in turn, to calculate the dielectric properties of porous polymers, in particular epoxy, recommend using an asymmetric model of the effective medium. Obviously, for each particular system, the most effective models should be selected individually.

## 6.7. Electrical conductivity of nanogranular metal-insulator composites

We should separately mention the electric transport properties of composite materials with particle sizes of dispersed electrically conductive phase (filler) of the order of 1-10 nm. In these cases, not only the type, chemical composition, relative concentration and size of the conducting granules and dielectric matrix, but also the phase state of the conducting and non-conducting phases (crystalline, amorphous, magnetic, superparamagnetic, etc.) and interphase boundaries, the charge state of the elements in the phases of nanoparticles, and so on should play a big role in the carrier transport.

First, we consider metal-insulator granular nanocomposites, where the insulator has a mineral origin (oxides, nitrides, metal carbides).

### 6.7.1. Methods for the preparation of granular metal-dielectric composites

The methods of obtaining such composites in the form of nanogranular metal-dielectric films are very diverse. To obtain a heterogeneous system that consists of nanosize regions of the metallic and dielectric phases, it is necessary that the atoms included in the composite should enter the substrate either

simultaneously or cyclically, but the cycle period should not exceed the time over which the film can form a continuous monophasic layer. These criteria are fully consistent with the method of deposition from the gas phase [18]. The main difficulty limiting the use of various technical modifications of this method is the need for sputtering or obtaining a dielectric on the substrate during growth. In this connection, the following methods are normally used: high-frequency magnetron sputtering of a composite target [19-22], high-frequency co-sputtering of dielectric and metallic targets [23], reactive sputtering of metal targets in an argon atmosphere with the addition of oxygen [23], successive high-frequency sputtering of metal and dielectric targets on rotating substrates [24], electron-beam co-sputtering of two targets [25], and ion-beam sputtering of composite targets [27, 28] (Figure 6.21a). Each of these methods has its advantages and disadvantages. One of the most successful methods is the latter, when plates of dielectric are placed on a metal plate with different distances between them (Figure 8.21b), that allows to obtain nanogranular composite film with a whole set of relative concentrations of metal and dielectric in one deposition cycle (at the same pressure components of the gas mixture, substrate temperature, etc. ).

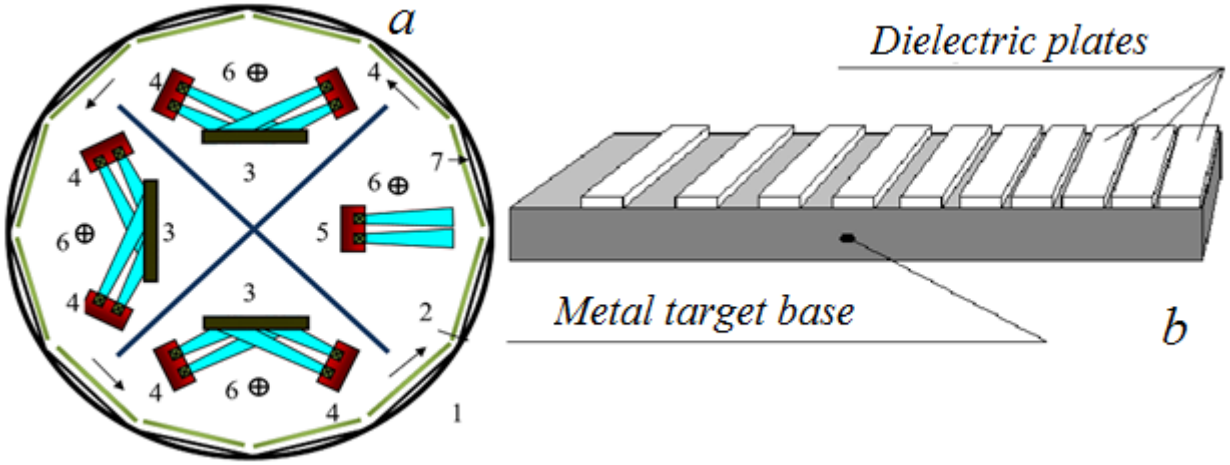



Fig. 6.21. (a) Multifunctional ion-beam deposition unit: 1 - vacuum chamber, 2 - substrate holder, 3 - water-cooled target, 4 - ion-beam atomizing source, 5 - source of ion etching, 6 - compensator, 7 - substrate; (b) Model of a composite metal alloy target with dielectric plates

### 6.7.2. Electrical properties of nanogranular metal-dielectric composites

The electrical properties of heterogeneous nanogranular systems are completely determined by their morphology, phase composition and, in this sense, the study of the electric transport properties of nanocomposites makes it possible to reveal the features of the structure, the effect of scale and concentration factors on the phase and structural transformations in them. On the other hand, the structure and morphology of nanogranular metal-dielectric composites determine their extraordinary electrical properties, the study of which allows us to expand our understanding of the mechanisms of carrier transport in a nanoscale heterogeneous medium.

There are several levels of electric transport consideration in granular nanocomposites. The first and most common (see above) approach separates all the metal-insulator nanocomposites by composition into two groups: with dielectric and metallic conductance (Figure 6.22). It considers changes in the resistivity as a result of the structural state of the system depending on the volume ratio of the dielectric or metallic phases, as well as under external influences. This approach is especially useful in analyzing the structural transformations in the composite during annealing.

The second level of nanocomposites consideration includes analysis of the changes in the electrical resistivity of a two-phase heterogeneous system in terms of the ratio of the volume fraction of the dielectric and conductive components. At this level, as discussed above, the concept of the percolation threshold is introduced, which divides the nanocomposites from the viewpoint of electrical conductivity into two groups. This approach is convenient in determining the position of the percolation threshold and in analyzing the concentration dependences of the electrical resistivity of metal-insulator composites of various compositions (Figure 6.22).

The third level of consideration is based on an analysis of the electrical transfer between nanoscale metal granules by tunneling (hopping) electrons through a dielectric layer and other phases (if they are formed, for example, as a core-shell structure inside a dielectric matrix).

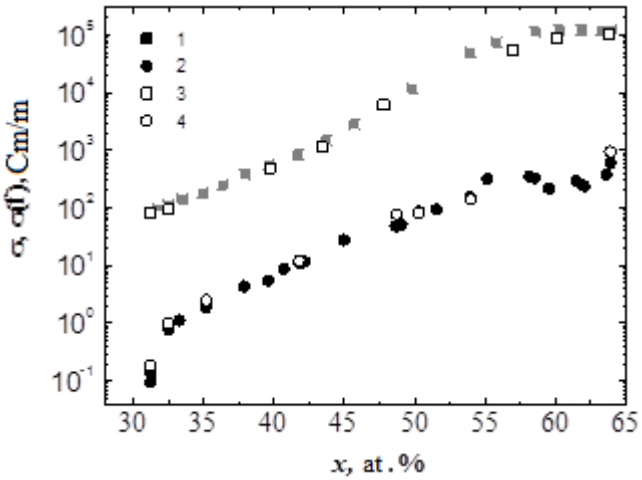
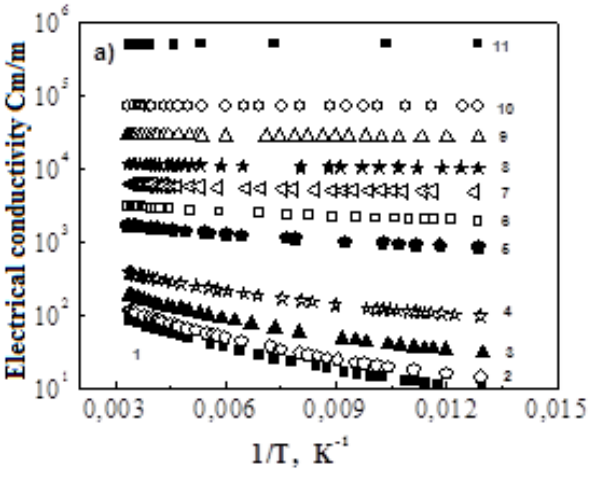
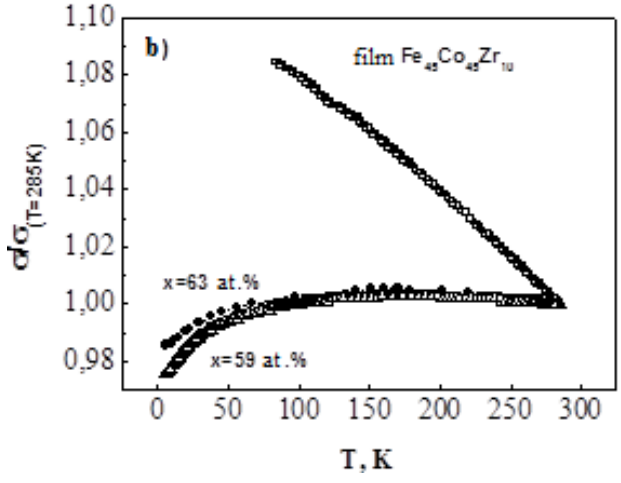



Fig. 6.22. An image of the concentration dependence of the resistivity (1, 2) and the specific admittance (3, 4) at 300 K in the metal-insulator $(FeCoZr)_x(Al_2O_3)_{1-x}$ nanocomposites obtained by ion beam sputtering of the composite target: 1, 3 - nanocomposites obtained under an argon atmosphere, 2, 4 - nanocomposites deposited in the argon-oxygen atmosphere.

Here we take into account the size of metallic granules and their shells, the distance between them in the dielectric matrix, the state of electron localization in each phase (shell, dielectric, metal) and the radius of electron localization in them. This approach allows us to describe not only the concentration, but also the temperature dependences of the electrical resistance of granular metal-dielectric nanocomposites (Figures 6.23a and 6.23b).




a



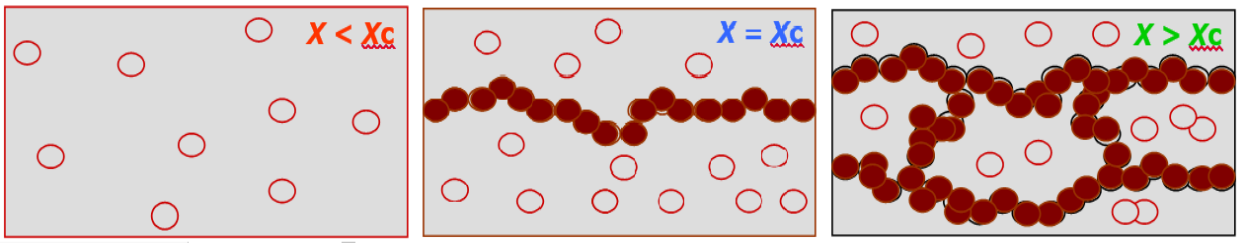

b

Fig. 6.23. Temperature dependences of the electrical resistivity of the metal-insulator nanocomposite before (a) and beyond (b) the percolation threshold: 1 - $(FeCoZr)_{31}(Al_2O_3)_{69}$; 2 - $(FeCoZr)_{33}(Al_2O_3)_{67}$; 3 - $(FeCoZr)_{35}(Al_2O_3)_{65}$; 4 - $(FeCoZr)_{38}(Al_2O_3)_{62}$; 5 - $(FeCoZr)_{43}(Al_2O_3)_{57}$; 6 - $(FeCoZr)_{46}(Al_2O_3)_{54}$; 7 - $(FeCoZr)_{48}(Al_2O_3)_{52}$; 8 - $(FeCoZr)_{50}(Al_2O_3)_{50}$; 9 - $(FeCoZr)_{52}(Al_2O_3)_{48}$; 10 - $(FeCoZr)_{63}(Al_2O_3)_{37}$; 11 - film $Fe_{45}Co_{45}Zr_{10}$ [29-32].

A schematic representation of the nanocomposites structure below and above the percolation threshold is shown in Fig. 6.24. It should be noted that in its pure form none of these approaches is used in interpreting the electrical properties of nanogranular composites.

The increase in the size and sphericity of the granules, as well as their more uniform distribution, realized by increasing the substrate temperature during sputtering and isothermal annealing, lead to a shift in the position of the percolation threshold of the nanocomposites toward an increase in the metal concentration.

We should note that the percolation threshold $x_c$ in metal-insulator nanocomposites is a structurally sensitive parameter, the position of which is influenced by various factors. Knowledge of the position of $x_c$ for known compositions of metallic granules and dielectric matrix and certain technological conditions is one of the methods for studying the morphology of complex composite materials.




*a* *b* *c*

Fig. 6.24. Schematic representation of the distribution of nanoparticles in a metal-insulator composite below the percolation threshold (*a*), at the percolation threshold when forming the first conductive cluster (*b*) and beyond the percolation threshold (*c*), when metallic granules form a net of current routes

Introduction of nonmetals (for example, boron), forming compounds with the material of a dielectric matrix (for example, borosilicate glass in the case of a matrix of silicon oxide), in the metallic phase shifts the position of the percolation threshold toward a larger metal content. Increasing the solubility of the chemical elements in dielectric matrix (Al instead of Si) in the metallic phase of the composites shifts the position of the percolation transition toward a smaller metal content.

When dielectric phase has a mineral origin (oxides, nitrides, metal carbides), a special class of metal-insulator nanocomposites can be assigned if each conducting nanoparticle is surrounded by its own oxide or nitride so that we have nanoparticles with core-shell structure inside the dielectric matrix. The formation of an additional shell at the surface of metallic granules (usually semiconducting or dielectric) occurs as a result of the introduction of reactive gases $N_2$ or $O_2$ during the deposition of the investigated composites. The presence of shell around metallic cluster usually leads to the shift of the threshold position for the formation of a highly conducting cluster towards increasing the metal content in the dielectric matrix. In this case, this threshold can not be called the percolation threshold in the sense described in Sections 6.2 and 6.3 (due to non-two-phase nature of the system). We propose to call this transition into the metallic regime of the

nanocomposite with the structure of filler nanoparticle as the core-shell the "threshold of tangency" of the nanoparticles.

### 6.7.3. Temperature dependences of the conductivity in amorphous granular nanocomposites

As noted above, the electrical properties of nanogranular composites depend radically on the ratio of the metallic and dielectric phases in the material, which leads to the existence of two fundamentally different conductivity regimes, metallic and nonmetallic [3, 18, 33-35].

The non-metallic (dielectric) conductivity regime is realized when the volume fraction of the filler (nanoparticles of pure metal or alloy) in the composite is below the percolation threshold (Fig. 6.24a). Two-phase material in this case consists of metallic nanogranules, which are electrically isolated from each other by dielectric interlayers. The dielectric regime is characterized by high values of the electrical resistivity, increasing by several orders of magnitude with a decrease in the proportion of metal in the composite from the percolation threshold (Fig. 6.24,) to $x$ = 0 % for a pure dielectric matrix (Figure 6.24a) [29-32]. When cooling, the resistivity of composites increases and in the temperature range 4.2-300 K the change in the resistivity $\rho$ reaches several orders of magnitude, characterized by a negative sign of the temperature coefficient of resistivity $TCR = d\rho(T)/dT$ (Fig. 6.23a).

The metallic conductance regime is realized when the volume fraction of the metallic phase in the material exceeds the percolation threshold $x_c$. In this case, conductive clusters and continuous metal channels (chains of mutually contacting granules with each other) are formed, penetrating all the material (Fig. 6.24c). This provides a predominantly metallic type of conductance with a positive *TCR* (Fig. 6.23b). Dielectric regions between the conducting channels increase the overall level of the electrical resistivity of the material, but do not affect the conduction mechanism as a whole. In this case, the material behaves like a metallic conductor, although its properties, which depend on the mean free path of an electron, vary considerably due to strong scattering at the boundaries of the granules and a decrease in the total fraction of metallic phase $x$ in the material. Usually near $x$ = 100 %, the electrical conductivity of such composites is several times smaller than the values characteristic of pure metals or metal alloys [18].

The study of the mechanisms of charge transport as a function of temperature provides information on the nature of the bonds between atoms both inside and between the structural elements of composites (metallic nanoparticles and dielectric or other nanostrata).

From the practical point of view, the most interesting is the study of the nonmetallic conductivity regime at different temperatures. To describe the electrical properties of nanocomposites, several conduction mechanisms have been proposed in this case, which are realized in different temperature ranges and depend on the presence of additional shells around the nanoparticles of the metallic filler.

For binary composites with a low concentration of the metal phase, the nanogranules are electrically isolated from each other in the bulk of the matrix, and therefore the electrical conductivity in such composites is mainly determined by the dielectric component, and the question of the formation of a temperature dependence in this case often remains open.

The experimental temperature dependences of the electrical conductivity of nanocomposites for this range of the high-conductivity phase concentrations in the low-temperature region in the most cases obey the so-called Mott low,

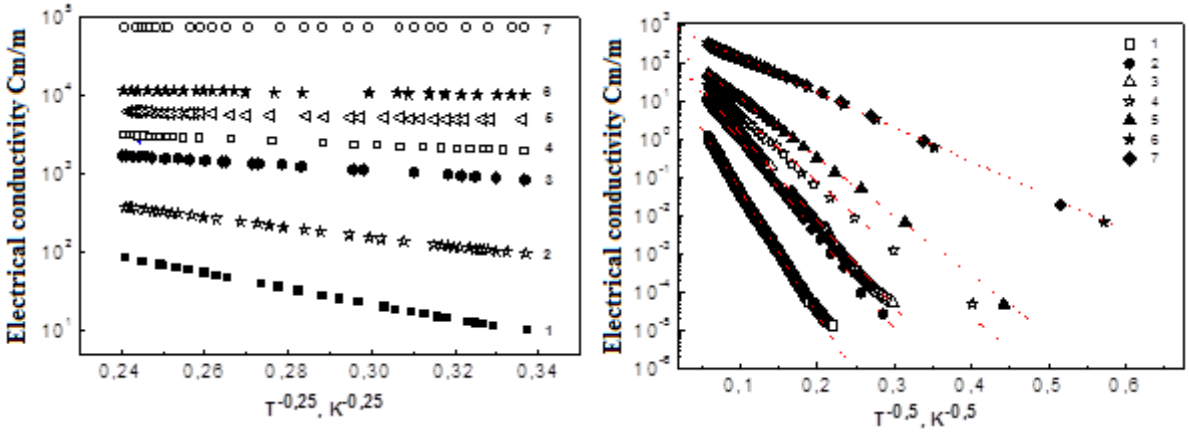
$$\sigma \propto \exp\left[-(T_o/T)^{-n}\right], \qquad (6.23)$$

which characterizes the conductance with variable activation energy (due to electron tunneling between metallic nanoparticles). The exponent $n$ in (6.23) usually takes two values - 0.25 and 0.5,

which are usually depends on the distribution of the density of localized states (DOS) of charge carriers in the vicinity of the Fermi level. As shown below, in nanogranular composite films of the $(FeCoZr)_x(Al_2O_3)_{1-x}$ [29-32], the *n* value depends on the gaseous medium in the vacuum chamber during ion beam sputtering of the target (see Section 6.7.1) - either pure argon (the so-called non-oxidized films) or a mixture of argon and oxygen (so-called oxidized films with the structure of nanoparticles of the core-shell type (Figure 6.25), where the shell is composed of intrinsic oxides of the metal phase [29- 32]).

Fig. 6.25. Schematic representation of the metal-insulator nanocomposite with the core-shell structure of metallic filler particles due to the oxidation of netallic core.

At low temperatures in a certain interval, it was experimentally found that in an unoxidized metal-insulator nanocomposite, the electrical conductivity changes with temperature in accordance with the law (6.23) for $n = 0.25$ (see Fig. 6.26a), which is valid in the case of hopping conductivity over localized states lying in a narrow energy band near the Fermi level.




a b

Fig. 6.26. Experimental temperature dependences of the electrical conductivity of $\sigma(T)$ for unoxidized (a) and oxidized (b) films $(FeCoZr)_x(Al_2O_3)_{1-x}$ [28-31]: 1 - $(FeCoZr)_{31}(Al_2O_3)_{69}$; 2 - $(FeCoZr)_{38}(Al_2O_3)_{62}$; 3 - $(FeCoZr)_{43}(Al_2O_3)_{57}$; 4 - $(FeCoZr)_{46}(Al_2O_3)_{54}$; 5 - $(FeCoZr)_{48}(Al_2O_3)_{52}$; 6 - $(FeCoZr)_{50}(Al_2O_3)_{50}$; 7 - $(FeCoZr)_{63}(Al_2O_3)_{37}$

The electric conductivity in this case can be described by the well-known Mott equation for hopping conductivity with a variable activation energy (VRH model) [36]:

$$\sigma(T) = \sigma_O \cdot \exp\left(-\frac{T_{0M}}{T}\right)^{1/4}, \quad (6.24)$$

where

$$\sigma_{0M} = e^2 \cdot R^2 \cdot \nu_{ph} \cdot \mathrm{g(E_F)}, \quad (6.25)$$

$$T_{0M} = \frac{16}{a^3 \cdot k \cdot g(E_F)}. \quad (6.26)$$

Here $e$ is the electron charge, $R$ is the jump length, $\nu_{ph}$ is the phonon frequency causing the electron jumps by the localized states, $T$ is the absolute temperature, $g(E_F)$ is DOS in the vicinity of the Fermi level, $a$ is the localization radius of the electron wave function, $k$ is the Boltzmann constant.

In the oxidized metal-insulator film nanocomposites, the temperature dependence of electrical conductivity follows the law (6.27) with $n$ = 0.5 (see Fig. 6.26*b*), which indicates the realization of VRH conductance of the Shklovsky-Efros type [36]. The temperature dependence of the conductivity in this case is described by the relations

$$\sigma(T) = \sigma_{0SE} \cdot \exp\left(-\frac{T_O}{T}\right)^{1/2}, \quad (6.27)$$

where

$$\sigma_{0SE} = e^2 \cdot R^2 \cdot \nu_{ph} \cdot \mathrm{g(E_F)}, \quad (6.28)$$

$$T_{0SE} = \frac{16}{a^3 \cdot k \cdot g(E_F)}. \quad (6.29)$$

Linearization of the temperature dependence of the conductivity for the studied x contents in nanocomposites in Fig. 6.26 in the Mott coordinates, $\ln(\sigma/\sigma_0) \propto (1/T)^n$ allows to confirm the VRH mechanism of carrier transport with the specific $T_0$ and $\sigma_0$ values, and extracted from them the effective characteristics of the electrons participating in the hopping transport (the density of localized states at the Fermi level $g(E_F)$, the localization radius $a$, etc.).

An example of the $g(E_F)$ calculation for several types of metal-insulator nanocomposites is shown in Fig. 6.27. If we extrapolate the *g(x)* dependence to the concentration of the metal phase corresponding to the percolation threshold, and then use relation

$$g(E_F)_{x_C} = g_0 \cdot x_C, \quad (6.30)$$

we can estimate the effective DOS at the Fermi level of metallic granules $g_0$[37]. As can be seen, the obtained values of DOS have very high magnitudes because the sources of the localized states can be not only the dangling bonds in the matrix (due to defects in dielectric matrix), but also the interface of the matrix-granule, as well as the electronic states of the granules themselves. With increasing the concentration of the metallic phase and approaching the percolation threshold, DOS tends to the values which are characteristic for amorphous metal alloys [38]. With an increase in the concentration of the dielectric phase, a decrease in the density of states at the Fermi level is observed, the value of which approximates the typical values for amorphous semiconductors obtained by sputtering from the gas phase [39].

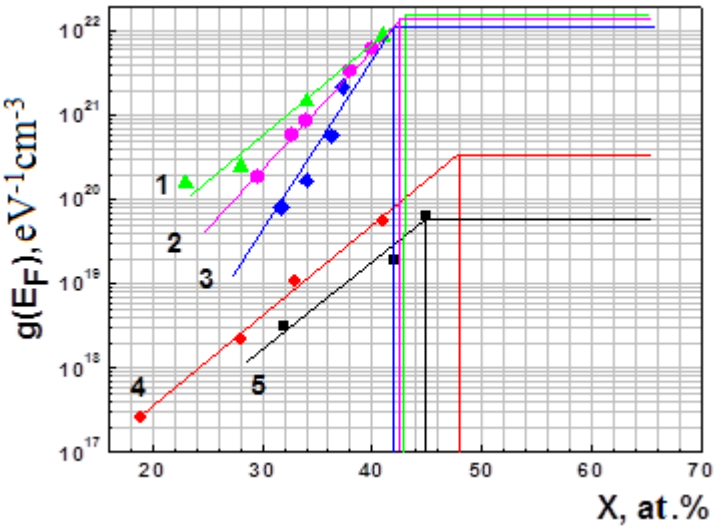



Fig. 6.27. The Fermi level DOS $g(E_F)$ dependences on the concentration of the metal phase in nanocomposites $(Fe_{45}Co_{45}Zr_{10})_x(SiO_2)_{100-x}$ (1), $(Co_{45}Fe_{45}Zr_{10})_x[Pb_{0,81}Sr_{0,04}(Na_{0,5}Bi_{0,5})_{0,15}(Zr_{0,575}Ti_{0,425})O_3]_{100-x}$ (2) $(Fe_{45}Co_{45}Zr_{10})x(Al_2O_3)_{100-x}$ (3), $(Co_{41}Fe_{39}B_{20})_x(SiO_2)_{100-x}$ (4) and $(Co_{84}Nb_{14}Ta_2)_x(SiO_2)_{100-x}$ (5)

A characteristic feature of the obtained results is the dependence of $g(E_F)$ on the *x* content: it grows in nanocomposites with granules in sequence CoNbTa → CoFeB → CoFeZr (Fig. 6.27). It should be noted that for composites with the same matrix, the effective DOS values have a very close slopes for the concentration of the metal phase dependences, while the composition of the granules leads to a significant change in the value of $g(E_F)$.

A number of studies have shown that some metallic-insulator nanocomposites exhibit a temperature conductivity for $T$ < 100 K as a law (6.27) with $n$ = 0.5. Such dependences, for example, were observed in Fe-$SiO_2$ nanocomposites [41, 42,43].

There are a number of alternative models explaining the law (6.24) with $n$ = 0.5. In the Sheng-Abeles model of thermally activated tunneling, it is assumed that charge transport is carried out by electrons tunneling directly from one granule to another through dielectric barriers [34, 40, 44,45]. According to this model, before the percolation threshold, the conductivity depends exponentially on the height and thickness of the dielectric barrier of layers separating metallic granules

$$\sigma \propto \exp(-2(2\pi/\hbar)(2m\phi)^{1/2}s) \ , \qquad (6.31)$$

where $\hbar$ is the Planck constant; $m$ is the effective mass of the electron; $\phi$ is the effective height of the barrier, $s$ is the thickness of the dielectric barrier, equal to the shortest distance between the boundaries of the metallic granules. The temperature dependence of the conductivity in this model is related to the change in the charge of the granules that arises when tunneling of electrons. Since the granules sizes $D$ in nanocomposites are small (< 10 nm), when one electron tunnels from one granule to another, the electrical neutrality of these granules is broken, so that the tunneling electron needs to overcome the energy barrier, the value of which is proportional to the energy of $E_C$ needed to create a dipole from the pair of electrically charged neighboring granules. Thus, in spite of the fact that tunneling is the mechanism of electric transport, only electrons thermally activated above the barrier due to the Coulomb interaction can participate in this process. This phenomenon is usually called the "Coulomb blockade". The number of thermally activated electrons possessing the required (Coulomb) energy is proportional to the Boltzmann factor [44,45]

$$N \sim \exp(-E_C/kT), \tag{6.32}$$

where $k$ is the Boltzmann constant; $T$ is the temperature. The Coulomb energy of the granule is assumed to be equal to

$$E_C = \frac{e^2}{\varepsilon} \frac{s}{D^2(1/2+S/D)}, \tag{6.33}$$

where $\varepsilon$ is the permittivity of the matrix; $D$ is the diameter of the granule. In the Sheng-Abeles model, it was postulated that the granules are spherical, and the minimum distance between them is proportional to the diameter of the granules, i.e. s/$D$ = *const* for a given ratio of the metallic and dielectric phases [34].

Taking these two processes into account (tunneling of electrons between granules and the need for an electron to acquire an energy $E$ such that $E_C \leq E$) allowed the following analytical expression for the electrical resistance of a nanogranular composite:

$$\rho = \rho_0 \left(\exp\left(2\sqrt{C/k_B T}\right)\right), \tag{6.34}$$

where

$$C = (2\pi/h)(2m\phi)^{1/2} sE_C \tag{6.35}$$

is the activation energy of the tunneling.

Despite a good qualitative description of the temperature dependence of the conductivity in the framework of the model of thermally activated tunneling in [41-43], the quantitative estimates obtained by several orders of magnitude with respect to certain experimentally determined values. In addition, when using different modifications of thermal activation tunneling models of the Sheng-Abeles type, the impossibility of explaining the power-law temperature dependence of the electrical resistivity of composites both in the range above 200 K (Figure 6.28) and on cooling to helium temperatures remained a serious issue [42-46]. Moreover, in a number of experiments [47, 48] there is an absolute disagreement with these models, expressed in a kind of "saturation" of the electrical resistivity (the absence of exponential growth of the electrical resistivity with decreasing temperature) when helium temperatures are reached, especially for metal-insulator composites with a metal phase concentration approaching the percolation threshold [49-54]. As shown in [55], the reason for this discrepancy is related to ignoring the role of the dielectric matrix as a medium that can also contribute to electronic transport.

The role of the dielectric matrix in the mechanisms of electric transport may be due to several reasons. The most important of them is its amorphization during deposition from the gas phase on cooled substrate [34, 40, 54], which causes the existence of a large number of localized states in the forbidden band of the insulator [56-60], over which a so-called hopping Mott conductivity (Figure 6.29) [55, 59, 60]. Another source of localized states in the dielectric matrix like silicon oxide is its non-stoichiometric composition in the form of oxygen depletion (the presence of oxygen vacancies) [61,62]. This leads to an increase in the role of dangling bonds of

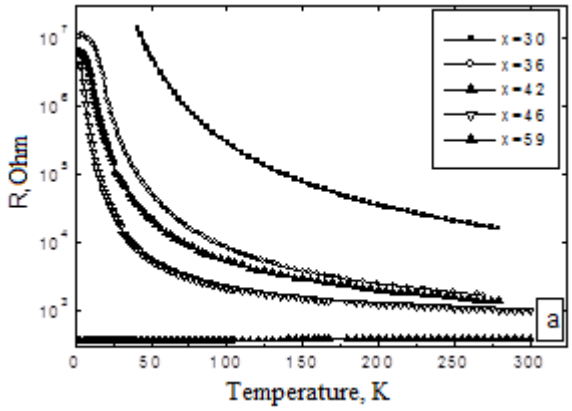



Fig. 6.28. Temperature dependence of the electrical resistivity of granular composites $(Co_{45}Fe_{45}Zr_{10})_x(SiO_n)_{100-x}$. The content of the metallic phase is shown in the figures in at.%.

silicon, giving local energy levels in the band gap of an amorphous dielectric, which can be either free or occupied by free charge carriers [63]. Since composites are obtained by deposition from the gas phase, the third possible cause of the appearance of localized states in the dielectric matrix, affecting the electrical conductivity of amorphous dielectrics, are metal atoms that are "dissolved" in the dielectric matrix and do not join the granules.

Thus, the higher the proportion of the metallic phase in the composite (the closer to the percolation threshold), the more localized states should be in its dielectric matrix. The results of the calculation have showed that with an increase of the metallic phase concentration in the composite, the average number of localized states between the granules decreases, and this dependence is practically linear.

It is clear from what has been said that the conductivity in such composites can not be uniquely determined only by tunneling of electrons between metallic granules, as is the case for the low concentration of the metallic phase ($x < 40$ at.%), Fig. 6.28 and 6.29. As the experiments show, the higher the metal concentration in the composite, the lower the temperatures, when the role of the dielectric matrix begins to appear. It is shown that, since the amorphous dielectric matrix of nanocomposites contains localized states, the most probable additional charge transfer mechanism is the Mott hopping conductivity over the matrix [55,59,60].

Attempts were made to construct more perfect models describing the electrical properties of granular composites near the percolation threshold. These approaches were based on calculations of the critical routes of conduction electrons passing through a random net of active resistors simulating a granular structure [64]. An estimate of the average values of the Coulomb energy, the distances between the granules and their diameters (these parameters are considered independently of each other) and the coordination number (the number of nearest granules between which tunneling is possible) make it possible to obtain a more accurate expression for the conductivity in granular composites [65] . The calculated values are commensurate with the experimental values, but the basic concept of the model of thermal tunneling remains unchanged. For example, according to the thermally activated VRH model of Shklovsky-Efros [35, 36, 59, 60], for nanogranular media, the law (6.25) with $n = 0.5$ can be due to the presence of a Coulomb gap in the density of localized states in the vicinity of the Fermi level when $g(E_F)$ in granules becomes zero according to the law $g(E_F) \propto (E - E_F)^2$. This gap is formed in disordered systems as an effect of the Coulomb interaction between the granules [66, 67]. In this case, it is essential to

have a sufficiently large random potential, which ensures the charge exchange of initially neutral granules in the ground state of the system [68], for example, due to the spreading of the granule size [69]. The gap is formed when the value of the random potential becomes comparable with the charge energy of the particles, as a consequence of which is the temperature limitation of the application of this model.

Thus, as shown by the analysis of the temperature dependences of the electrical resistivity of granular nanocomposites, electron transport by localized states (hopping conductivity) is possible along with tunneling of electrons between metallic granules. It should be emphasized that the VRH conductivity in amorphous dielectrics is nothing more than the tunneling of an electron from a localized state to an unoccupied state without activation into a delocalized zone [3, 35, 70].

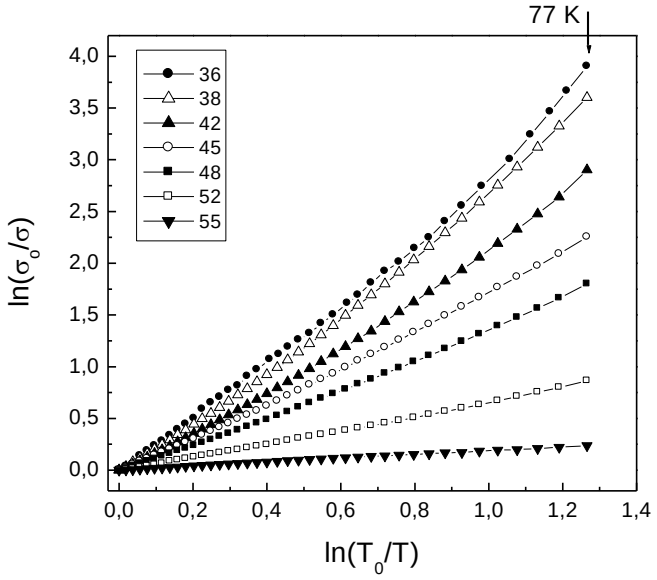



Fig. 6.29. Temperature dependences of the normalized conductivity of the samples $(Co_{41}Fe_{39}B_{20})_x(SiO_n)_{100-x}$ at different concentrations of the metal phase *x* in at.%, which is shown in the inset. $\sigma_0$ is the value of electrical conductivity at a temperature $T_0$ = 300 K

Tunneling (hopping) by localized states can be both elastic and inelastic. In the case of granular composites, inelastic tunneling is more significant than for a single tunnel barrier between two metallic granules. This is due to such a characteristic feature of nanoscale metal granules as the discreteness of the energy levels due to a small number of interacting atoms.

In the energy region near the Fermi level, where the main effects of electron interactions are realized, the average distance between the levels is defined as $\delta = 1/N(0)V$, where $V$ is the volume of the granule, $N(0)$ is the single-spin energy density of states at the Fermi level for the unit volume [3, 35]. In the case of a 10 nm for the size of metal particle, when it contains about $10^5$ conduction electrons, the difference between neighboring energy levels corresponds to about $8{,}63 \cdot 10^{-4}$ eV [71]. With a decrease in the diameter of the granules, this difference between the levels will be even more. Therefore, the probability that the energy of the electron in the granule at the energy level $E_i$ corresponds to the energy of the free localized state will be extremely low. Consequently, the transfer of an electron by localized states can only be inelastic, accompanied by a change in its wave vector and energy.

To explain deflection mechanisms of electric transport between the laws of Sheng-Abeles, Mott, Shklovsky-Efros, etc. and experimental temperature dependences of the electrical resistivity in the range from liquid helium to room temperatures a theoretical model of inelastic tunneling through the amorphous dielectric layers was developed [49 , 50, 72, 73-75]. According to this model, the observed experimental temperature dependences of the conductivity were explained by resonant tunneling through a chain of localized states in the amorphous insulating layer between the tunnel junctions. By virtue of a sufficiently large number of atoms in the granules ($N \sim 1000$), this theoretical model can be applied to determine the nature of the electron transportr between the metal particles in the granular structures. This also contributes to the fact that the Coulomb energy of the granules may result in the formation of one-dimensional quasi-stable conduction channels.

The sources of localized states are usually considered to be structural defects in the matrix and in the granule/matrix interfaces, which lead to a sharp increase in the tunneling transparency of the barrier. According to the model [74, 75], the processes of inelastic resonance tunneling in channels containing localized states near the Fermi level with an energy spread of the order of $kT$ becomes to play the main role in the conduction mechanism. The temperature dependence of the conductivity in a channel containing $n$ impurities has a power-law form

$$\sigma_n = P\left(\frac{\Lambda^2}{\rho_0 c^5}\right)^{(n-1)/(n+1)} \frac{(ga^2n^2l)^n T^{\gamma_n} E^{\beta_n}}{al} \exp\left[\frac{-2l}{a(n+1)}\right], \tag{6.36}$$

where $a$ is the radius of the localized state; $l$ is the average distance between the granules; $\gamma_n = n - 2/(n+1)$; $\beta_n = 2n/(n+1)$; $P$ is the coefficient; $\Lambda$ is the constant of the deformation potential; $\rho_0$ is the density of the matrix material; $c$ is the speed of sound; $g$ is the density of localized states; $E$ is the depth of occurrence of the localized state in the barrier region. The average electrical conductivity between the granules is determined by the sum

$$\sigma^{(gr)} = \sum_n \sigma_n. \tag{6.37}$$

With increasing temperature, inelastic channels with a large number of impurities are included. It was shown in [74, 75] that there exists a temperature interval $[T_n, T_{n+1}]$ in which the single term $\sigma_n$ gives the largest contribution to (6.34). In this temperature range, the temperature dependence of the electrical conductivity $\sigma^{(gr)}$ can be approximated by a power-law dependence with the power $\gamma_n$ at $T > T^*$, where $T^*$ is given by the relation

$$\ln(T^* gal^2) \approx -\left(\frac{l}{a}\right)^{1/3}. \tag{6.38}$$

In the expression (6.38), it is necessary to take into account several channels, so that the conductivity of $\sigma^{(gr)}$ will be determined by the contribution of channels with the number $n$ near the average number $\langle n \rangle$ of localized states between the granules

$$\ln \sigma^{(gr)} \infty - \left[\frac{2l}{a} \ln\left(\frac{1}{gal^2T}\right)\right]^{1/2}. \tag{6.39}$$

At certain radii of the localized states $a$ and the distances between the granules $l$, falling in the range of values near $gal^2T \approx 1$, in the first approximation of the power expansion by $1/T$ in (6.39) $\sigma^{(gr)}$, the temperature dependence (6.25) is observed with $n = 0{,}5$.

Increasing the distance between the granules leads to an increase in the number of channels and the number of impurities in the channels. For $n \to \infty$, the total conductivity passes through all

channels from the mode of resonant tunneling to the hopping conductivity regime, determined by the Mott law (6.27) with n = 0.25.

In the granular structure below the percolation threshold, the total conductivity of the structure $\sigma$ is determined by the tunneling conductivity between the granules $\sigma^{(gr)}$, or, if the granules form bounded conducting clusters, then the conductivity of the clusters and the conductivity between the clusters $\sigma^{(gr)}$. Since the conductivity of the clusters is much higher than the tunneling conductivity, in the first approximation we assume that the conductivity $\sigma$ of the granular structure in the temperature interval $[T_n, T_{n+1}]$ will have a power dependence on the temperature, which is determined by $\sigma^{(gr)}$ with $n = \langle n \rangle$ is the number of localized states averaged over the whole structure in the tunnel channels between the granules. Approximating the temperature dependences of the conductivity, obtained experimentally, by power-law dependences with a power of $\gamma$ and taking into account (6.36), it is possible to determine the average number of localized states $\langle n \rangle$ that participate in electron transport in the granular structure at a given temperature [76]

$$\langle n \rangle = \frac{1}{2}\left[\gamma - 1 + (\gamma^2 + 2\gamma + 9)^{1/2}\right]. \qquad (6.40)$$

Fig. 6.30 shows the temperature dependences of the normalized conductivity of composites $(Co_{41}Fe_{39}B_{20})_x(SiO_n)_{100-x}$ at different concentrations of the metal phase in semilogarithmic coordinates. The analysis of the obtained experimental dependences have shown that in a wide range of temperatures (~ 200-300 K) these dependences really satisfy the power law. This made it possible to determine the value of the degree $\gamma$ of the temperature dependences of the conductivity and calculate from (6.37) the average number of localized states $\langle n \rangle$ in the tunnel channels between isolated conducting granule clusters (Fig. 6.30).

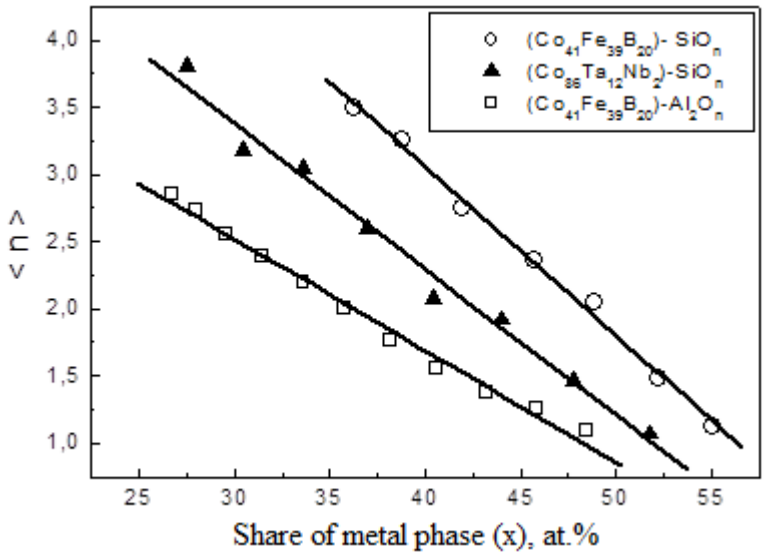



Fig. 6.30. The concentration dependence of the average number of localized states $\langle n \rangle$, taking part in the transport of an electron between two isolated granules [76]

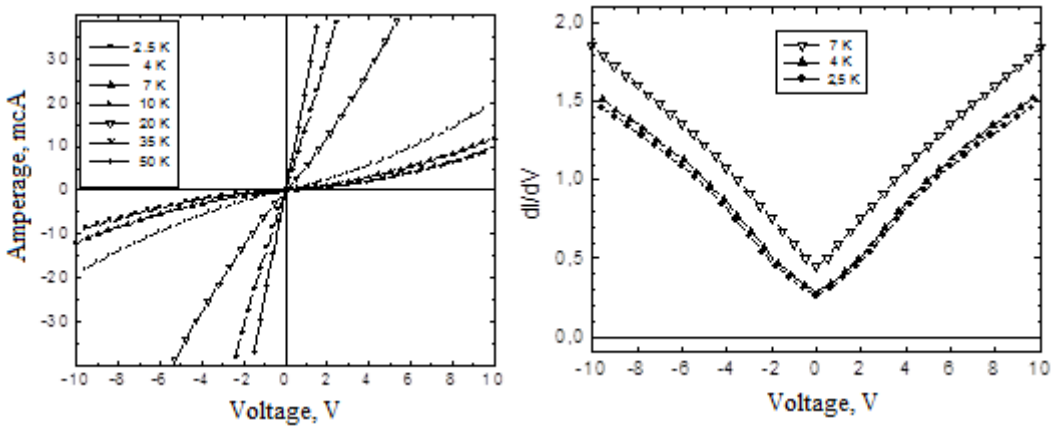
Analysis of the temperature dependences of the electrical resistivity for a number of nanocomposites has shown that transition from a power law or laws of the type (6.24), (6.28) or (6.34) to the ordinary exponential law is observed at temperatures above 150-200 K, and sometimes above room temperature

$$\rho = \rho_0 \exp\left(\frac{w}{kT}\right), \tag{6.41}$$

where $w$ is the activation energy of the jump, whose magnitude is of the order of the width of the zone of localized states [77]. For example, from the slope of the temperature dependences for the composites $(Co_{45}Fe_{45}Zr_{10})_x(Al_2O_3)_{100-x}$, shown in Fig. 6.15, it is possible to obtain activation energies $w$ = 0,030 ± 0,005 eV for $x \approx 34$ at. % (Curve 1), $w$ = 0,066 ± 0,007 eV for $x \approx 28$ at. % (Curve 2), and $w$ = 0,09 ± 0,01 eV for $x \approx 23$ at. % (Curve 3). Similar dependences were observed in nanocomposites $(Co_{45}Fe_{45}Zr_{10})_x(Al_2O_3)_{100-x}$ after step annealing, see paragraph 8.7.5.

As indicated above, studies of a number of granular composites [72] show the presence of a saturation region of the electrical resistivity at low temperatures (Figure 6.28). These experimental facts contradict the known models of electric transport and in [72] they were attributed to the physical features of the composites, namely, to their nanostructured character and thermal activation tunneling mechanism of carrier transport.

It was found that the temperature at which the exponential growth of the electrical resistivity is interrupted is weakly correlated with the electrical resistivity of the samples. On the other hand, the value of this temperature increases proportionally to the increase in the fraction of the metallic phase in nanocomposites [75]. In addition, it was shown that the observed "saturation" of the electrical resistivity of composites is accompanied by features such as the highly nonmonotonic character of the I-V characteristics, and also the significant dependence of the type of I-Vs on temperature (Figures 6.30 - 6.32) [75].



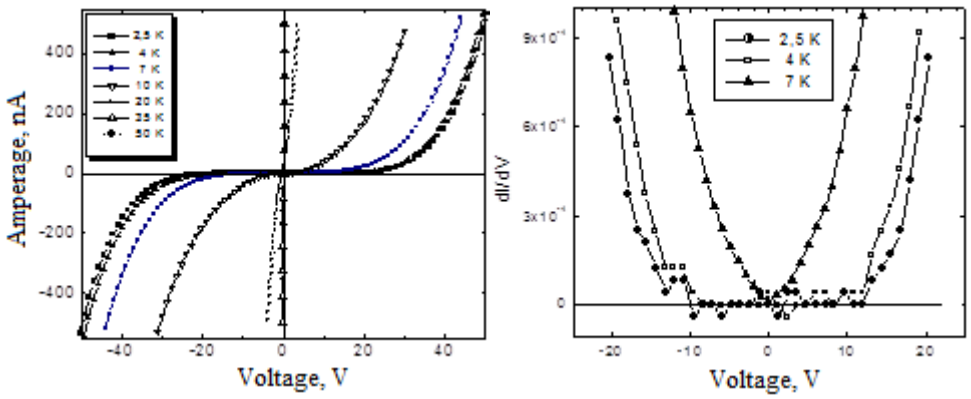

*a* *b*

Fig. 6.31. I-Vs of nanogranular composite systems $(Co_{41}Fe_{39}B_{20})_{46}(SiOn)_{54}$, measured at different temperatures (*a*), and the current derivative with respect to voltage (*b*)

Note two characteristic features in the dependences obtained are obvious:

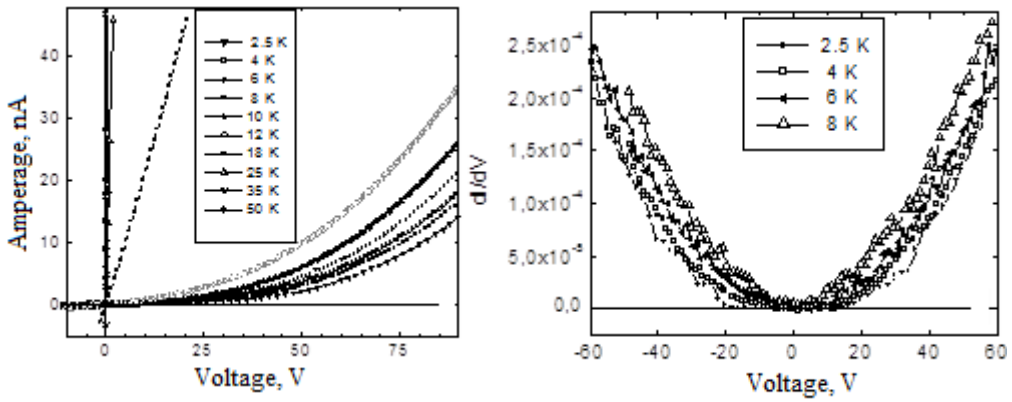
1. In the low-temperature region (< ~15 K), the I-Vs characteristics are extremely nonlinear and there is a range of voltage values (lower than a certain threshold value) at which the current through the composite practically does not flow. This follows from an analysis of the magnitude of the *dI/dV* derivatives, showing that there is an interval of voltage values for which the electric current through the sample really does not pass, i.e. the conductivity is completely absent.

2. Both the shape and the parameters of the I-Vs depend to a large extent on the fraction of the metallic phase contained in the composite. Indeed, if a composite is close to the percolation threshold (Figure 6.31), the conductivity in the low-voltage region (1-2 V) is low, but nonzero. An increase of dielectric phase in the material leads to the appearance of a range of voltage values at which the conductivity of the composites becomes zero (Figures 6.32 and 6.33). Moreover, for the same temperature value, the voltage threshold, at which conductivity is nonzero, is proportional to the fraction of the dielectric phase. Thus, for $T$ = 2.5 K, the threshold voltage is 10-12 V in the composite $(Co_{41}Fe_{39}B_{20})_{43.5}(SiOn)_{56.5}$ and 16-18 V for $(Co_{41}Fe_{39}B_{20})_{36}(SiOn)_{64}$. With an increase in the fraction of the dielectric phase (in the comparative analysis of the I-V characteristics for different composites), the temperature at which conductance through the sample take places increases.




*a* *b*

Fig. 6.32. I-Vs of nanogranular composites $(Co_{41}Fe_{39}B_{20})_{43.5}(SiO_n)_{56.5}$, measured at different temperatures (*a*), and the current derivative with respect to voltage (*b*)




*a* *b*

Fig. 6.33. The positive branches of the I-Vs for the nanogranular composite $(Co_{41}Fe_{39}B_{20})_{36}(SiO_n)_{64}$, measured at different temperatures (*a*), and the current derivative with respect to the voltage (*b*)

The results obtained can be explained with the help of ideas about the Coulomb blockade. In the thermal activation tunneling model of Sheng-Abeles, the exponential growth of the electrical resistivity in the low-temperature region is attributed to the fact that the number of electrons with

sufficient energy to overcome the Coulomb blockade decreases exponentially with temperature in accordance with the expression (6.41) [40, 44, 45, 53]. When cooling the composite to temperatures close to 0 K, the number of carriers capable of participating in the electric transport also tends to zero. In other words, when the value of $kT$ is smaller than the Coulomb activation energy of tunneling ($kT < C$), the electron is unable to leave one granule and tunnel into the other, i.e. the process of electric transport is practically not realized. With increasing temperature, the value of $kT$ increases and accordingly the number of electrons thermally excited so that they are able to participate in the corporate charge transfer through the composite, i.e. the conductivity rises. This is fully illustrated by the results presented in Fig. 6.34, showing the values of the current flowing through the composites of the system $(Co_{41}Fe_{39}B_{20})_x(SiO_n)_{100-x}$, at a constant value of the applied voltage equal to 10 V.

It is obvious that the metallic phase content in the composite affects the magnitude of the current, since just it ultimately determines the size of the granules and, consequently, the value of the Coulomb energy. For the sample $(Co_{41}Fe_{39}B_{20})_{46}(SiO_n)_{54}$, which composition is close to the percolation threshold, below 10 V there is no sharp decrease in current (by several orders of magnitude). In this composite, the granule size is known to be maximal within the framework of this system (~ 7 nm) and granules are not absolutely isolated in the dielectric matrix, but form clusters consisting of several granules. For such a structure, the values of the Coulomb activation energy will be small, and even at T = 2.5 K the conductivity of such a composite will be different from zero (Figure 6.34), since $kT$ is greater than the activation energy. For composites with a smaller fraction of the metal phase, on the contrary, when approaching low temperatures, the current decreases sharply, approaching practically zero at T = 2.5 and 4 K (Figures 6.32 and 6.33). This is the consequence of the Coulomb blockade. Apparently, the results obtained are a consequence of a small scattering of granules diameters (small dispersion of activation energy values, respectively) and complete electrical isolation of the granules from each other.

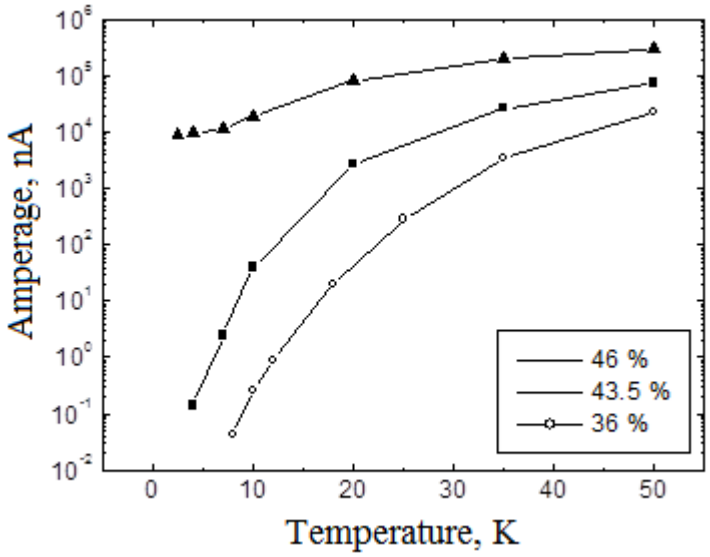



Fig. 6.34. Dependence of the current flowing through the composites $(Co_{41}Fe_{39}B_{20})_x(SiO_n)_{100-x}$, on the temperature at a constant value of the voltage U = 10 V applied to the sample. The concentration of the metallic phase of the composites ($x$) is shown in the figure in at.%.

The activation of electrons above the Coulomb barrier can be realized not only by thermal energy, but also by the energy of the electric field. The potential difference that actually arises

between two neighboring granules, when an external field is applied to the sample, can be interpreted as a displacement of the Fermi levels of two neighboring granules. Moreover, this displacement can be equal to or even exceed the height of the Coulomb barrier. Under such conditions, there is no need for thermal activation of electrons. This means that tunneling can occur at arbitrarily low temperatures if the electric field strength is high enough and the field activation of the tunneling takes place [78]. The transition to field activation, or the appearance of an additional component in the electrical conductivity of composites due to field activation, can occur in principle at any temperature. Obviously, if the condition $kT \leq e\Delta V$ is satisfied ($e$ is the charge of the electron, and $\Delta V$ is the voltage drop between two adjacent isolated granules), the transfer will be determined to a greater extent by an electric field (field operation) than by thermal activation of electrons. In this case, the thermally activated conductivity model will not be fully adequate to the processes taking place in the composite.

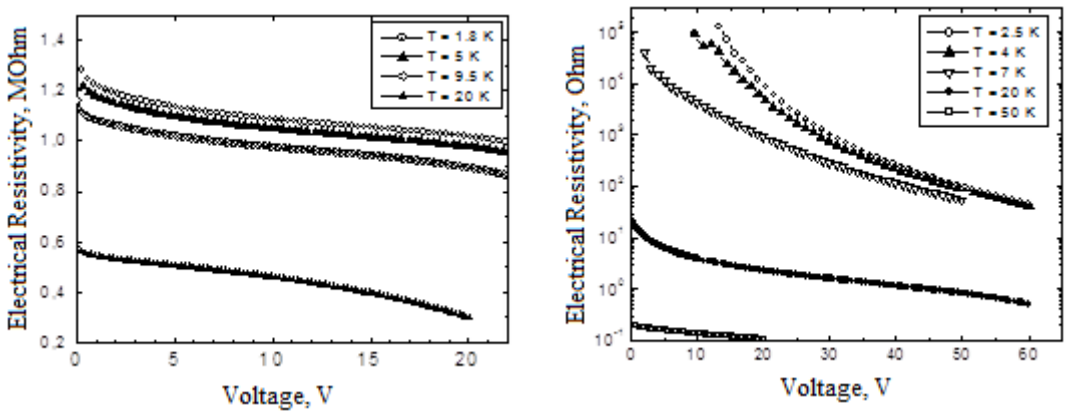
### 6.7.4. The influence of the measurement regimes on the temperature dependences of the electrical resistivity of amorphous nanogranular composites

As noted in the previous section, the nonlinearity of the I-V characteristics of nanogranular composites at low temperatures is interlaced by the transition from the thermal activation conductivity to the field-activation mode. This means that an increase in the value of the voltage applied to the sample at a constant temperature leads to a significant increase in current through the sample, since the conductivity in this case is initiated by an electric field. Thus, the form of the temperature dependence of the electrical resistance of nanocomposites at low temperatures will depend strongly on the DC/AC measurement voltage regime (constant or not) over the entire temperature range. For homogeneous materials, there is no fundamental difference between these two modes, however, in the case of granular nanocomposites, the mode selection determines the form of the I-V dependence.

Let us consider the effect of the measurement regime on the type of the obtained dependence od the electrical resistivity of a nanocomposite on temperature.

In the process of measuring the electrical resistance in the regime of constant current, cooling of the samples leads to an increase in the electrical resistivity due to the thermal activation mechanism of electric transport. In this case, the voltage applied to the sample will also increase with decreasing temperature (see Figure 6.35), since the flowing current remains constant. It follows from the above data that a decrease in temperature from 50 to 4 K leads to an increase in the voltage on the sample by almost 400 times. In this case, the voltage drop $\Delta V$ between the adjacent dielectric grains becomes much larger than $kT/e$.

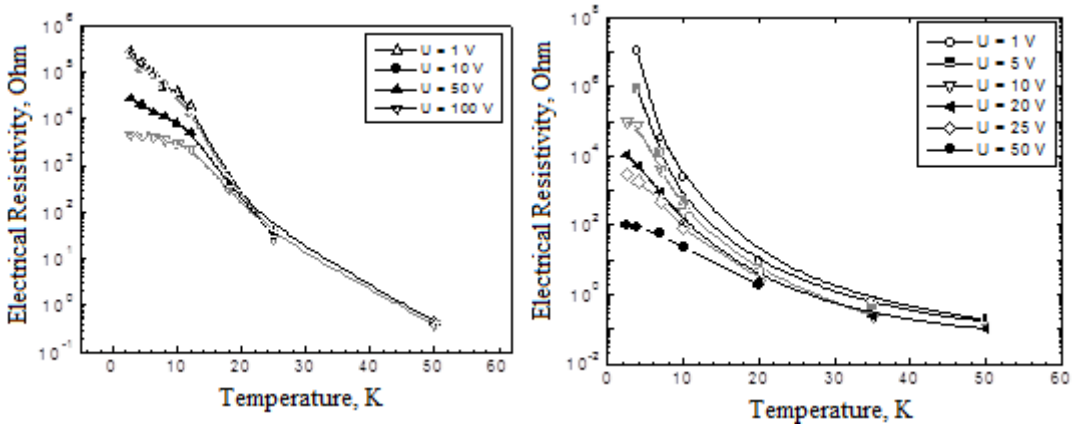
When the condition $kT/e < \Delta V$ is satisfied, the current flow regime becomes high-field [34, 44, 45, 78]. In this case, the electrons acquire the necessary energy sufficient to overcome the Coulomb




*a* *b*

Fig. 6.35. The dependence of the electrical resistivity in granular composites $(Co_{41}Fe_{39}B_{20})_{42}(SiO_n)_{36}$ (*a*) and $(Co_{41}Fe_{39}B_{20})_{36}(SiO_n)_{54}$ (*b*) on the voltage applied to the sample. Dependences are obtained as a result of recalculation of the corresponding I-V characteristics.

blockade not due to thermal activation, but owing to excitation by an electric field, as a result of which the conductivity turn out to be higher, than it could be in the case of a low-field regime. Consequently, the saturation electrical resistance observed when the cooling of the samples down to ~10 K is attributed to the transition to conductivity carried out by electrons excited by an electric field (see Fig. 6.36). It is obvious that in the low-temperature region, when a small voltage $e\Delta V < kT$ is applied, the Coulomb blockade regime is realized, and the electrical resistivity of the samples increases significantly with temperature decrease. In the case when the voltage applied to the samples during the measurements is large (tens of Volts), the Coulomb blockade does not occur and the growth rate of the electrical resistivity upon cooling is much less.




*a* *b*

Fig. 6.36. The temperature dependences of the electrical resistivity in granular composites $(Co_{41}Fe_{39}B_{20})_{42}(SiO_n)_{36}$ (*a*) and $(Co_{41}Fe_{39}B_{20})_{36}(SiO_n)_{64}$ (*b*) for different values of the voltages applied to the measured sample

Numerical estimates confirm that in real modes of measuring the temperature dependence of the electrical resistivity in nanocomposites at a constant current, a situation can occur in which $kT < \Delta Ve$. Suppose for simplicity that in our measurements the equipotential lines in the volume

of the composite are parallel to the plane of the electrodes located at a distance of 0.2 mm from each other. It is assumed that the granules are completely isolated from each other (composite is below the percolation threshold), and in each equipotential layer there is only one layer of granules. With an average granule diameter of 5 nm, the distance between them in the nanocomposite $(Co_{41}Fe_{39}B_{20})_{36}(SiO_n)_{64}$ corresponds to 3 nm (these values are obtained from SEM micrographs). If a voltage of about 10 V is applied between the electrodes of the sample with such geometrical parameters, then the voltage drop between neighboring granules will be ~0.4 mV.

Then from the equality $\Delta Ve = kT$ it follows that for a given $\Delta V$ the transition from the thermally activated to the high-field conductivity regime should be carried out at 5 K. In other words, if the voltage at which the measurement is performed is about 10 V, then the thermal activation mode of conductivity is realized at T > 5 K. Indeed, referring to the data presented in Fig. 6.32*b*, we can see that at temperatures below 5 K and voltages not exceeding 10 V, the current flowing through the sample is less than 0.05 nA, i.e. practically does not leak. This means that in this case the Coulomb blockade regime is realized. If the voltage is increased at the same temperature values, then a sharp increase in conductivity is observed, due to the field activation of the charge carriers.

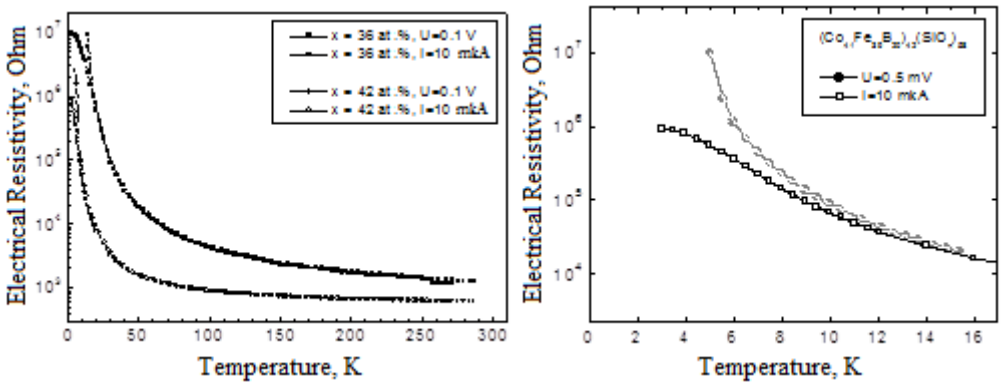
At $T > 5$ K (see Fig. 6.32*b*), the conductivity differs from zero even at the minimum values of voltage becouse the thermal energy is sufficient to overcome the Coulomb blockade. To realise the Coulomb blockade and the associated sharp growth of the electrical resistance in the composite, it is necessary to provide relatively high values of the voltage (tens or more Volts). If we measure electrical properties of composites with such voltage values at room temperature, then the samples will be simply destroyed because of the large densities of the electric current flowing through the thin-film material. Actually, therefore, when studying the temperature dependence of composites in the constant-voltage regime, the need to preserve the sample determines the use of low values of the electric voltage (less than one Volt), which at helium temperatures leads to a Coulomb blockade.

Let us now examine the mode of the electrical resistivity temperature dependence mesurement in the constant current regime, which leads to a sharp increase in the voltage applied to the sample (see 6.32 and 6.33). The lower the temperature, the higher the electrical resistance and the higher the voltage is needed to stabilize the current. Therefore, the cooling of the measurement object ultimately leads to the fact that, from the usual thermal activation mode of conductivity, it passes into the high-field regime. A consequence of this is the appearance of a saturation region on the temperature dependences of the electrical resistivity at T < 20 K.

Confirmation of the described model on the effect of the measurement regime on the actual temperature behavior of the electrical resistance in the samples of $(Co_{41}Fe_{39}B_{20})_x(SiO_n)_{100-x}$ is shown in Fig. 6.37. In fixed voltage mode, the selected value of V = 0.1V provides the Coulomb blockade mode (see also Figures 6.32 and 6.33), and in this mode the electrical resistance of composites really increases monotonically with cooling. Stabilization of the electric current value at a 10 μA requires voltages of the order of tens to hundreds of Volts at low temperatures ($T < 20$ K). It is obvious that at such values of the voltages in the composites the high-field conductivity dominates and electrical resistivity go to saturation.

Thus, in the case when the electrical resistivity of composites is measured in the regime of current stabilization, when cooling to liquid helium temperatures, the activation mechanism of the charge carriers is changed. At relatively high temperatures, a thermal activation mechanism is

realized and the temperature dependence of the electrical resistivity, taken in the regime of current stabilization, coincides with the dependence taken in the regime of voltage stabilization. At low




*a* *b*

Fig. 6.37. The temperature dependences of the electrical resistivity in granular composites $(Co_{41}Fe_{39}B_{20})_x(SiO_n)_{100-x}$, measured in the regime of current (*a*) or voltage stabilization (*b*)

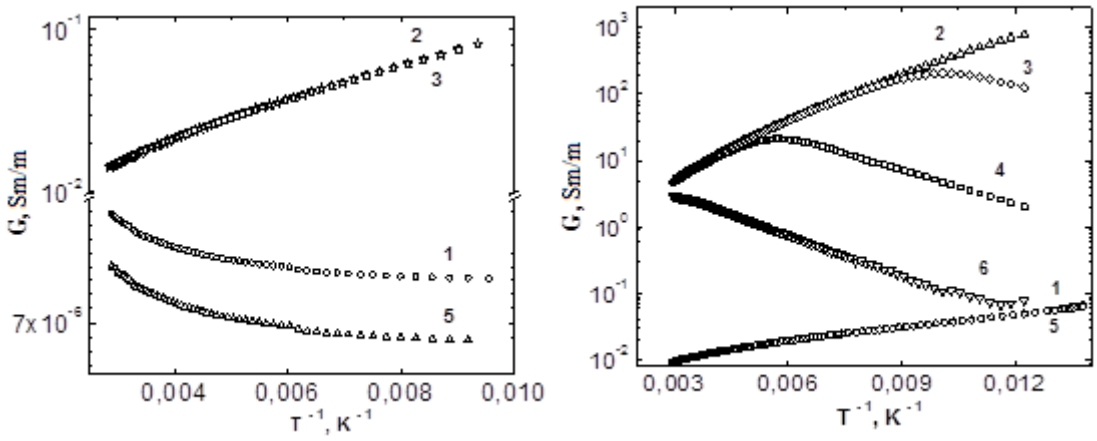
temperatures, the field mechanism of activation of electrons begins to dominate and the number of such electrons proves to be greater than if the thermal activation mechanism was realized. As a consequence, the conductivity of the samples decreases to a much lesser degree with respect to the thermal activation regime. In practice, as follows from Fig. 6.37, there is a discrepancy between the curves of the temperature dependence of the electrical resistance, taken in the regime of current or voltage stabilization.

### 6.7.5. Reactive impedance of inductive type in metal-insulator nanocomposites

In Section 6.7.3 we have made the detailed analysis of the temperature and concentration dependences of the electric conductivity in unoxidized and oxidized metal-insulator film nanocomposites in DC regime, which allowed to reveal the basic mechanisms and regularities of the transport of charge carriers. But the greater practical importance for electronics is the study of electric transport alternating current regime (impedance, admittance, phase shift, etc.). In describing the above nanocomposites by the method of impedance spectroscopy, it was shown that an inductive component of the impedance appears in the equivalent circuit of some composites beyond the percolation threshold (see 6.5). When the dimensions of the filler particles are changed to nanometer sizes, much of metal-insulator nanocomposites show the predominance of the inductive component over the capacitive in the reactive part of the impedance even below the percolation threshold.

In particular, such behavior is typical when there are shells of intrinsic oxides around each metallic nanoparticle inside the dielectric matrix [79]. This behavior has been termed the "negative capacitance" effect. Let us examine in more detail the behavior of the frequency and temperature dependences of the total impedance, as well as its real and reactive components in nanogranular composite films obtained by ion-beam sputtering of composite targets from the iron-cobalt-zirconium alloy and various types of dielectrics, followed by stepwise annealing ($CaF_2$, $Al_2O_3$, PZT, etc.).

As the experiments [31, 80-86] showed, the dependences of the real part of conductivity (admittance) $G(f,T)$ on frequency and temperature in oxidized and non-oxidized films $(FeCoZr)_x(Al_2O_3)_{1-x}$ had the following characteristic features. Figure 6.38 shows that low-




*a* *b*

Fig. 6.38. Experimental temperature dependences of the real part of the admittance $G(T)$ for unoxidized (*a*) and oxidized (*b*) films [87]: 1 - $(FeCoZr)_{55}(Al_2O_3)_{45}$ at a frequency of 1 kHz; 2 - $(FeCoZr)_{31}(Al_2O_3)_{69}$ at a frequency of 1 kHz; 3 - $(FeCoZr)_{31}(Al_2O_3)_{69}$ at a frequency of 10 kHz; 4 - $(FeCoZr)_{31}(Al_2O_3)_{69}$ at a frequency of 100 kHz; 5- $(FeCoZr)_{55}(Al_2O_3)_{45}$ at a frequency of 960 kHz; 6 - $(FeCoZr)_{31}(Al_2O_3)_{69}$ at a frequency of 960 kHz

frequency (at $f < 5$ kHz) temperature dependences of $G(T)$ in the Arrhenius coordinates in general show trends similar to the above-described temperature dependences of the electrical conductivity $G(T)$ in Fig. 6.15 in DC regime. The only difference was that for the unoxidized nanocomposite films, the slope of the $lg[\sigma] - (1/T)$ curves varied with $x > x_c \approx 45$at% [87], confirming the formation of a continuous percolation cluster with metallic conductivity when passing through the percolation threshold $x_c$.

The second peculiarity was that the behavior of the admittance at high frequencies ($f > 10$ kHz) in the samples of oxidized films (curves 4 and 6 in Fig. 6.27*b*) [87] followed the well-known Mott law [60] for hopping conductivity on an alternating current:

$$G \sim j_r \sim f^{\alpha}, \qquad (6.42)$$

where $j_r$ is the real part of the electric current density, $\alpha$ is the exponent with a value of $\leq 2$.

The third feature of the curves $G(f)$ in the high frequency range at temperatures of 77 - 300 K, as shown in Figure 6.28, is to observe the effect of "negative capacitance" which may be interpreted as the presence of the equivalent circuit of substitution films efficient inductive contribution $L$.

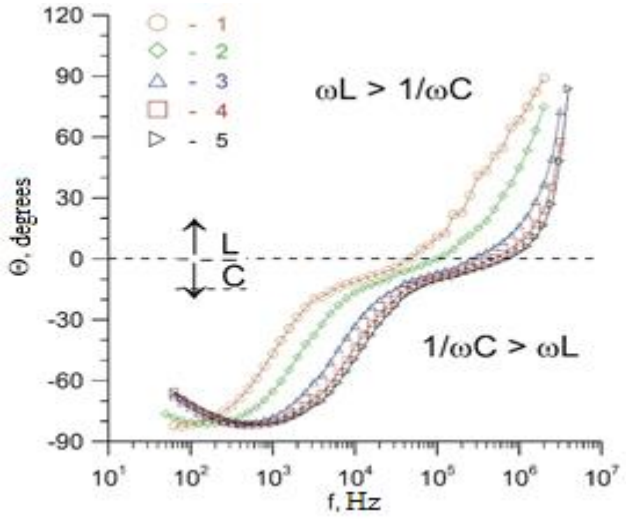



Fig. 6.39. Typical frequency dependences of the phase angle $\Theta$ for the oxidized film $(FeCoZr)_{43}(Al_2O_3)_{57}$ after annealing at 448 K for the temperatures T = 77 K (1), 123 K (2), 173 K (3), 233K (4 ) and 273 K (5) [79,87]

It should be noted that similar dependences are observed both in oxidized and unoxidized films $(FeCoZr)_x(Al_2O_3)_{100-x}$. The largest value of $L$ near room temperature was observed for oxidized films $(FeCoZr)_x(Al_2O_3)_{100-x}$ with $x$ = 38-50 at.% after annealing in the temperature range $T$ = 448 - 873 K [79].

According to the methods of processing the results of impedance spectroscopy of heterogeneous media, described in detail in [7], each type of frequency dependence of nanocomposites impedance corresponds to its equivalent circuit, that is, an electric circuit having the same dependence of impedance on frequency as the sample under study. The form of the equivalent circuit for replacing the film can be determined from the hodograph of the sample impedance, i.e. from the dependence of the imaginary part of the impedance of the composite on its real component, and also on the character of the dependence of the total impedance on the frequency and the sign of the phase shift between the current and the voltage. The study of the obtained frequency dependences of the real part of the impedance shows [31, 79, 80-86, 89] that equivalent substitution circuits for the samples under investigation do not correspond to the ideal limiting cases considered in [7], but include inductive components. This feature shows that the presence of $L$ below the percolation threshold in metal-insulator nanocomposites is not typical for binary composites consisting of micron-sized particles.

As in the case of nanocomposites with micron sizes of metallic particles, the inductive contribution of $L$ in equivalent substitution circuits for samples beyond the percolation threshold (with $x > x_c$) can be explained by the formation of a highly conducting network of metallic nanogranules in unoxidized films, which permeates the entire composite, and nanogranules with a "core-shell " structure in the oxidized films. In this case, the current flows through the regions of the smallest R, which, perhaps, provides the spiral-like character of the current paths. A schematic representation of complex self-crossing trajectories, as an analogue of a "coilless inductor", is shown in Fig. 6.29 [90].

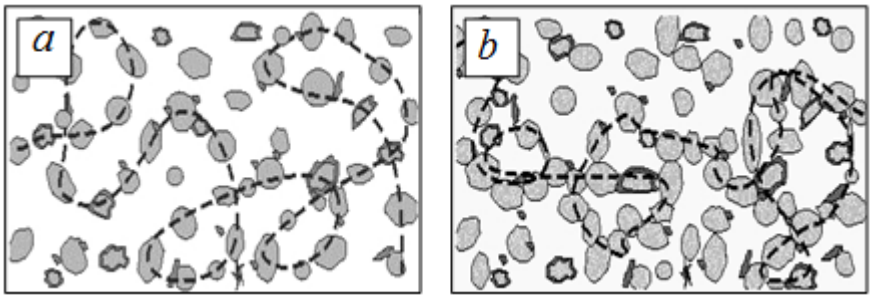



Fig. 6.40. Schematic representation of conductive routes (dashed lines) in nanocomposite films below (*a*) and beyond (b) the percolation threshold [90]

However, such a model can not explain the effect of "negative capacitance" in nanocomposites with $x < x_c$. As follows from the original Mott model [60], for the hopping conductance on alternating current, the exponent $\alpha$ in the ratio (6.39) for disordered crystalline substances should be equal to 0.8. However, in many experiments it was shown that the observed exponent $\alpha$ often exceeds this value, reaching the values $0{,}8 < \alpha \leq 2$ ([79, 83-88] and the literature references therein).

Below we present the improved model of hopping conductivity on alternating current which is based on taking into account the frequency dependence of the exponent $\alpha$ in the Mott ratio (6.39). This model is based on the ideas presented in [79, 80-86, 89, 91]. It allows one to explain the occurrence of the "negative capacitance" effect in nanostructured metal-dielectric films on the basis of violation of electrical neutrality for two neighboring metallic nanoparticles. As a result of the tunneling (hopping) of the electron between them and the polarization of the surrounding nanoparticle matrix happens. A schematic representation of the change in the state of metallic nanoparticles due to electron tunneling between neighboring nanoparticles is shown in Fig. 6.30.

In accordance with these ideas, the improved model is based on a number of basic provisions that take into account the following conditions and stages of the hopping between nanoparticles:

1. Nanoparticles in a composite film are considered as closely located neutral potential wells, in which electrons capable of jumping can be localized (Fig. 6.41*a*);

2. In the absence of an electric field between potential wells, tunneling (jumps) of electrons with identical probabilities in all directions can occur. The probability of a jump per unit time $p$ is a function of the temperature $T$, the distance between the wells $r$ and the activation energy of the jumps $\varepsilon$, and also depends on the concentration of the nanoparticles (wells);

3. In addition, it is assumed that the temperature region under consideration, in which the hopping conductivity on an alternating current is investigated, lies above the one where a VRH conductivity of the Mott type is realized at a constant current (since experiments were carried out at $T > 77$ K);

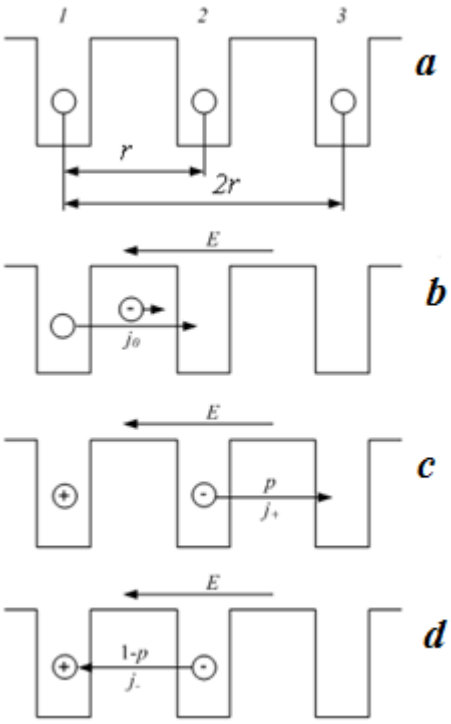



Fig. 6.41. A schematic representation of the state of potential wells (metallic nanoparticles) as a result of electron tunneling between neighboring nanoparticles [90, 91]

4. An external alternating electric field $E$, is defined as

$$E=E_0 \cdot \sin \omega t, \tag{8.43}$$

where $\omega = 2\pi f$ is the circular (cyclic) frequency of the alternating current, $E_0$ is the weak amplitude of the electric field strength, which leads to the appearance of a current. It does not change the probability of jumps per unit time $p(T, \alpha, \Delta E)$ (see, for example, [92]) from one potential well to another. Such a field leads only to the asymmetry of the jumps, which means the appearance of a component of the current density $j_0$, which coincides in phase with the compulsory electric field (Fig. 6.41b).

Let the application of a weak alternating field (6.43) causes an electron jump from well 1 to well 2 (Figure 6.41b) during the first half-period of field $E$, which leads to the appearance of the first real part of the electric current $j_0$, which coincides with the field $E$ and is defined by formula

$$j_0 = GE = GE_0 \sin \omega t. \tag{6.44}$$

After the electron jump from the neutral well 1 to the neutral well 2, a dipole arises (Figure 6.41*c*) and the electron remains in the well 2 for a certain time $\tau$, called the average lifetime of the electron in the well. After this time, two variants of the next electron jumps are possible. In the first of them, an electron with probability $p$ can jump to the next well 3 (in the direction of the field $E$, Figure 6.30*c*). For the second variant, the electron with probability $(1 - p)$ can return from well 2 to well 1 (in the direction opposite to field $E$, figure 6.41*d*).

If the electron can not return from well 2 to well 1 and jump into well 3, a second current density component appears

$$j_+ = \sigma E_0 p \sin[\omega(t-\tau)]. \quad (6.45)$$

When implementing the second variant, the electron returns from the second well to the first one after the expiration of the time τ, which leads to the appearance of the third component of the current density:

$$j_- = -\sigma E_0 (1-p) \sin[\omega(t-\tau)]. \quad (6.46)$$

This type of jump causes the disappearance of the dipole. The total current density will be described by the relation:

$$j = j_0 + j_+ + j_- = G\{\sin \omega t + p \sin[\omega(t-\tau)] - (1-p) \sin[\omega(t-\tau)]\}. \quad (6.47)$$

Since the components $j_+$ and $j_-$ are phase-shifted with respect to $j_0$, the total current density has both real and imaginary components.

The processes described above for hopping electron transport in a nanocomposite lead to an increase in the real component of the current density, which for the parallel equivalent circuit has the form

$$j_r = GE_0 \sin[\omega t(1 - cos\omega\tau + 2pcos\omega\tau)]. \quad (6.48)$$

Analysis of the formula (6.48) have shown that in the low-frequency region the current density ($j_{LF}$) does not depend on the frequency, and its magnitude is

$$j_r = j_{LF} = 2pGE_0 \sin \omega t. \quad (6.49)$$

In the high-frequency region, the current density ($j_{HF}$) also does not depend on the frequency *f* and is described by the relation

$$j_r = j_{HF} \cong 2pGE_0 \sin \omega t. \quad (6.50)$$

In the intermediate frequency range, a frequency dependence of the current density appears, which can be described by a formula of the type

$$j_r \sim f^{\alpha(f)}, \quad (6.51)$$

where, in contrast to the Mott law (6.42) [60], the exponent $\alpha$, as will be shown below, is a function of the frequency $\alpha(f)$.

We note that the above arguments are valid for the temperature region $T < 500$ K, when $p(T)\cdot\tau << 1$, i.e. only a small fraction of the potential wells are in a charged state.

It was shown in Ref. [93] that in the system of interacting potential wells the lifetime $\tau$ of an electron in the well is not a constant for all wells. Its value should depend on the distance between the wells and the dimensions of the nanoparticles. The authors of [93] showed that the distribution of the time $\tau$ can be normal or close to it. However, in a normal distribution, negative values of $\tau$ should appear, meaning an impossible scenario of electron jumps, when the electron returns from the first well to the first well before its jump from the first to the second, that is impossible. In this connection, as follows from physical considerations that the probability density $p$ exists only for positive values of the times $\tau$. This condition, in particular, corresponds to the Landau distribution in the Moyale approximation [94]:

$$f_L(\tau) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{t-\tau}{2\sigma} - \frac{1}{2}\exp\left(\frac{t-\tau}{\sigma}\right)\right), \tag{6.52}$$

where $\sigma$ is the standard deviation from the average lifetime of the electron in the well $\tau$.

Further studies to calculate the real component of the current density by the formula (6.45) were performed by numerical methods using computer simulation using the Landau distribution (6.49). Examples of the results of calculating the real part of the current density $j_r$ from the product $(f\cdot\tau)$ for given values of the electron hopping probability are shown in Fig. 6.31 for the jump probability values $p$ = 0.01; 0.001 and 0.0001.

Using the dependences $\alpha(f\cdot\tau)$ for different values of the jump probabilities $p$, the maximum values of the frequency coefficient $\alpha_M$ were determined for different values of the hopping probabilities $p$, which is presented in Fig. 6.44.

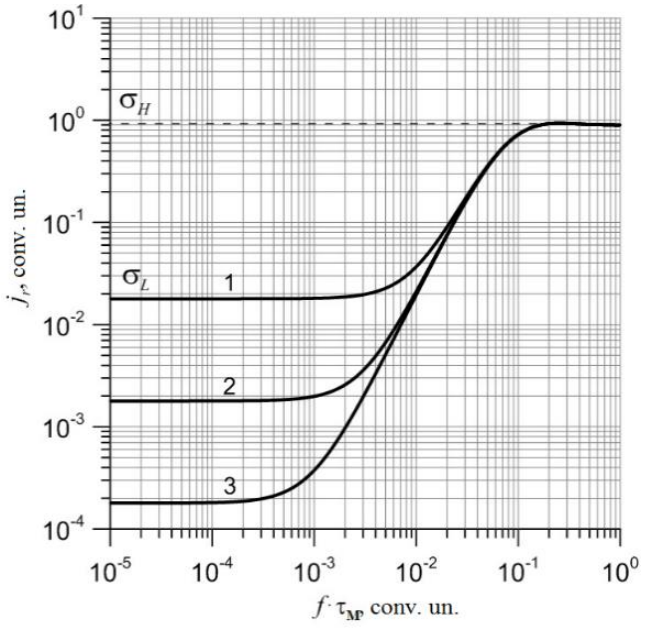


Fig. 6.42. The calculated dependence of the real component of current density $j_r$ on the product of frequency $f$ and time $\tau_M$: 1 – $p$ = 0,01; 2 – $p$ = 0,001; 3 – $p$ = 0,0001. Computer simulation for the Landau distribution [77, 93].

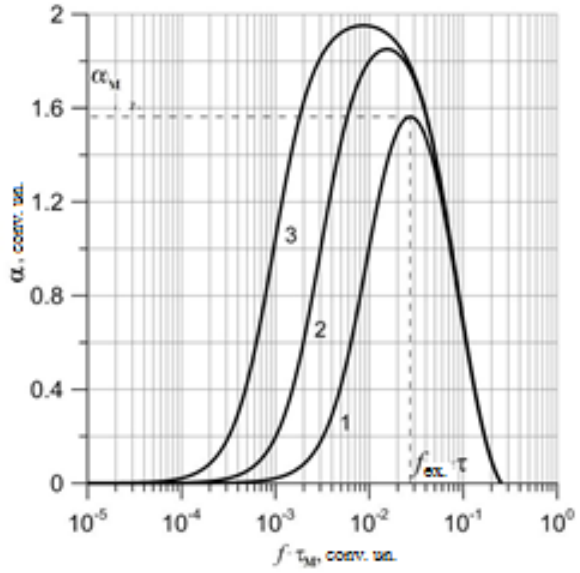



Fig. 6.43. The calculated dependence of the frequency exponent $\alpha$ as a function of the product of the frequency $f$ and the time $\tau_M$ [79, 95]: 1 – $p$ = 0,01; 2 – $p$ = 0,001; 3 – $p$ = 0,0001

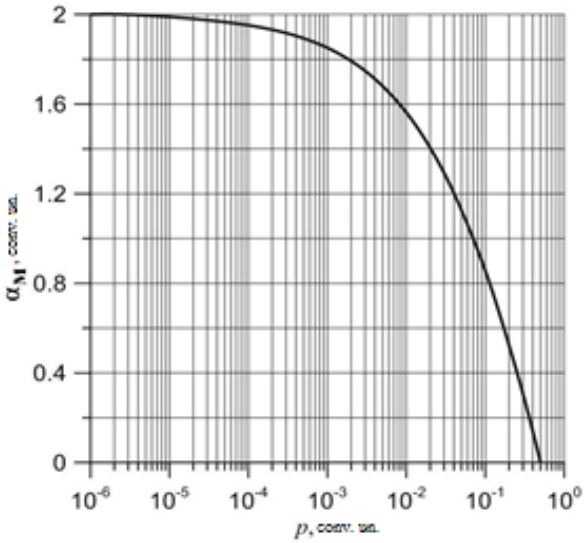



Fig. 6.44. The calculated dependence of the maximum value of the frequency coefficient $\alpha_M$ on the value of the hopping probability $p$ [79, 95]

As can be seen from Fig. 6.44, for small values of the hopping probability $p$, the values of the frequency coefficient $\alpha_M$ are close to 2. This corresponds to the situation when a jump occurs between neutral potential wells. As a result of such a jump, the wells receive opposite charges (and electric dipole arises), and the electric field of the dipole causes the mobile charge to return to the first well.

Small values of $\alpha$ correspond to large values of the hopping probability $p$. When the probability value increases to $p = 0.5$, the current density ceases to depend on frequency and the value of $\alpha$ becomes equal to 0. This situation can arise, for example, when a charge is exchanged between a neutral and a charged potential well.

The values of the hopping probability $p$ in (8.45) can be estimated from the experimental dependences $G(f)$, an example of which for an oxidized FeCoZr- $Al_2O_3$ nanocomposite is shown in Figure 6.45.

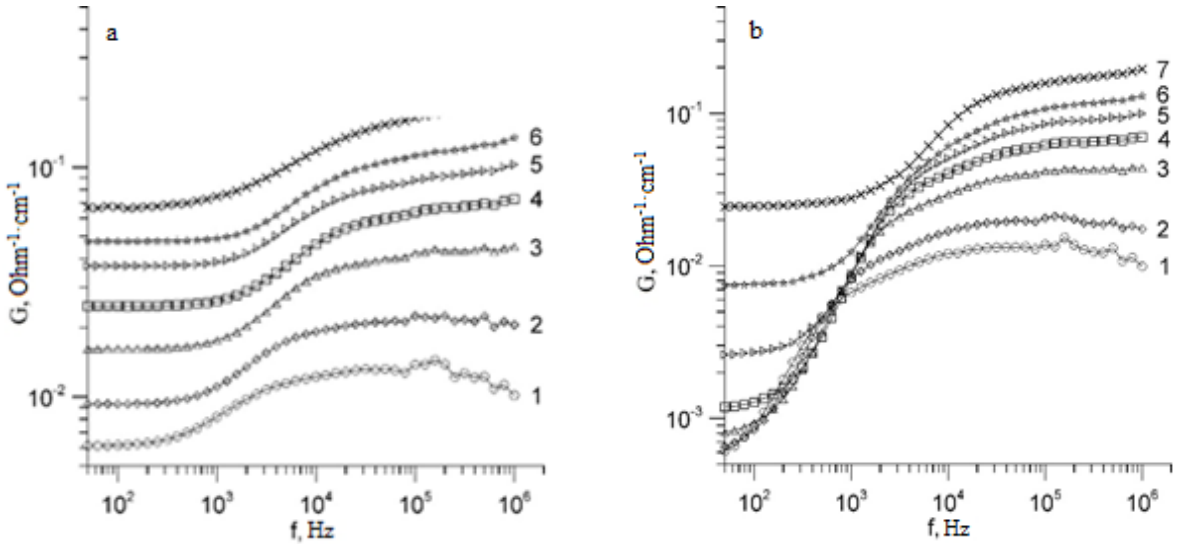



Fig. 6.45. Frequency dependences of the real part of admittance $G$ for the oxidized film $(FeCoZr)_{43}(Al_2O_3)_{57}$ before (*a*) and after (*b*) annealing at 548 K for temperatures: 1 - $T$ = 79 K; 2 - 103 K; 3 - 133 K; 4 - 163 K; 5 - 193 K; 6 - 238 K [80, 96]

From the analysis of the relation (8.45), it follows that the value of $p$ is determined by the following formula:

$$p = \frac{\sigma(f \to 0)}{2\sigma(f \to \infty)}. \tag{6.53}$$

This means that for each temperature, the value of $p$ can be estimated from the relation:

$$p = \frac{G_{LF}}{2G_{HF}}, \tag{6.54}$$

where $G_{LF}$ and $G_{HF}$ are the values of the conductivity at the low and high frequencies, respectively.

The presence of low- and high-frequency "shelves" on the experimental curves of $G(f)$ in Fig. 6.34 allows us to use the relations (6.50) or (6.51) for estimating the values of the probability density $p$.

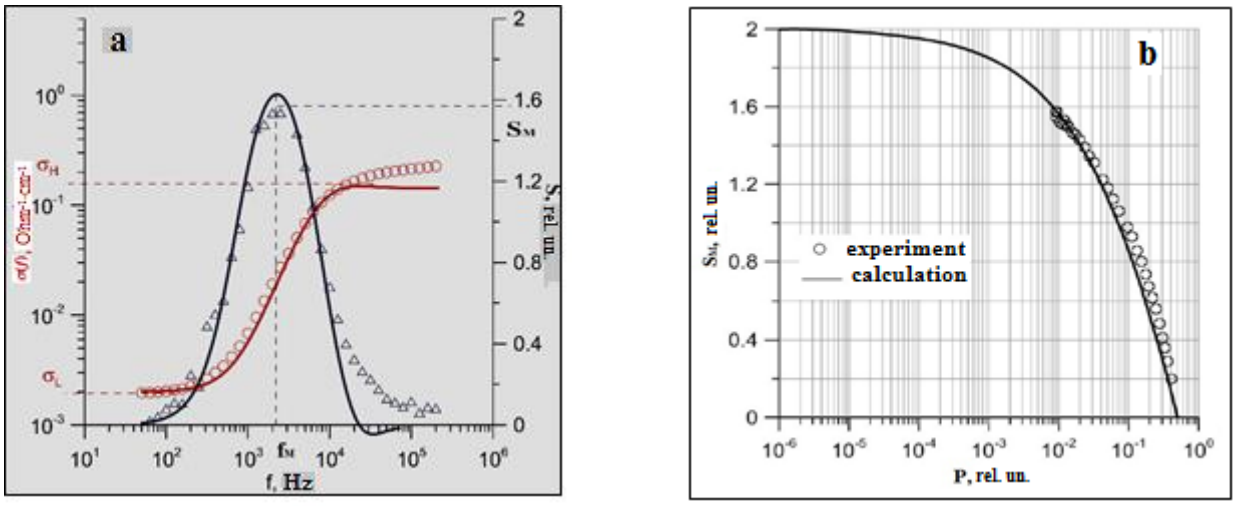



Fig. 6.46. The frequency dependences of the real part of the admittance $G(f)$, the frequency factor $\alpha(f)$ (a) and the maximum value of the frequency coefficient $\alpha_M$ on the probability $p$ (b) for the oxidized film $(FeCoZr)_{50}(Al_2O_3)_{50}$ obtained from experiments (points) and calculations by the model (curves) for measuring $T$ = 103 K and an annealing temperature of 548 K. The solid curve in Fig. 6 was obtained by modeling for $p = 9 \cdot 10^{-3}$ and $\tau = 9 \cdot 10^{-5}$ s [96]

Summarizing the main results that follow from the model described above, we can state that the interaction of electrons located on charged nanoparticles and surrounded by a layer of a polarized dielectric matrix leads to a redistribution of the probability density $p$ for each subsequent electron jump for some characteristic time $\tau$ (the lifetime of the electron at nanoparticle or the lifetime of an electric dipole). In this case, the dependence $G(f)$ has the form of the Mott law (6.39), where, however, the exponent $\alpha$ depends on the frequency (it should be called *the frequency factor*) and is determined by the distribution of the probability density $p$ of jumps. It follows from Fig. 6.46, the results of modeling and experimental estimation of the dependences $G(f)$, $\alpha(f)$ and $\alpha_M(p)$ show very good agreement.

According to the described model, the dependences $G(f)$, calculated (Figure 6.43 and the corresponding curve in Figure 8.46a) and observed in the experiments (Figure 6.46 and the curve in Figure 6.45), have a sigma-like character, which makes the relation (6.53) very useful for the simplified estimate of $p$.

As shown in [79], this behavior of $G(f)$ is valid for all types of nanocomposites studied in this paper, in which an inductive contribution to the impedance (admittance) was recorded. The presence of the latter is most clearly manifested in the form of a positive angular phase shift $\theta > 0$ in Fig. 6.39 for frequency dependences above a certain frequency [79]. As can be seen from this figure, the curves of $\theta(f)$ directly indicate that below some frequency $f_{min}$, the phase shift angles are negative, and the higher the angles are positive. The last (current delay as compared with the applied voltage) corresponds to the effect of "negative capacitance" or an inductive contribution to the impedance at high frequencies. Let us clarify this statement.

In the case of nanocomposites under investigation, the electron jump from one neutral nanoparticle to the second one can occur during the Maxwellian relaxation time $\approx 10^{-13}$ s by one of two ways. First, by tunneling, and secondly, by hopping over a barrier created by a dielectric layer (a matrix or nanoparticles own nanoparticles) separating the particles of the metal phase. In both cases, after the jump, a pair of positively and negatively charged nanoparticles arises forming

an electric dipole (Figure 6.41). The formation of a dipole after the capture of an electron by a nanoparticle leads to an increase in the lifetime of an electron $\tau$ on it, i.e. delay the next jump. This delay time can be determined from relation

$$f \leq \frac{1}{2\tau} \approx f_{min}, \tag{6.55}$$

where, as noted above, $f_{min}$ is determined by the point of intersection of the frequency axis of the curve $\theta(f)$ (Fig. 6.39). The evaluation shows that in the studied nanocomposites the values of $\tau$ lie in the range $10^{-6}$ - $10^{-3}$ s. Consequently, for frequencies $f > f_{min}$, the delay in the phase of the next jump back from the well 2 to the well 1 is $2\pi f\tau$, i.e. becomes more than $2\pi$, that creates the possibility of a positive phase shift $\theta$.

Thus, this hopping conductance model makes it possible to explain the appearance of an effective inductive contribution of $L$ to the impedance of nanocomposites. Moreover, a comparative analysis of the dependence of the impedance real part on frequency of the electric field and temperature for unoxidized and oxidized films allows one to interpret high values of $L$ (up to 20 μH /μm3) at room temperature for the oxidized films, as a result of the formation of nanoparticles with a core-shell structure. In accordance with the proposed model, this effect can be explained by the fact that the formation of oxide shells, firstly, stabilizes the nanosize of particles, and secondly, increases the polarization of the dielectric around metallic nanoparticles in hopping electron transport [97-99].

In addition, the described model of hopping conductance makes it possible to estimate the activation energies of the hopping transport and to calculate the thickness of the dielectric layer $r$ separating metallic nanoparticles.

The potential energy of a pair of charges (dipoles) on the first and second wells, measured in eV, has the form:

$$U_1 = \frac{e}{4\pi\varepsilon_0\varepsilon_r r}, \tag{6.56}$$

where $e$ is the electron charge, $\varepsilon_0$ = 8.85·10-12 F/m is the permittivity of the vacuum, $\varepsilon_r$ is the relative permittivity of the matrix material (for $Al_2O_3$ it is equal to 10), $r$ is the average jump length of the electron equivalent to the thickness of the dielectric gap separating metallic nanoparticles.

The probability of a backward jump of an electron from the second well to the first one after the expiration of the time $\tau$ (jump against the field due to atomic vibration energy) can be estimated from expression

$$(1 - p) = A\exp\left(\frac{U_1}{k_B T}\right), \tag{6.57}$$

where $A$ is the numerical coefficient, $U_1$ is the height of the barrier.

In the case when an electron jumps along the field from well 2 to well 3, the probability of this event is determined by the relation:

$$p = A\exp\left(-\frac{U_1}{k_B T}\right), \tag{6.58}$$

where the barrier is much lower and it can be assumed that

$$U_2 = \frac{e}{4\pi\varepsilon_0\varepsilon_r 2r} \approx \frac{U_1}{2}. \tag{6.59}$$

It follows that

$$p = \frac{p}{(1-p)+p} = \frac{A\exp\left(-\frac{U_1}{2k_BT}\right)}{A\exp\left(\frac{U_1}{k_BT}\right)+A\exp\left(-\frac{U_1}{2k_BT}\right)} = \frac{\exp\left(-\frac{U_1}{2k_BT}\right)}{\exp\left(\frac{U_1}{k_BT}\right)+\exp\left(-\frac{U_1}{2k_BT}\right)}.$$

After the transformations, we obtain the following formula:

$$p \approx \exp\left(-\frac{3}{2}\frac{U_1}{k_BT}\right), \tag{6.60}$$

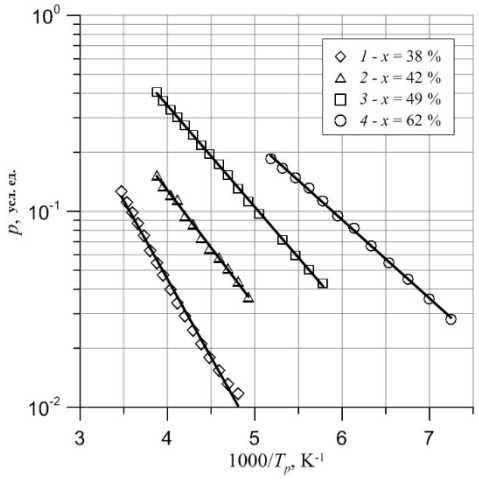
where the energy of the thermal activation of the jump $\Delta E_p$ for the probability $p$ is defined as:

$$\Delta E_p = \frac{3}{2}U_1. \tag{6.61}$$

So, in the region of sufficiently high temperatures, where this model is valid (issue 3 of the postulates of the model formulated at the beginning of this section), the value of the probability $p$ can be written in the form of the relation:

$$p \sim \exp\left(-\frac{\Delta E_p}{k_BT}\right). \tag{6.62}$$

It follows from (6.62) that in the Arrhenius scale the dependence $\lg p = f\left(\frac{1000}{T_p}\right)$ must be linear, which allows one to determine the activation energy of the hopping probability $\Delta E_p$ from the slope of the lines.



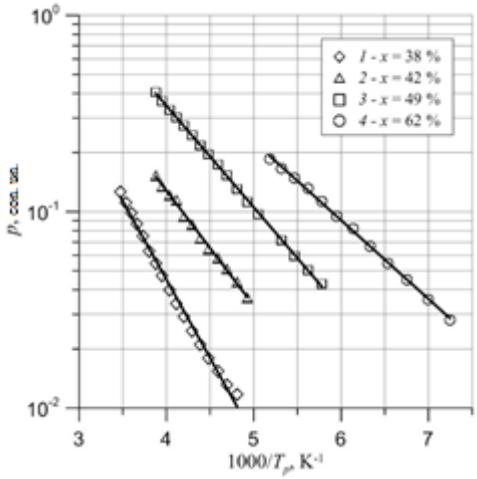



Fig. 6.47. Dependences of the hopping probability $p$, determined from the experimental curves of $\sigma(f)$ on the inverse temperature, for the oxidized films $(FeCoZr)_x(Al_2O_3)_{100-x}$ with different metallic phase content $x$ [79]

The curves shown in Fig. 6.46 for the dependence of $p$ on the reciprocal temperature, determined from the experimental dependences of $G(f)$, show that the activation energies of hopping transfer $\Delta E_p$ for the presented range of compositions $x$ in the samples studied vary from 0.08 eV (for $x$ = 62 at.%) to 0,16 eV (at $x$ = 38 at.%). This coincides with the data obtained from the relation (6.41) in [87, 89, 90, 101] in the study of DC conductivity.

Using experimentally determined activation energy $\Delta E_p$ and formula (6.53), it is possible to determine the thickness of the dielectric layer $r$ separating metallic nanoparticles. The calculation performed in accordance with this model for the film $(FeCoZr)_{62}(Al_2O_3)_{38}$ ($\Delta E$ = 0.08 eV) showed that $r \approx 2.6$ nm [101] that for the films with $x > x_c$ can be considered as the doubled thickness of oxide shells and correlates well with transmission electron microscopy data.

**Bibliography**

1. Yu.Yu. Tarasevich: Percolation: THEORY, APPLICATIONS, ALGORITHMS, URSS, Moscow, 2002

2. Broadbent S. K., Hammersley J. M. Percolation processes I. Crystals and mazes//Proc. of Phil. Soc. – 1957. — V.53.— P. 629-641.

3. Shklovskii B.I., Efros A.L. The theory of percolation and conductivity of highly inhomogeneous media. // UFN - 1975. – V.117. - No. 3. - P.401-436.

4. Grimmet G. Percolation. Berlin: Springer-Verlag, 1999. —444 p.

5. Stauffer D., Aharony A. Perkolationstheorie: eine Einführung. —Weinnheim: VCH–Verlagsgesellschaft, 1995. —202 p.

6. Frisch H. L.. Hammersley J. M. , Welsh D. J. A. Monte Carlo Estimates of Percolation Probabilities for Various Lattices // Phys. Rev.— 1962.— V.126.— P.949 – 951.

7. Poklonskii N.A., Gorbachuk N.I.. Fundamentals of Impedance Spectroscopy of Composites; Lecture Course. Minsk: BSU, 2005. 130 p.

8. Efros A.L., Shklovskii B.I. Critical behavior of conductivity and dielectric constant the metal-non-metal transition threshold // Phys. stat. sol. (b). —1976. —V. 76, No 2. —P. 475—485.
9. Lukyanets, S.P. A model of macroscopically inhomogeneous mixtures of an ideal conductor near the percolation threshold / S.P. Luk'yanets, A.A. Snarskii // Soviet Solid State. 1988 - V. 94, P. 301.306.
10. Morozovskii, A.E. Multiscaling in randomly inhomogeneous media: effective conductivity, relative spectral density of 1/f noise, and the following moments: A.E. Morozovskii, A.A. Snarskii // Soviet Solid State 1992. V. 102, P. 683-692.
11. Electrical conductivity of composite materials based on finely dispersed silicon near the insulator-metal transition / N.A. Poklonsky, N.I. Gorbachuk, I.V. Pototsky, D.A. Trofimchuk // Inorganic Materials .- 2004.- V. 40 , No. 11.- P. 1293-1298.
12. Concentration dependence of the electrical conductivity of composites in the high-frequency region / I.A. Chmutin, N.G. Ryvkina, A. T. Ponomarenko, V.G. Shevchenko // High-molecular compounds. Ser. A.- 1996.- V. 38, No. 2.- P. 291.296 .
13. Gorbachuk N.I., Gurin V.S., Poklonsky N.A. Influence of moisture content on electrical conductivity of xerogels $SiO_2$ / LiCl // Physics and chemistry of glass. -2001. -V. 27, C. 762-771.
14. Bruggeman D. A. G. Berechnung verschiedener physikalischer Konstanten von heterogenen Substanzen I. Dielektrizitätskonstanten und Leitfähigkeiten der Mischkörper aus isotropen Substanzen // Ann. Phys.— 1935.— V.416, № 7.— S. 636-664.
15. Sen P. N., Scala C., Cohen M. H. A self-similar model for sedimentary rocks with application to the dielectric constant of fused glass beads // Geophysics.— 1981. № 46.— P. 781–785.
16. Brosseau C. Dielectric properties of polymer-liquid composite systems // Polymer News.— 1995. № 20.— P. 270-273
17. Brosseau C. Dielectric properties of polymer-liquid composite systems // J. Appl. Phys.— 1994. № 75.— P. 672-674.
18. Neugebauer C.A. Resistivity of Cermet Films Containing Oxides of Silicon // Thin Solid Films.-1970.-V.6.-P.443-447.
19. Omata Y., Sakakima H. Thermal stability of softmagnetic properties of Co-(Nb,Ta)-(Zr,Hf) films with high saturation magnetization // Transactions on magnetics. - 1987. - V.23, № 5. - P. 1005-1008.
20. Guzman J.I., Koeppe P.V., Kryder M.N. Magnetic properties of sputtered bilayer and laminated CoZr/$SiO_2$ thin films // Transactions on magnetics. - 1988. - V.24, № 6. - P. 1095 – 1101.

21. Choh K.K., Judy J.H. The effects of an applied magnetic field on the magnetic properties of rf - sputtered amorphous CoZrNb thin films // Transactions on magnetics. - 1987. - V.23, - №5. - P. 965 - 969.

22. Gurumrugan K., Mangalaraj D., Narayandass K. Magnetron sputtered transparent conducting CdO thin films // J. of Electronic Materials. - 1996. - V.25, - № 4. - P.2011-2023.

23. Sankar S., Berkowitz A.E., Smith D.J. Spin-dependent transport of Co-$SiO_2$ granular films approaching percolation // Phys.Rev.B.-2000.-V.62, - N.21.- P.14273-14278.

24. J.S.Moodera, E.F.Gallagher, K.Robinson, J.Nowak Optimum tunnel barrier in ferromagnetic–insulator–ferromagnetic tunneling structures // Appl.Phys.Lett. -1997.-V.70.-P.3050-3068.

25. Mitani S., Fujimori H., Ohnuma S. Spin-dependent tunneling phenomena in insulating granular systems // JMMM.-1997.-V.165.-P.141-148.

26. N.Kobayashi, S.Ohnuma, T.Masumoto, H.Fujimori (Fe-Co)-(Mg-fluoride) insulating nanogranular system with enhanced tunnel-type giant magnetoresistance // J.Appl.Phys.- 2001.-V.90, - N.8.- P.4159-4162.

27. Kalinin Yu.E., Sitnikov A.V., Stognei O.V., Zolotukhin I.V., Neretin P.V Electrical properties and giant magnetoresistance of CoFeB-$SiO_2$ amorphous granular composites // Mat. Scien. and Engin. -2001.- A304-306.- P.941-945.

28. Zolotukhin I.V., Neretin P.V., Kalinin Yu.E., Stogney O.V., Sitnikov A.V. Electrical resistance of amorphous nanocomposites CoTaNb + $SiO_2$ // Alternative energy and ecology -2002.- № 2.-P.7-14.

29. Fedotova, J. FeCoZr-$Al_2O_3$ granular nanocomposite films with tailored structural, electric, magnetotransport and magnetic properties / Advances in nanoscale magnetism / springer proceedings in physics: ed. B. Aktas and F.Mikailov. – Berlin, 2008. – vol. 122. – p. 231–267.

30. A.M. Saad, A.K. Fedotov, I.A. Svito, J.A. Fedotova, B.V. Andrievsky, Yu.E. Kalinin, V.V. Fedotova, V. Malyutina-Bronskaya, A.A. Patryn, A.V. Mazanik, A.V. Sitnikov. Impedance and magnetization of cofezr nanoclusters embedded into alumina matrix / J. Alloys Compds. – 2006. – vol.423. – p. 186–180.

31. A.M.Saad, A.V. Mazanik, Yu.E. Kalinin, J.A. Fedotova, A.K. Fedotov, S. Wrotek, A.V.Sitnikov, I.A.Svito. Structure and electrical properties of cofezr-aluminium oxide nanocomposite films / // rev. adv. mater. sci. – 2004. – vol. 8. – p. 34–40.

32. A.M.Saad,, A.K. Fedotov, J.A. Fedotova, I.A. Svito, B.V. Andrievsky, Yu.E. Kalinin, V.V. Fedotova, V. Malyutina-Bronskaya, A.A. Patryn, A.V. Mazanik, A.V. Sitnikov. Characterization of $(Fe_{0.45}Co_{0.45}Zr_{0.10})_x(Al_2O_3)_{1-x}$ nanocomposite films applicable as spintronic materials // phys. stat. sol. c – 2006. – vol. 3, № 5.– p. 1283–1290.

33. Gittleman J.L., Goldstain Y., Bozowski S. Magnetic properties of Granular Nikel Films // Physical Review B.-1972. - V.B5. - P.3609-3621.

34. Abeles B., Sheng P., Coutts M.D., Arie Y. // Advances in Physics. - 1975. - V.24. - P.407-461.

35. Efros A.L., Shklovski B.I. Conduction of nanostructured metall – insulator // Phys. Stat. Solid. B. – 1976. - № 76. - P. 475 - 490.

36. Mott N., Davis E. Electronic processes in non-crystalline substances // M .: Mir, 1974. - 472 p.

37. Yu.E.Kalinin, A.T.Ponomarenko, V.G.Shevchenko, A.V.Sitnikov, O.V.Stognei, O.Figovsky I.V.Zolotukhin, J. Metal-dielectric nanocomposites with amorphous structure// Nanostructured Polymers and Nanocomposites.- 2006.- V.2.- N.1.- P.23-34.

38. P. Aldhafen. Metallic glass. // M .: Mir, - 1986. - Part II. - P. 328-378.

39. W. Heywang, W. Birkholz, R. Ainzinger, L. Hanke, K. Kempter, A. Schneller. Amorphous and polycrystalline semiconductors. / M .: Mir, - 1987. - P.23.

40. Sheng P., Abeles B. and Arie Y. Hopping conductivity in granular Metals // Phys.Rev.Lett.- 1973.-V.31, N.1.-P.44-47.

41. Honda S., Okada T. and Nawate M. Tunneling giant magnetoresistance in Fe-$SiO_2$ multilayered and alloyed films // JMMM.-1997.- V.165.-P. 153-156.

42. B.A. Aronzon, A.E. Varfolomeev, D.Yu. Kovalev et al. Conductivity, magnetoresistance and Hall effect in granular Fe-SiO2 films // Soviet Solid State.-1999. - V.41. - P.944-950.

43. Z. Zhang, Chengxian Li, Chao Li, Shihui Ge Giant magnetoresistance of Co-Al-O insulating granular films deposited at various substrate temperatures // JMMM.-1999. –V.198-199 .– P.30-32.

44. Abeles B., Cohen R.W., Cullen G.W. Enhancement of Superconductivity in Metal Films // Phys. Rev. Lett.-1966. - V.17. - P.632-634.

45. S.Mitani, S. Takahashi, K.Takahashi et al. Enhanced magnetoresistance in insulating granular-systems: Evidence for Higher-order tunneling // Phys.Rev.Lett.-1998.-V.81. - N.13. - P.2799-2802.

46. S.Barzilai, Y.Goldstain, I.Balberg and J.S.Helman Magnetic and transport properties of granular cobalt films // Phys.Rev.B. - 1981. - V.23. - N.4. - P.1809-1817.

47. Kalinin Yu.E., Remizov A.N., Sitnikov A.V. Electrical properties of amorphous nanocomposites $(Co_{45}Fe_{45}Zr_{10})x(Al_2O_3)_{1-X}$// Soviet Solid State. - 2004. - V. 46. - B.11. - P.2076-2082.

48. Frisch H. L.. Hammersley J. M. , Welsh D. J. A. Monte Carlo Estimates of Percolation Probabilities for Various Lattices // Phys. Rev. – 1962. - V.126. N3. - P.949 – 951.

49. Fujimori H., Mitani S., Takanashi K. Giant magnetoresistance in insulating granular films and planar tunneling junctions // Mat.Sci.Eng.A.-1999.-V.A267.-P.184-192.

50. Sankar S., Berkowitz A.E., Smith D.J. Spin-dependent transport of Co-$SiO_2$ granular films approaching percolation // Phys.Rev.B.-2000.-V.62, N.21.- P.14273-14278.

51. Z. Zhang, Chengxian Li, Chao Li, Shihui Ge. Giant magnetoresistance of Co-Al-O insulating granular films deposited at various substrate temperatures. JMMM.-1999.–V.198-199.– P.30-32.

52. S.Mitani, S. Takahashi, K.Takahashi et al. Enhanced magnetoresistance in insulating granular-systems: Evidence for Higher-order tunneling. Phys.Rev.Lett.-1998.-V.81, N.13.-P.2799-2802.

53. S.Barzilai, Y.Goldstain, I.Balberg and J.S.Helman Magnetic and transport properties of granular cobalt films. Phys.Rev.B.-1981.-V.23, N.4-P.1809-1817.

54. DA Zakgeym, I.V. Rozhansky, I.P. Smirnova, S.A. Gurevich Temperature dependence of the conductivity of composite films Cu: $SiO_2$ experiment and numerical simulation. Soviet Solid State. 2000-V.118, C.637-646.

55. O.V. Stogney, Yu.E. Kalinin, A.V. Sitnikov, I.V. Zolotukhin, A.V. Slyusarev. Resistive and magnetoresistive properties of granular amorphous composites CoFeB-$SiO_n$. Physics of Metals and Metallography.-2001.-V. 91, C.24-31.

56. Helman J.S., Abeles B. Tunneling of Spin-Polarized Electrons and Magnetoresistance in Granular Ni Films // Phys. Rev. Lett.-1976.-V.37, N.21.-P.1429-1433.

57. Anderson P.W. Absence of Diffusion in Certain Random Lattices // Phys.Rev.-1958.-V.109.-P.1492-1505.

58. Hughes R.C. Electronic and ionic charge carriers in irradiated single crystal and fused quartz // Rad.Effects.-1975.-V.26.-P.225-235.

59. Mott N.F. Electronic properties of vitreous silicon dioxide // Physics of $SiO_2$ and Its Interfaces.– Pergamon Press. –1978.-P.1-13.

60. Mott N., Davis E. Electronic processes in non-crystalline substances. M .: Mir, 1982. Vol. 1, 368 p.

61. Zakis Yu.R. The Simplest Thermal Defects in Glasses. Physics and Chemistry of Glass-Forming Systems. -Riga, 1980.-issue 7.-P.3-36.

62. Intrinsic defects in fused silica / A.N.Shendrik, A.Silin, L.Skuja et al. // Proc.$11^{th}$ Int. Congr. Glass. –Prague, CVTS-Dum Techn.-1977.-V.1.-P.13-21.

63. Silin A.R., Trukhin A.N. Point defects and elementary excitations in crystalline and glassy $SiO_2$. - Riga: Zinatne, 1985. 244 p.

64. Sheng P., Klafter J. Hopping Conductivity in Granular Disordered Systems // Phys. Rev. B. - 1983. - V.27. - P.2583–2586.

65. Lin C.-H., Wu G.Y. Hopping conduction in granular metals // Physica B.- 2000. - V.B 279. - P.341-346.

66. Möbius A., Richter M., Drittler B. Coulomb gap in two- and three-dimensional systems: Simulation results for large samples. Phys. Rev. B.-1992.–V.45, N.20.-P.11568-11579.

67. Cuevas E., Ortuno M., Ruiz J. Ground state of granular metals. Phys. Rev. Let.-1993.-V.71, N.12.-P.1871-1874.

68. D.A. Zakgeym, I.V. Rozhansky, I.P. Smirnova, S.A. Gurevich. Temperature dependence of the conductivity of composite films Cu:$SiO_2$ experiment and numerical simulation . Soviet Phys. JETP. 2000. V.118. C.637-646.

69. Meilikhov E.Z. Thermally activated conductivity and current-voltage characteristic of the dielectric phase of granular metals. Soviet Phys. JETP.-1999. V.115, P.1484-1496.

70. A.V. Tartakovskii, M.V. Fistul, M.E. Raih, I.M. Ruzin. Hopping conductivity in metal-semiconductor-metal contacts / // Soviet Solid State. 1986.-V.21, C.603-608.

71. Petrov Yu.I. Physics of small particles.-Moscow: Nauka, 1982.-359 p.

72. Stogney O.V., Kalinin Yu.E., Sitnikov A.V. Low-temperature behavior of electrical resistivity in granular metal-insulator composites. // Systemic problems of quality, mathematical modeling, information, electronic and laser technologies: Proceedings of the International Conference and the Russian Scientific School .- M .: Radio and Communication, 2002.- P.7 (R. III).-P.76-80.

73. Miyazak T., Tezuka N. Giant magnetic tunneling effect in Fe/$Al_2O_3$/Fe junction // JMMM. - 1995. - V.139. -P.L231-L234.

74. K. Sumiyama, K.Suzuki, S.A.Makhlouf et al. Structural evolution and magnetic properties of nano-granular metallic alloys. J.Non-Cryst.Solids.-1995.-V.192-193.-P.539-545.

75. C.Laurent, D.Mauri, E.Kay and S.S.Parkin Magnetic properties of granular Co-polimer films. J.Appl.Phys. 1989. V.65. P.2017-2020.

76. J.Hesse, H.Bremers, O.Hupe et al. Different susceptibilities of nanosized single-domain particles derived from magnetization measurements. JMMM. 2000. V.212. P.153-167.

77. Dormann J.L. et al. From pure superparamagnetic regime to glass collective state of magnetic moments in $\alpha$-$Fe_2O_3$ nanoparticle assemblies. JMMM. 1998.-V.187. P.L139-L144.

78. B.Abeles, P. Sheng M.D.Coutts and Y. Arie. Structural and electrical properties of granular metal films. Advances in Physics.-1975. V.24. P.407-461.

79. A. Fedotova, J. Transport ładunków i właściwości strukturalne wybranych nanokompozytów metal-dielektryk (Транспорт носителей и структурные свойства отдельных нанокомпозитов металл-диэлектрик) / J.Fedotova, T.Koltunowicz, P.Zhukowski: pod red. P.Zhukowskiego. – Lublin: Politechnika Lubelska, 2012. – 193 s.

80. P. Zhukowski, T. Koltunowicz, J. Partyka, J.A.Fedotova, A.V. Larkin. Hopping conductivity of metal-dielectric nanocomposites produced by means of magnetron sputtering with the application of oxygen and argon ions. Vacuum. 2009. Vol.83. P.S280−S283.

81. A.M.Saad, B.V. Andrievsky, A.Fedotov, J.A.Fedotova, T.Figielski, Yu.E. Kalinin, V. Malyutina-Bronskaya, A.V. Mazanik, A.A. Patryn, A.V. Sitnikov, I.A. Svito. AC and DC carrier transport in $(FeCoZr)_x(Al_2O_3)_{100-x}$ nanocomposite films for spintronic applications. Proceedings of the 1st International workshop on semiconductor nanocrystals (SEMINANO 2005), Budapest, September 10-12, 2005: in Vol. 2. Semiconductor nanocrystals; ed. B.Püdör, Zs.J. Horvoth, P.Basa. – Budapest, 2005. Vol. 2. P. 321–324.

82. P. Zhukowski, T.N.Koltunowicz, J.A.Fedotova, A.V. Larkin. An effect of annealing on electric properties of nanocomposites $(CoFeZr)_x(Al_2O_3)_{100-x}$ produced by magnetron sputtering in the atmosphere of argon and oxygen beyond the percolation threshold. Electrical Review. 2010. № 7. P. 157–159.

83. A.V.Larkin, A.Saad, A.K.Fedotov, J.A.Fedotova, A.Patryn, B.Andriyevskyy, Z.Czapla, S.Dacko, T.N.Koltunowicz, P.Wegierek. Imaginary part of admittance in $(Fe_{0.45}Co_{0.45}Zr_{0.10})_x(Al_2O_3)_{1-x}$ nanocomposites. Electrical Review. 2010. №7. P. 202–204.

84. A.V. Lar'kin, Yu.A.Fedotova, B.V.Andrievsky, A.A.Patrin, A.K.Fedotov, Yu.E.Kalininin, A.V.Sitnikov. Conductivity on alternating current of nanocomposites FeCoZr-$Al_2O_3$ on the metal side of the metal-insulator transition. Reports of the international scientific conference "Actual problems of solid state physics", Minsk, October 23-26, 2007: 3 Volumes. Editors: V.M.Oleokhnovich [and others .]. Minsk, 2007. V. 1. P.145-147.

85. Saad, A. Influence of sputtering ambient on impedance and magneto-impedance of soft magnetic CoFeZr nanoparticles embedded in alumina matrix / A. Saad, J. Fedotova, A. Larkin, Yu. Kalinin, A. Sitnikov, V. Fedotova, Yu. Ilyashuk, A. Fedotov // Proceedings of Fourth Seeheim Conference on magnetism (SCM 2010), Frankfurt, Germany, March 28 − April 1, 2010. − P.276−277.

86. B.Andryevski, A.Partyn, Z.Czapla, S.Dacko, J.A.Fedotova, A.V.Larkin. A.K.Fedotov. Impedance spectroscopy study of amorphous composites $(FeCoZr)_x(Al_2O_3)_{1-x}$. Electrical Review. 2008. Vol.3. P.117–120.

87. A.M.Saad, A.K.Fedotov, J.A.Fedotova, I.A. Svito, B.V. Andrievsky, Yu.E. Kalinin, V.V. Fedotova, V. Malyutina-Bronskaya, A.A. Patryn, A.V. Mazanik, A.V. Sitnikov. Characterization of $(Fe_{0.45}Co_{0.45}Zr_{0.10})_x(Al_2O_3)_{1-x}$ nanocomposite films applicable as spintronic materials. Phys. Stat. Sol. C. 2006. Vol. 3, P. 1283–1290.

88. P.Zhukowski, T.N.Koltunowicz, P.Wegierek, J.A.Fedotova, A.K.Fedotov, A.V.Larkin. Formation of non-coil inductance in nanocomposites $(Fe_{0,45}Co_{0,45}Zr_{0,10})_x(Al_2O_3)_{1-x}$ manufactured by ion-beam sputtering of complex target in Ar+O atmosphere. Acta Physica Polonica A. 2011. Vol.120. P.43−45.

89. A.M.Saad, V.A. Kalaev, J.A.Fedotova, K.A. Sitnikov, A.V. Sitnikov, Yu.E.Kalinin, A.K. Fedotov, V.A. Svito. Structure and magnetic properties of nanogranular composites CoFeZr-alumina. Rev. Adv. Mater. Sci. – 2007. – Vol. 14. – P. 14–34.

90. Fedotova, J. $(FeCoZr)_x(Al_2O_3)_{1-x}$ granular nanocomposite films with tailored structural, electric, magnetotransport and magnetic properties. Advances in nanoscale magnetism. Springer proceedings in physics: ed. Baktas and F Mikailov. Berlin, 2008. vol. 122. P. 231–267.

91. P.Zukowski, T.Koltunowicz, J.Partyka, P.Węgierek, M.Kolasik, A.Wdowiak, A.V.Larkin, J.Fedotova, A.K.Fedotov, F.F.Komarov, L.A.Vlasukova. Model of hopping recharging at alternating and direct currents and its experimental verification for nanostructures prepared with ion techniques. Proceedings of 5th International Conference New Electrical and Electronic Technologies and Their Industrial Implementation ("NEET 2007"), Zakopane, Poland, June 12–15, 2007. Lublin University of Technology; ed: T.Koltunowicz. Lublin, 2007. P.141.

92. J. Przewoznik, Cz. Kapusta, M.Milosavljevic, J. V. Kasiuk, J. Zukrowski, M. Sikora, A.A. Maximenko, D. Szepietowska, K.P. Homewood. Magnetoresistance in FeCoZr-$Al_2O_3$ nanocomposite films containing “metal core-oxide shell” nanogranules. J.Phys. D: Appl. Phys. 2011. Vol.44. P. 495001-1-495001-12.

93. E. Thirumal [et al.]. Magnetic, electric, and dielectric properties of FeCo alloy nanoparticles dispersed in amorphous matrix. Phys. Stat. Sol. A. 2010. Vol. 207 P. 2505-2510.

94. Nowak, R. Statystyka dla fizyków, R.Nowak. Warszawa: Państwowe wydawnictwo naukowe, 2002. 664 s.

95. P.Zukowski, T.Koltunowicz, J.Partyka, P.Wegierek, M.Kolasik, A.Wdowiak, A.V.Larkin, J.Fedotova, A.K.Fedotov, F.F.Komarov, L.A.Vlasukova. Model of conductivity hopping and its verification for nanostructures prepared with ion techniques. Electrical Review. 2008. Vol.. –P.247–250.

96. A.V. Larkin, I.A. Svito, Yu.A. Fedotova, A.K. Fedotov. Conductivity on a direct and alternating current of film metal-insulator nanocomposites. Proc. of BSU. Ser. 1: Phys., M., Inform. 2009. No. 1. P.26-37.

97. P.Zhukowski, T.N.Koltunowicz, P.Wegierek, J.A.Fedotova, A.K.Fedotov, A.V.Larkin Formation of non-coil inductance in nanocomposites $(Fe_{0,45}Co_{0,45}Zr_{0,10})_x(Al_2O_3)_{1-x}$ manufactured by ion-beam sputtering of complex target in Ar+O atmosphere. Acta Physica Polonica A. 2011. Vol.120, P.43−45.

98. T.N.Koltunowicz, P.Zhukowski, J.A.Fedotova, A.V.Larkin. Inductive-type properties of $(Co_{45}Fe_{45}Zr_{10})_x(Al_2O_3)_{100-x}$ nanocomposites produced by the ion-beam sputtering in the argon and oxygen ambient. J. Nano Electron. Phys. 2012. Vol. 4. P.01002-1–01002-4.

99. P.Zhukowski, T.Koltunowicz, J.Partyka, P.Wegierek, J.Fedotova, A.K.Fedotov, A.V.Larkin. The effect of annealing on electrical properties of $(FeCoZr)_x+(Al_2O_3)_{1+x}$ nanocomposites. Electrical Review. 2008. Vol.3. P.244–247.

100. A.M.Saad, A.K.Fedotov, I.A. Svito, J.A.Fedotova, B.V. Andrievsky, Yu.E. Kalinin, V.V. Fedotova, V. Malyutina-Bronskaya, A.A. Patryn, A.V. Mazanik, A.V. Sitnikov. Impedance and magnetization of CoFeZr nanoclusters embedded into alumina matrix. J. Alloys Compds. 2006. Vol.423. P. 186–180.

101. P. Zhukowski, T.N.Koltunowicz, J.A.Fedotova, A.V. Larkin An effect of annealing on electric properties of nanocomposites $(CoFeZr)_x(Al_2O_3)_{100-x}$ produced by magnetron sputtering in the atmosphere of argon and oxygen beyond the percolation threshold. Electrical Review. 2010. № 7. P. 157–159.

## 7. Magnetotransport and Magnetism in Nanocomposite and Multilayered Materials
### 7.1. Introduction

Heterogeneous and patterned nanostructured films with non-planar, i.e. perpendicular, magnetic anisotropy (PMA) of oriented nanogranules or nanoarrays are considered as promising materials for designing high density recording media [1, 2], systems for visualization of magnetic fields distribution and magnetic tunneling junction sensors [3]. Among the most convenient methods to produce magnetically oriented planar structures are template-assisted synthesis of magnetic nanoarrays with shape anisotropy [4, 5, 6] and formation of Pt, Pd-3*d* metal compounds with $L_{10}$ structure that exhibits high magnetocrystalline anisotropy [7].

Nowadays the attempts to fabricate such media for data storage with density as high as 1 Tbit/in$^2$, and the bit size around 20 nm$^2$ face strong limitation because of superparamagnetic (SP) effect which is undesirable thermal fluctuations of magnetic moments when grain size falls down to a few nm [2]. The formation of percolated perpendicular magnetic media containing exchange-coupled magnetic films with densely distributed artificially engineered pinning centers (antidots) [5] enables to overcome this restriction.

In this context, fabrication of antidot arrays in [3d (Co, Fe) – 5d (Pt, Pd)]n films by deposition onto porous substrates, can be an effective method to tailor the magnetization reversal mechanism with simultaneous stabilization of the orientation of magnetization at room temperature due to the strong domain-wall pinning. Magnetic anisotropy of such nanoporous films is very sensitive to topographic features of templates like anodized $Al_2O_3$ used as a substrate, especially at length scales close to domain-wall width and exchange length. It is known that complex morphology of nanoporous films results in the deterioration of intrinsic perpendicular anisotropy due to the local misalignment of easy magnetization axis at the proximity of nanopore rims. At the same time, anodized porous $TiO_2$ has the structure of open-end nanotube arrays with expanded flat interpore areas that could promote the conservation of PMA in films despite of their porous structure.

It should be added that recent results reveal perpendicular magnetic anisotropy also in granular composite films that contain metallic nanoparticles embedded into non-magnetic (insulating) matrices [8, 9]. Such an intrinsic growth-induced magnetic anisotropy is quite unusual phenomenon for granular films and was reported only for a few systems like granular $Co_xAg_{1-x}$ , Co-Zr-O and Co-$Al_2O_3$ films, Co-rich Co-Ag and Co-Cu granular materials as well as for (Fe-Ni-Co)-fluoropolymer films. It is generally accepted that the origin of perpendicular magnetic anisotropy in granular films is related to the preferential arrangement of metallic nanoparticles in the direction perpendicular to the film surface resulting in the growth of columnar-like nanoparticles or nanoparticle agglomerations. However, the driving force and the physical nature of such specifically shaped columnar metallic structures growing inside non-magnetic (insulating) matrices are still under dispute.

Present paper summarizes our latest results related to the correlation between structure and phase composition of studied films and the effect of perpendicular (non-planar) magnetic anisotropy observed in two types of materials: CoPd nanoporous multilayered (ML) films fabricated on $Al_2O_3$ and $TiO_2$ templates and granular metal-insulator films like FeCoZr-$CaF_2$ and FeCoZr-$Al_2O_3$.

### 7.2 Experimental

The films were deposited by dc ion sputtering, using 2 keV argon ion gun, at 0.28 nm/s, to a thickness 1-6 μm onto Al uncooled substrates ($T_{sub}$ ~373 K). The Ar pressure was 67 mPa and base pressure in the chamber was 0.26 Pa. Sputtering targets were composed of $Fe_{45}Co_{45}Zr_{10}$ alloy plates, covered with strips of $CaF_2$ that were pasted on top. In this way, the metal/insulator volume ratio in the deposited films could be varied by changing the relative area coverage of the

constituents. Composition of $(FeCoZr)_x(CaF_2)_{100-x}$ ($29 \le x \le 73$ at. %) films was verified by energy dispersive X-ray spectroscopy (EDX) in a scanning electron microscope and by Rutherford backscattering spectrometry (RBS).

Granular nanocomposite films $(Fe_{45}Co_{45}Zr_{10})_x(CaF_2)_{100-x}$ ($x$ = 58, 73 at.%) were irradiated by 167 MeV $Xe^{26+}$ ions with fluence $D$ in the range $8 \cdot 10^{12}$-$1 \cdot 10^{13}$ ion/cm$^2$ generated by IC-100 heavy ion cyclic accelerator (JINR, Dubna). Orientation of ion beam is along the normal to the films plane.

The fabrication of nanoporous films is a two-step process consisting of the preparation of porous $Al_2O_3$ and $TiO_2$ templates and subsequent deposition of Co/Pd multilayers. Details of templates fabrication are discussed in our recent papers [10, 11].

Multilayers (MLs) of Co/Pd were deposited on the surface of porous $TiO_2$ and $Si/SiO_2$ wafers by thermal evaporation in ultra-high vacuum with pressure $10^{-7}$ Pa. Simultaneous deposition on $Si/SiO_2$ wafers and templates of porous $TiO_2$ permits to compare the effect of porous $TiO_2$ morphology on magnetic properties of films with respect to flat multilayers. The thickness of the layers was monitored during the deposition process by quartz crystal microbalance.

The surface morphology of initial $TiO_2$ templates and Co/Pd multilayers on $TiO_2$ was checked by scanning electron microscope (SEM) HITACHI S-4800 at voltage of 15 kV.

Microstructural studies of granular films were done by transmission electron microscopy (TEM) at normal and high resolution (HRTEM), using Philips EM400T and Philips CM200 microscopes, operated at 120 kV and 200 kV, respectively. The films deposited on Al substrates, prepared by ion-beam thinning for plan or cross-sectional view were analysed.

X-ray reflectivity (XRR) and X-ray diffraction (XRD) of thin CoPd films were performed with two-circle laboratory diffractometer (Panalytical X'Pert Pro) equipped with PW3050/60 vertical goniometer and PW3383/00 X-ray tube (2.2 kW, Cu = 1.54056 Å) were applied for verification of structure and phase composition of Co/Pd films. XRR results were fitted using X'Pert Reflectivity software.

The magnetization (μ) of granular films were measured in the temperature range from 2 K up to 300 K and in magnetic fields $H$ up to 90 kOe using the Vibrating Sample Magnetometer (VSM) option of the Quantum Design Physical Property Measurement System (PPMS). Magnetization (M) of CoPd MLs was measured at room temperature in magnetic field H up to 20 kOe using MPMS-XL SQUID (Quantum Design). During measurements magnetic field was applied in parallel ($H_{\parallel}$) and perpendicular ($H_{\perp}$) direction to the film surface.

The $^{57}$Fe transmission and conversion electron Mössbauer spectroscopy measurements were performed using a conventional constant acceleration type spectrometer with a 20 mCi $^{57}$Co in Rh source moving at room temperature. Values of center shift $\delta$ are related to $\alpha$–Fe at RT.

Peculiarities of magnetic anisotropy in complex composite systems were analyzed by magnetic resonance spectroscopy at room temperature using continuous wave X-band (microwave frequency 9.32 GHz) electron spin resonance (ESR) spectrometer Varian E112 with modulation of magnetic field at frequency 100 kHz (the absorption signals were detected as its first derivative in the spectrum) [12]. The features of magnetic anisotropy were characterized in terms of components of demagnetizing tensor $N_{IP}$ and $N_{OOP}$ and the first order effective uniaxial anisotropy constant $K_{eff}$.

## 7.3 Template-assisted CoPd antidot arrays with perpendicular magnetic anisotropy

### 7.3.1 CoPd porous multilayered films fabricated on porous $Al_2O_3$ template with highly ordered nanopores distribution

SEM images of initial AAO templates reveal porous films with perfectly ordered hexagonal arrangement of nanopores (see Figure 7.1). The size of fabricated antidots is equal to initial pore

diameter of used AAO templates. SEM images for CoPd antidots prepared by deposition of CoPd MLs over AAO template are shown in Figure 7.1.

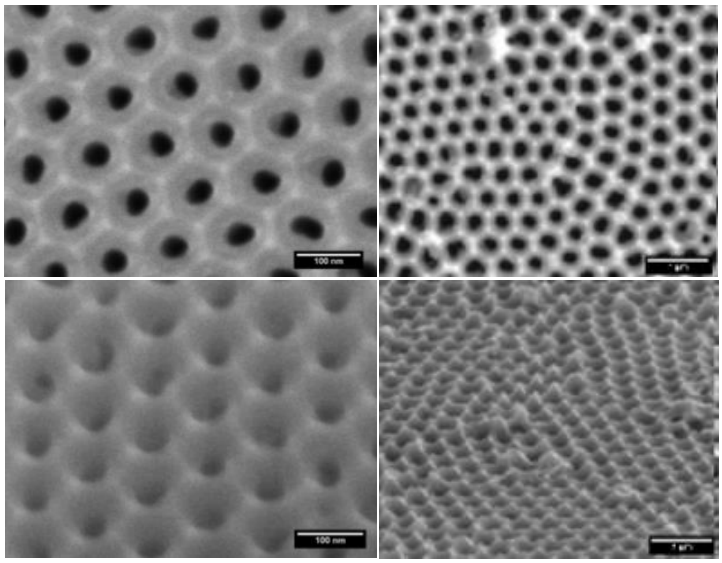


Figure 7.1. SEM images of initial AAO templates (a, b) and CoPd MLs deposited over the AAO templates (c, d) obtained at angles of $90^0$ (a, c) and $30^0$ (b, d) with respect to the sample surface.

Due to imperfect round shape of nanopores the dimensions of pores were characterized with two parameters, Feret's diameter $D_F$ and equivalent diameter $D_E$. $D_F$ is the longest distance between any two points along the selection boundary, while $D_E$ is calculated from the mean area of the pores using the assumption of perfect circular pores. Estimated from SEM images values of morphological features in the initial AAO template and as deposited CoPd MLs are summarized in Table 7.1.

Table 7.1. Nanopore diameters $D_E$ and $D_F$ for AAO template and Co/Pd MLs deposited on AAO.

| Sample | $D_E$, nm | $D_F$, nm |
|---|---|---|
| AAO template | 177±4 | 186±4 |
| Co/Pd porous ML film | 174±5 | 185±5 |

XRR analysis of continuous Si/$SiO_2$//Pd/[$Co_{0.3\ nm}$/$Pd_{0.55\ nm}$]/Pd MLs (not shown here) revealed that *R* values of Co and Pd interfaces are generally larger than the nominal thickness of films. This should be associated with intermixing of Co and Pd layers during the deposition resulting in the formation of $Co_xPd_{1-x}$ quasi – alloy [5, 6]. Therefore we assumed that the whole CoPd multilayer constitutes $Co_xPd_{1-x}$ quasi – alloy with thickness that equals to the thickness of 15 deposited CoPd bilayers.

Experimental XRD patterns of continuous and porous CoPd MLs are presented in Figure 2. The patterns of continuous MLs show only broad overlapped reflections that were deconvoluted by pseudo-Voigt functions into two contributions with position of peaks at $2\theta \approx 40^0$ for Pd buffer/cap, and $2\theta = 40.7^0 - 41.1^0$ for CoPd MLs. Lattice parameters for both components are calculated based on approach assuming the formation of $Co_xPd_{1-x}$ quasi-alloy with cubic fcc structure. The obtained values (see Table 7.2) well correspond to (111) peak of Pd and (111) peak of $Co_xPd_{1-x}$ quasi-alloy [13]. Similar lattice constants *a* calculated for continuous and porous films indicate very close composition of $Co_xPd_{1-x}$.

From the width of Pd and CoPd reflections in XRD patterns of continuous and porous films the coherent scattering length $D^{coh}$ was estimated with Scherrer equation. It is seen that $<D^{coh}>$ for porous CoPd is 17-18 nm and 8-12 nm, respectively. One can see that the coherent scattering length for antidots does not change in comparison to continuous films [14].

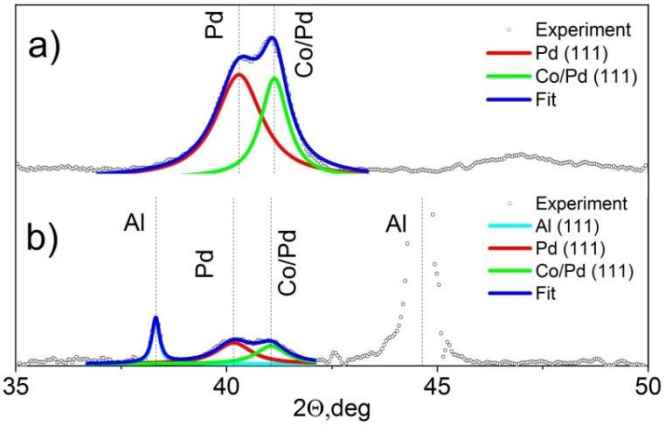



Figure 7.2. XRD patters of continuous Si//$SiO_2$//[$Co_{0.3\ nm}$/$Pd_{0.55\ nm}$]/Pd (a) and porous $Al_2O_3$//Pd/[$Co_{0.3\ nm}$/$Pd_{0.55\ nm}$]/Pd (b) films [14].

Table 7.2. Structural data obtained by XRD analysis of CoPd continuous and porous films: *a* – lattice parameter, $D^{coh}$ – the coherent scattering length corresponding to grain size [14].

| Film | *a* (Å) | $D^{coh}$ (nm) | *a* (Å) | $D^{coh}$ (nm) |
|---|---|---|---|---|
| | Pd | | $Co_xPd_{100-x}$ | |
| Continuous | 3.877±0.002 | 7.1±0.5 | 3.801±0.002 | 12±2 |
| Porous | 3.888±0.005 | 9.5±1.4 | 3.808±0.005 | 11±3 |

Figure 7.3 shows magnetization loops *M(H)* for continuous and porous CoPd ML film. The values of coercive field ($H_C$), remanent magnetization ($M_r$) and saturation magnetization ($M_S$) (or

magnetization at $H_{max}$ = 2.0 kOe) determined from these curves are summarized in Table 3. The magnetization loops of CoPd MLs were normalized to the mass of Co.

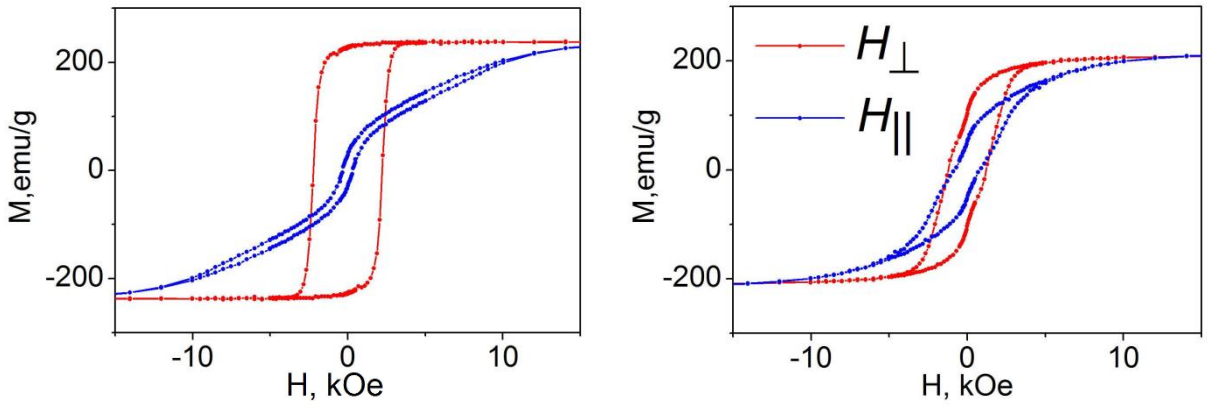



Figure 7.3. Magnetization loops *M(H)* for reference (a) and porous (b) CoPd MLs [14].

Almost perfect rectangular shape of *M(H)* dependence for reference CoPd films for perpendicular geometry shows that easy axis of magnetization is along the normal to the surface with squareness of the loop $M_r/M_S \approx 0.97$. Additionally, the film shows very large perpendicular coercive field $H_{\perp C}$ equal to 2.2 kOe, demonstrating its prominent PMA. Noticeably, PMA is conserved in CoPd antidots, but with reduced values of $M_r/M_S$ (up to 0.5) and coercive field $H_C$ of approximately 1.2 kOe. Some deterioration of PMA is also confirmed by *M(H)* curves for continuous and nanoporous CoPd films in hard magnetization direction which demonstrated enhanced contribution of in-plane magnetic anisotropy at the expense of PMA. It can be explained by hemispherical distributi on of the material on the U-shaped surface of $Al_2O_3$ templates with advanced morphology.

Table 7.3. Magnetic properties of continuous and porous CoPd films. $H_C$ – coercive field, $M_r$ – remanent magnetization, $M_S$ – saturation magnetization [14].

| Samples | $M_S$, emu/g | $H_C$, kOe | $M_R/M_S$ | $H_C$, kOe | $M_R/M_S$ |
|---|---|---|---|---|---|
| | | $H_{\perp}$ | | $H_{\parallel}$ | |
| continuous | 23.5 | 0.3 | 0.13 | 2.2 | 0.97 |
| porous | 21.0 | 0.8 | 0.26 | 1.2 | 0.50 |

### 7.3.2 CoPd porous multilayered films fabricated on nanotubular $TiO_2$ templates

The $TiO_2$ template is shown in Figure 4 as top-view SEM image and images obtained for electron beam tilted at $45^0$ and $30^0$ angles. Anodized $TiO_2$ has tubular structure with quite homogeneous distribution of nanopores. Argon plasma etching of template surface led to the efficient smoothing of surface relief as compared to the initial relief observed immediately after anodization.

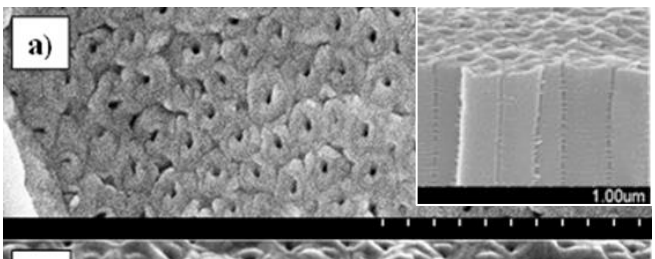


Figure 7.4. SEM images of Pd/[$Co_{0.3\ nm}$/$Pd_{0.55\ nm}$]/Pd ML film deposited on porous $TiO_2$ (a) top-view image; (b) image taken at an angle of $60^0$; and (c) at an angle of $45^0$ with respect to sample surface. The inset illustrate nanotubular structure of porous $TiO_2$ template.

The analysis of SEM images exhibited specific features of template morphology. In particular, Figure 7.4*b* shows that the top of each nanotube consists of separate cells with mildly declining crater-like shape with the pore inside with sharp outer edges. To estimate size of morphological features of $TiO_2$ (*D*, *d* and *w*) the top-view SEM image of the template was analyzed by ImageJ platform. Because of inaccurate spherical shape of the nanotubes and pores their dimensions were described by two Feret's diameters (outer $D_F$ and inner $d_F$) [15] which are defined as the longest distance between any two points along the selected boundary. For further considerations the equivalent diameters (outer $D_E$ and inner $d_E$) were calculated from the mean area of the nanotubes and pores with the assumption of their spherical shape. Distributions of $D_E$, $d_E$ and $w_E$ are shown in Figures 7.1 *d*, *e* and *f*, respectively. Estimation of Feret's diameters of the template gave the values of $D_F = 256\pm28$ nm and $d_F = 50\pm14$ nm, greater than values of equivalent diameters $D_E = 214 \pm 22$ nm and $d_E = 32\pm7$ nm. Value of interpore distance $w_E$ was $153\pm36$ nm.

Experimental XRD patterns of continuous and porous multilayers measured in the angular range 2θ from 25 deg to 60 deg are shown in Figure 7.5 (for clarity in the limited range from $36^o$ to $43^o$). The patterns of continuous multilayers show only broad overlapped reflections that were deconvoluted by pseudo-Voigt functions to two contributions with position of peaks at $2\theta = 40.23^0$ for Pd buffer/cap, and $41.23^0$ for the alloy. Assuming the cubic *fcc* structure of Pd and CoPd alloy with [13] we calculated the lattice parameters of both components. The values (see Table 2) well correspond to (111) peak of Pd [27] and (111) peak of CoPd alloy [13]. The alloy-like Co environment is in agreement with detailed X-ray absorption analysis (EXAFS) of short range atomic structure and interatomic distances in Co/Pd MLs [13, 16].

XRD patterns of porous films besides of reflections for Ti substrate foil (the porous $TiO_2$ is amorphous), which overlap with peaks of interest, reveal shoulder at right side of Ti peaks and broadening much beyond the instrumental linewidth of the diffractometer. On the basis of data for reference sample this shoulder was deconvoluted with pseudo-Voigt functions into two lines reflecting contributions of Pd buffer/cap and CoPd alloy. During the fitting the position and areal contribution of Pd peak is fixed at the value obtained from XRD patterns for reference film. The lattice parameter *a* of CoPd alloy is shown in Table 7.4. From the width of Pd and CoPd alloy

reflections in XRD patterns of porous films we deduced from Scherrer formula the coherent scattering length $D^{coh}$ which corresponds to grain sizes.

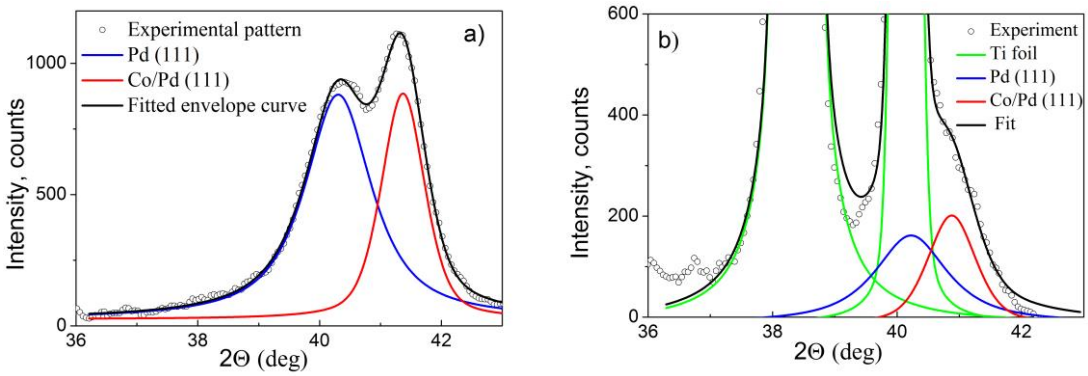



Figure 7.5. XRD patters of continuous $Si//SiO_2//[Co_{0.3\ nm}/Pd_{0.5\ nm}]/Pd$ (a) films and porous $TiO_2//Pd/[Co_{0.3\ nm}/Pd_{0.55\ nm}]/Pd$ (b) films

Table 7.4. Structural data obtained by X-ray diffraction analysis of continuous and porous films: $a$ – lattice parameter, $D^{coh}$ – the size of regions of coherent scattering. Parameters notified with* is fixed during fitting procedure.

| Film | Pd | | $Co_xPd_{100-x}$ | |
|---|---|---|---|---|
| | $a$ (Å) | $D^{coh}$ (nm) | $a$ (Å) | $D^{coh}$ (nm) |
| Continuous | 3.871 ± 0.005 | 6.9 ± 0.3 | 3.776 ± 0.005 | 11.2 ±0.8 |
| Porous | 3.871* | 6.9* | 3.819± 0.019 | 10.3± 5.9 |

Figure 7.6 shows magnetization loops *M(H)*, normalized to the mass of Co, for continuous and porous $Pd/[Co_{0.3\ nm}/Pd_{0.55\ nm}]/Pd$ films. Almost perfect rectangular form of *M(H)* dependencies measured in perpendicular geometry shows that easy magnetization axis for continuous Co/Pd film lies along the normal to its surface. Herewith continuous film shows high $H_C$ = 2200 kOe in perpendicular direction ($H_\perp$) and squareness of the loop $M_r/M_S \approx 0.96$, evidencing its prominent perpendicular magnetic anisotropy. However, the ($H_\parallel$) loop for the film shows hysteretic behavior characterized with the S-shape which is the signature of in-plane anisotropy contribution typically originating from the shape anisotropy of thin film.

As demonstrated in Fig. 7.6b, perpendicular magnetic anisotropy is conserved in the nanoporous film, but with a reduced value of $M_r/M_S$ = 0.71 and an enhanced coercive field $H_C$ = 2.7 kOe for $H_\perp$ loop associated with the pinning of magnetic moments on nanopore edges [1, 4].

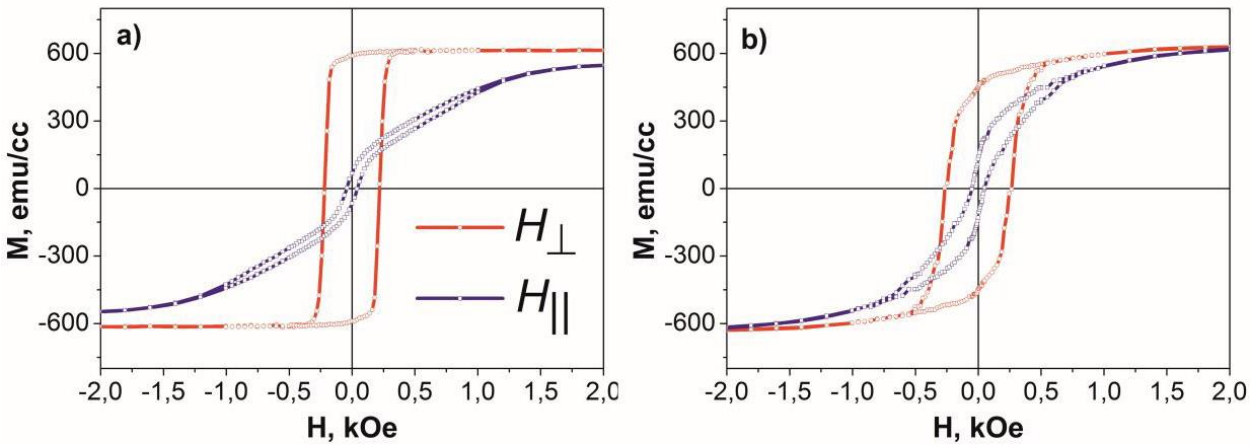



Figure 7.6. Magnetization loops *M(H)* measured in perpendicular ($H\perp$) and parallel ($H\parallel$) geometries for the continuous $Si/SiO_2$//Pd/[$Co_{0.3\ nm}$/$Pd_{0.55\ nm}$] Pd (a) and porous $TiO_2$//Pd/[$Co_{0.3\ nm}$/$Pd_{0.55\ nm}$]/Pd film (b).

Generally, in case of continuous film with well-defined hard and easy magnetization axis, value of effective anisotropy constant $K_{eff}$ could be calculated by formula $K_{eff} = H_K \cdot M_S/2$. In this case value of effective perpendicular anisotropy field $H_K$ could be estimated directly from $M(H_{//})$ loop as the field $H$ required for saturation of magnetization along hard magnetization axis. However, Figure 7.6 clearly reveals that $M(H_{//})$ curves for studied continuous and porous Pd/[$Co_{0.3\ nm}$/$Pd_{0.55\ nm}$]/Pd films are not ideally linear typical for magnetic orientation along hard magnetic axis thus drawing to some uncertainty in estimation of $H_K$ values from $M(H_{//})$ curve.

For that reason effective perpendicular magnetic anisotropy in both films should be more accurately characterized as the difference of areas under $M(H_\perp)$ and $M(H_{//})$ loops [17]. Calculated values of $K_{eff}$ and $H_K$ for continuous and porous films are summarized in the Table 7.5.

Table 7.5. Values of effective anisotropy field $H_k$ for continuous and porous Pd/Co/Pd films as well as magnetocrystalline anisotropy field $H_a$. $H_d$ – magnetostatic shape anisotropy (demagnetizing) field, $M_S^{Vol}$ – volume magnetization of film in $H_\perp$ direction, $K_{eff}$ – effective anisotropy constant, $K_U$ – uniaxial (magnetocrystalline) anisotropy constant

| Film | $K_{eff}\cdot10^{-6}$, erg/cm$^3$ | $H_k$, kOe | $H_d$, kOe | $H_a$, kOe | $K_U\cdot10^{-6}$, erg/cm$^3$ |
|---|---|---|---|---|---|
| Continuous | 3.9 | 12.8 | 7.7 | 20.5 | 6.3 |
| Porous | 1.7 | 5.4 | — | — | — |

It should be noted that effective anisotropy field $H_K = H_a - H_d$, where $H_d$ is magnetostatic shape anisotropy (demagnetizing field), $H_a$ is magnetocrystolline anisotropy. For continuous infinite thin films $H_d = -4\pi\cdot M_S^{Vol}$, where $M_S^{Vol}$ is volume magnetization of film as a whole [18]. Calculation of $M_S^{Vol}$ values is performed assuming thickness of CoPd films and neglecting capping and buffer Pd layers. Evaluated values of $M_S^{Vol}$, $H_d$ and $H_a$ (for continuous film only) are presented in Table 7.5.

After that value of magnetocrystalline anisotropy constant $K_U$ for continuous film was calculated by formula $K_U = M_S^{Vol}(H_K + 4\pi M_S^{Vol})/2$ [18], accounting for the value of $H_a = H_K + H_d$ presented in the Table 7.5. Estimated value of $K_U \approx 6.3\cdot10^6$ erg/cm$^3$ is consistent with previously reported for continuous Co/Pd ML films [19]. It should be emphasized that value of $H_d$ for porous film cannot be calculated reliably basing on available experimental data thus making impossible determination of correspondent $K_U$.

In order to clarify the possible mechanisms of the magnetization reversal in the continuous and porous films, the dependencies of the coercive field *HC* on the angle ϕ between the direction of easy magnetization and the applied field *H* are presented in Figure 7.7.

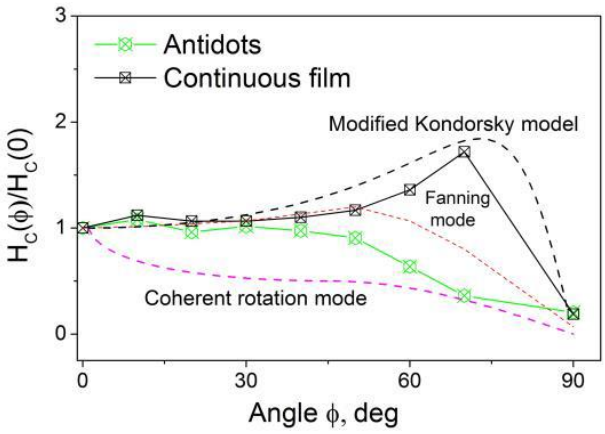



Figure 7.7. Coercive fields $H_C$ as a function of an angle ϕ between the applied field *H* and the easy magnetization axis for the antidots and continuous films. Theoretical *HC*(ϕ) dependencies according to a modified Kondorsky model [20] and the rotational mechanisms [21] are shown on the plot by dashed black and dashed red color, respectively.

For the continuous film we observed the maximum for *HC(ϕ)* at ϕ equal to approximately 70-75 deg. Such shape is typical for the magnetization reversal considered in a modified Kondorsky model, which combines the magnetization reversal by domain wall propagation with the rotational mechanism [20]. The coercive field *HC* for the porous film is constant in the ϕ range of 0-50 degrees, and then it slowly decreases. Such dependence is similar to that characteristic for the magnetization reversal by the incoherent/coherent rotation of magnetic moments in the nanostructures [21]. Alternatively, experimentally obtained curves could be interpreted as a combination of a domain-wall motion and the rotational mechanisms describing the magnetization reversal in the interpore areas of the film and areas around the nanopore edges, respectively [2].

The method of magnetic resonance spectroscopy allows determination the peculiarities of magnetic anisotropy in complex composite systems. The uniqueness of method of numerical analysis of magnetic resonance spectra is the opportunity to calculate not only the gyromagnetic ratio but also components of demagnetizing tensor and the effective uniaxial anisotropy constant $K_U$. This helps to understand the nature of magnetic anisotropy of a particular specimen [22, 23]. Magnetic properties of CoPd continuous and porous multilayered films fabricated on nanotubular $TiO_2$ templates were analyzed by magnetic resonance spectroscopy at RT. The orientation of samples in an external magnetic field was characterized by the polar angle $\phi_H$ between the normal to the surface of a sample and magnetic field *H*. The $\phi_H$ value of 0° corresponds to the direction of external magnetic field perpendicular to the surface of a sample labeled hereinafter as OOP (out-of-plane). Similarly the direction of external magnetic field parallel to the surface of a sample ($\phi_H$ = 90°) is labeled as IP (in-plane).

Magnetic resonance spectra of $TiO_2$//Pd/[$Co_{0.3\ nm}$/$Pd_{1\ nm}$]/Pd film at perpendicular and in-plane directions of applied magnetic field represented on Figure 7.8. The signals are not saturating with the microwave power increase that denote the ferromagnetic nature of absorption. As one can see in case of perpendicular orientation besides the main ferromagnetic absorption line signals of modes spin wave resonance (SWR) are observed indicating the spin waves propagation towards the normal to the surface.

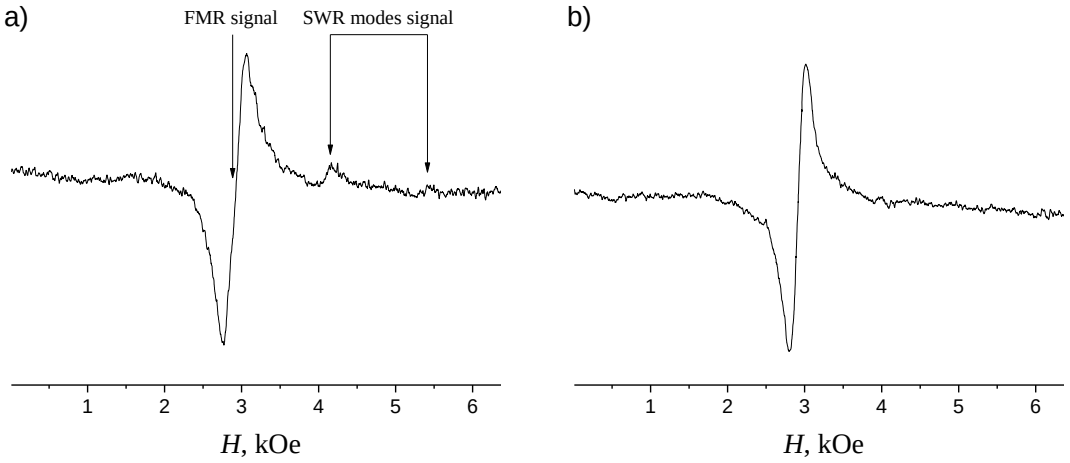



Figure 7.8. Magnetic resonance spectra of $TiO_2$//Pd/[$Co_{0.3\ nm}$/$Pd_{1\ nm}$]/Pd film at perpendicular (a) and in-plane (b) directions of magnetic field.

Numerical analysis of magnetic resonance spectra of CoPd multilayered films based on Landau – Lifshitz magnetization dynamics and the energy dispersion relationship [24] was made using the Smit-Beljers-Suhl formula for a resonance condition [25]:

$$\left(hf/g\mu_B\right)^2 = \left.\frac{\left(E_{\theta\theta}E_{\phi\phi} - E_{\theta\phi}^2\right)}{\left(M\cdot \mathrm{Sin}\theta\right)^2}\right|_{\theta=\theta_0,\ \phi=\phi_0} \tag{7.1}$$

where $h$ – Planck constant, $f$ – microwave frequency, $\mu_B$ – Bohr magneton, $M$ – magnetization of a sample, $E_{\theta\theta}$, $E_{\phi\phi}$ и $E_{\theta\phi}$ – the second derivatives of the sample free energy density $E$ with respect to angles ($\theta_M$, $\phi_M$) of magnetization in spherical coordinate system (see Fig. 7.1) calculated at ($\theta_0$, $\phi_0$) corresponding to the equilibrium magnetization state. The equilibrium state of magnetization is defined by the condition of minimum of free energy density of the system: ($\partial E/\partial\theta = 0$, $\partial E/\partial\phi = 0$).

The free energy density $E$ of a magnetized system with magnetization $M$ in external magnetic field $B$ is considered as the sum of the Zeeman energy, the demagnetization energy and the uniaxial anisotropy energy:

$$E = E_Z + E_d + E_{ua}, \tag{7.2}$$

Assuming that

$$\begin{aligned} E_Z &= -\mathbf{M}\cdot\mathbf{H} = M\cdot H\left(\mathrm{Sin}\theta_M\cdot\mathrm{Sin}\theta_H\cdot\mathrm{Cos}(\phi-\phi_B) + \mathrm{Cos}\theta_M\cdot\mathrm{Cos}\theta_H\right), \\ E_d &= -\frac{1}{2}\mathbf{M}\cdot\mathbf{H}_d = -\frac{1}{2}M\cdot H_d\cdot\mathrm{Cos}^2\theta_M, \\ E_U &= K_U\cdot\mathrm{Cos}^2\theta_M = \frac{1}{2}M\cdot\frac{2\cdot K_U}{M_S}\mathrm{Cos}^2\theta_M \end{aligned} \tag{7.3}$$

Where $M_S$ – saturation magnetization of the sample, $H_d$ – demagnetizing field caused by the shape anisotropy of magnetic sample, $K_U$ – the first order effective uniaxial anisotropy constant, $\theta_M$, $\theta_H$ – angles between normal to the surface and the magnetization, magnetic field, respectively.

The results of MR spectroscopy reveal that for the in-plane orientations of magnetic field the value of resonant magnetic field $H_r$ does not depend on angle $\phi$. This fact indicates that magnetic properties are isotropic in ($x$, $y$) plane ($z$-axis is collinear with the normal to sample surface). In this case in expression (7.3) the angle $\phi = \phi_B$. The demagnetizing field $H_d$ for isotropic in ($x$, $y$) plane film is proportional to magnetization of a sample and is defined by components of the tensor of demagnetizing factor [26]: $H_d = (N_{IP} - N_{OOP})M$. It should be noted that

demagnetizing tensor is normalized to unity so as it is impossible to induce the $H_d$ greater than magnetization of a sample: $4\pi = 2 \cdot N_{IP} + N_{OOP}$.

Substituting expression (7.3) for the free energy $E$ in the relationship (7.1) we obtain the following equation for the resonant condition of absorption of microwave field by film in an external magnetic field $H_r$:

$$\begin{aligned}\left(hf/g\mu_B\right)^2 = &\left(H_r \cdot \mathrm{Cos}(\theta_B - \theta_M) + \left((N_{IP} - N_{OOP})M + 2 \cdot K_U / M_S\right) \cdot \mathrm{Cos}(2\theta_M)\right) \times \\ &\times \left(H_r \cdot \mathrm{Cos}(\theta_B - \theta_M) + \left((N_{IP} - N_{OOP})M + 2 \cdot K_U / M_S\right) \cdot \mathrm{Cos}^2(\theta_M)\right).\end{aligned} \tag{7.4}$$

Where the value of angle $\theta_M$ is defined by the condition of minimum of free energy density of the system [24]:

$$H \cdot \mathrm{Sin}(\theta_B - \theta_M) + \left((N_{IP} - N_{OOP})M - 2 \cdot K_U / M_S\right) \cdot \mathrm{Sin}\theta_M \cdot \mathrm{Cos}\theta_M = 0. \tag{7.5}$$

In order to determine the values of components of demagnetizing tensor $N_{IP}$ and $N_{OOP}$ and the first order effective uniaxial anisotropy constant $K_U$ magnetic resonance spectra were registered at perpendicular ($\theta_H = 0$, $H_r = H_r^{OOP}$) and in-plane ($\theta_H = \pi/2$, $H_r = H_r^{IP}$) orientations of the samples. Thus solving the system of equation using expressions (7.4) and (7.5) we determined the values $N_{IP}$, $N_{OOP}$ and $K_U$ represented in Table 7.6.

Table 7.6. Values of components of demagnetizing tensor ($N_{IP}$, $N_{OOP}$), the uniaxial anisotropy constant ($K_U$) and demagnetizing field $H_d$ for continuous and porous CoPd films.

| Sample | g-factor | $2 \cdot N_{IP} / 4\pi$ | $N_{OOP} / 4\pi$ | $K_U$, erg/cm$^3$ | $H_d$, kOe |
|---|---|---|---|---|---|
| Continues film [$Co_{0.3\ nm}/Pd_{1\ nm}$] | 2.16 | -0.01 | 1.0 | $1.7 \cdot 10^6$ | 6.55 |
| Porous film [$Co_{0.3\ nm}/Pd_{1\ nm}$] | | 0.53 | 0.47 | $1.6 \cdot 10^6$ | 1.59 |
| Continues film [$Co_{0.3\ nm}/Pd_{0.55\ nm}$] | | -0.03 | 1.0 | $2.1 \cdot 10^6$ | 7.18 |
| Porous film [$Co_{0.3\ nm}/Pd_{0.55\ nm}$] | | 0.59 | 0.41 | $1.6 \cdot 10^6$ | 0.96 |

According to values of components of demagnetizing tensor represented in table 6 CoPd continuous multilayered films can be considered as infinitely thin films with $H_d = 4\pi M_S$. At the same time the porosity of multilayered CoPd films leads to alternation of demagnetizing tensor with the decrease of $H_d$ that can be easily calculated using estimated values of $N_{IP}$ and $N_{OOP}$. As it is known the value of $K_U$ in multilayered films is determined by the thickness and material of layers [18]. This can explain a very small difference between the calculated values of $K_U$ (see table 7.6).

The calculated values of the uniaxial anisotropy constant $K_U$ are in good agreement with those represented in [18] for similar CoPd multilayered films. Thus we believe that described method of determination of components of demagnetizing tensor ($N_{IP}$, $N_{OOP}$) and the uniaxial anisotropy constant $K_U$ using magnetic resonance spectroscopy and SQUID magnetometry data allows calculation of these parameters quite accurately.

## 7.4 Growth-induced non-planar magnetic anisotropy in FeCoZr-$CaF_2$ nanogranular films: structural and magnetic characterization

As was previously reported, oriented nanostructures inducing PMA could be successfully grown in the films by ion-beam sputtering techniques at specific deposition regimes [8, 9]. Complex analysis of microstructure and magnetic properties of $(FeCoZr)_x(CaF_2)_{100-x}$ nanocomposite films was carried out for verification of correlations between synthesis conditions, types of magnetic nanostructures formed and an observed effect of PMA.

The results of cross-sectional TEM analysis of $x = 57$ and $x = 73$ films are shown on Figures 7u.9(a), 7.9(b) and 7.9(c), 7.9(d), respectively. They clearly illustrate the columnar nature of the films. Metallic columns are seen as dark and dark grey areas, and the insulating as bright areas.

The *bcc* α-FeCo(Zr) and *fcc* $CaF_2$ phases were identified by SAED (inset in Figure 7.9(a)) and FFT (inset in Figure 7.9(d)). The columns start to grow nearly perpendicular to the substrate surface (Figure 7.9(c)), but as can be seen in both (a) and (c) they tend to incline with increasing distance from the substrate and close to the top of the film their departure from the surface normal is of 20°. Nevertheless, the columns stretch in their length over the whole film thickness, from the interface with the substrate up to the top surface. Inset in Figure 7.9(c) shows that the metallic columns consist of agglomerated nanograins, approximately 3-5 nm in size, one sitting on top of another. However, the high-resolution images in (b) and (d) show elongation of these nanograins along the column lengths. Also, the metallic columns are denser in the sample containing higher concentration of the metallic phase.

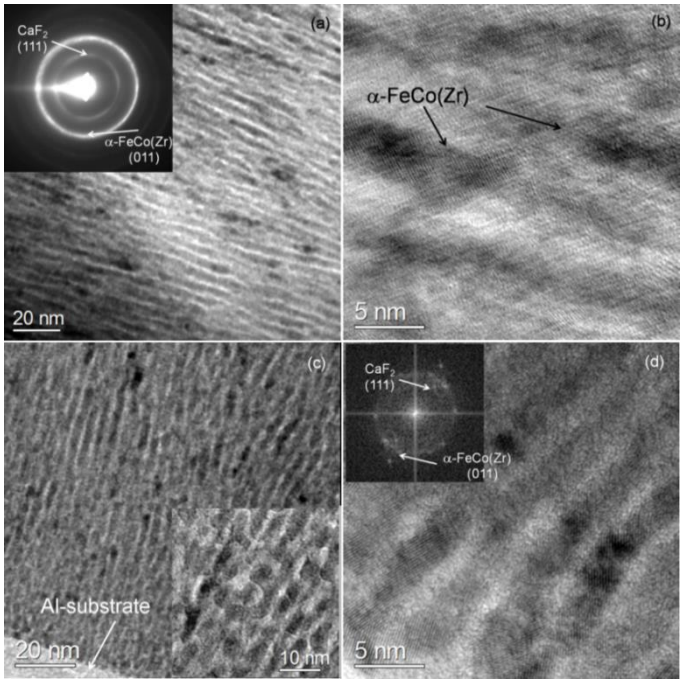



Figure 7.9. Cross-sectional TEM analysis of $(FeCoZr)_{57}(CaF_2)_{43}$ (a), (b) and $(FeCoZr)_{73}(CaF_2)_{27}$ films (c), (d). Insets in (a) and (d) show SAED and FFT patterns taken from the corresponding films.

Columnar shape of metallic nanoparticles with preferential growth orientation in the film normal direction influences on magnetic properties of material. Mössbauer spectra of the films with elongated nanoparticles ($x = 57$ and $x = 73$) demonstrate specific ratio of line intensities (Figure 7.10a). At $x = 73$ magnetically split spectrum reveals the ratio of line intensities in the sextet corresponding to α-FeCo(Zr) phase $h_1 : h_2 : h_3 : h_4 : h_5 : h_6$ amounting to $3 : \sim 0.3 : 1 : 1 : \sim 0.3 : 3$ instead of $3 : 2 : 1 : 1 : 2 : 3$ occurring for random orientation of magnetic moments. The coefficient defined as $K = h_2/h_3$ equals 0.3 for the $x = 73$ sample at RT. Its value is practically unchanged also at $T = 78$ K and reflects preferential orientation of α-FeCo(Zr) nanoparticles magnetic moments in the direction near to the film normal.

The angle θ between normal to the film surface and the magnetic moments in nanograins, estimated from the relationship $\theta = \arccos\{[(4-K)/(4+K)]^{1/2}\}$ [27] equals ~24° for $x = 73$ and ~29° for $x = 57$ that correlates well with cross-sectional TEM images.

Magnetization curves $\mu(H)$ of $(FeCoZr)_{73}(CaF_2)_{27}$ film measured at different temperatures in film plane direction are shown in Figure 7.10b. They demonstrate almost linear field dependence up to saturation that corresponds to out of plane easy magnetization axis. Anisotropy field, $H_a$,

corresponding to a hard axis saturation is about 1.5 kOe which his much higher than saturation of isotropic α-FeCo(Zr) film (~0.3-0.6 kOe [28]). The pronounced linear field dependence up to anisotropy field, $H_a$, points out domination of the easy axis of magnetization lies out-of-plain in these films. Recalling a columnar character of the α-FeCo(Zr) nanograins evidenced from TEM analysis which grow nearly perpendicular to the substrate surface one can try to estimate the tilting angle θ between the normal to the films surface and the mean direction of magnetic moments (the easy axis) using relation $\theta = 90^o - \arccos(M_r/M_s)$, where $M_r$ – is remanence, $M_s$ – is saturation magnetization at $H_a$. For $x = 73$ film characterized at 2 K by the derived θ angle equals 5° that proves close to film normal orientation of nanoparticles magnetic momants.

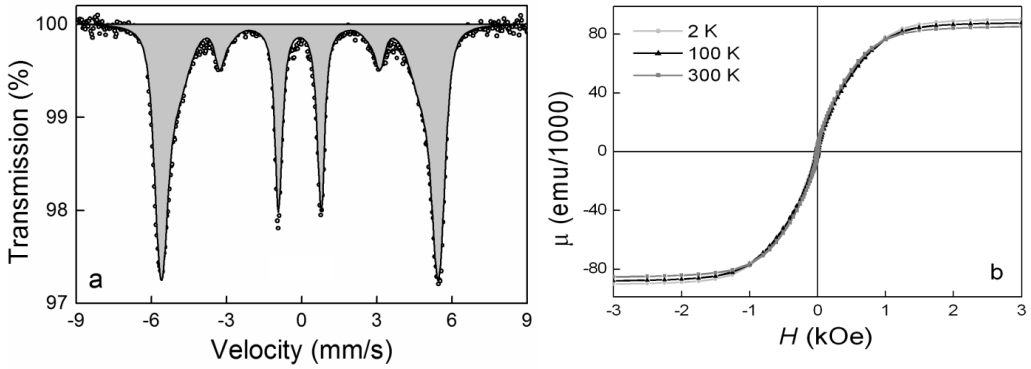



Figure 7.10. RT Mössbauer spectra (a) and magnetization curves μ(*H*) at different temperatures (b) of $(FeCoZr)_{73}(CaF_2)_{27}$ film.

In order to minimize deviations of particles long-axes orientation from films normal as well as to homogenate the size of magnetic nanocolumns and their distribution inside matrix irradiation of $(FeCoZr)_x(CaF_2)_{100-x}$ films demonstrating PMA was carried out with heavy Xe ions in the direction of the films normal.

Mössbauer spectra of $x = 57$ and $x = 73$ films irradiated with fluence $D = 10^{13}$ ion/cm$^2$ are presented in Figure 7.11. Irradiation of $x = 57$ film with $D = 10^{13}$ ion/cm$^2$ leads to the sextet lines broadening, whereas θ decreases from 29 to 24º. The latter indicates the enhancement of PMA as compared with initial film. Sextet broadening is the result of some structural disordering in α-FeCo(Zr) phase after irradiation. Mössbauer spectra of $x = 73$ film after irradiation with $D = 10^{13}$ ion/cm$^2$ presented in Figure 7.11b also shows some PMA enhancement according to the relative line intensities, similarly to the case of $x = 57$ sample irradiated with the same dose. However the decrease of θ is low (from 24º for initial film to 21º for irradiated specimen).

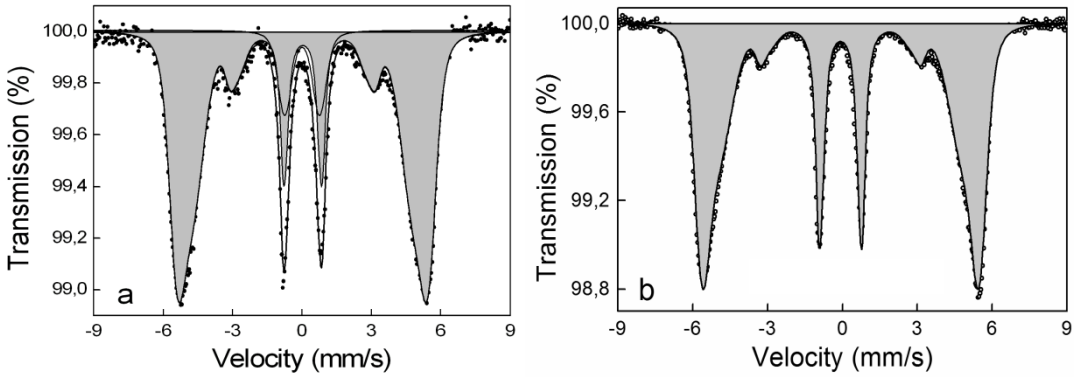



Figure 7.11. Mössbauer spectra of $(FeCoZr)_{57}(CaF_2)_{43}$ (a) and $(FeCoZr)_{73}(CaF_2)_{27}$ (b) films irradiated with the fluence $D = 1 \cdot 10^{13}$ ion/cm$^2$.

Magnetization curves μ(*H*) of $x = 73$ film irradiated by $Xe^+$ ions with fluence $D = 8 \cdot 10^{12}$ ion/cm$^2$ measured in film plane and in film normal direction are presented in Figure

7.12. This regime of irradiation leads to the maximal changes of the film magnetic anisotropy, according to the results of vibrating magnetometry.

As can be seen in Figure 7.12a, irradiation leads to the essential increase of $H_a$ value which exceeds 2 kOe at 3 K as compared to $H_a \sim 1.5$ kOe for the initial film. The shape of $\mu(H)$ curves also changes. Namely, after irradiation, magnetization curves consist of two almost linear regions joining at $H_a$. Linear increase of $\mu$ at low fields up to $H_a$ evidences that such orientation of magnetic field corresponds to the hard magnetization direction. The increase of $H_a$ after irradiation can, therefore, be attributed to the enhancement of perpendicular magnetic anisotropy in the film. It is associated with ion-induced refinement of metallic nanocolumns arrangement inside matrix perpendicularly to the surface when single ion tracks "cut" nanoparticles that deviate from normal orientation. In contrast to that, irradiation of the $x = 73$ film with higher fluence $D = 1 \cdot 10^{13}$ ion/cm$^2$ does not lead to such significant changes in the shape of $\mu(H)$ dependences. However, $H_a$ also increases after irradiation and its value is comparable with that obtained after irradiation with $D = 8 \cdot 10^{12}$ ion/cm$^2$, especially in the case of low temperature $\mu(H)$ curves ($T = 2$-100 K).

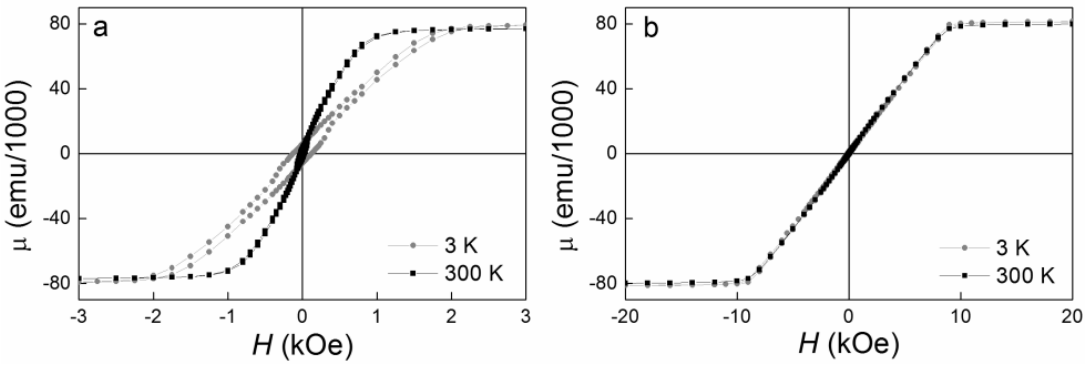



Figure 7.12. $\mu(H)$ dependences of $(FeCoZr)_{73}(CaF_2)_{27}$ film irradiated by Xe ions with the fluence $D = 8 \cdot 10^{12}$ ion/cm$^2$ measured in film plane (a) and in film normal direction (b).

Dependences of $\mu(H)$ for $x = 73$ film after irradiation with $D = 8 \cdot 10^{12}$ ion/cm$^2$ measured in film normal direction are presented in Figure 7.12b. The main peculiarity of $\mu(H)$ curves as compared to the initial film is the decrease of demagnetizing field $H_d$ corresponding to $\mu$ saturation in this detection of $H$ from 12 kOe down to 9 kOe. High $H_d$ parameter is usually observed in the normal direction to the film plane for films with high value of magnetization. The highest $H_d$ value corresponds then to uniform ferromagnetic film. Thus, a decrease of $H_d$ after irradiation indicates decrease of planar contribution to the magnetic anisotropy due to the shape of film with high volume fraction of ferromagnetic material (about 0.59). A possible reason for that is partial destruction of ferromagnetic nanoparticles by single ion tracks that causes formation of additional non-magnetic barriers between elongated nanoparticles preventing their magnetic interaction.

## 7.5 Summary

Co/Pd multilayered antidot arrays with perfect hexagonal ordering were fabricated by deposition of films on AAO templates with $D_E \approx 180$ nm. Phase analysis of as-deposited continuous and porous Co/Pd MLs indicated formation of fcc $Co_xPd_{1-x}$ quasi-alloy with coherence length $D^{coh} \approx 12$ nm.

Porous CoPd MLs revealed conservation of PMA with $H_C = 1.2$ kOe and $M_R/M_S \approx 0.5$, slightly deteriorated with respect to continuous CoPd film due to advanced (U-shaped) surface morphology of CoPd antidot arrays.

For the first time, the nanoporous Co/Pd films with a perpendicular magnetic anisotropy are fabricated by the deposition on the templates of tubular nanoporous TiO2. SEM images of

Co/Pd films show that Co/Pd film reflects the porous structure of the template with a very similar diameter of the pores. It indicates that morphology of the porous film inherits the geometry of the initial TiO2 template with a cone-like shape of the cells with well-defined pores and good periodicity. Phase composition studies for the continuous and porous films show the formation of the CoPd alloy.

The dependence $HC(\phi)$ for the continuous film with large perpendicular magnetic anisotropy favors the reversal magnetization by the domain walls propagation. The coercivity *HC* and remanence *Mr* values for the porous Co/Pd film show that the perpendicular magnetic anisotropy is conserved, but we observed a small deviation from the perfect rectangular shape of the hysteresis loops. This is reflected in the decrease of *Mr*/*MS* value and an essential positive shift of the nucleation field *Hn*. An additional indicator of the anisotropy change for the porous film in comparison to the continuous one is a significant increase of *HC* for the in-plane loops. The $HC(\phi)$ dependence for the nanoporous film can be explained by the rotational mechanism of the magnetization reversal present in the ensemble of exchange-coupled single domain nanograins. Granular nanostructure of the porous films, containing nanograins with sizes below a single-domain limit is manifested by the estimation of $D^{coh}$=10-11 nm.

Using numerical analysis of magnetic resonance spectra data of CoPd multilayered films on the templates of tubular nanoporous TiO2 it was shown that the porosity leads to alternation of demagnetizing tensor with the decrease of demagnetizing field lowering the influence of magnetostatic shape anisotropy and enhancing the PMA. The calculated values of the uniaxial anisotropy constant $K_U$ are about $2 \cdot 10^6$ erg/cm$^3$ that is in a good agreement with those represented in [18] for similar CoPd multilayered films. The similarity of $K_U$ values for continuous and porous CoPd films can be explained by the fact that $K_U$ in multilayered films is determined mainly by the thickness and material of layers.

It should be emphasized that despite some deterioration of the perpendicular magnetic anisotropy in the Co/Pd film on the tubular $TiO_2$ template, the value of the coercive field $H_C$ is essentially larger than values reported for the majority of porous films, such as Co/Pd [2], deposited on the porous $Al_2O_3$, as well as for Co/Pt grown on the porous $ZrO_2$ [1, 4].

A systematic analysis of structure, phase composition and magnetic properties of granular $(FeCoZr)_x(CaF_2)_{100-x}$ films shows that at a high alloy content $x$ = 58-73 at.% nanoparticles demonstrate a columnar structure, with regularly distributed isolated alloy and matrix columns, a few nm wide and stretching in length from the substrate to the top of the films. Orientation of columns is nearly perpendicular to the film plane at the substrate interface, but tilts by up to ~20° off normal. The alloy columns consist of *bcc* α-FeCo(Zr) nanocrystalline grains (3-5 nm), slightly elongated along the column. Such a columnar structure explains non-planar shape induced magnetic anisotropy in the films, observed already below the electric percolation threshold ($x$ ~58). The effect is more pronounced with increasing the alloy content (up to $x$ = 73). The best alignment of easy magnetization axes of nanograins corresponds to an average angle 5–20° with respect to the normal to the film surface.

The origin of columnar growth is attributed to the deposition conditions, where the elevated substrate temperature allows for a sufficient atomic mobility and the two phases – α-FeCo(Zr) and $CaF_2$ – can separate and form such a structure.

Irradiation of nanocomposite $(FeCoZr)_x(CaF_2)_{100-x}$ films ($x$ = 58-73 at.%) demonstrating perpendicular magnetic anisotropy by heavy Xe ions with fluences $D = 8 \cdot 10^{12}$-$1 \cdot 10^{13}$ ion/cm$^2$ allows enhancement of anisotropic properties of the films, namely, to decrease in deviations of particles magnetic moments orientation with respect to films normal from 24-29° to 21-24°, increase in perpendicular magnetic anisotropy field $H_a$ from 1,5 kOe to 2 kOe and decrease in demagnetizing field $H_d$ from 12 to 9 kOe. These changes are associated with ion-induced refinement of metallic nanocolumns arrangement inside matrix perpendicularly to the surface when single ion tracks “cut” nanoparticles that deviate from normal orientation, as well as with formation of additional non-magnetic barriers between elongated nanoparticles preventing their magnetic interaction.

**References:**


1. M.Grobis, C. Schulze, M.Faustini, D.Grosso et al, Appl.Phys.Lett.98 (2011) 192504.
2. M.T. Rahman, N. Shams and C.H. Lai, Nanotechnology 19 (2008) 325302.
3. M.T. Rahman, N.N. Shams, C.H. Lai, J. Fidler, D. Suess, Phys. Rev. B 81 (2010) 014418.
4. V. Neu, C .Schulze, M.Faustini, J.Lee, D.Makarov et al, Nanotechnology 24 (2013) 145702.
5. S.K. Kim, Y.M. Koo, V.A. Chernov, J.B. Kortright et al, Phys. Rev. B 62 (2000) 3025.
6. S.K. Kim and S.C. Shindoi, J. Appl. Phys. 89, No. 5 (2001) 3055.
7. O. Hellwig, T. Hauet, T. Thomson, E. Dobisz, J. D. Risner-Jamtgaard et al., Appl. Phys. Lett. 95 (2009) 232505.
8. S.M. Ryabchenko, A.A. Timopheev, V.M. Kalita, O.V. Stognei, A.V. Sitnikov, J. Appl. Phys. 109 (2011) 043903-1.
9. J.V. Kasiuk, J.A. Fedotova, J. Przewoznik, J. Zukrowski, M. Sikora, Cz.Kapusta, A.Grce, M.Milosavljević, J. Appl. Phys. 116, Iss. 4 (2014) 044301.
10. A. Maximenko, J. Fedotova, M. Marszalek, O. Kupreeva, S. Zavadski. Proc. Intern. Conf. ”Nanomeeting-2015”. World Scientific, Singapore. – 2015. – P. 132-135.
11. A. Maximenko, J. Fedotova, M. Marszalek, Y. Zabila, A. Zarzycki et al, Proc. Intern. Conf. ”Nanomeeting-2015”. World Scientific, Singapore. – 2015. – P. 350-353.
12. Weil, J. A. Electron paramagnetic resonance – Elementary theory and practical applications / J. A. Weil, J. R. Bolton, J. E. Wertz. – New York: Wiley-Interscience, 1994. – 592 p.
13. J. Carrey, E.A. Berkowitz, W.F. Egelhoff jr., D.J. Smith, Appl.Phys.Lett. 83 (2003) 5259.
14. A. Maximenko, J. Fedotova, M. Marszałek, A. Zarzycki, Y. Zabila, J. Magn. Magn. Mater. 400 (2016) 200.
15. M. L. Yan, N. Powers, D. J. Sellmyer, J. Appl. Phys. 93 (2003) 8292.
16. P. Caro, A. Cebollada, D. Ravelosona, F. Briones, et al, J. Appl. Phys. 81 (8) (1997) 24.
17. D. Navas, M. Hernández-Vélez, M. Vázquez, W. Lee, K. Nielsch. Appl Phys Lett 90 (2007) 192501.
18. M.T. Johnson, P.J.H. Bloemen, F.J.A. den Broeder, J.J. de Vries, Rep. Prog. Phys. 59 (1996) 1409.
19. K.Yakushiji, T.Saruya, H.Kubota, A.Fukushima, T.Nagahama et al, Appl. Phys. Lett. 97 (2010) 232508.
20. N.P.Suponev, R.M.Grechishkin, M.B.Lyakhova, Yu.E.Pushkar, 157J. Magn.Magn. Mater. (1996) 158 376.
21. G.A.Milne, D.J.Dunlop, J. Geophys. Res. 111 (2006) B12S08.
22. J.-M.L. Beaujour, W. Chen, K. Krycka, C.-C. Kao, J.Z. Sun, A.D. Kent, The European Phys. J. B 59 (4) (2007) 475.
23. Ferromagnetic Resonance - Theory and Applications. Ed. by O.Yalçın. – InTech (2013) 46p.
24. L.D. Landau and E.M. Lifshits, Electrodynamics of Continuous Media, 2nd edition, Oxford, England: Pergamon Press, 1984.
25. J. Smith, H. G. Beljers, Philips Res. Repts. 10 (1955) 113.
26. B.D.Cullity, C.D.Graham, Introduction to Magnetic Materials. John Wiley & Sons : IEEE Press, Hoboken, New Jersey 2009. –544 p.
27. F. Bodker, M. F. Hansen, C. Bender Koch et al, Phys. Rev. B 61 (2000) 6826.
28. J. Fedotova, J. Kasiuk, J. Przewoznik, Cz. Kapusta et al, J.Alloys Compds. 509 (2011) 9869.



*Acknowledgements*

We acknowledge the support from the Belarusian State program “Functional materials”, projects 1.16 (2011-2013), 1.37 (2014-2015), 2.4.8 (2014-2015), National Science Centre (NCN,

Poland) grant 2014/13/N/ST8/00731, the Serbian Ministry of Education, Science and Technological Development, project OI 171023, EU COST Action MP0903 Nanoalloys, Belarusian Foundation for Fundamental Research (project F14M-029), Mianowski Fund (Warszawa).

Special gratitude to Prof. Yu. Kalinin and Dr. Hab. A. Sitnikov for providing the samples. Special gratitude for Dr. Hab. V. Skuratov for samples irradiation.

## 8. Nanoscale Materials and Structures for Spintronics
## 8.1.The basic mechanisms of charge curriers transport in solids
### 8.1.1 Diffusion mechanism of conductivity and weak localization mode

With the simplest ideas about the diffusion motion of free charge carriers in the conduction band under the action of an electric field through a crystal, its conductivity over a wide temperature range is described by the classical Drude formula.

$$\sigma(T) = \frac{ne^2\tau_p}{m*}, \tag{8.1}$$

where: $e$ – is the electron charge; $m^*$– is the effective masse; $n$ – is the free electrons concentration; $\tau_p$ – is the electron time of relaxation; $T$ – is the temperature.

However, Drude's formula does not allow to explain a number of anomalies in the temperature dependence of the conductivity of disordered electronic systems near the insulatot-to-metal transition, in particular, when approaching the transition from its "metallic" side. In this case, the diffusion motion of carriers from one electric contact to another is accompanied by quantum interference of electron waves propagating along a closed trajectory in two opposite directions. When quantum interference is taken into account, the carrier diffusion coefficient decreases in comparison with that for classical analysis, i. e., the conductivity decreases (it is the so-called processes of weak localization). Moreover, the influence of interference processes increases with decreasing dimensionality of the system, since the formation of self-intersecting trajectories increases in this case.

An accurate calculation based on the localization concept of Taules [1] showed that the temperature dependence of the quantum corrections to the classical conductivity due to this process in three-dimensional (3D) and two-dimensional (2D) conductors can be written as:

$$\Delta\sigma_L{}^{3D}(T) = \frac{e^2}{2\pi^2\hbar}\left(D\tau_\varphi\right)^{-1/2}, \tag{8.2}$$

$$\Delta\sigma_L{}^{2D}(T) = \frac{e^2}{2\pi^2\hbar}\ln\left(\frac{\tau_\varphi}{\tau_p}\right), \tag{8.3}$$

where $\hbar$ – is the Planck constant; $D$ – is the diffusion coefficient; $\tau_\varphi$ – is the electron phase-breaking time.

The value of $\tau_\varphi$ is determined by non-elastic or quasi-elastic electron scattering: electron-electron, electron-phonon interaction or spin-flip scattering. At that $\tau_\varphi \propto T^{-P}$ and $p > 1$. The exponents are: $p = 3/2$ and $p = 2$ for an electron-electron scattering in the cases $k_B T < \hbar/\tau_p$ and $k_B T > \hbar/\tau_p$, respectively. In the case of an electron-phonon scattering mechanism $p = 3$. Here $k_B$ – is the Boltzmann constant and $T$ – is the absolute temperature.

The introduction of a new relaxation time allows us to introduce a new characteristic length (the phase-breaking length) of the wave function at which the phase of the wave function is conserved:

$$L_L = \left(D\tau_\varphi\right)^{1/2}. \tag{8.4}$$

A number of peculiarities in the interaction of electrons are also manifested in disordered electronic systems. If previously it was assumed that, in accordance with the theory of the Landau electron Fermi liquid, allowance for the disordered motion of carriers in the impurity field does

not lead to significant changes in the theory itself, in this case the probability of electron-electron interaction increases because the electron in such a conductor passes the interaction region not "ballistically," but diffusively, multiple-dissipating elastically on impurities. As a result, the probability of the interference processes in the region characterized by the coherence $L_I$ length increases

$$L_I = \left( \frac{D\hbar}{k_B T} \right)^{1/2} . \tag{8.5}$$

Corresponding to 3D and 2D cases, quantum corrections to the classical Drude conductivity due to the enhancement of the electron-electron interaction can be written in the form:

,

$$\Delta\sigma_I{}^{3D}(T) = \frac{1{,}3e^2}{4\pi^2\hbar\sqrt{2}} \left( \frac{4}{3} - \frac{3}{2}F \right) \sqrt{\frac{k_B T}{\hbar D}} , \tag{8.6}$$

$$\Delta\sigma_I{}^{2D}(T) = (1-F)\frac{e^2}{2\pi^2\hbar} \ln\left( \frac{T}{T_0} \right) , \tag{8.7}$$

$$F = \int_0^{2\pi} \frac{d\Theta}{2\left(1 + k_F / K^*\right)} \tag{8.8}$$

where $k_F$ – is the Fermi wave vector; $K^*$ – is the inverse screening length; $\theta$ – is the angle of change in the direction of the wave vector.

Thus, the classical Drude formula does not adequately describe the low-temperature diffusion transport of a charge in weakly disordered crystals in which the temperature dependence of the electrical conductivity can be determined by the quantum interference of electronic waves, and the increasing of the electron-electron interaction. Both these processes depend significantly on the dimensionality of the electronic system.

### 8.1.2 Hopping conductivity

In a lightly doped semiconductors when average distance between impurities with concentration $N_d$ is more than Bohr radius $a_B$ ( $N_d a_B{}^3 << 1$), at low temperatures conductivity do not tends to zero. It has resulting value due to movement (hopping) of the electrons through the localized into forbidden band electronic states. For lightly doped and lightly compensated semiconductors at not very low temperatures dependence of the conductivity can be described by activation mechanism with constant activation energies. It is so-called nearest neighbour hopping regime:

$$\sigma(T) = \sigma_1 \exp\left( -\frac{\varepsilon_1}{k_B T} \right) + \sigma_2 \exp\left( -\frac{\varepsilon_2}{k_B T} \right) + \sigma_3 \exp\left( -\frac{\varepsilon_3}{k_B T} \right) , \tag{8.9}$$

Where first term corresponds to the band conductivity with activation energy $\varepsilon_1$, threds term corresponds to hopping conductivity through impurity states with a constant activation energy $\varepsilon_3$, and, finally, second term corresponds activation mechanism of conductivity ($\varepsilon_2$ – conductivity), when charge curriers move through single charged donor impurities.

Two approximations are used in order to describe electron transport properties on insulating side of the insulator-to-metal transition: i. Mott's conception of minimal metallic conductivity; ii. stimulated by phonons hopping of electron between localized states because of

presence of exponentionally small overlapping of wave functions of nearest neighour states or electron tunneling through localized states near the Fermi level.

Supposing in the first case transfer of electrons onto mobility edge the conductivity can be described by the expression:

$$\sigma(T) = \sigma_{\min} \exp\left(\frac{-(\varepsilon_0 - \varepsilon_F)}{k_B T}\right), \qquad \sigma_{\min} = g^2 \frac{e^2}{2\hbar^2 a_0}. \tag{8.10}$$

where $\varepsilon_0$ – is the energy which divide localized and extended states; it is so-called mobility edge; $\varepsilon_F$ – is the Fermi energy; g – is the constant of electrons interaction; $a_0$ – is the lattice constant;

$\sigma_{\min}$ – is the value of minimal conductivity in the point transition metal-insulator.

For the first time the theory of thermally activated conductivity was developed by Mott [2]. He used model of equivalent net of resistance by Miller and Abrahams. Supposing, that density of electron states near the Fermy level is constant and radius of localized states does not depend on energy he obtained the expression for temperature dependence of resistance in the regime of strong localization:

$$R = R_0 \exp\left(\frac{T_0}{T}\right)^p, \tag{8.11}$$

where $p = 1/(1 + D)$; $D = 1, 2, 3$ – dimension of a hopping transport; $T_0$ – is the constant, depending on localization radius and density of states in the vicinity of the Fermi level; This constant is labeled $T_0 = T_M$.

Since, unlike the previously mentioned hopping transport, when the temperature changes, the average hopping length changes. That is why such a conductivity mechanism is called a variable range hopping.

Allowance for the Coulomb interaction between localized electrons, carried out by Shklovskii and Efros [3], showed that a gap appears in the density of states in the vicinity of the Fermi level. A change in the density of states in the case of hopping transport leads to the fact that the exponent in the formula (11) p = 1/2, and the constant $T_0 = T_{Sh\text{-}E}$ is determined by the localization radius and dielectric constant of the crystal. The effect of the presence of the Coulomb gap on the hopping transport is manifested at temperatures when the width of the gap exceeds the energy of the thermal fluctuations, i.e., as the temperature is lowered, a transition from the Mott's conductivity mechanism to the Shklovskii-Efros mechanism can be observed.

It should be mention that temperature dependence of the form (8.11) with exponent p = 1/2 is also characteristic for the case when charge transport is carried out by tunneling electrons between conducting clusters in an insulating matrix [4]. The constant $T_0$ in this case is determined by the distance between the clusters, the height of the potential barrier, and the energy necessary for the formation of oppositely charged clusters.

Strongly doped semiconductors are characterized by diffusion mechanism of charge transfer. However, in the presence of strong compensation, a transition to activation conductivity is possible, which is essentially Anderson transition. If, in the case of weak doping and partial compensation, the temperature dependence of the resistance of semiconductors can be described by the first term of the formula (9), where the quantity $\varepsilon_1$, in general, will no longer represent the impurity activation energy and will depend on the temperature, then in case of strong compensation due to a significant change in the relief of the conduction and valence bands, the temperature dependence is determined by the competition of conductivity due to tunneling through potential barriers of various sizes with conductivity due to the carriers, activated at the percolation level.

By the level of percolation is usually understood the energy, which the charge carrier can pass through the sample in a classical way, without tunneling through the barrier, but by rounding it. It is clear that at high temperatures the activation conductivity is more preferable, while at law temperature – tunneling of the electrons between potential wells dominants.

### 8.1.2 Spin-dependent conductivity

According to the first Mott's statement, the conductivity of a metal is the sum of the independent conductivities of electrons with spin up and down:

$$\sigma = \sigma_{\uparrow} + \sigma_{\downarrow} \tag{8.12}$$

Within each conduction channel, it is determined by various factors. To illustrate their role, we write the Drude's formula for conductivity in the following form:

$$\sigma_D = (e^2 k_F^2/\pi\hbar 6)\cdot l. \tag{8.13}$$

where $\sigma_D$ – is the Drude's conductivity per spin, $e^2/\pi\hbar \approx 0.387\cdot10^{-4}$ Ом$^{-1}$ – is the quantum of the spin conduction, $k_F$ is the Fermi wave vector and l is the mean free path, determined by the relaxation time $\tau$ and the electron velocity on the Fermi surface, $v_F$, i.e. $l = v_F\tau$.

The Drude formula is applicable only to free electrons, but its use provides an excellent qualitative understanding of the factors that determine spin-dependent conductivity. Since the conductivity is determined by the electrons on the Fermi surface, and according to the Pauli principle, electrons below the Fermi level can not receive energy in the electric field, since all states at higher energies are occupied. Therefore, only electrons on the Fermi surface can contribute to the electric current. The mean free path depends on the velocity on the Fermi surface and on the relaxation time. The latter can be determined from the Fermi golden rule:

$$\tau^{-1} = 2\pi/\hbar <V_{CT}^2> g(\varepsilon_F), \tag{8.14}$$

where $<V_{CT}^2>$ - is the mean value of the scattering potential and $g(\varepsilon_F)$ - is the density of electronic states on the Fermi surface for the corresponding spin.

All quantities in these expressions are spin-dependent, but at the same time the reason of such dependence is different. So, the wave vector $k_F$ and the velocity $v_F$ on the Fermi surface are the intrinsic characteristics of the metal and are completely determined by its band structure. In ferromagnetic metals, these quantities are different for electrons with spin direction up and down. The density of electronic states on the Fermi surface $g(\varepsilon_F)$, is also determined by the band structure, which depends on the direction of the spin. At the same time, the scattering potential in formula (8.14) is not an intrinsic property of the metal, but is determined by defects scattering, impurities or phonons. It is clear that the scattering potential can be either spin-dependent or spin-independent. If, for one orientation of the spin, the scattering potential of magnetic or nonmagnetic atoms is one, and for the other differs, strong spin-dependent scattering can be expected. However, it should be noted, that the relaxation time in formula (8.13) is determined by averaging the square of the scattering potential. In this case, different types of scattering centers can give the spin-dependence of its averaged value. In this case, the spin-dependent band structure becomes decisive and gives the main contribution to the spin-dependence of the mean free path and, consequently, the conductivity.

Let us briefly consider the band structure of nonmagnetic (Cu, Ag, Au) and magnetic (Co, Fe, Ni) metals, the combinations of which are most often used in the study of the giant magnetoresistive effect. Since the spin-orbit interaction in *3d* transition metals is weak, the band structure for electrons with spin up and down can be treated independently. The transition metals have *4s*, *4p*, and *3d* valence electrons, which differ in their orbital angular momentum. The states

*4s* and *4p* form a *sp*-conduction band, in which electrons have a high velocity, a low density of states and, consequently, a large mean free path, i.e. we can assume that they are responsible for the conductivity of *3d* metals. At the same time, the *d*-band is localized in a relatively narrow energy interval and is characterized by a high density of states and a low electron velocity. In the energy interval, where the *sp* and *d* - bands overlap, they can no longer be regarded as independent due to strong *sp-d* hybridization, which substantially modifies the band structure. This naturally leads to a significant change in the properties of free electrons, manifested in a decrease in their velocity.

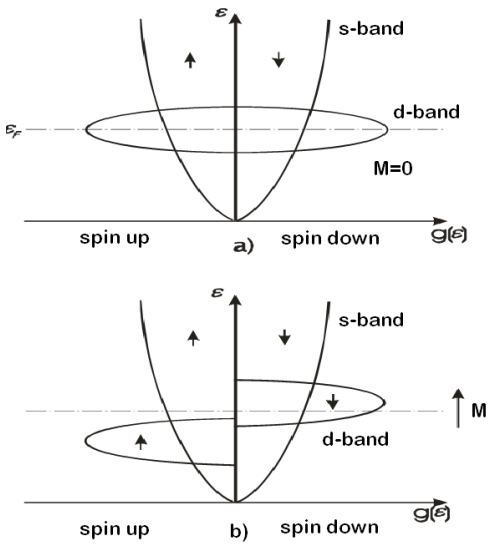



Fig. 8.1. Schematic presentation of the band structure nonmagnetic (a) and magnetic (b) metal

Schematic presentation of a band structure nonmagnetic and magnetic metal is shown in Fig. 8.1. One can see that for ferromagnetic *3d* metals *d*-band splited because of exchange interaction. It is well known that *d*-electrons repulsive because of Coulomb interaction. It leads they have antiparallel spins and occupied the same orbital. Nevertheless, for *d*-electrons lower energy level with parallel spins is more preferable. That is why Coulomb repulsion in accordance with Pauli printsip leads to appearance exchange interaction and spontaneous magnetic moment. While when the spin direction of all electrons is the same, the kinetic energy of system will be increased. This energy can be more than width of *d*-band. These two tendency can be balanced and ferromagnetic ordering became the most preferable. For a system with ferromagnetic ordering well known Stoner's criterium must be valid $j \cdot g(\varepsilon_F) > 1$, where $j$ – is the exchange integral. Because of exchange splitting of the *d*-band a number of occupied states for electrons with different spin direction is different resulting to non-equal magnetic moment which is equal $2{,}2\mu_B$, $1{,}7\mu_B$ and $0{,}6\mu_B$ for Fe, Co and Ni correspondingly.

Conductivity depends on the Fermi level position with respect to *d*-band. In the case of Cu *d* – band is fully occupied and Fermi level is inside of *sp*-band. Since velocity of electrons in *sp*– band is big and density of state is low the scattering probability is low. One can conclude that coper is a good conductor.

On the other hand, for ferromagnetic metal, for example cobalt, because of *d*-band splitting the spin-up subband is fully occupied while the spin down subband is occupied partially. This case Fermi level is inside *sp*-band for spin-up electrons, but in to *d*-band for spin down electrons. It leads to different values of conductivity for spin-up and spin-down bands. For spin-up electrons the situation is similar as for Cu: conductivity determines by *sp* –electrons and it is high enough. At the same time, the conductivity of spin down electrons is not completely determined by *sp*-electrons because of strong *sp-d* hybridization mixing *sp* and *d*-states leading to strong spin asymmetry conductivity for Co.

## 8.2. Magnetoresistive effect for different mechanisms of conductivity
### 8.2.1 Diffusion mechanism of conductivity. Classical magnetic fields

The term magnetoresistance or magnetoresistive effect refers to the change in resistance in the direction of the applied electric field. The simplest ideas about the change in the charge transport in a crystal in the presence of an external magnetic field (magnetotransport) can be obtained from an analysis of the motion of a charged particle in electric and magnetic fields. The magnetic field dependences of the kinetic coefficients are found from the solution of the Boltzmann kinetic equation, which in the presence of external electric and magnetic fields leads to the following expressions for the magnetoresistive effect:

$$\frac{\Delta R}{R_0} = \frac{R(B) - R(0)}{R(0)} = b_r (\mu_H B)^2 , \quad (8.15)$$

where $b_r$ – is the coefficient of magnetoresistance; $\mu_H$ – is the Hall mobility; $B$ – is the magnetic field induction; $R(0)$ – is the resistance without a magnetic field B = 0; $R(B)$ – is the resistance in a magnetic field with induction B.

Expression (8.15) was obtained under the assumption of a power-law dependence of the momentum relaxation time $\tau_p$ on the energy

$$\tau_p = \tau_{p0} \left( \frac{\varepsilon}{k_B T} \right)^r . \quad (8.16)$$

Under these assumptions, the magnetoresistance coefficient may be written in the form:

$$b_r = \frac{\Gamma(5/2 + 3r)\Gamma(5/2 + r)}{\left[\Gamma(5/2 + 2r)\right]^2} - 1 , \quad (8.17)$$

and the resistance in the presence of magnetic field with induction $B$ may be written in the form:

$$R_B = R_0(1 + b_r \beta^2) \quad (8.18)$$

where $\Gamma(x)$ – is the gamma function and $\beta = \mu B$.

Since change of the resistance in a magnetic field is due to Lorenz fource action on a moving electron the effect is called Lorentz magnetoresistive effect. It is positive and $b_r > 0$. As the field increases, the magnetoresistance coefficient (8.17) decreases as $B^{-2}$ with the magnetic field increasing and there is corresponding decrease in the rate of growth of the resistance and saturation value of the resistance in strong magnetic field. The saturation magnetoresistance value depends on scattering mechanism of the charge carriers. Table 8.1 shows the value of exponent $r$ in the expression (8.16) and value of the magnetoresistance coefficient $b_r$ for some values of $r$.

Table 8.1 – Exponents $r$ and value of the magnetoresistance coefficient $b_r$ for different mechanisms of charge carriers scattering.

| № | Scattering mechanism | $r$ | $b_r$ |
|---|---|---|---|
| 1 | Acoustic-mode, deformation potential | 1/2 | 0,274 |
| 2 | Ionized impurities | 3/2 | 0,578 |
| 3 | Neutral impurities | 0 | 0 |
| 4 | Acoustic-mode, piezoelectric potential | -1/2 | 0,087 |

| | | | |
|---|---|---|---|
| 5 | Dislocations | 1/2 | 0,274 |

### 8.2.2 Quantizing magnetic fields

Since kinetic Boltzmann's equation is quasiclassical it is applicable in the region of weak magnetic fields only when cyclotron radius $R_c$ more than electron mean free path $l_p$

$$R_c >> l_p;\ \ \omega_c \tau_p << 1;\ \mu B << 1, \tag{8.19}$$

where $\omega_c = eB/m^*$ is the cyclotron frequency.

Such magnetic fields are called classically weak. If the inverse (8.19) relations are satisfied, that is, when the carrier makes several revolutions along the cyclotron orbit before scattering into another state, then such fields are called classically strong. If the cyclotron rotation energy exceeds the energy, which corresponds to the energy distance between the Landau levels ($\hbar\omega_c >> k_B T$) and the thermal broadening of the level, the magnetic fields are called quantizing. The energy spectrum of an electron in a crystal in the quantum limit with allowance for spin can be written in the form:

$$\varepsilon = \frac{\hbar^2 k_z^{\ 2}}{2m*} + \hbar\omega_c (n_L + 1/2) + m_s \mu_B g B \ , \tag{8.20}$$

where $\kappa_z$ – is the wave vector of the electron in the magnetic field direction; $n_L$ – is the oscillator quantum number (Landau level number); $m_s = \pm ½$ – is the spin quantum number; $\mu_B$ – is the Bohr magneton; $g$ –is the spectroscopic splitting factor.

Quantization of the energy spectrum of charge carriers in a magnetic field leads to a drastic change the dependence of the density of states on the energy

$$g(\varepsilon) = \frac{m^{*3/2}}{\sqrt{2}\pi^2\hbar^3}\frac{\hbar\omega_c}{2}\sum_{n=0}^{n_{\max}}\left[\varepsilon - \left(n_L + 1/2\right)\hbar\omega_c + m_s\mu_B g H\right]^{-1/2} . \tag{8.21}$$

Thus, at the Landau level it turns into infinity and falls off as $g(\varepsilon) \propto \varepsilon^{-1/2}$ while remaining proportional to the magnetic field, when it is away from it.

In quantizing magnetic fields, not only the energy spectrum of electrons changes radically, but also the nature of their interaction with the crystal lattice. Thus, with a parallel orientation of the electric and magnetic fields the scattering leads to transitions between Landau's levels and a nonzero longitudinal magnetoresistance even for spherical surfaces of constant energy. In crossed fields, the current is due to collisions, since in the absence of scattering the current component in the direction of the electric field is zero.

The most general theoretical consideration of MRE in the quantum limit by solving the quantum equation of motion for the density matrix was carried out by Adams and Holstein [5]. For a number of mechanisms of carrier charge scattering, they obtained a power dependence of the positive MRE on the magnetic field and temperature for degenerate and nondegenerate semiconductors

$$\frac{\Delta R}{R_0} = B^{\alpha} T^{\beta} \ . \tag{8.22}$$

These exponents are listed in the table 8.2.

It should be note that, in contrast to the Lorentzian magnetoresistive effect in classical fields in the quantum limit the quantity of α depends on the scattering mechanism and varies over a wide range. For example, α = 0 for the scattering of carriers by impurity ions, and α = 11/2 for

the case of scattering by acoustic phonons in a degenerate electron gas at low temperatures. The oscillating density of states in the quantum limit leads to nonmonotonic dependences of the kinetic coefficients on the magnetic field, which manifest themselves as the Shubnikov-de Haas effect in degenerate materials, magnetophonon and spin-magnetophonon resonances, regardless of degeneracy.

Table 8.2. Magnetoresistive effect in the quantum limit for different scattering mechanisms

| Scattering mechanism | Non-generated semiconductor | | | | Strongly generated semiconductor | | | |
|---|---|---|---|---|---|---|---|---|
| | $\alpha_\perp$ | $\alpha_\parallel$ | $\eta_\perp$ | $\eta_\parallel$ | $\alpha_\perp$ | $\alpha_\parallel$ | $\eta_\perp$ | $\eta_\parallel$ |
| Acoustic-mode, deformation potential (low temp) | 5/2 | -3/2 | 3/2 | -1/2 | 11/2 | 0 | 5/2 | 0 |
| Acoustic-mode, deformation potential (high temp.) | 2 | -1/2 | 1 | ½ | 5 | 1 | 2 | 1 |
| Point defects | 2 | -3/2 | 1 | -1/2 | 5 | 0 | 2 | 0 |
| Acoustic-mode, piezoelectric potential (Low temp.) | 3/2 | -3/2 | 1/2 | -1/2 | 9/2 | 0 | 3/2 | 0 |
| Acoustic-mode, piezoelectric potential (high temp.) | 1 | -1/2 | 0 | ½ | 4 | 1 | 1 | 1 |
| Optic-mode lattice vibration (high temp.) | 1 | -1/2 | 0 | ½ | 4 | 1 | 1 | 1 |
| Ionized impurities | 0 | -3/2 | 0 | -3/2 | 3 | 0 | 3 | 1 |

Finally, we can schematically represent magnetoresistive effect dependence on magnetic field for diffusion conductivity mechanism a following. Parabolic dependence with different coefficient (depending on scattering mechanism) in a weak classical magnetic field, saturation MR effect in a strong classical magnetic field, and finally power law in the region of quantizing magnetic field with different exponents value, which varies from zero to 5,5. Schematically this plot is shown in Fig. 8.2.

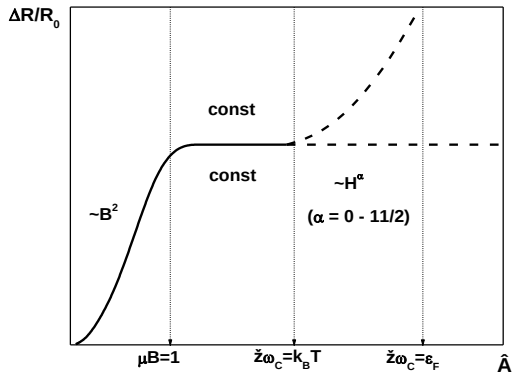



Fig. 8.2. Schematic presentation of magnetoresistive effect dependence on magnetic field.

Contrary Lorentz magnetoresistive effect in classical magnetic fields, in the quantum limit value of $\alpha$ depends on scattering mechanism. For example, $\alpha = 0$ for ionized impurity scattering and $\alpha = 11/2$ for acoustic phonon scattering in the degenerated electron gas at low temperatures. Moreover, oscillating in a quantum limit density of states (formulae (8.21)) leads to nonmonotonic magnetic field dependence electrical and optical properties electronic systems in a strong magnetic field: magnetoresistance oscillation, magnetic permeability oscillation, magnetooptic absorption, etc.

### 8.2.3 Magnetoresistive effect in weak localization mode

In degenerate electron gas Lorentz magnetoresistive effect is very small. This means that in degenerate semiconductors and metals the magnetoresistance is zero, to a first approximation. It maximum value is of order $\Delta R/R_0 \sim (k_B T/\varepsilon_F)$. Consequently, in a weak disordered electronic system on the metallic side of the insulator-to-metal transition change of the resistance in a magnetic field will be determined by change of quantum corrections to Drude conductivity. For the processes of weak localization external magnetic field destroy quantum interference of electron waves and decrease localization probability. As a result the resistance of the weakly disordered electron system in the magnetic field decreases, i. e. magnetoresistive effect is negative. While the magnetic field increases electron-electron interaction because of change of a screaning length, leading to positive MR effect. Calculations for both quantum corrections to conductivity with different dimension in the magnetic field lead to the result:

$$\Delta\sigma_q{}^D(B) = a_q \frac{e^2}{2\pi^2\hbar} \Phi_q{}^D \Big/ l_q{}^{D-2}, \tag{8.23}$$

where $q$ means the weak localization ($q \equiv L$), or electron-electron interactuion ($q \equiv I$ ); $a_q$ is the constant of electron-electron interaction.

Values of all coefficients, characteristic length $l_q^{D-2}$, as well as functions $\Phi_q^D$ for the processes of weak localization and electron-electron interaction are presented as a tables in [6].

In the framework of the theory of weak localization it is possible to explane not only negative magnetoresistance, but also change the sign of the MR from positive to negative in increasing magnetic field. Take part into consideration spin-orbit interaction and scattering of electrons by magnetic impurities the expression for magnetoresistive effect in the regime of weak localization can be written [7]:

$$\Delta\sigma(B) = \left(\frac{e^2}{2\pi^2\hbar}\right)\left\{\Psi\left(\frac{1}{2}+\frac{1}{\tau_p a}\right) - \Psi\left(\frac{1}{2}+\frac{1}{\tau_1 a}\right) + \frac{1}{2}\left[\Psi\left(\frac{1}{2}+\frac{1}{\tau_2 a}\right) - \Psi\left(\frac{1}{2}+\frac{1}{\tau_3 a}\right)\right]\right\}, \tag{8.24}$$

,

where and $\Psi(x)$ – is the digamma function and

$$a = \frac{4DeB}{\hbar}$$

Characteristic relaxation times $\tau_1, \tau_2, \tau_3$ is determined from following equations:

$$\frac{1}{\tau_1} = \frac{2}{\tau_{so}} + \frac{2}{\tau_s} + \frac{1}{\tau_\varphi}; \tag{8.25}$$

$$\frac{1}{\tau_2} = \frac{4}{\tau_s} + \frac{1}{\tau_\varphi}; \tag{8.26}$$

$$\frac{1}{\tau_3} = \frac{2}{\tau_s} + \frac{4}{\tau_{so}} + \frac{1}{\tau_\varphi}, \tag{8.27}$$

where $\tau_{so}$ – is the spin-orbit relaxation time; $\tau_s$ – is the magnetic relaxation time.

### 8.2.4 Magnetoresistive effect in strong localization mode

A characteristic feature of the regime of strong localization is the hopping mechanism of conductivity and its huge value of positive magnetoresistive effect, which exponentially depends on the strength of the magnetic field. The strong increase in resistance is due to the fact that, the magnetic field does not change the trajectory of the charge carrier, but compresses the wave functions of the impurity electrons in the transverse direction, converting them in the simplest case into cigar-shaped ones. In this case, the overlap of the "tails" of the wave functions of neighboring impurities on average decreases sharply, which leads to an exponential increase in the resistance in a magnetic field.

The calculation of the magnetic-field dependence of the magnetoresistance for the hopping conduction mechanism was first carried out by Mikoshiba [8]. Somewhat later the same calculation using the percolation method was made by Shklovskii [9]. It was shown, that there are two different regions of the values of the magnetic field. However, in both regions, the results of calculations are significantly different. The weak-field region, which can be determined as (($L_B > a/(N_Д a^3)^{1/6}$ и $N_Д a^3 << 1$, $L_B = \hbar/eB$ – *magnetic length* ) Micoshiba and Schlovsky expression can be written:

$$\rho_3(B) = \rho_3(0)\exp\left(t\frac{ae^2}{N_Д c^2\hbar^2}B^2\right) \tag{8.28}$$

where Schlovsky exponent is equal t = 0,036 and Micoshiba one - t = 0,083.

According to Micoshiba calculations in the strong-field region the magnetoresistance can be described by formulas

$$\rho_3(B) = \rho_3(0)\exp\left(\frac{eB}{4N_Д^{2/3}c\hbar}\right), \tag{8.29}$$

$$\rho_3(B) = \rho_3(0)\exp\left\{\zeta\frac{2^{1/4}m^{1/4}e^{1/2}\varepsilon_B^{1/4}B^{1/2}}{c^{1/2}\hbar N_Д^{1/2}}\right\} \tag{8.30}$$

where $\zeta = 0{,}92$ and $\varepsilon_B$ – is the shallow donor ionization energy.

One can see they have different magnetic field and impurity concentration dependenses. It deals with a different assumption concerning the hopping angle between the magnetic field direction and direction of the electron hopping.

In the magnetic field interval $B_0 < B < 100B_0$ $\varepsilon_B \sim B^{1/3}$ and expression (8.30) can be rewritten accordingly Schlovsky by formula

:

$$\rho_3(B) = \rho_3(0)\exp\left(constB^m\right), \tag{8.31}$$

where exponent $m$ a little more than $m = 0{,}5$; $B_0$ – is the magnetic field when magnetic length is equal to Bohr radius $l_B = a_B$.

It is necessary to emphasize that in the regime of hopping conductivity not only positive but also negative magnetoresistive effect is predicted theoretically. The first explanation of negative magnetoresistance on the insulating side of the insulator-to-metal transition was considered as a consequence of a change in the characteristic temperature $T_0$ in a magnetic field. As a result of a simple calculation exponentional negative magnetoresistance was predicted.

### 8.2.5 Geometrical and dimensional dependence of the magnetoresistive effect

When considering the magnetoresistive effect on the basis of the Boltzmann kinetic equation, the magnitude of the magnetoresistance of a finite size sample depends not only on the strength of the magnetic field, but also on the shape and dimensions of the sample. For the simplest forms, for example, for a rectangular parallelepiped, detailed calculations are made that relate the magnitude of the magnetoresistance to the ratio of the length *l* of the sample to its width *b* ($u = l / b$). As example for rectangular parallelepiped we can write this dependence for three cases only:

$$\Delta\rho / \rho_0(\beta, u) = \Delta\rho / \rho_0(\infty, \beta) g(u, \beta), \tag{8.32}$$

where

$$g(u,\beta) = \begin{cases} 1, & u = \infty \\ \sqrt{(1+\beta^2)}, & u = 1 \\ 1+\beta^2, & u = 0 \end{cases}.$$

The graphs of the functions *g (β)* for different *u* are shown in Fig. 8.3. The limiting case *u* = 0 (unlimited, infinitely wide sample) can be realized in practice by taking a circular disk with one electrode in the center and the other on the peripheral circle (the so-called Corbino disk, Fig. 8.5 a). The electric field in the disk has only a radial component, and the current lines are the form of logarithmic spirals that intersect the radii of the disk at the Hall angle. Another limiting case u = ∞ can be realized by taking a long thin rod (Fig. 8.5 b.). It is clear that the physical magnetoresistive effect, in which the magnetoresistance coefficient characterizes the mechanism of interaction of charge carriers with the crystal lattice, can be measured on a long thin rod, when the geometric component of the magnetoresistance can be neglected.

If the relaxation time does not depend on the energy, i.e. the exponent *r* in formula (8.16) is equal to zero, then the value of the magnetoresistance of rectangular parallelepiped in a wide magnetic field strength range and the ratio *l / b* can be determined from the following formula:

$$\frac{\Delta\rho}{\rho_0} = \frac{\sqrt{1+\beta^2}\,[1 + f(u)(\sqrt{1+\beta^2} - 1)]}{1 + f(u)(\sqrt{1+\beta^2} - 1)} - 1, \tag{8.33}$$

where

$$f(u) = \frac{2G(u) - 1}{1 - u},$$

and

$$G(u) = \frac{16}{\pi^3}\frac{1}{U}\sum_{k=0}^{\infty}\frac{1}{2k+1} th(2k+1)\frac{\pi}{2}U$$

Fig's 8.3 and 8.4 show the *g* and *G*-factors on *β* and *U*, correspondingly.

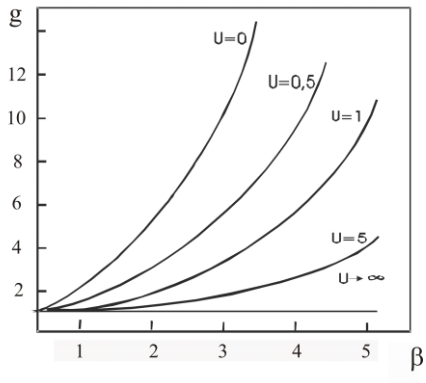



Fig. 8.3. Dependence of the Kurt-Lippman *g*-factor on the magnetic field for different values of *U*

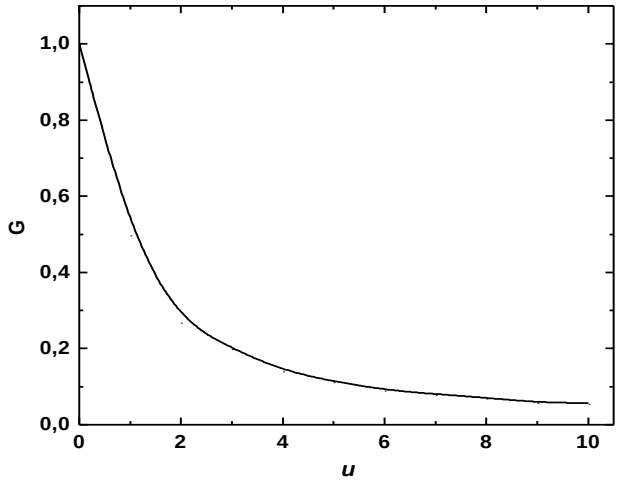



Fig. 8.4. G-factor dependence on length-to-width ratio.

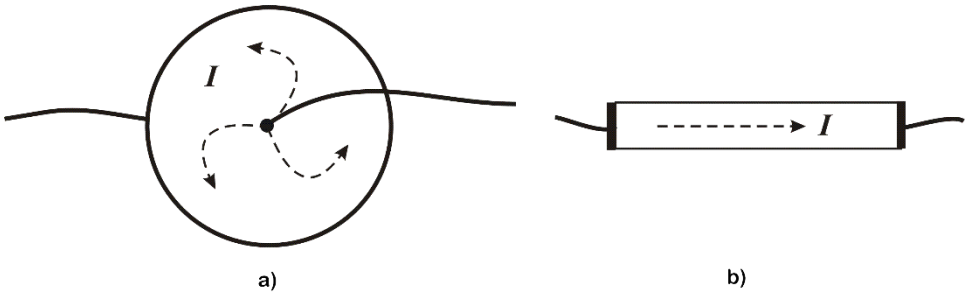



Fig. 8.5. Corbino disk (a) and long rectangular parallelepiped (b)

## 8.3 Peculiarities of the magnetoresistive effect in solids with magnetic ordering

The solids with magnetic ordering have spontaneous magnetization. That is why there is an additional MR effect, which deals with a resistance anisotropy due to the spontaneous magnetization. In such media resistivity components can be devided on two kinds: spontaneous and ordinary.

$$\rho_{ij}(B) = \rho_{ij} + \rho^{0}_{ij}(B), \qquad (8.34)$$

where $\rho_{ij}$ – is the spontaneous coefficients and $\rho^0_{ij}(B)$ – is the "ordinary" coefficients.

Extrapolating the resistance to zero magnetic field we can obtained $\rho$ when the current is parallel to magnetization $\rho_{\parallel}$ and when current is perpendicular to magnetization $\rho_{\perp}$. Schematic representation magnetic field dependence of the resistance of ferromagnetic is presented in fig. 8.6.

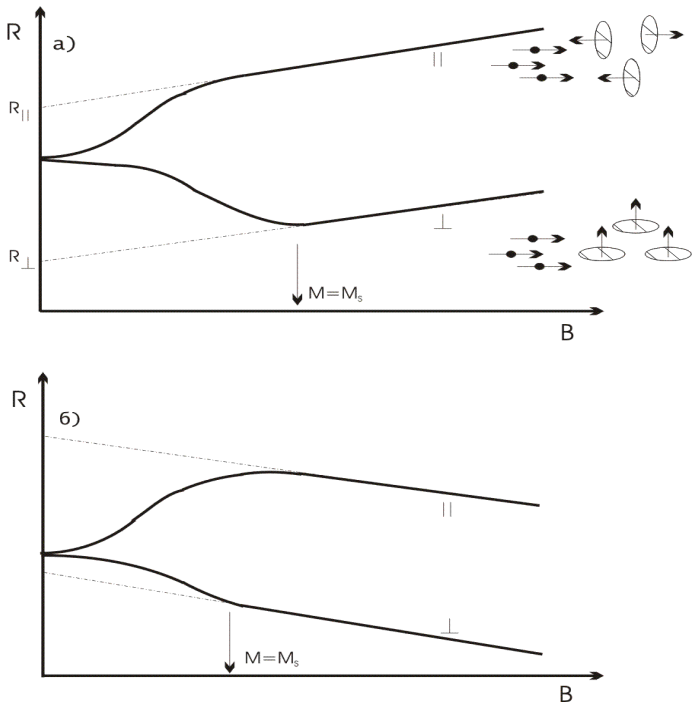



Fig. 8.6. Schematic representation magnetic field dependence the resistance of solids with magnetic ordering

$\rho_{\parallel}$ and $\rho_{\perp}$ reflects reorientation of magnetization with respect to electric current direction. When magnetization reach saturation anisotropic resistance saturates too. An increase or decrease of the magnetoresistive effect at $M > Ms$ can be deal with positive component of Lorentnz magnetoresistance or negative component dealing with spin-dependent scattering.

If $\theta$ is angle between current and magnetization direction the resistivity can be written:

$$\rho_{B=0} = (\rho_{\parallel} + 2\rho_{\perp})/3 + (\cos^2\theta - ⅓)(\rho_{\parallel} - \rho_{\perp}). \quad (8.35)$$

Difference between $\rho_{\parallel}$ and $\rho_{\perp}$ is so-cooled an anisotropic magnetoresistance $\Delta\rho = \rho_{\parallel} - \rho_{\perp}$. It is clear, that at temperatures higher that *Tc* this effect vanishes. Phenomenological theory of the anisotropic magnetoresistance predicts, that $\rho_{\parallel} > \rho_{\perp}$, but in some measurements opposite relation was observed.

Microscopic theory of the AMR was developed by Mott. He assumed, that *s*- electrons are responsible for flowing current and electrons with spin-up and spin-down direction give additive contribution into conductivity. In the framework of so-cooled two-current model he obtained next expression for conductivity

$$\sigma = ne^2/m[1/(1/\tau_{SS} + 1/\tau_{S\uparrow,d\uparrow}) + 1/(1/\tau_{SS} + 1/\tau_{S\downarrow d\downarrow})], \quad (8.36)$$

where $n \equiv n_{S\uparrow} = n_{S\downarrow}$, $\tau_{SS} \equiv \tau_{S\uparrow,S\uparrow} = \tau_{S\downarrow S\downarrow}$. Aarrows indicate the *s* and *d* -electrons spin direction

Formula (8.36) can be rewritten as:

$$1/\rho = 1/(\rho_{SS} + \rho_{S\uparrow,d\uparrow}) + 1/(\rho_{SS} + \rho_{S\downarrow,d\downarrow}), \quad (8.37)$$

and presented as two currents Mott's model.

Band diagram of ferromagnetic material at different temperatures: higher and lower that Tc, as well as two current Mott's model is showed in fig. 8.7 a and b.

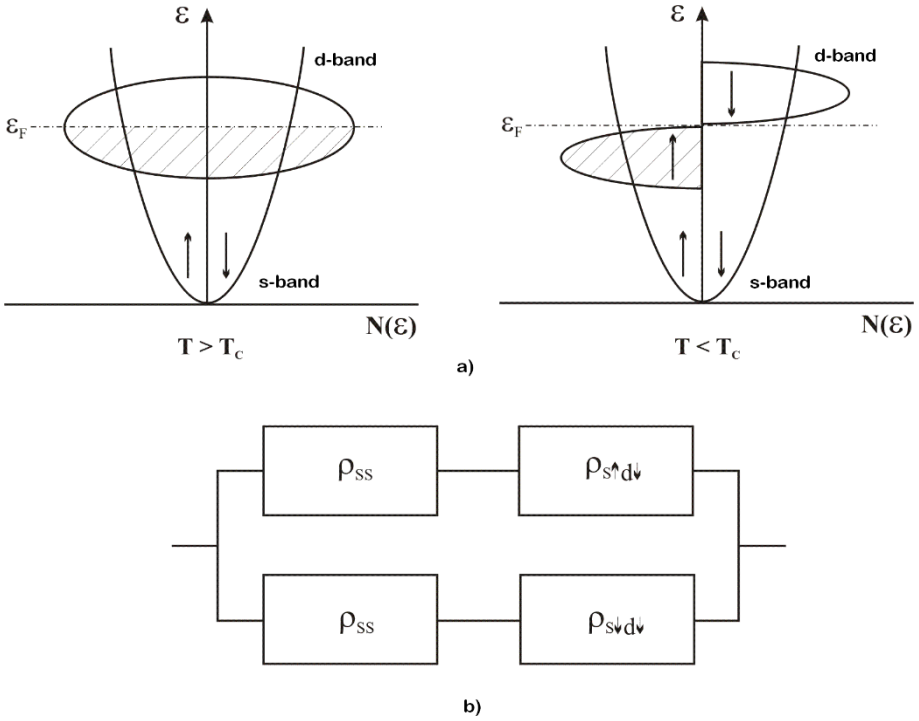



Fig. 8.7. Band diagram of ferromagnetic material when *T*>*Tc* (left panel) and *T*<*Tc* (right panel) (a) and an equivalent scheme of Mott's two current model (b).

One can see that the resistance of ferromagnetic material decrease with increase magnetic ordering since density of states of *d*-band tends to zero ($\rho_{S\uparrow,d\uparrow} = 0$) and spin-up *s*-electrons can't be scattered into *d*-band. Phenomenological theory based on symmetry of crystal and anisotropy of the resistance of solids with magnetic ordering predicts that the resistance is higher when the current is parallel to the sample magnetization direction compare the case when it perpendicular to the magnetization.

In conclusion it should be note additional scattering mechanism in magnetic media which deal with presence of the domain walls. In mesoscopic electronic systems this scattering leads to peculiarities of the MR effect due to domain walls movement at magnetization increase.

## 8.4 Giant magnetoresistive effect and structures for its observation

Like any magnetoresistive effect, the giant magnetoresistance (GMR) is a change in the resistance of a substance under the action of an external magnetic field. It was discovered as a significant decrease in the electrical resistance of the multilayer structure of Fe/Cr. This effect is much larger than the usual Lorentz and anisotropic magnetoresistance, and was therefore called "giant." The change in the resistance of multilayer structures is observed because the external magnetic field orientates the magnetic moments of successive magnetic layers in one direction, as schematically shown in Fig. 8.8. In the absence of a magnetic field, the magnetization of the ferromagnetic layers is antiparallel. The application of the magnetic field directs the magnetic moments in one direction and increases the magnetization of the layers to saturation, which leads to a decrease in the resistance of multilayer structures. The resistance decreases from $R_{AP}$ to $R_P$ as shown in Fig. 8.8 a.

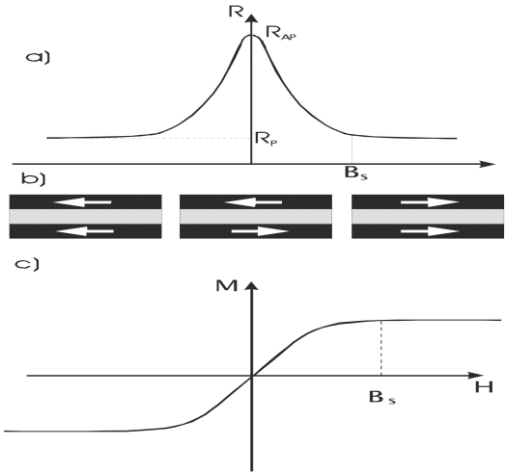



Fig. 8.8. Schematic representation of the GMR effect. The change in the resistance of a magnetic multilayer structure in an external magnetic field (a). The direction of magnetization of the three-layer structure (arrows) in different fields (b). The magnetization curve of the structure (c).

It should be noted, that the presence of antiferromagnetic interlayer interaction, however, is not a necessary condition for observing GMR. Antiferromagnetic ordering can be obtained by using successive ferromagnetic layers with different coercive forces. In this case, the magnetic moments of the soft magnetic and hard layers are reoriented in a different magnetic field, thereby creating an interval of magnetic fields in which they have an antiparallel orientation and the resistance of the layers is large. The magnetic granular medium is also a system with GMR. In such a system, the ferromagnetic clusters are in a nonmagnetic metal matrix. Randomly oriented magnetic moments can be ordered by a magnetic field, which also leads to a decrease in resistance. The various structures possessing GMR are shown in Fig. 8.9.

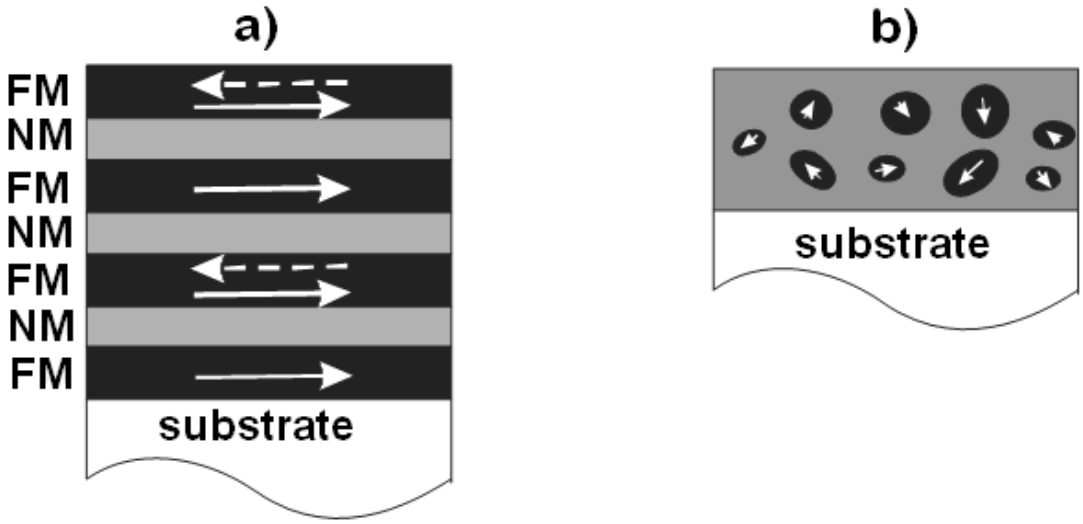



Fig. 8.9. Structures for the observation of GMR. Magnetosensitive multilayer structure FM-NM-FM (a), and inhomogeneous (granular) ferromagnetic film (b).

### 8.5 Resistive model of giant magnetoresistance

Physical nature of GMR can be easily understand from simple four layers model, which consists from two magnetic and two nonmagnetic layers. Inside of every magnetic layers electron spin can be parallel or antiparallel to the direction of magnetization of the layer. According to the resistive model, each metal layer (and each interface) is considered as an independent resistance (resistor). Within each conduction channel, the resistors are connected in parallel or in series, depending on the relationship between the mean free path and the thickness of the layer. If the mean free path is shorter than the thickness of the layer, then each layer conducts the current independently and the resistors must be connected in parallel. Obviously, under such assumptions, the resistance of the parallel and antiparallel configurations is the same and, consequently, the GMR is zero. This means that for the observation of the SMS, the mean free path must be sufficiently large. The latter corresponds to a qualitative model of GMR, based on the possibility of electrons to pass freely through the layer, feeling the direction of the magnetization of successive ferromagnetic layers. Under such assumptions, the mean free path should be comparable with the thickness of the layer. In this case the scattering probability inside the layer will be the sum of the scattering probabilities within each layer and at each interface. Therefore, within the given spin channel, the impedance is the sum of the resistances of each layer and each surface, i.e. the resistors are connected in series.

Fig. 8.10 shows this structure with parallel (a) and antiparallel (b) magnetization orientation of these layers. The resistance for the spin-up and spin-down electrons in these layers is different and they labeled as $R_{\uparrow}$ and $R_{\downarrow}$.

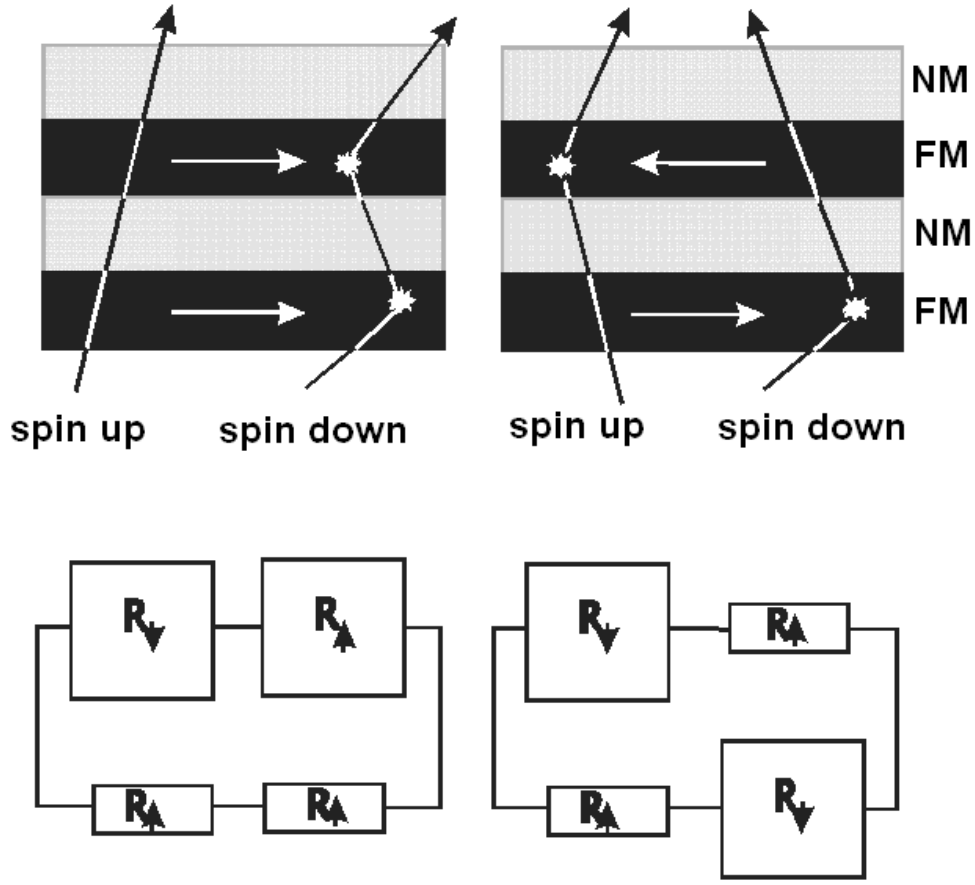



Fig. 8.10. Schematic representation of electronic transport in a multilayer structure with parallel (a) and antiparallel (b) magnetization of the layers. An equivalent circuit is shown below for resistance within the two-current series resistive model.

The resistance of two layers structure, which consists from magnetic and nonmagnetic layers with thickness $d_{nm}$ and $d_{fm}$ for two spins channels can be written:

$$R_{\uparrow\downarrow} = \rho_{NM} d_{NM} + \rho_{\uparrow\downarrow} d_{FM} , \tag{8.38}$$

Total resistance for ferromagnetically ordering layers will be:

$$R_P = N \cdot (R_\uparrow R_\downarrow / R_\uparrow + R_\downarrow), \tag{8.39}$$

and for antiferromagnet ordering

$$R_{AP} = N \cdot (R_\uparrow + R_\downarrow / 2) \tag{8.40}$$

Thus, for MR effect we can obtain next expression

$$\Delta R/R = (R_{AP} - R_P)/R_P = (R_\downarrow - R_\uparrow)^2 / 4 R_\downarrow R_\uparrow \tag{8.41}$$

Supposing that resistivity of nonmagnetic layer is small, compare with ferromagnetic layer expression for GMR can be written.

$$\Delta R/R = (\rho_\downarrow - \rho_\uparrow)^2 / 4 \rho_\downarrow \rho_\uparrow = (\alpha - 1)^2 / 4\alpha, \tag{8.42}$$

where α is relation spin-down and spin-up resistivity. It is the parameter of spin anisotropy, GMR is equal to zero when α =1.

It is also important to note the importance of taking into account the finite resistance of the non-magnetic layer. With its allowance, the GMR value will be determined by the expression:

$$\Delta R / R = (\alpha\text{-}1)\ 2/4\ (\alpha + p d_{nm} / d_{fm})\ (1 + p d_{nm} / d_{fm}), \qquad (8.43)$$

where $p = \rho_{nm} / \rho_{\uparrow}$.

It is easy to see from formula (8.43) that for a given value of α, the GMR will increase with decreasing $pd_{nm}/d_{fm}$. Therefore, in order to obtain a large GMR value, it is important to have a low resistance of a non-magnetic metal. Depending on the $d_{nm}$ thickness, the GMR monotonically decreases and at large thicknesses it is proportional to $1 / (d_{nm})$. It was experimentally found that the GMR decreases exponentially with increasing $d_{nm}$. The reason for this is that a simple model of series resistances is not applicable when $d_{nm}$ is larger than the mean free path. In this case, models for describing GMR should be more complex.

In the GMR model considered above, it was assumed that the resistance in parallel geometry is always less than in the antiparallel one. In most cases, this is true and GMR is usually called "normal". However, if the multilayer structure consists of different ferromagnetic layers, then the GMR can be inverse, i.e. positive. In this case, the GMR will be determined:

$$\Delta R/R = (\alpha_1\text{-}1)\ (\alpha_2\text{-}1) / (\alpha_1\ (1 + q) + \alpha_2\ (1 + q\text{-}1)), \qquad (8.44)$$

where $\alpha_1$ and $\alpha_2$ are the asymmetry parameters of two different ferromagnetic layers, i.e. $\alpha_1 = \rho_{\downarrow}^{(1)}/\rho_{\uparrow}^{(1)}$ and $\alpha_2 = \rho_{\downarrow}^{(2)}/\rho_{\uparrow}^{(2)}$ and *q* is the ratio of the resistances of electrons with spin up of two ferromagnetic layers, i.e. $q = p_{\uparrow}^{(1)}/p_{\uparrow}^{(2)}$. It follows from formula (8.44) that when two ferromagnetic layers have different asymmetries in the resistance, i.e. $\alpha_1 > 1$ and $\alpha_2 < 1$ or vice versa, then the GMR will have an inverse sign, i.e. the structure resistance in a magnetic field will increase.

A multilayer structure for observing inverse GMR and the processes of scattering of electrons in it is shown in Fig. 8.11.

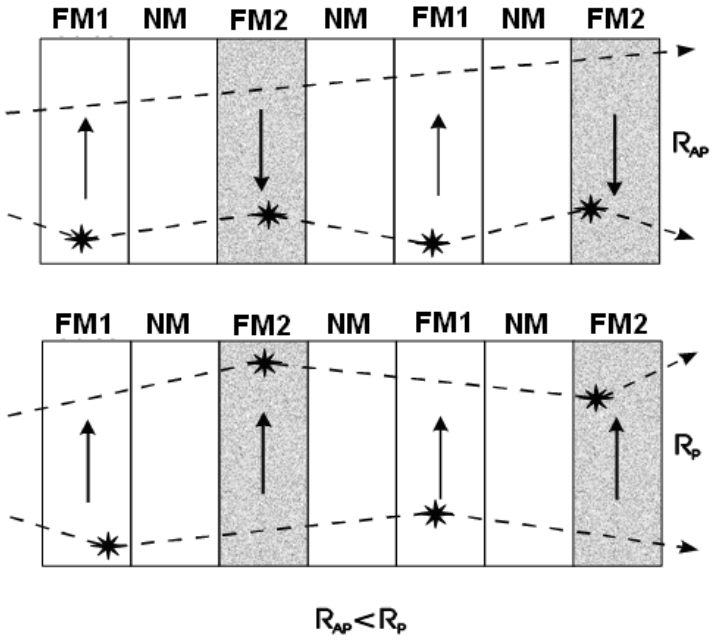



Fig. 8.11. Schematic presentation of the multilayers structure with inverse giant magnetoresistive effect. Stars show scattering of the electrons.

Fig. 8.12 shows typical magnetoresistance (a) and magnetization (b) for magnetic multilayers and inhomogeneous magnetic media. Arrows show the direction of the magnetic field change. One can see the hysteresis magnetization as well as magnetoresistive effect. Magnetoresistive effect reaches maximum value in saturation magnetic field $B_S$ and it is equal to zero in coercive force magnetic field $B_C$.

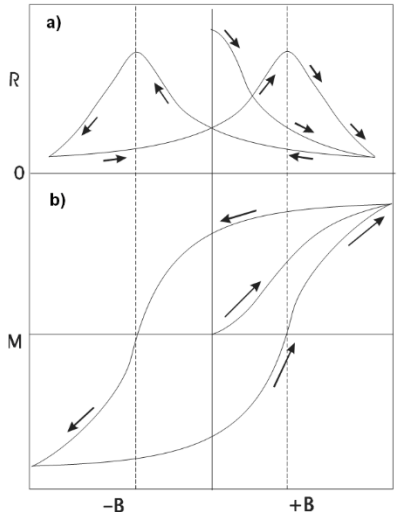



Рис. 8.12. Schematic presentation of the magnetoresistance (a) and magnetization (b) inhomogeneous magnetic media.

## 8.6 Tunneling magnetoresistive effect

The tunnel effect is an example that confirms the wave nature of an electron. If in classical physics the particle is completely reflected from the barrier, when its energy is less than the height of the barrier, then in quantum mechanics there is a nonzero probability of finding a particle on the other side of the barrier. Figure 8.13 a illustrates tunneling through a plane barrier of thickness *d*, with height $\varepsilon_b$, when the electron energy $\varepsilon_F$ smaller than the height of the barrier. Figure 8.13b shows possible energy diagrams of tunnel junctions, and in Fig. 8.13c - tunneling mechanisms: direct tunneling through the barrier N = 0; resonant tunneling through a localized state N = 1; and inelastic tunneling N> 1 or "Variable range hopping".

Let us consider a tunnel junction ferromagnetic-insulator (not a magnet)- ferromagnet, shown in Fig. 8.14 a. In the absence of a magnetic field, ferromagnetic electrodes have the opposite direction of magnetization. The *d*-electron band of the contacts is splited by the exchange interaction as shown in Fig. 8.14 b. In this antiparallel (AP) configuration of magnetized contacts, it is possible to tunnel electrons with spin up from a large number of states to a small number and vice versa, as indicated by the arrows in Fig. 8.14 b. The superposition of the magnetic field leads to a parallel orientation (P) of the magnetization of the ferromagnetic electrodes (Fig. 8.15a). The band diagram for this case is shown in Fig. 8.15 b. In this configuration of the contact magnetizations (P), spin-up electrons tunnel from a large number of states to a large number, and electrons with spin down - from a small number of states to a small one. This leads to the fact that the tunneling resistance for the parallel and antiparallel configurations of the magnetization of the contacts is different. The change in the resistance of such a structure considered above in the reorientation of the magnetization in an external magnetic field is called the tunnel magnetoresistance (TMR). If, for example, the coercive force of the two ferromagnetic contacts is different, and their magnetization configuration varies from antiparallel to parallel and again to antiparallel when the magnetic field is applied, then TMR changes as schematically shown in Fig. 8.15.

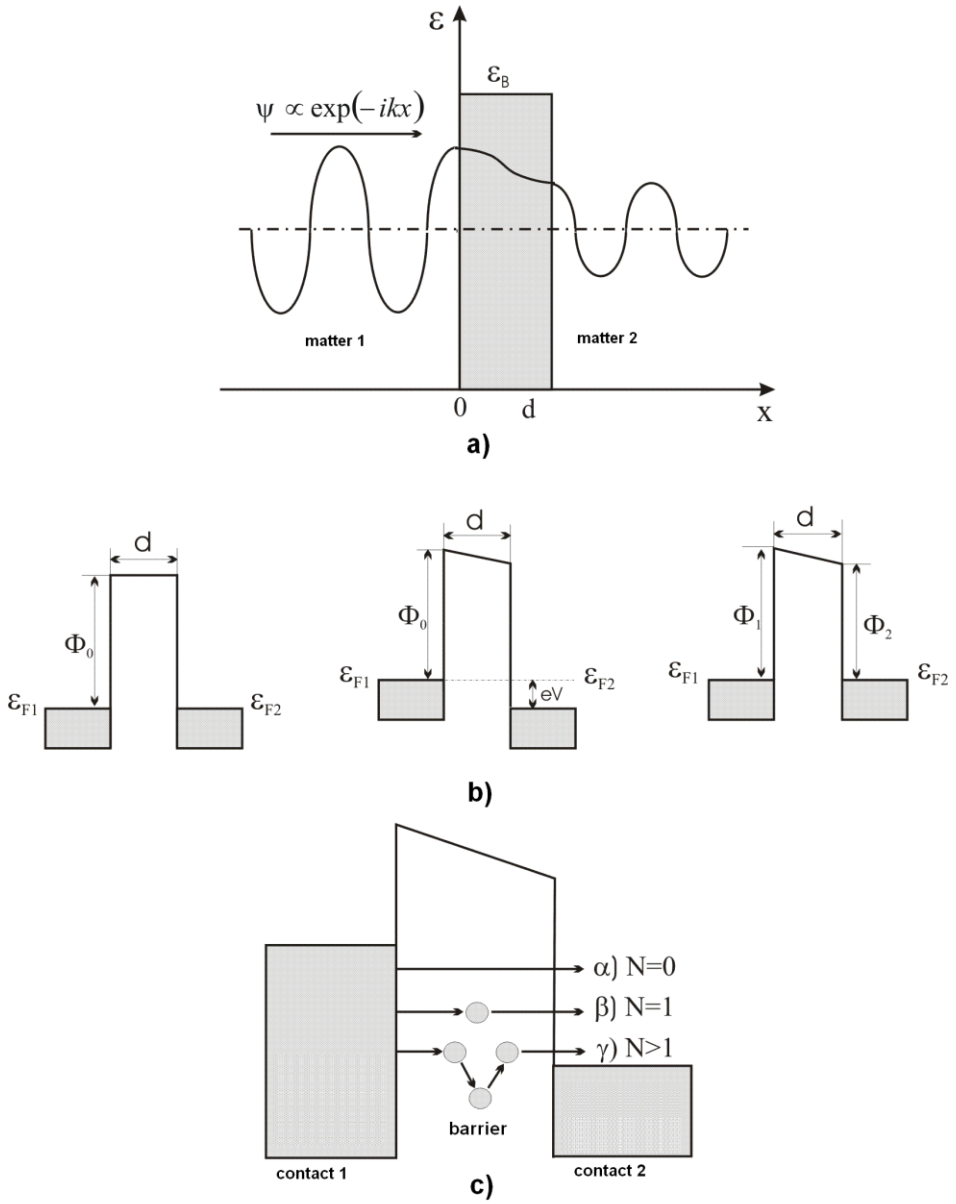



Fig. 8.13. Schematic representation of the wave function in tunneling (a), the energy diagram of tunnel junctions (b), the tunneling process (c). N = 0 - direct tunneling, N = 1 - resonant tunneling through the localized state, N□2 - inelastic tunneling with variable range hopping ("variable range hopping").

The first model of the tunneling magnetoresistive effect was developed by Yulira [10]. When the potential difference V is superimposed on the tunnel junction, the energy diagram changes, as shown in Fig. 8.13 b. Conductivity of the contact at displacement V will be:

$$G \equiv dI/dV \sim g_1(\varepsilon_F)\, g_2(\varepsilon_F), \tag{8.45}$$

where $g_1$ and $g_2$ are the density of states in the electrodes 1 and 2, respectively. Let us determine the degree of spin polarization for electrons with spin up in electrode i as $\alpha_i = N_{i\uparrow}/(N_{i\uparrow} + N_{i\downarrow})$.

The polarization coefficient of the spins in the i electrode will be

$$P_i = 2\alpha_i - 1. \tag{8.46}$$

Then the conductivity for parallel GP and antiparallel GAP configurations can be written:

$$G_P \sim \alpha_1\alpha_2 + (1 - \alpha_1)(1 - \alpha_2) = \tfrac{1}{2}(1 + P_1P_2) \tag{8.47}$$

and

$$G_{AP} \sim \alpha_1(1 - \alpha_2) + (1 - \alpha_1)\alpha_2 = \tfrac{1}{2}(1 - P_1P_2), \tag{8.48}$$

and for a magnetoresistance normalized to $R_P$ and $R_{AP}$, we get:

$$\Delta R/R_P \equiv (R_{AP} - R_P)/R_P = (1/G_{AP} - 1/G_P)G_P == 2P_1P_2/(1 - P_1P_2) \tag{8.49}$$

$$\Delta R/R_{AP} \equiv (R_A - R_P)/R_A = (1/G_{AP} - 1/G_P)G_{AP} == 2P_1P_2/(1 + P_1P_2). \tag{8.50}$$

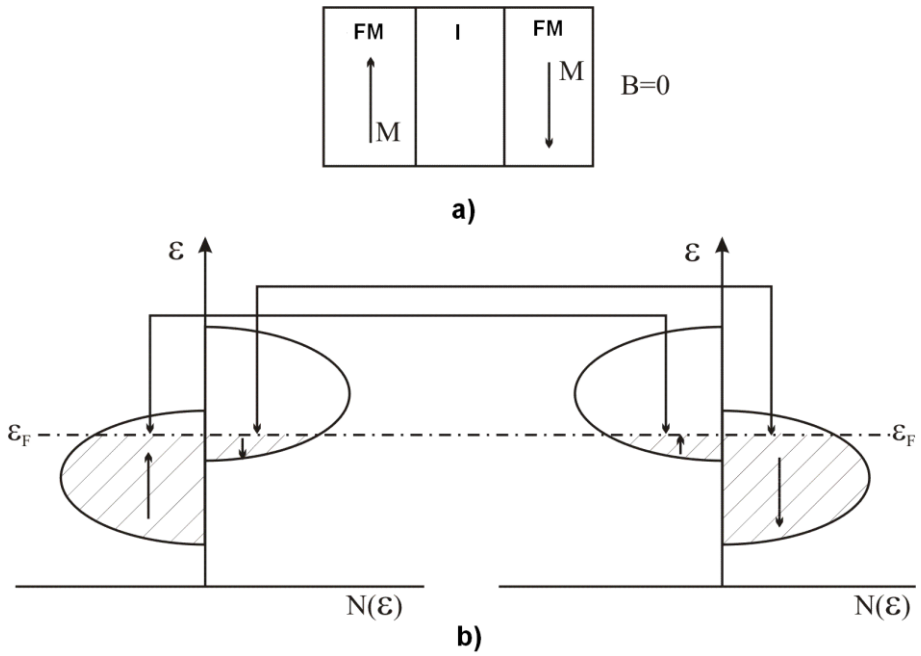



Fig. 8.14. Tunnel contact FM-I-FM (a) and the energy structure of its *d*-band (b) under antiferromagnetic exchange interaction (B = 0)

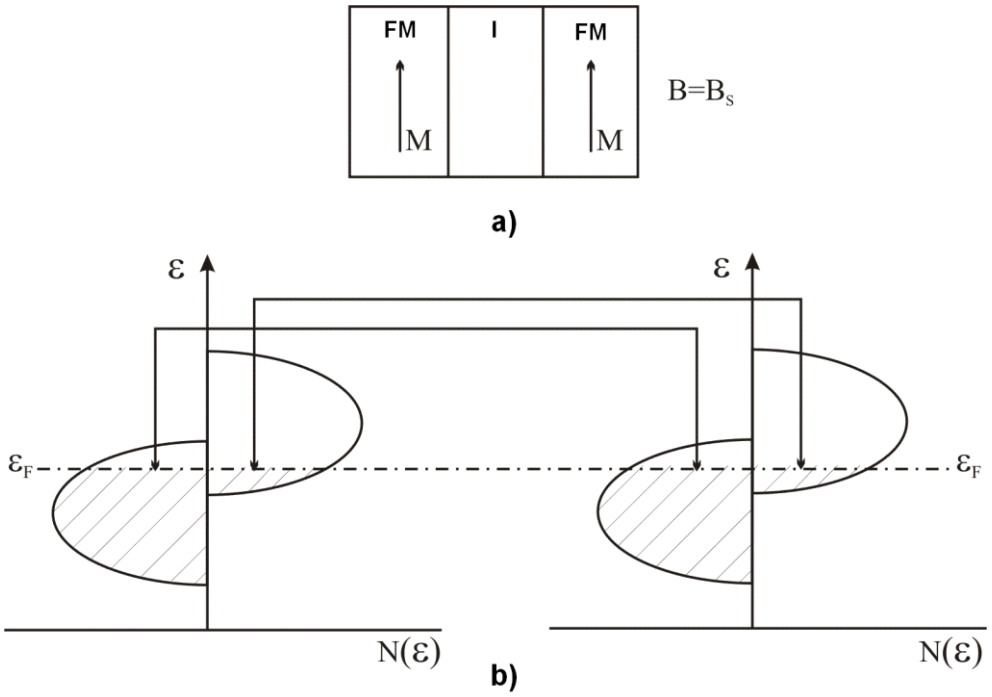



Fig. 8.15. Tunnel contact FM-I-FM (a) and the structure of its *d*-band (b) for ferromagnetic pairing ($B = B_S$)

For the case of granular magnetically ordered media, ferromagnetic clusters must have an antiferromagnetic ordering. The maximum resistance of such a system will be when the macroscopic magnetic moment is zero. For this configuration, the total conductivity of $G_P$ is:

$$G_P \sim 1/2G_P + 1/2G_{AP} \sim 1/4(1+P_1P_2) + 1/4(1-P_1P_2) = ½ \qquad (8.51)$$

Then the normalized to the maximum value of TMR can be written:

$$\Delta R / R_P \equiv (R_{AP}-R_P) / R_P = (1 / G_P - 1/G_{AP})G_P = P_1P_2/(1 + P_1P_2) \qquad (8.52)$$

In the model considered above, TMR is exclusively determined by the degree of polarization of the electron spin. In reality, a number of other factors can play an important role, the most important of which is the tunneling of free electrons.

## 8.7 The colossal magnetoresistive effect

The magnetoresistive effect, called the colossal magnetoresistance (CMR), is found in lanthanum manganites, for example, $La_{1-X}Sr_XMnO_3$. In a certain sense, the CMR can be classified as a negative magnetoresistive effect due to a decrease in scattering by spin fluctuations, as described above. However, in such compounds the reason for the resistance of the spin due to scattering by spin fluctuations is different in comparison with ordinary ferromagnetic metals. In these compounds, conductivity is realized by jumping electrons between $Mn^{3+}$ and $Mn^{4+}$ ions, as discussed for the exponential magnetoresistance in the hopping conductivity regime, which is most typical for doped semiconductors at low temperatures. Moreover, this jump is possible when the magnetic moments of two Mn atoms are parallel, i.e. in the ferromagnetic state. In other words, we can say that a metal-insulator phase transition is observed at the Curie temperature. As the temperature increases from zero to the Curie temperature, the scattering by spin fluctuations increases. In this case, the density of electronic states at the Fermi level decreases, which leads to a maximum of resistance near the Curie temperature, which can be easily changed by applying an

external magnetic field. In Fig. 8.16 a shows the temperature dependence of the CMR for the compound $La_{0,75}Ca_{0,25}MnO_3$. At temperatures above the Curie temperature ($T_C \approx 250$ K), the number of free electrons increases and the resistance decreases again. The temperature dependence of the resistance of this compound is shown in Fig. 8.16 b.

In an external magnetic field, the ferromagnetic ordering increases, which leads to an increase in the hopping probability and, consequently, to a decrease in resistance. Therefore, this negative magnetoresistance is very large in the vicinity of the Curie temperature and, as noted above, is called colossal. We note that the CMR is usually observed in strong magnetic fields of the order of unity of Tesla, which naturally hinders its use in the development of magnetosensitive sensors.

At low temperatures, when the ferromagnetic ordering is sufficiently large and can not be changed by the external magnetic field of the CMR, it becomes negligibly small. At the same time, in manganites (at least in polycrystalline samples), another type of magnetoresistive effect can be observed at low temperatures. This type of magnetoresistance appears in weak fields and is due to the mechanism of tunneling. In this case, the electrons tunnel from grain to grain through the interface, and accordingly the tunnel resistance decreases, as the external field orders the magnetization of the individual domains.

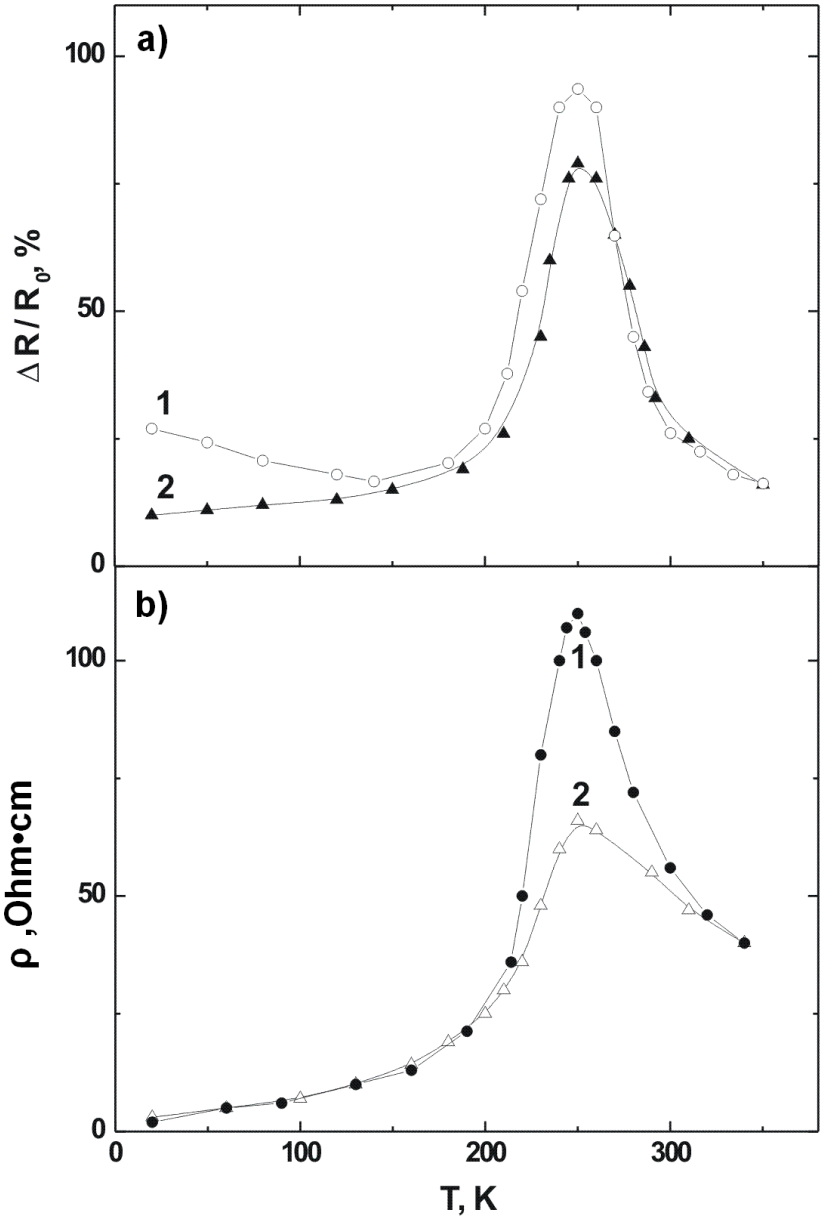


Fig. 8.16. Temperature dependence of the CMC La0.75Ca0.25MnO3 (a) and its resistivity (b) without a magnetic field (1) and in a magnetic field B = 4 T (2).

### 8.8 The spin-valve system

The spin-valve device can also be realized on the basis of a giant magnetoresistive effect in a multilayer magnetically ordered medium with a fixation of the direction of magnetization of one of the layers by means of an antiferromagnet located next to it. The structure of such a device is shown in Fig. 8.17 a, the magnetization hysteresis loops and the magnetoresistive effect in Fig. 8.17 b and c, respectively. The mechanism of changing the magnetization of such a structure with a change in the direction of the magnetic field to the opposite is easy to understand on the basis of considerations considered earlier. Obviously, in the case of the direction of the external magnetic field in the magnetization direction of the fixed layer, the resistance of the structure decreases very rapidly. In the figure, this is the direction -B. At the same time, in the opposite direction of the magnetic field, a parallel orientation of the magnetization of all layers along the direction of the external field is achieved in much larger fields. In addition, when the magnetic field is reversed both in magnetization and in the magnetoresistive effect, hysteresis phenomena are observed.

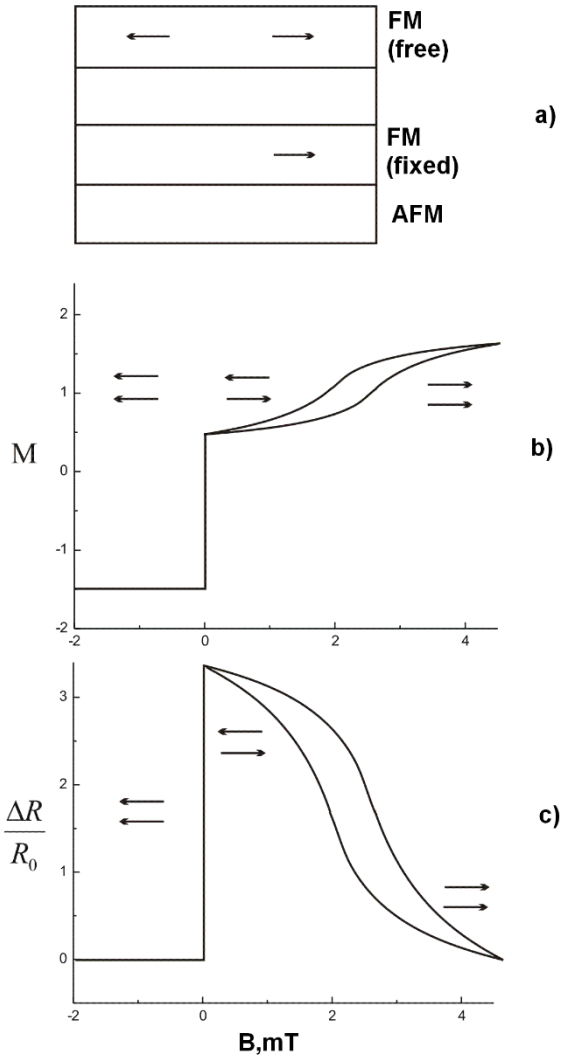



Fig. 8.17. Schematic diagram of the spin-valve device (a), hysteresis loop of magnetization (b) and GMR of the spin-valve (c) at room temperature

From the above example, it is easy to see the difference between the spin-valve system based on tunneling and giant magnetoresistive effects. Integration of such spin-sensitive devices into traditional electronics based on semiconductor materials will undoubtedly provide an opportunity to obtain fundamentally new functional devices whose importance at the present time is even difficult to predict.

**References**


1. Thouless, D. J. Electrons in disordered systems and the theory of localization/ D. J. Thouless // Phys. Rev. B. – 1974. – Vol. 13. – P. 93 – 98.

2. Мотт, Н. Электронные процессы в некристаллических веществах / Н. Мотт, Э. Дэвис. – 2 – е изд., перераб. и доп. в 2 – х томах. – М.: Мир, 1982. – 664 с.1.

3. Шкловский, Б. И. Электронные свойства легированных полупроводников/ Б. И. Шкловский, Ф. Л. Эфрос. – М.: Наука, 1979. – 416 с.

4. Abeles, B. Structural and electrical properties of granular metal films /
B. Abeles, P. Sheng, M. Coutts, Y. Arie // Adv. Phys. – 1975. – Vol. 24. – P. 407 – 461.

5. Adams, E. N. Quantum theory of transverse galvanomagnetic phenomena /
E. N. Adams, T. D. Holstein // J. Phys. Chem. Sol. – 1959. – Vol. 10, № 4. – P. 254 – 276.

6. Полянская, Т. А. Квантовые поправки к проводимости в полупроводниках с двухмерным и трехмерным электронным газом / Т. А. Полянская, Ю. В. Шмарцев // ФТП. – 1989. – Т. 23, № 1. – С. 3 – 31.

7. Hikami, S. Spin – orbit interaction and magnetoresistance in the two dimensional random system / S. Hikami, A. L. Larkin, J. Nagajka // Progr. Theor. Phys. – 1980. – Vol. 63, № 2. – P. 707 – 710.

8. Mikoshiba, N. Strong – Field magnetoresistance of impurity condaction in n–type Germanium / N. Mikoshiba // Phys. Rev. – 1962. – Vol. 127, № 6. – P. 1962 – 1969.

9. Шкловский, Б. И. Прыжковая проводимость полупроводников в сильном магнитном поле / Б. И. Шкловский // ЖЭТФ. –1971. – Т. 61, № 5. – С. 2033 – 2040.

10. Jullire, M. Tunneling between ferromagnetic films / M. Jullire // Physics Letters. – 1975. – Vol. 54 A, № 3. – P. 225 – 226.

## 9. Fluorescent quantum dots for bioimaging

The nanoparticles or quantum dots with cadmium selenide core and zinc sulfide shell are assigned to nanoparticles of type I, which are characterized by the predominant localization of nonequilibrium charge carriers in the core. Due to the quantum size effect CdSe/ZnS nanoparticles with sizes ranging from 2 to 4 nm overlap almost the whole visible range according to the position of the absorption spectrum edge and the luminescence band maximum. The main advantages of these inorganic fluorophores over the molecular one ─ exceptional photostability and multiplexing ─ promote to the high interest in its practical installation in biochemical and biomedical applications. The subject of the chapter 12 is the basic mechanisms of synthesis and modification of the CdSe/ZnS nanoparticle surface, its application in the biomedical diagnostics.

### 9.1 Semiconductor fluorescent nanoparticles CdSe/ZnS: high temperature synthesis, optical properties.

The highest potential application in biomedical analysis based on optical imaging (immuno-fluorescence analysis, optical microscopy histology, etc.) have a cadmium selenide nanocrystals, capped by passivating shell of zinc sulfide - composite nanocrystal of core/shell type CdSe/ZnS. Fluorescent probes are widespread in the study of cell biology. Organic fluorophores traditionally used in this research field, but they have several disadvantages, such as photobleaching, broad and overlapping emission bands. These disadvantages limit the use of organic fluorophores in applications requiring prolonged light excitation or multicolor response.

The inorganic nanoparticles are photostable fluorescent sources, the size of which determines the emission wavelength. High brightness and stability to photobleaching allows their use in low-intensity laser excitation unlimited time. This makes such fluorescent labels are particularly attractive for high-resolution cell imaging techniques and creation of three-dimensional (3D) and four-dimensional (4-D) reconstruction of images over time [1]. High brightness is also important to detect a signal from the single fluorescent probes and increase the number of bio-medical analyzes. Tuning the wavelength and the ability to distinguish between signal from the nanoparticles of different sizes open the perspective ща data collection from several labeled molecules under study.

First of all let us consider the distinction between the nanocrystal and nanoparticle. According to Murray et al [2] nanocrystals include nanoparticles with well-defined shape and size, as well as having a well-identified crystalline phase and surface chemistry. The nanoparticles are called nanoparticles if the type of crystal structure can not be identified due to, for example, its domain or amorphous structure. Nanoparticles or nanocrystals are monodisperse if the standard deviation of its sizes $\sigma \leq 5\%$. For nanocrystals with sizes up to 15 nm such standard deviation is associated with the characteristic size of the lattice constant.

The mechanism of the synthesis of monodisperse colloids was studied as early as 1950 [3]. It has been shown that for the synthesis of monodisperse colloidal particles and control of their size it is necessary the nucleation rate was much greater than the rate of particle growth. For this, the rapid injection of precursor into the working volume is used in such amounts that the concentration of nuclei formed instantaneously exceeded supersaturation threshold (Fig. 9.1).

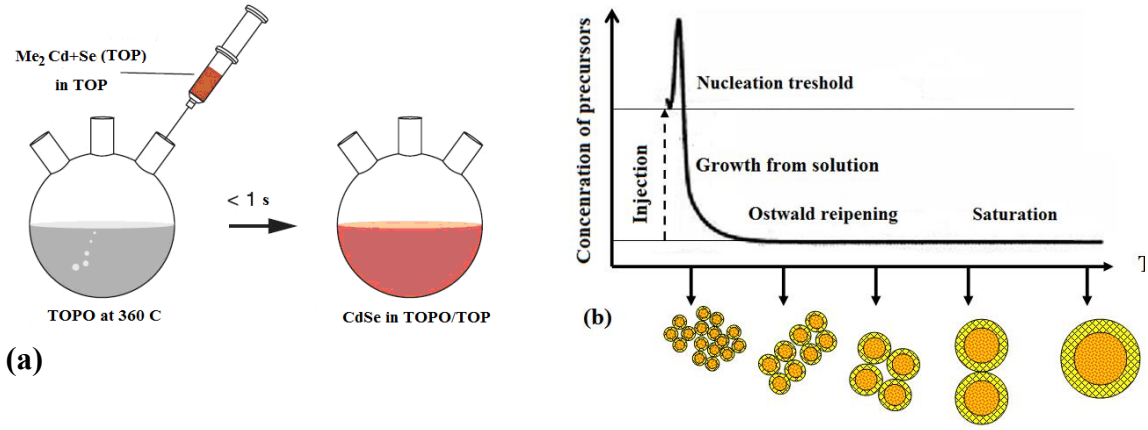



Figure 9.1 - Rapid injection of precursor in high temperature synthesis of monodisperse nanocrystals (a) and a typical plot of precursor concentration versus time (b).

At the initial stage of growth the distribution of nanocrystal size is determined by the time during which the nuclei are formed and begin to grow. For monodisperse reaction product it is important that the proportion of nanocrystals, which have time to start growing during nucleation will be small.

The following slow phase of growth occurs by the mechanism of Ostwald ripening, the particles of a polydisperse system must coarse due to the dissolution of smaller and growth of larger particles. The surface tension dependence on the particle size is developed in this mechanism. The concentration of the solution in contact with the small particle is higher than the concentration of the solution being in equilibrium with the larger particle. The high surface energy of small particles promotes their dissolution and formation of larger particles from dissolved residues. The average size of the nanocrystals during Ostwald ripening increases with a decrease in their number. The nanocrystals can be removed from the working volume at any stage of the slow growth. In this case the standard deviation is within $10 < \sigma$ <15%, subsequent size selection procedure using centrifugation allows to select fractions of monodispersed nanocrystals with a standard deviation of about 5%.

The organometallic (dimethyl or diethylcadmium, diethyl zinc) compounds in coordinating solvents such alkylphosphines or alkylphosphines oxides at temperatures of about 350 $^{\circ}$C ae used aAs a precursor source of group II elements of Mendeleev's Table for the synthesis of semiconductor nanocrystals of CdSe/ZnS. Organophosphines chalcogenides are used as precursors and sources of elements of VI group of the periodic tables of sulfur and selenium.

A classic example of the synthesis of CdSe nanocrystals in a solution of amphiphilic oxide tris-n-phosphine (Otofa) is reported by Sighal and co-authors [4]. In these solutions, the hydrophilic groups phosphine oxide coordinate the Cd atoms on the surface of the nanocrystals while hydrophobic alkyl chains form a densely packed coating surface to stabilize nanocrystals against aggregation (Fig. 9.2).

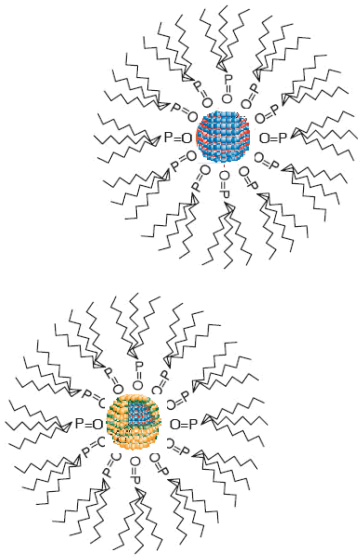


Figure 9.2 - The trioctylphosphine oxide moieties on the surface of CdSe nanocrystals (left) and on the surface of the of CdSe/ZnS nanocrystals (right).

Nonpolar alkyl chains at the surface of the nanocrystals are exposed to the solvent (Fig. 9.2), thus providing the nanocrystal solubility in organic solvents such as hexane or chloroform. To a nonpolar solution of nanocrystals is added the polar (e.g. methanol), with following centrifugation in order to remove from product solution the free TOPO ligands or precursor residues. Increasing the polarity of the solvent results in the precipitation of small nanocrystals. One can separate the mixture of nanocrystals with different sizes if repeat this process many times, increasing polarity of the solvent. This method is called the size-selective precipitation and it provides nanocrystals size distribution with a standard deviation of less than 5%.

About 80% of the atoms in CdSe nanocrystals are on the surface, as a result of high surface to volume ratio for such small sizes. Optical properties and electronic structure at such high surface to volume ratio greatly depends on the surface. The dangling bonds on the surface serve as a centers of charge carrier traps (surface-related trap states) that open the nonradiative recombination channels which in turn leads to a decrease of the luminescence quantum yield. Capping the surface of the nanocrystals with organic ligands contributes not only to stabilize them in solution, but also a slight increase in the luminescence quantum yield. The luminescence quantum yield can be significantly increased by surface capping of inorganic shell which "heal" the surface defects such as dangling bonds or other defects such as Tamm and/or Shockley.

Technique of inorganic shell epitaxial growth on the surface of nanocrystals is limited by a few requirements. Firstly, the synthesized core should already be stable to the shell synthesis conditions. Second, the surface energy of the core and deposited shell should be close enough to the heterogeneous nucleation barrier for shell on the core was lower than homogeneous nucleation of new nanocrystals from the shell material. Third, the shell atom must be easily introduced to the surface of the core.

Usually already synthesized size-selected CdSe core is transferred to a working tank containing solvent and stabilizer. The mixture is heated to the desired temperature and precursors introduced into it for the growth of the shell. The rate of shell precursor introduction should be lower than the injection speed for the nucleation of new nanocrystals from the shell material.

ZnS capping of CdSe nanocrystals lead to a significant increase in luminescence quantum yield of ZnS-capped CdSe/ nanocrystals compared to uncapped CdSe nanocrystals [5]. The luminescence quantum yield of ZnS-capped CdSe nanocrystals reaches 85%, in contrast to the highest possible luminescence quantum yield of TOPO-coated CdSe nanocrystals of about 15%. The observed increase in the quantum yield is due to the removal of surface defects ("traps") on the surface of the nanocrystals. This process is called electronic passivation. Nanocrystals CdSe/ZnS nanocrystals referred to as type I, since the band gap of the zinc sulfide is larger than

the bandgap of cadmium selenide. In addition the length of Zn-S bond in the shell close to the Cd-Se bond length in the core.

For CdSe nanocrystals surface defects occur in a wide and weak emission bands at the 700-800 nm due to radiative transitions from the local surface states, the levels of which lie in the forbidden zone (Fig. 9.3).

Electronic surface passivation reduces the number of charge trapping centers ("charge traps") and leads to the disappearance of the bands associated with transitions from these states (Fig. 9.3). It is yet another effect - reducing the probability of photo-oxidation CdSe core capped with zinc sulfide in addition to the effects associated with the increase of luminescence quantum yield. Thus, the capping of CdSe core with ZnS shell leads to an increase in the luminescence quantum yield of up to ten times, and also promotes increased resistance of composite nanocrystals to photooxidation and environmental effects.

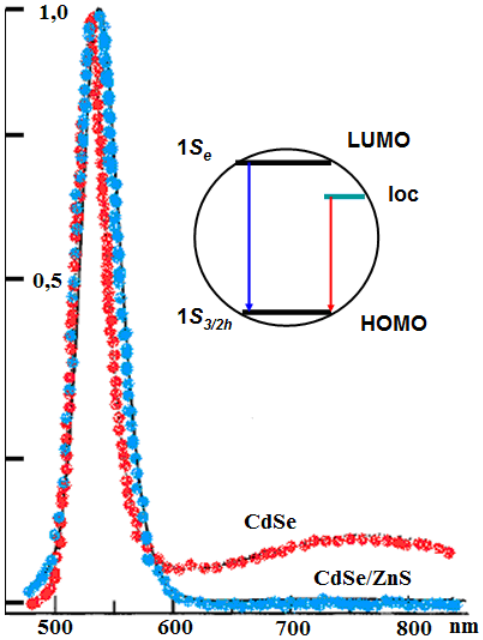



Figure 9.3 - Luminescence spectra CdSe and CdSe/ZnS nanocrystals and a diagram of the electronic radiative transitions between quantum states of the electron $1S_e$ (LUMO) and holes $1S_{h3/2}$ (HOMO), corresponding to the fundamental luminescence band with a maximum of about 540 nm and a radiative transition from the local charge trap level loc, corresponding to a weak band in the spectrum of uncoated nanocrystals with a maximum of about 750 nm.

CdSe/ZnS nanocrystals are the subject of the most careful study [6] and the most popular in biomedical applications. Because of the large difference in the size of the band gap between the cadmium selenide (1.74 eV) and zinc sulphide (3.61 eV), an exciton is well localized in thecore, as it is typical for type I nanoparticles.

The movement of charge carriers in nanocrystals is limited in all three dimensions, so in the context of optical properties they are often called quantum dots (QDs). Quantum size effect is manifested in the dependence of the discrete spectrum of energy states in nanocrystals on size, and the density of electronic states modification determines the shape of the absorption spectrum of the crystals (Fig. 9.4, a). There is a band corresponding to a fundamental transfer $1S_e \rightarrow 1S_{h3/2}$ at the edge of the absorption spectrum (inset to Fig. 9.3). With a decrease in the size of CT edge of the absorption spectrum is shifted to shorter wavelengths (Fig. 9.4, a). Maximum of quantum dot luminescence band also shifts to shorter wavelengths with decreasing their size (Fig. 9.4, b).

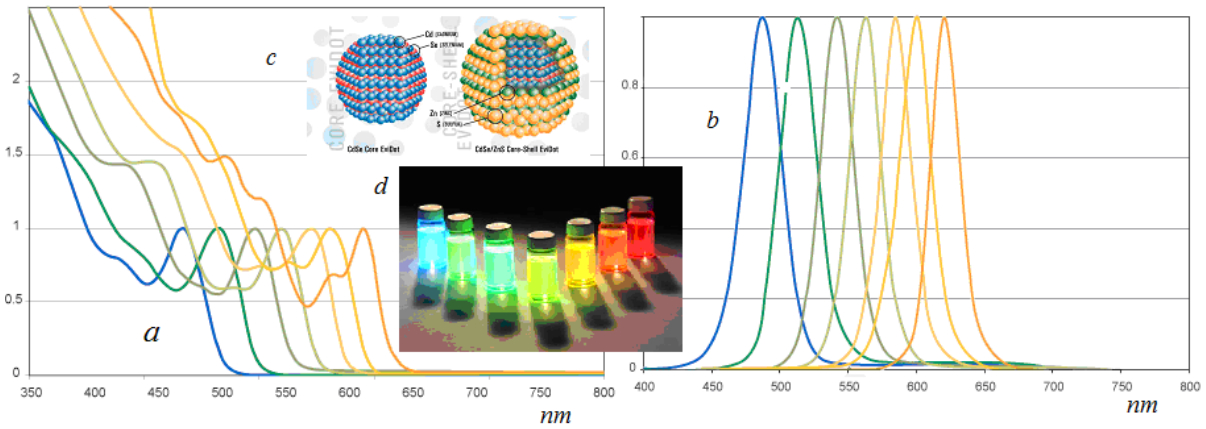



Figure 9.4 - The absorption (a) and luminescence (b) spectra of core/shell CdSe/ZnS quantum dots (c), obtained as result of high-temperature colloidal synthesis and size selection (d) (Evident Technologies web [7]).

These distinctions are clearly visible to the naked eye for a series of QD samples ranging in size from 5 to 2 nm (Fig. 9.4, d). Wide tuning of absorption spectrum edge and luminescence band maximum position across the visible range can be achieved by varying the size and composition of the quantum dots (Figure 9.4, c). This allows simultaneously obtain the optically distinguishable response from several different molecules or events. For example, small CdSe nanocrystals (~ 2 nm) emit in the range from 495 to 515 nm, whereas larger CdSe nanocrystals (~ 5 nm) emit in the region between 605 and 630 nm [8].

Nanocrystals have unique optical spectra and very differ from organic fluorophores. (Fig. 9.5). The organic dyes generally have narrow absorption spectra and thus can only be excited within a narrow wavelength window (Fig. 9.5, spectrum 1). Furthermore, they have an asymmetrical emission band with a characteristic broadening in the long wavelength side (Fig. 9.5, spectrum 2). In contrast, the quantum dots have a broad absorption spectrum, allowing use a broad range of excitation wavelengths (Fig. 9.5, spectrum 3), and their emission spectra are narrow and symmetrical (fig. 9.5 spectrum 4) [8].

Consequently, multicolor quantum dots (Fig. 9.4, d) of various sizes can be excited at one wavelength in the short wavelength region while their emission spectra minimally overlap.

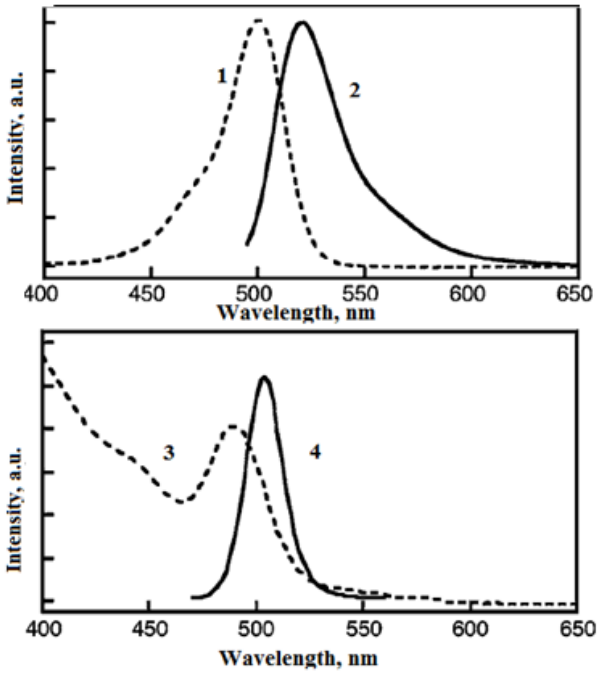


Figure 9.5 - Excitation (1, 3) and luminescence (2, 4) spectra of typical organic chromophore fluorescein (top) and a typical nanocrystal CdS/ZnS (bottom). The luminescence excitation wave length for fluorescein is 476 nm and for nanocrystals - 355 nm [8].

QD is also very stable light sources due to their inorganic and crystalline structure, which makes them more resistant to photobleaching than organic dye molecules (**Error! Bookmark not defined.**, 9]). This advantage has been repeatedly demonstrated in several experiments for labeling biological molecules where photostability of quantum dots compared with conventional fluorophores, such as rhodamine, fluorescein, and others.

Increased photostability of quantum dots gives a high potential as labels for applications where it is necessary to deal with images of thick cells and tissues for prolonged periods of time or a signal from several collecting optical sections without damaging the sample. Furthermore QD have a higher two-photon absorption cross section than organic fluorophores [10]), which makes them promising markers in studies of thick specimens and in vivo imaging using multiphoton excitation.

In fairness it should be noted that there is a distinctive feature of the QD, which in some cases is the disadvantage. It is about the nature of intermittence luminescence of single QD. Indeed QD randomly experience transitions between emitting and non-emitting states. Intermittence emission is observed only in experiment with single QD. Therefore, such a disadvantage restricts the applications of QD in detection of single molecules.

In some cases there is a spontaneous increase in brightness, which is called QD brightening. Sometimes this QD property makes their use problematic in systems of photons accounts or quantitative assessments. Both phenomena intermittence and brightening luminescence can be explained by the mobility of charge carriers in the surface trap sites.

## 9.2 Application of CdSe/ZnS quantum dots in biomedical visualization (cells, tumors, etc.).

Progress in the field of synthesis and optimization of quantum dots as fluorescent markers for biological application opens up broad prospects for the study of cellular and molecular structures, the tracking cell clones monitoring physiological events in living cells, cell motility measurement and tracking cells in vivo.

However, QD, synthesized by the method described above, are hydrophobic. The first step in life science application of nanocrystals is the modification of its surface to give them a hydrophilic property. This surface modification is called solubilization.

### 9.2.1 The functionalization of nanocrystals

Various solubilization methods using surfactants ligands give the surface of QD not only hydrophilic properties but also functional properties that allows conjugate them with biological molecules. Such solubilization of the QD surface is called functionalization. There are many protocols of nanocrystal functionalization, making them biocompatibility, but all these protocols can be grouped into three main groups or strategies [11, 12].

The first group includes substitution of TOPO molecules by bifunctional ligands on thr ZnS shell surface. The structure of these ligands should have one functional group which has high affinity for the ZnS surface, such as a thiol group. Other functional group must be ionogenic, for example such as hydroxyl, carboxyl or amino. As a result of substitution thiol group binds to the surface of nanocrystals to form a ZnS-S-thiol ligand bond, and the ionogenic group is exposed to the environment, provides hydrophilicity and stabilize the nanocrystals in water solutions. The procedure is quite simple, but in some cases, instead of simple thiols and mercaptans used polydentate thiols having multiple thiol groups to eliminate the dynamic instability of ZnS-thiol bond. In addition, the carboxyl group poorly stabilizes the nanocrystals in an alkaline medium. Bifunctional ligands, such as cysteine and cysteamine can also be a good substitute for TOPO

ligands on the nanocrystal surface. However nanocrystals functionalized with aminothiol need the stabilization at low pH.

In aqueous media, carboxyl and amino groups are ionized to form carboxyl anion and amino cation. Therefore nanocrystals solubilazed with mercaptoacetic acid and cysteamine can be called anionic and cationic quantum dots, respectively. Dimensions thereby solubilized nanocrystals substantially unchanged unlike nanocrystals solubilized by the second group of protocols.

The second group includes the formation of the polymerizable silica shell functionalized with polar groups. This method is called surface silanization. As a result of such a solubilizing strategy nanocrystals as potential fluorescent labels size increases. Since the encapsulation of quantum dot with block copolymer phospholipid shell their size increases from 4.8 before encapsulation to 20-30 nm in encapsulated form. Such fluorescent capsules are stable in a wide range of pH values, but their size makes it difficult to diffuse into the cells.

The third group of protocols used amphiphilic copolymers which are fixed by hydrophobic interactions with TOPO ligands on the nanocrystal surface while hydrophilic moieties carried out subsequent coupling with biological molecules in an aqueous medium. The main processes described above as strategies of nanocrystal solubilization and functionalization are schematically shown in Figure 9.6.

Free carboxyl and amino groups on the modified nanocrystal surface are used for binding (conjugation) with biomolecules by means of a peptide bond formation (Fig. 9.6). Phospholipids, ditiotriitol and oligomers also can be used as surfactants for functionalization of nanocrystals [1].

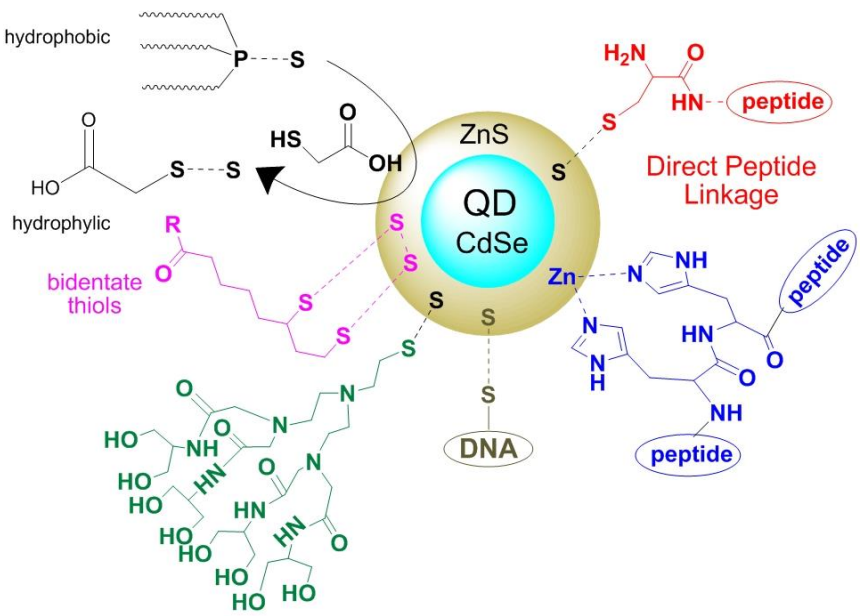



Figure 9.6 Strategies of nanocrystal functionalization to give hydrophilic properties to their surface and the subsequent binding to biological targets (Wikipedia [13]).

Conjugation schemes of nanocrystals to proteins, biological receptors and other biological targets are also divided into three categories, each of which has its advantages and disadvantages. The first category includes peptide bond forming reaction between the quantum dots and the amino acid residues of the proteins, utilized of 1-ethyl-3- (3-dimethylaminopropyl) carbodiimide hydrochloride (EDC) as catalyst. The second category is the direct reaction of thiolated DNA, thiolated peptide or polyhistidine residues having sulfur atoms with sulfur of ZnS nanocrystal shell surface (Fig. 9.6). The third category includes physical adsorption due to electrostatic interaction between the anionic QD with a specially engineered proteins having positively charged domains.

The luminescence quantum yield of these QD-proteins conjugates even higher than that of unbound QD [14].

Conjugation thiol-coated alkyl-COOH ligands nanocrystals using EDC often leads to the formation of aggregates due to low stability in acidic and neutral buffers. If the same methodology apply to encapsulated nanocrystals with free -COOH groups on the capsule surface too large conjugates with uncontrolled number of proteins associated with a single nanocrystal one can obtained. Furthermore, this chemical reception provides nonspecific aggregation of different proteins on the encapsulated nanocrystal surface. However, there are commercial samples of QD with ~ 20 protein streptavidin conjugates, which can be used for subsequent conjugation with biotinylated antibodies.

Greater success has been achieved by Russian scientists under the supervision of Nabiev to develop a strategy of nanocrystal functionalization with proteins and peptides [15]. Despite the above noted problems related to unnecessary increase in the size of fluorescent labels, functionalized proteins, they were obtained ultra-small (10-12 nm) tags based on CdSe/ZnS nanocrystals and single domain antibodies derived from Lama Alpacas which have high stoichiometric orientation in a specific antigen recognition and controlled number of binding sites in solutions [16].

### 9.2.2 Applications of quantum dots for biological imaging.

Let us list all the qualities of the QD, which make them efficient fluorescent labels in biomedical research: high luminescence quantum yield and a high molar extinction coefficient; a wide range of absorption and narrow (half-width of 20-30 nm) luminescence band; large effective Stokes shift between the edge of the absorption band and luminescence bands maximum whose positions can be adjusted by QD size; high resistance against photo- and chemical degradation.

The most comprehensive reviews of the first studies on the use of fluorescent QD for cell image can be found in Alivizatos [1] and Medintz [11] papers.

Today the range of practical introductions of fluorescent nanocrystals is very extensive in the life sciences. These include first of all the labeling and visualization of tumor tissues [17], biosensorics [18], flow cytometry [19], photodynamic therapy [20] and targeted drug delivery [21].

The directions of these works can be conventionally divided into two levels of quantum emitter penetration in biological systems – *in vitro*, i.e. in laboratory conditions and *in vivo* in living cells. *In vitro* study include immunohistochemistry, based on the reaction of the antigen-antibody recognition. Nanocrystals labeled antibodies able to specifically bind to antigens and produce a luminescent response of the localization of these molecules. Furthermore nanocrystal labeled receptor ligands provide information on signaling pathways in cells. *In vitro* study include imaging cell line processes and their mobility. A number of model problems and corresponding analytic methods, such as *in situ* hybridization, and Forster resonance transfer of electronic excitation energy can also be solved using fluorescent probes.

*In vivo* studies include a wide range of tasks associated with the visualization and localization processes in living organisms, and are held in large quantities in with laboratory mice. A major breakthrough in this area can be considered a new method in surgery -image guaided surgery, ie imaging in surgery. Consider the most famous and striking examples in this application field of science.

In the journal Science, two articles were published in 1998, which, perhaps, were the starting point to demonstrate the applications of fluorescent quantum dots for biological problems. In the first article Alivizatos et al [8] for the first time demonstrated the benefits of a wide range of QD excitation. Stable fluorescent images of the mice fibroblast cells labeled with green and red quantum dots (dual emission image) was obtained with a single excitation source.

In the second article and Chan and Nie [22] showed that labeled with CdSe/ZnS QD transferrin was subjected to receptor mediated endocytosis in cultivated HeLa cells. In fact it has been shown that transferrin associated with quantum dots able to perform their biological function, it is recognized by cellular receptors and penetrates the cell. Nanocrystals that act as fluorescent labels were compared with rhodamine 6G. Under identical conditions nanocrystals 100 times more photostability, and one CdSe/ZnS nanocrystal emits as 20 molecules of rhodamine 6G.

Figure 9.7 shows schematically how the method developed by the authors [23] work to scan molecular profile based multicolor/multicycle and in situ visualization of single cells. The method is named M3P method in accordance with the abbreviation of *multicolour multicycle molecular profiling*, due to the fact that it is based on the use of fluorescent signals simultaneously obtained from several emitters (*multicolor*). It allows in multiple cycles (*multicycle*) to determine the molecular profile (*molecular profiling*) of individual cells without destroying their morphology.

Panel *a* in Figure 9.7 shows the specific recognition of nanocrystal functionalised antibody and antigen in one-step reaction. Protein *Staphylococcus aureus* (*SpA*) is immobilized on nanocrystal through a covalent bond. Conjugate Qd-Spa binding with the antibody (Ab), such as immunoglobulin (IgG) is specific. Moreover thanks to the advantageous stoichiometry of the Qd-SpA-Ab complex it does not react with other antibodies or conjugates of quantum dots. Due to this fact Qd conjugate of different colors can be mixed after specific binding with different antibodies in cocktail (panel *b* in Fig. 9.7).

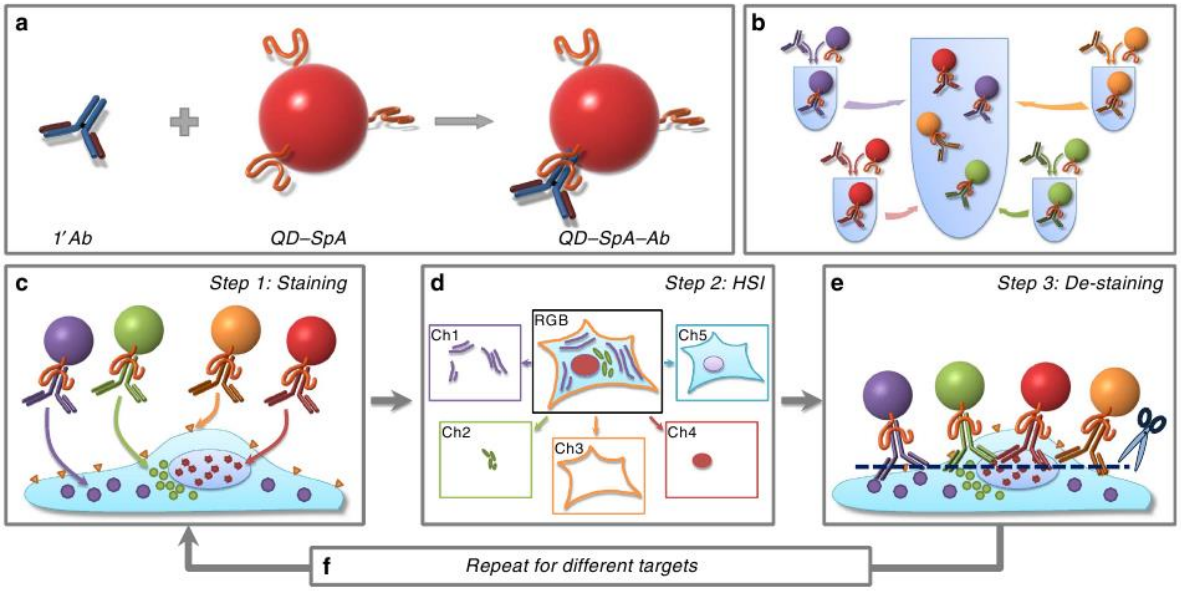



Figure 9.7 Multicolor and multicycle platform on the base of functionalized nanocrystals to read the molecular profile of individual cells, recognizing up to 100 different molecules in one cell without breaking their morphology ([23] open access).

Panel *c* in Figure 9.7 shows a one-step staining of cell culture using the resulting cocktail. It is the first stage of molecular cell profiling. Panel *d* shows schematically a fluorescent image of immiscible nanocrystals of different colors, each color is the specific area of the cell culture. It is the second stage of quantitative evaluation molecular cell profiling.

The hyperspectral imaging (HI) technique is used for fluorescent image registration. HI camera record spectral signal from each pixel in the image. Then the identical signal can be isolated in a separate channel, and its spatial location is represented as schematically shown in Figure 9.7, *d*.

The third step of this method involves the removal of labels (panel *e* in Fig. 9.7) for subsequent staining cycles in which other molecular targets in the cell should be identified (panel *f* Fig. 9.7). The advantages of this method should include the ability to scan multiple fluorescence images of the same cell without breaking the morphology, as evidenced by Figure 9.8.

Figure 9.8 *a* shows a fluorescent image of HeLa cells fixed in formalin after parallel staining targeted on five molecular targets in, namely the protein Ki-67, HSP90, Lamin A, Cox-4 and b-tubulin (left to right in Figure 9.8, *b*, and *c*). For this purpose, the conjugates of quantum dots with a fluorescence maximum at wavelengths of 525, 545, 565, 585 and 605 nm, are used (left to right in Figure 9.8, *b*, and *c*). The difference in the position of fluorescence maximum only 20 nm can not be identified by the human eye in RGB semitones. It is known that Lamin A localized mainly in the cell membrane, HSP90 in the cytoplasm and Cox-4 in mitochondria.

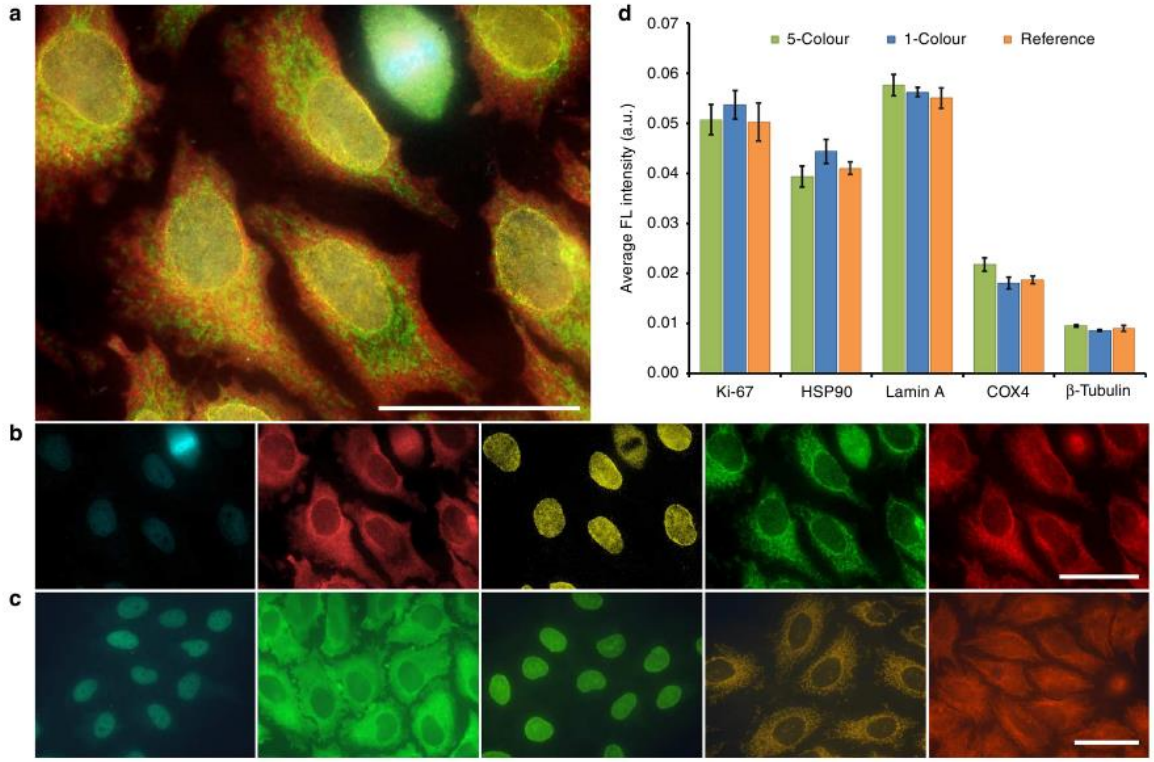



Figure 9.8 Fluorescence images of parallel multiplexed staining with five pre-assembled QD-SpA-Ab probes. (Zrazhevskiy and Gao 2013, open access).

As shown in Figure 9.8, conjugates of quantum dots with different emission wavelength to form a kind of spatial distribution (Fig. 9.8 *a–c*) in intra- and intercellular space, depending on the localization of molecular targets, with which the conjugate of a certain color is bounded. Application of HI camera and digital image processing allows to estimate the relative expression level of each protein target in cells (Fig. 9.8, *d*).

The data on the relative expression level of different proteins are obtained both by individual staining cells experiments only with one specific conjugate (panel *c* and "1 color" on the panel *d*, Fig. 9.8) and in experiments with parallel staining conjugates simultaneously for five different targets (panel *b* and "5 color" on the panel *d*, fig. 9.8). The authors emphasize that such a quality image when scanning a wide spectral range with a narrow bandwidth for organic chromophores can not be due to their photobleaching.

The manifold of molecular biological processes in the cell provided a large arsenal of protein, full library which can be read by M3P repeating cycles profiling with new and new sets of QD-Abs conjugat on the same cell by subjecting it to washing after each read cycle. De-staining is carry out by a special buffer that prevents the formation of precipitation, easily breaks the bonds of the antigen-antibody and quenches the fluorescence of quantum dots trace remnants.

Another interesting application is related to the high sensitivity of quantum dots to the polarity of their environment. The fluorescence spectra from a set of points in histological sections of colon tissue with different levels of pathology that were stained with hydrophilic semiconductor CdSe/ZnS nanocrystals that were modified by a unique method at the phase interface of immiscible liquids [24]. The shifts in the fluorescence spectra that were recorded for stained cells at different

pathologies are described using the electrochromism of the nanoparticles. Aqueous solutions of the CdSe/ZnS nanocrystals with different pH values are used as systems that model the acidity of a biological medium. It has been shown [25] that the shifts of the fluorescence bands of the CdSe/ZnS nanocrystals are caused by a change in the local electrical field that is induced by solvated ions near their surface at different pH values of the solutions. The application of the CdSe/ZnS nanocrystals as nanoprobes for the local pH in biological tissue is discussed in [25] in the context of this model.

When the fluorescent images of the colorectal benign neoplasms (adenoma) and the fluorescence spectra that are obtained from different points in these images are considered, the maximum of the fluorescence band of the CdSe/ZnS nanocrystals in the glands is at 591 nm, whereas in the stromal environment it is 581 nm (Figure 9.9), i.e., the bathochromic shift of the fluorescence band in these glands, when compared to the healthy glands, is 3–4 nm, while the stroma has an 8-to 9-nm hypsochromic shift observed. In the stromal environment of the glands of the benign neoplasms, the fluorescence intensity of the CdSe/ZnS nanocrystals is significantly smaller than for the glands. It should be noted that the tendencies that are observed for the changes in the parameters of the fluorescence of the CdSe/ZnS nanocrystals do not depend on the size of the nanocrystals: similar regularities are observed for sections that are stained with CdSe/ZnS nanocrystals with sizes of 3.2 and 4.0 nm.

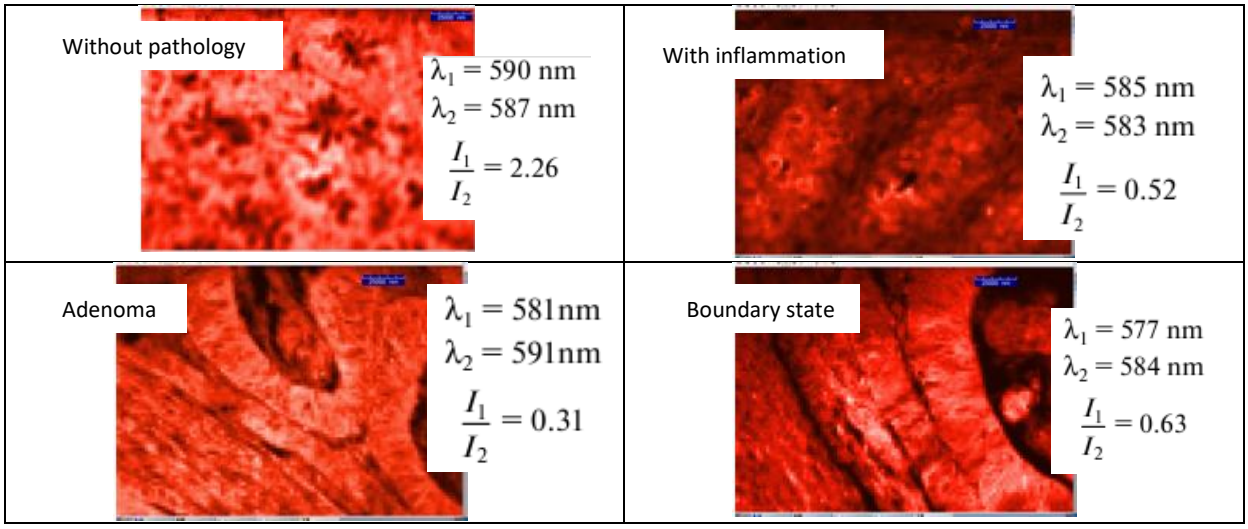



Figure 9.9 - Fluorescence images of histological sections with different levels of the pathology that are stained with the hydrophilic CdSe/ZnS nanocrystals and the positions of the maxima of the fluorescence bands of the glands ($\lambda_2$) and stromal environment ($\lambda_1$), and the ratio of their integral fluorescence intensities.

Spectra that are recorded in the regions in the immediate vicinity of the cancer neoplasms (hereinafter in the text, they are referred to as “boundary tissues”) should be separated into an individual group. These tissues have a characteristic hypsochromic shift of the maximum of the fluorescence band of the nanocrystals both in the glands and in the stromal environment by 3 and 13 nm, respectively, when compared to the epithelial cells without pathology (Fig. 9.9). Cancer cells differ from benign tumors. The maximum of the fluorescence band of the NPs in the stroma is at 586 nm, and in the glands it is at 580 nm. The fluorescence intensity of the glands and stromal environment is the same.

The semiconductor CdSe/ZnS nanocrystals that are modified by the original method at the interface of immiscible liquids can be successfully used as contrasting agents for investigation of the morphology of the histological sections of the colon tissue that has different levels of the pathology. In addition to by obtaining the contrast, which is necessary for analysis of the cell morphology, the cells can be differentiated by the position of the fluorescence band of the NPs. The position of the fluorescence band of the NPs depends on the pH in the aqueous solutions. The

possible physical mechanisms of the effect of the pH on the parameters of the fluorescence of the hydrophilic CdSe/ZnS nanocrystals is the Stark effect, which is caused by electrostatic interactions with the ions of the double Helmholtz double layer near the surface of the nanocrystals. This model has demonstrated that the Stark spectra of the CdSe/ZnS nanocrystals that are immobilized on the colon cells correlate with the Stark spectra of the CdSe/ZnS nanocrystals. Thus, the hydrophilic CdSe/ZnS nanocrystals that are obtained at the phase interface can be used not only as contrasting agents for biopsy, but at the same time as fluorescent probes of the local pH. The value of the local pH, in turn, can be proposed as an independent digital parameter for the differentiation of the pathology in the biological tissue.

It can be concluded from the above said that the spectral position of the maximum and the intensity of the fluorescence of the histological sections of the colon that are stained with the CdSe/ZnS nanocrystals localized in the glands and in the stroma are sensitive to a change in the local pH of the cells.

We gave here examples of the use of nanocrystals with specific [23] and nonspecific binding [25] to biological tissues. These examples well demonstrate that with both types of binding, the nanocrystals respond either to the presence of specific biomarkers or to a change in the local potential near the cells.

**References**


1. Alivisatos A. P., Gu W., Larabell C. /Quantum dots as cellular probes // Annu. Rev. Biomed. Eng. 2005. 7:55–76.
2. C. B. Murray and C. R. Kagan, M. G. Bawendi Synthesis and characterization of monodisperse nanocrystals and close-packed nanocrystal assemblies// Annu. Rev. Mater. Sci. 2000. 30:545–610.
3. V. K. La Mer, R.H. Dinegar Theory, Production and Mechanism of Formation of Monodispersed Hydrosols// Joutn Amer Chem Soc/ 1950.- Vol. 72.- № 11.- P. 4847-4854.
4. Anupam Singhal, Hans C. Fischer, Johnson Wong, and Warren C. W. Chan Biomedical applications of semiconductor quantum dots In: Biological Nanostructures and Applications of Nanostructures In: Biology Electrical, Mechanical, and Optical Properties Series Edited by Michael A. Stroscio and Mitra Dutta Kluwer Academic Publisherse 2005 P. 37-48.
5. Hines MA, Guyot-Sionnest P. Synthesis and characterization of strongly luminescing ZnS-capped CdSe nanocrystals 1996. J. Phys. Chem. 100:468–71
6. Dabbousi, B. O., J. Rodriguez-Viejo, F. V. Mikulec, J. R. Heine, H. Mattoussi, R. Ober, K. F. Jensen, and M. G. Bawendi. "(CdSe)ZnS Core-Shell Quantum Dots: Synthesis and Characterization of a Size Scries of Highly Luminescent Nanocrystallites." The Journal of Physical Chemistry B 101 (November 1997): 9463-9475.
7. "CdSe/ZnS Core Shell Evidots in Toluene." Evident Tcchnologies.Web.
8. Bruchez M, Moronne M, Gin P, Weiss S, Alivisatos AP. 1998. Semiconductor nanocrystals as fluorescent biological labels. Science 281:2013-16.
9. Michalet X, Pinaud F, Lacoste TD, Dahan M, Bruchez M, et al. 2001. Properties of fluorescent semiconductor nanocrystals and their applications to biological labeling. Single Mol 2:261-76.
10. Larson DR, Zipfel WR, Williams RM, Clark SW, Bruchez MP, et al. 2003. Water-soluble quantum dots for multiphoton fluorescence imaging in vivo. Science 300:1434-36.
11. L. Medintz, H. T. Uyeda, E.R. Goldman, H. Mattoussi/ Quantum dot bioconjugates for imaging, labelling and sensing// Mature Materials. Vol. 4.- 2005.- 435-446.
12. R Bilan, F Fleury, I Nabiev, A Sukhanova Quantum dot surface chemistry and functionalization for cell targeting and imaging / Bioconjugate chemistry 26 (4), 609-624.
13. https://en.wikipedia.org/wiki/Cell-penetrating_peptide
14. Mattoussi, H. et al. Self-assembly of CdSe-ZnS quantum dot bioconjugates using an engineered recombinant protein. J. Am. Chem. Soc. 122, 12142–12150 (2000).

15. Molecular interaction of proteins and peptides with nanoparticles AA Shemetov, I Nabiev, A Sukhanova ACS nano 2012 6 (6), 4585-4602.
16. A Sukhanova, et al Oriented conjugates of single-domain antibodies and quantum dots: toward a new generation of ultrasmall diagnostic nanoprobes, Nanomedicine: nanotechnology, biology and medicine.- 2012 8 (4), 516-525.
17. Baba, K., and Nishida, K. (2012) Single-molecule tracking in living cells using single quantum dot applications. Theranostics 2, 655−667.
18. Li, J., and Zhu, J. J. (2013) Quantum dots for fluorescent biosensing and bio-imaging applications. Analyst 138, 2506−2515.
19. Akinfieva, O., Nabiev, I., and Sukhanova, A. (2013) New directions in quantum dot-based cytometry detection of cancer serum markers and tumor cells. Crit. Rev. Oncol. Hematol. 86,1−14.
20. Qi, Z. D., Li, D. W., Jiang, P., Jiang, F. L., Li, Y. S., Liu, Y., Wong, W. K., and Cheah, K. W. (2011) Biocompatible CdSe quantum dot-based photosensitizer under two-photon excitation for photodynamic therapy. J. Mater. Chem. 21, 2455−2458.
21. Probst, C. E., Zrazhevskiy, P., Bagalkot, V., and Gao, X. (2013) Quantum dots as a platform for nanoparticle drug delivery vehicle design. Adv. Drug Delivery Rev. 65, 703−718.
22. W. C. W. Chan, Sh. Nie Quantum dot bioconjugates for ultrasensitive nonisotopic detection// Science.- 1998.- 281.- 2016-2016.
23. P. Zrazhevskiy, X. Gao Quantum dot imaging platform for single-cell molecular profiling// Nature communications.- 2013.-4:1619.
24. N. D. Strekal and S. A. Maskevich, BY Patent No. 17735 (2013).
25. Motevicha I. G., Strekal N. D., Shulga A. V., Maskevich S. A. (2018) CdSe/ZnS Fluorescent Nanoparticles as Nanoprobes of Local pH in Diagnostics of Oncology// Optics and Spectroscopy. 124 (5): 637–643.

## 10. Plasmonic nanomaterials for photonics, biochemistry and quantum technology

The objects of chapter 10 are the nanomaterials on the base of noble metals silver and gold, at the surface of which the two types of waves associated with conduction electron collective vibrations are developed under the light ─ propagating surface plasmon polaritons and localized plasmons. The classification of materials with plasmon properties depending on the morphology and type of the excited plasmons ─ plasmon crystals, plasmon resonance particles, plasmon films, as well as its application in surface plasmon resonance spectroscopy, plasmon-coupled fluorescence, surface-enhanced Raman spectroscopy are presented.

## 10.1 The basic concepts of optoplasmonics

### 10.1.1Surface plasmon polaritons on a flat surface between metal and dielectric

Let us consider the solution of Maxwell's equations for the case when an electromagnetic wave is incident on the surface of the metal-dielectric. The normal to the plane of incidence coincides with the direction of the axis $x_3$. The field of the incident wave is a p-polarized (Fig. 10.1, *a*).

Maxwell's equation solutions for this case is represented as [1]

$$\vec{H}(\vec{x},t) = (0,A,0)\exp(ikx_1 - k_3^{(1)}x_3 - i\omega t)$$

$$\vec{E}(\vec{x},t) = -A\frac{c}{i\omega\varepsilon_1}(k_3^{(1)},0,ik)\exp(ikx_1 - k_3^{(1)}x_3 - i\omega t) \qquad (10.1)$$

in the dielectric region above the interface surface at $x_3 > 0$ and

$$\vec{H}(\vec{x},t) = (0,B,0)\exp(ikx_1 + k_3^{(m)}x_3 - i\omega t)$$

$$\vec{E}(\vec{x},t) = -B\frac{c}{i\omega\varepsilon(\omega)}(k_3^{(m)},0,ik)\exp(ikx_1 + k_3^{(m)}x_3 - i\omega t) \qquad (10.2)$$

in the metal under the interface surface at $x_3 < 0$.

The wave vectors $k_3^{(1)}$ describing the propagation of the field in the dielectric in the direction of $x_3$, and $k_3^{(m)}$ describing the propagation the field in the metal in the direction of - $x_3$ determine the field decay, respectively, in the dielectric and the metal with increasing distance from the interface (Fig. 10.1, *b*). The real parts of these vectors should be positive for the system of equations (10.1) and (10.2) describe the electromagnetic field localized at the metal - dielectric at $x_3 = 0$.

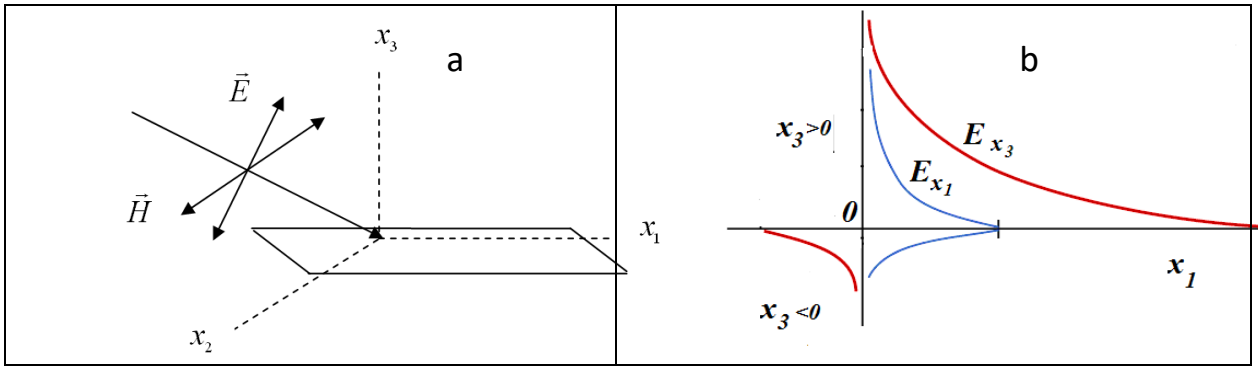



Figure 10.1 A plane electromagnetic wave (a) and decay of the transverse $E_{x3}$ and longitudinal $E_{x1}$ field components (b) near interface between the metal ($x_3 < 0$) and dielectric ($x_3 > 0$)

From the boundary conditions at $x_3 = 0$ it follows that

$$A\frac{k_3^{(1)}}{\varepsilon_1} = -B\frac{k_3^{m}}{\varepsilon(\omega)}$$

A = B, . (10.3)

Nontrivial solutions ($A \neq 0$ и $B \neq 0$) in accordance with (10.3) would be responsible the dispersion relation between the frequency of the *p*-polarized light with its wave vector *k*:

$$\frac{k_3^{(m)}}{k_3^{(1)}} = -\frac{\varepsilon(\omega)}{\varepsilon_1} \quad (10.4)$$

Surface electromagnetic waves, whose amplitudes are determined by the expressions (10.1) and (10.2), and the dispersion – by the expression (10.4) are called *surface plasmon polaritons (PPP)*. From the physics point of view it is the *photons that interact with collective electron oscillations.*

The wave vectors $k_3^{(1)}$ and $k_3^{(m)}$ must be real and positive, because they correspond to waves propagating along the surface in the neighboring environments. Consequently, the ratio of (10.4) can be sutisfied, if dielectric constants $\varepsilon(\omega)$ and $\varepsilon_1$ have different signs. For metals $\varepsilon(\omega)$ takes a negative value in the frequency region where the wave vector is imaginary, that indicates the attenuation of such waves in the metal and their existence only at the metal surface. Unlike planar electromagnetic waves propagating in the bulk and transverse in nature, PPP are partially transverse and partially longitudinal waves of TM type, t. e. contain the components of the electric field along the direction $x_1$ and $x_3$ (figure 10.1, b), as well as the magnetic field along the direction $x_2$, which follows from the solutions (10.1) and (10.2). The longitudinal component of the electric field determines the propagation the field in both environments from the surface with wave vectors $k_3^{(1)}$ and $k_3^{(m)}$. Distribution of PPP field along the surface is characterized by the transverse component of the field and the wave vector $k_{PPP}$, expression of which can be obtained on the basis of the dispersion relation (10.4) in the form

$$k_{PPP} = \frac{\omega}{c}\sqrt{\frac{\varepsilon_1 \cdot \varepsilon(\omega)}{\varepsilon_1 + \varepsilon(\omega)}} \quad (10.5)$$

From the dispersion relation (Fig. 10.2) it follows that at low frequencies surface modes similar to light, ie. they are manifested mainly as *polaritons* and have a wave-like nature in this area. As frequency increase the dispersion curve PPP moves away from the curve that describes the dispersion of light, approaching an asymptote, which corresponds to the plasmon resonance frequency which is equal to frequency of pure "longitudinal" excitation in the electron gas - the frequency of the *localized surface plasmon*

$\omega_{ps} = \omega / \sqrt{2}$ ~(5÷10)·$10^{15}$ $c^{-1}$.

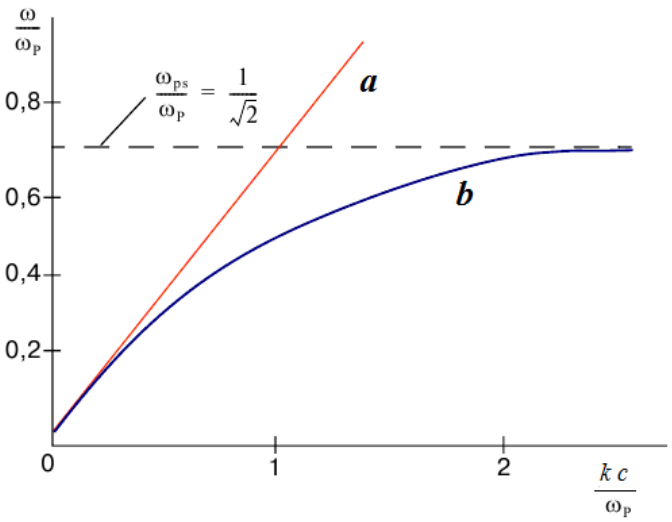



Figure 10.2 - The dispersion relation for electromagnetic waves in vacuum (a) and surface plasmon polaritons (b).

This occurs when the relative permeability of the metal and dielectric have the same value but different signs. This case corresponds to an infinite wavenumber $k \to \infty$, as it follows from (10.5). The electrostatic surface oscillations in the form of *nonpropagating* collective oscillations of electron plasma - *plasmons*, which are the limiting case when the PPP at $k \to \infty$ arise in resonance conditions.

For the visible and infrared range dispersion curves of photons and PPP are similar, but PPP wavelength slightly longer than photon wavelength at the same frequency. This means that transformation of PPP into light is not possible. Special conditions are necessary to create to excite PPP (Fig. 10.2).

From the solutions of Maxwell's equations also should, at the metal - dielectric interface it is impossible to excite PPP with use of s - polarized (TE) plane wave. Tools and systems for excitation PPP, shown schematically in Figure 10.3, have been used in specific analytical methods, which we will consider further. First, let us consider the already mentioned type of plasmon excitations - localized plasmons.

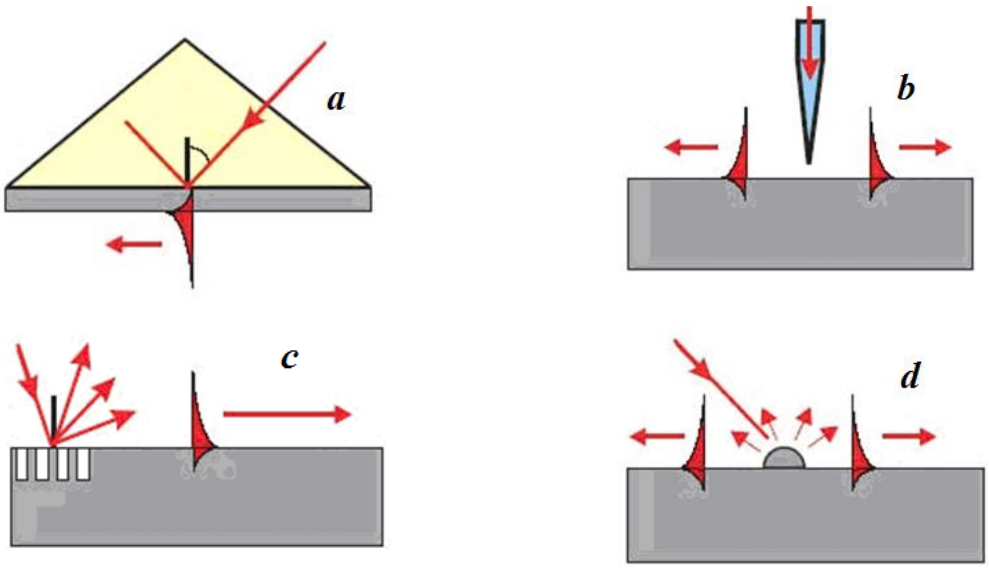


Figure 10.3 – Schematic representation of excitation PPP with p-polarization light using a prism in Kretschmann geometry (a), using an optical near-field microscope probe (b), at diffraction on a lattice (c) and on the surface roughness (d)

### 10.1.2 Localized plasmons

In addition to propagating surface plasmon polariton and plasmons, there are localized excitation, which is called *localized surface plasmon* (LPP). The roughness in the form of protrusions, grains or islands of metal particles (Figure 10.3, d) should exist on a smooth surface of the metal - dielectric interface for the LPP occurrence.

The expression for the LPP frequencies can be obtained by solving the Laplace equation with the appropriate boundary conditions. The electrostatic approximation, not taking into account the effect of the delay is applicable, if the characteristic of the particle size a lot smaller than the wavelength corresponding to the frequency of the LP. The electrostatic approximation, not taking into account the effect of the retardation is applicable, if the characteristic of the particle size *a* lot smaller than the wavelength $\lambda$ corresponding to the frequency of the LP, $a << \lambda$. For example, in the case of a metal sphere of radius *R*, located at the coordinate origin in a medium with dielectric constant $\varepsilon_0$, solution of the Laplace equation for the electrostatic potential $\phi(R,\theta,\varphi)$ within the sphere can be written as

$$\phi(R,\theta,\varphi) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} a_{lm} r^{l} Y_{lm}(\theta,\varphi), \quad 0 \le r \le R, \tag{10.6}$$

where $Y_{lm}(\theta,\varphi)$ – eigenfunctions of the Laplace operator.

Similarly, the solution of the Laplace equation for electrostatic potential outside the sphere

$$\phi(R,\theta,\varphi) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} b_{lm} \frac{1}{r^{l+1}} Y_{lm}(\theta,\varphi), \quad r \ge R. \tag{10.7}$$

The account of the boundary conditions on the surface of the sphere, namely the continuity of the potential $\phi$ and its derivative $\partial\phi/\partial r$, gives the dispersion equation for the LPP frequencies as

$$\frac{\varepsilon(\omega)}{\varepsilon_0} + \frac{l+1}{l} = 0. \tag{10.8}$$

For metals, whose dielectric constant is defined within the Drude model, the dispersion solution of equation (10.8) can be obtained in the form of

$$\omega_l = \omega_p \left[\frac{l}{\varepsilon_0 (l+1) + l}\right]^{1/2}, \qquad l = 1,2,3,... \tag{10.9}$$

Here the numbers $l$ make sense of the quantum numbers of the LP angular momentum. For small spheres a significant contribution is made only dipole modes for which the number of $l = 1.$ With particle size increasing the noticeable contribution from multipole excitations of higher order may be, so that at $l \to \infty$ the LP frequency coincide with PPP frequencies $\omega_\infty = \omega_p / (\varepsilon_0 + 1)^{1/2}$ at flat surface ( $R \to \infty$ ).

Mie scattering theory gives an accurate analytical solution for the case of plane waves scattered by the conductive particle size *a*. According to MI, the spectrum of the scattered light particle is a superposition of the so-called partial waves coming from the center of the sphere, each of these spherical waves, is characterized by angular momentum $l$ as LPP excitation.

For ellipsoid shaped particles there are three sets of LPP frequencies corresponding to the three axes of the ellipsoid [2].Three frequency of dipole LPP in a small spheroid can be calculated according to the formula for the dispersion in the form

$$\frac{\varepsilon(\omega)}{\varepsilon_0} = 1 - \frac{1}{L_i},$$

where $L_i$ - the so-called depolarization factor that depends only on the ellipsoid shape (i = 1,2,3 denotes the coordinate axes associated with the axes of symmetry of the ellipsoid). In case of a sphere $L_i$ = 1/3. The area in which LPP modes lie corresponds to the same area for the bulk plasmons frequency, for which the metal permittivity is negative.

In metallic nanoparticles quantum size effects do not have such apparent development as in the semiconductor nanoparticles. However, in the Mie theory amplitude, phase and polarization of the light scattered by particles depends on the ratio $2\pi a/\lambda$, where *a* is particle size and $\lambda$ wavelength of scattered light. Expression (10.9) not contains explicitly the particle size but excitation LPP with large angular momentum $l$ can only be realized for larger particles. Therefore, when it comes to the impact of size on the metal particle properties, in particular this applies to no internal quantum characteristics of charge carriers, but response of conductive particle as a whole. Such characteristics are microscopic cross-sections of scattering and absorption of light.

Polarizability components corresponding to the axes of the ellipsoid (i = 1, 2, 3) for particles of an ellipsoidal shape can be written within the electrostatic approximation [3] as

$$\alpha_i = \frac{4\pi}{3} abc \frac{\varepsilon_m - \varepsilon_e}{\varepsilon_e + L_i(\varepsilon_m - \varepsilon_i)}, \qquad (10.10)$$

where $\varepsilon_m$ and $\varepsilon_e$ - the dielectric constant of the metal and the environment, $L_i$ - a constant shape or depolarization; *a*, *b*, and *c* - semiaxes of the ellipse.

It is easily seen that for a spherical particle, for which $a = b = c = r$ and $L_i = 1/3$ for all directions, the formula (10.10) gives the expression

$$\alpha = 4\pi r^3 \frac{\varepsilon_m - \varepsilon_e}{\varepsilon_m + 2\varepsilon_e}. \qquad (10.11)$$

Despite the apparent simplicity of the electrostatic approximation, it can be easily discovered that metal nanoparticles properties depend on its shape and the dielectric functions of metal and environment.

The disadvantage of the electrostatic approach is the inability to examine the scattering and dissipation of energy processes in the far field at a distance $r >> \lambda$ in the so-called radiation or wave zone.

Another approach to describe the surface plasmons in ellipsoidal particles is an oscillating dipole which is localized in the particles center [4]. In the frame of the dipole approximation processes are considered in the near and far field – at distances much smaller and much higher than, the wavelength of light.

The spectral response to the LPP resonant excitation is determined by processes of plasmons decay, including radiation and non-radiative dissipation of excitation energy. Indeed, LPP can effectively delay by radiating photons [5]. Non-radiative relaxation of the LPP is associated with the excitation of electron-hole pairs, and the processes of electron-electron, electron-phonon scattering and scattering on the surface defects.

It should be noted that the PPP and LPP are the various types of electronic excitations. PPP are characterized by a dispersion curve (10.3) and are propagating surface modes. In contrast, the LPP (another called by electromagnetic resonances of the surface shape) limited by peculiarities of surface structure. They are characterized by a discrete set of frequencies, depending on the size, shape and dielectric constant of parts on the surface in which they are located. LPP can be

resonantly excited by light at an appropriate frequency and polarization, regardless of the direction of the wave vector of the incident light. Furthermore, they are damped effectively with the emission of photons.

In contrast, the PPP can be excited only if its frequency and wave vector coincide with the frequency and wave vector of the incident light. PPP can also be excited by LPP (Figure 10.3, d) and effectively decay with LPP excitation. As a result, there may be a significant increase in the scattering PPP by surface defects, if the frequency and of LPP and PPP are close [6]. On the other hand, the excitation of PPP is an effective channel of LPP decay on a metal surface [7].

## 10.2 Basics of plasmon resonance spectroscopy and plasmone-coupled fluorescence

The transformation of light to PPP with a prism method (Figure 10.2, a) is based on the phenomenon of frustrated total reflection at the incidence of p-polarized light onto the interface of two media from the optically denser medium. The angle of incidence θ is chosen from the known condition in optics

$$\theta > \arcsin(\varepsilon_1 / \varepsilon_3)^{1/2}, \quad \varepsilon_1 < \varepsilon_3, \tag{10.12}$$

where $\varepsilon_3$ - the dielectric constant of the prism material. Method exists in two versions: the geometry of Otto and Kretschmann geometry.

Let us consider the principle surface plasmon resonance (SPR) spectroscopy and sensing for the study of intermolecular reactions [8]. The prism with a high refractive index (Figure 10.4, *1*) with deposited on its surface a thin gold film (Figure 10.4, *2*) is commonly used to excite SPR. The opposite side of the film expose to the medium to be analyzed (Figure 10.4, *3*) representing the flow system.

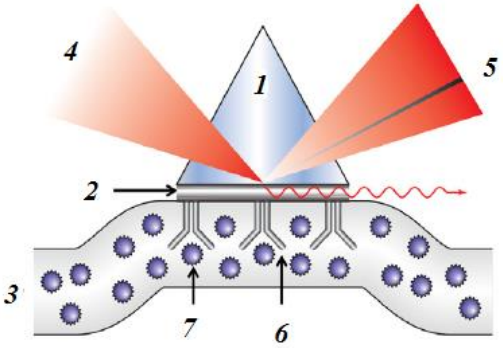



Figure 10.4 - Registration scheme in surface plasmon resonance spectroscopy and sensing

Surface plasmons are excited in the gold film under light beam (Figure 10.4, *4*), which is directed to the film surface through prism at an appropriate (10.12) angle. Light reflected from the film through a prism is displayed (Figure 10.4, *5*) for analysis. The monolayer of antibodies (Figure 10.4, *6*), that can specifically bind antigens (Figure 10.4, *7*) from the solution is deposited onto the surface of gold film, washed by the analyte solution in flow cell. Under the specific binding of antigens and antibodies the properties of monolayer on the gold film change and accordingly change the parameters such as the resonant angle $\theta$, the resonant frequency of the PPP, as well as the intensity or the phase of the reflected light.

The peptides or nucleic acid with ligand-receptor sites can be immobilized on the surface of the gold film instead of antibodies. If the analyzed objects are macromolecules (with mass > 5 kDa) their binding to the modified surface of the sensor leads to a significant change in the refractive index near the film. This assay format is called direct.

If the analyte has multiple potential binding sites with receptor molecules after complex formation the free valencies had remained on the solid phase. The free valences may be complexed with the receptor in solution, which leads to an additional increase in signal due to increasing array

of molecules bounded to the surface. As a result the analyte in the complex is confined between two or more receptor molecules. This format is called the "sandwich" and is more sensitive detection than direct format.

The polarization of the reflected light also can be used to analyze changes in the SPR. Methods that detect these changes are known as ellipsometry. The p-wave component of the light undergoes significant changes of amplitude and phase, whereas the s-component remains virtually unchanged under SPR. Thus, the ellipsometer signal reflects two processes: polarization contrast, associated with a change in the difference between the amplitudes of *s*- and *p*-component and phase contrast associated with their phase shift.

Nanoparticles of noble metals are often used as amplifiers in SPR sensors. An important advantage of metal nanoparticles is a dual mechanism of amplification PPR - by increasing the weight deposited on surface, and due to the interaction of surface plasmons in the metal film sensor and localized plasmons in the nanoparticle, ie, combination of surface plasmon resonance and localized plasmon resonance. Efficacy of double resonance depends on the distance between the nanoparticle and the surface, the size and shape of the nanoparticle. The most widely used gold nanoparticles due to its high stability, well established methods of functionalization, as well as the possibility of obtaining a controlled particle size and shape.

Figure 10.5 shows schematically what a signal can be detected in the sensing, based on PPR. If PPP excitation is carried out at a fixed wavelength, then the recorded parameters can be change in reflectance $\Delta R$ and the change in angle $\Delta\theta$, corresponding to minimum in spectrum of the reflected light caused by the sorption of the analyte (Fig. 10.5 *a*). If, conversely, the light transmittance measured at a fixed angle, the recorded changes in the wavelength of maximum spectral transmittance change $\Delta\lambda$ and dielectric constant $\Delta\varepsilon$ or $\Delta T$ at a certain wavelength (Fig. 10.5, *b*).

A special kind of luminescence it is possible to observe if the analyte immobilized at the surface of gold (or silver) film is the fluorescent molecule. This kind of emission differs from the fluorescence emission of the free molecules by spatial anisotropy and polarization. Let us consider the main features of this emission.

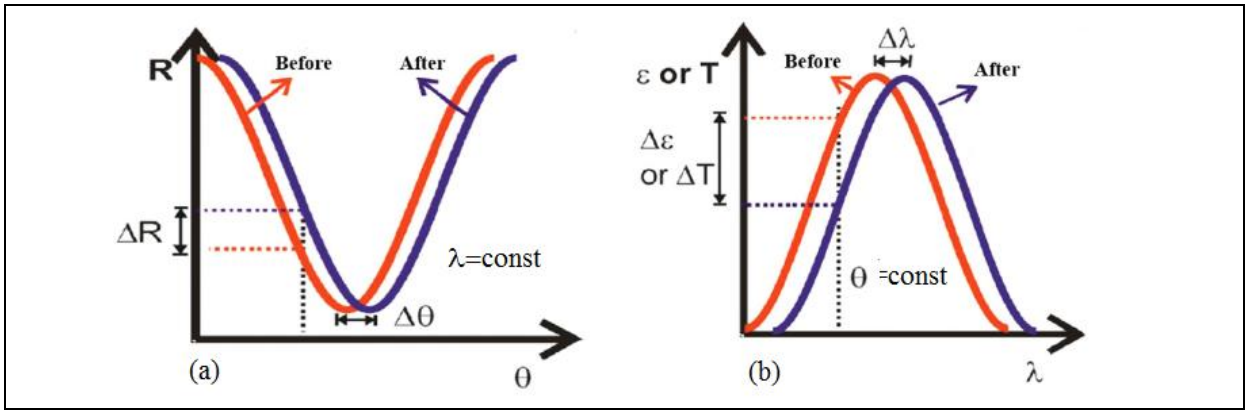



Figure 10.5- Principles of signal registration in PPR based sensorics.

The system of gold (or silver) film (Fig. 10.6, *1*, *a* and *b*) on the surface of cylindrical prism (10.6, *2*, *a* and *b*) in Kretschmann (Fig. 10.6 *a*) or reverse Kretschmann geometry (Fig. 10.6, *b*) may be used for observing the directional emission. Kretschmann geometry (KG) or reverse Kretschmann geometry (RKG) differ in the way of analyte immobilization – on the back side of gold film in KG (Figure 10.6, *3*, *a*) or between gold film prism in the RKG (Figure 10.6, *3*, *b*). Excitation light from an external source is supplied through a prism in KG (Fig. 10.6, *4*, *a*) or through the film in RKG (Fig. 10.6, *4*, *b*). Surface plasmons are excited resonantly and propagated on both surfaces of the gold film, as in the KG and in the RKG (Fig. 10.6, *6*, *a* and *b*).

The deep minimum of the angular distribution of reflectance (fig. 10.6, 5, *c*) correspond to resonance angle $\theta$ and indicate on PPP excitation. At this angle the PPP wave vectors $k_{PPP}$ (see (10.6) coincide with projection of wave vector of light, passed through the prism, namely following condition is satisfied

$$k_{PPP} = k_o n_p \sin\theta, \tag{10.13}$$

where $k_0$ and $n_p$ are the wave vector of light from internal source and prism refraction index, correspondingly.

Fully *p*-polarized luminescence is observed at angles corresponding to the minima in reflection coefficient of light (fig. 10.6, 7, a, b and c). Such directional and polarized luminescence can not be attributed to a separate fluorophores or plasmon polaritons, so it got the name of the *surface plasmon-coupled emission* (SPCE) [9]. This term was introduced by Lakowicz and reflects the essence of the observed phenomenon. To date, the rigorous theory of SPCE does not exist, but there is an understanding, accepted by many authors.

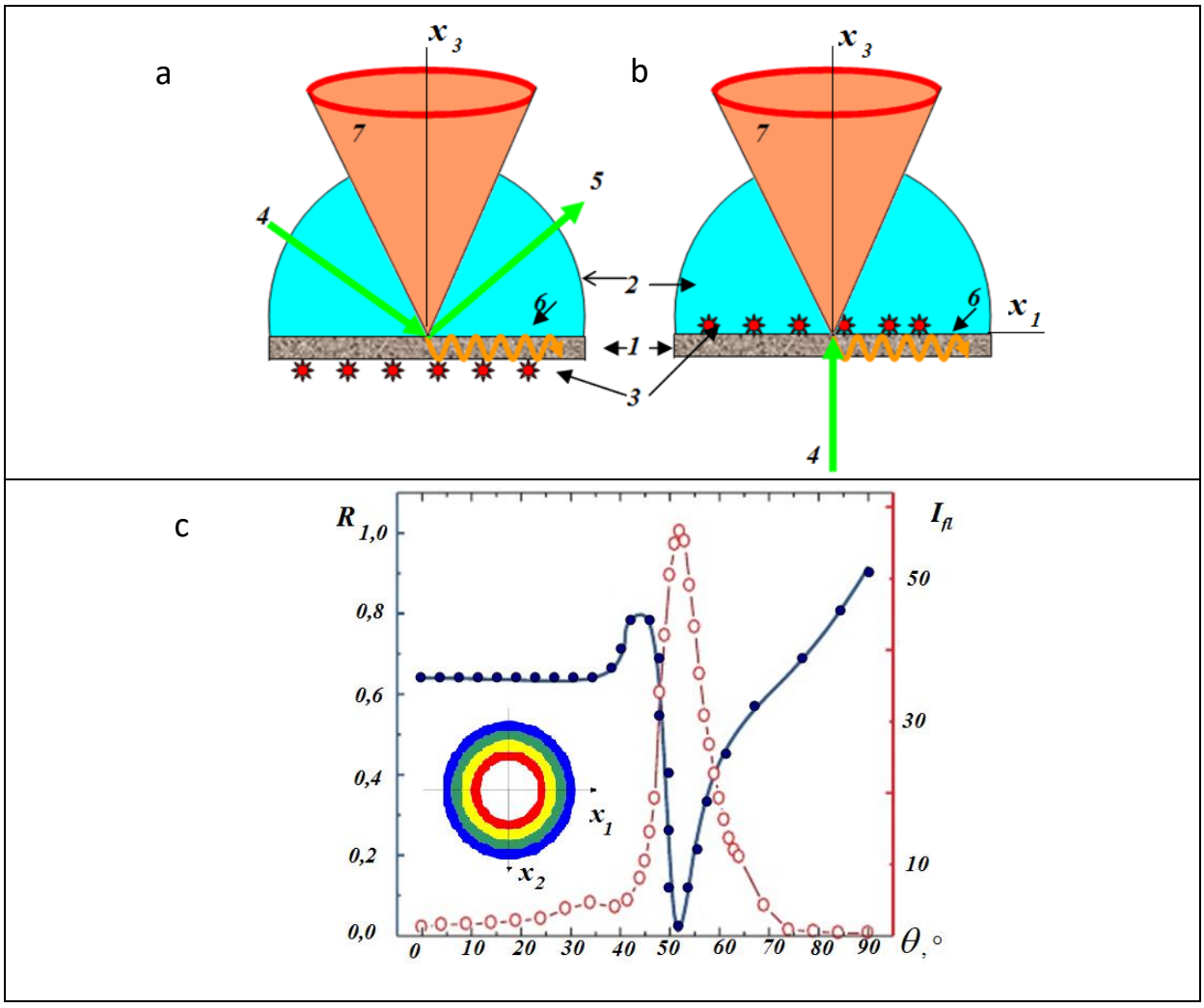



Figure 10.6 - Surface plasmon resonance and surface plasmon-coupled emission on a prism in Kretchman (a) and reverse Kretchman (b) configuration. Angular distribution of reflection index (5) and SPCE (7) of a typical fluorophore-doped polymer film spin-cast to a gold film-coated substrate attached to a prism (c). Inset to (c): Multicolor distribution of SPCE from mixture of analyte fluorophore with different wavelength of fluorescence maxima.

The best way to present this understanding is the RKG example. In the case of RKG, when the condition (10.12) known not to hold, it is believed that the PPP are excited by fluorophore

molecules. Typical Foerster radius, where the efficiency of the resonant dipole-dipole energy transfer from molecules to the plasmons is 50% equal to 5 nm. Fluorophores should be distanced from the surface of plasmon film in the range of 20 - 200 nm, which eliminates mechanism of resonant energy transfer between the excited fluorophores and surface plasmon. However, plasmons emit at angles which are determined by optical properties of the metal film and adsorbate. The spectral composition of the SPCE is the same as that of fluorophores immobilized on a substrate. However, this emission is fully *p*-polarized, even when the random orientation of the fluorophores on the substrate, and regardless of the exciting light polarization. SPCE can be excited both with unpolarized light, and in both types of *s*- or *p*- polarization in RKG. Really at normal incidence of the light there is no selected plane of incidence. This SPCE property indicates that it is not fluorophore emission.

The detailed mechanism of the phenomenon of plasmon radiation remains unclear. To date, one of the mysterious question remains – adsorbed molecule or plasmonic systems are responsible for a detectable optical response. However, in the Lakowichz group new plasmon-based method for DNA *in situ* hybridization [10], was proposed in 2003. This rapid practical introduction of SPCE is explained by fact the method is a modification of well-known SPR method for the fluorescent targets (Fig. 10.4.). The SPR method is known to be registered PPP band shift at the specific binding of molecular chromophores in the adsorption layer at the metal surface.

It is believed that in the near field of the film between the plasmon excited fluorophores and SPP, a strong coupling exist, as a result of which the plasmons emit in the far field SPCE. An important characteristic of the plasmon resonance is that the resonance angle is highly dependent on the wavelength (inset to Figure 10.6). The angular separation of the signal from fluorophores mixture is determined by the metal and dielectric constant of the prism in accordance with the expression (10.12).

PPR spectroscopy is not limited by Kretschmann configuration, as mentioned above. Let us recall (Figure 10.3), for exciting PPP, you can use the four basic ways - Kretschmann prism (figure 10.3, *a*.), the probe optical microscope (figure 10.3, *b*), a diffraction grating (figure 10.3, *c*) and inhomogeneity in the form of nanoparticles or roughness (Fig. 10.3, *d*).

Chromophore molecules located close to the metal surface, may interact with plasmon polaritons by field. If the surface is flat, these modes can not be reradiated as photons. If, however, the surface has a roughness, the PPP can be scattered as photons that leads to optical radiation from a preferred direction and intrinsic polarization [11]. Light scattered by the PPP near grating plasmon crystal is completely p-polarized and is emitted within the angles $\theta$, satisfying angular momentum conservation:

$$k_0 \sin\theta = \pm k_{PPP} \pm \frac{2\pi n}{d}, \qquad (10.14)$$

where $k_0$ and $k_{PPP}$ – the wave vectors of the incident light and the surface PPP, $d$ – the grating period, $n$ - an integer.

Plasmonic crystals can be applied in compact optical circuits and miniaturized electromagnetic devices [12], SERS substrates [13, 14], chemical and biological sensors [15], photovoltaics [16], imaging systems, tunable filters [17], infrared (IR) sources [18], etc.

At the junction of plasmonics and biomedicine and not so long ago, a new branch of science and technology - *neuroplasmonics*. The study of brain function, as we know, is important not only for medicine but also for the design of new neuro-computers and robotics. Super perfect neural system present a billions of individual neurons connected by synaptic connections with hundreds or thousands of other neurons. There are functional region in the field depth of the brain such as the thalamus, hypothalamus and others in this network. To visualize the neural processes and neural signal detection methods such as computer tomography, electroencephalography, thermal imaging, magnetic resonance tomography, positron emission tomography and SPR spectroscopy and fiber optic techniques. Fiber optic techniques, in turn, is divided into two main groups - the visualization of internal and external optical signal. The most common technique at the present

time in the study of brain function based on detection of the electrical signal. However, measurement of the electrical signal by the electrodes does not give a satisfactory spatial and temporal resolution.

Optical signal of neurons registration has a number of advantages over the registration of electrical signals. These advantages include: the absence of electrical noise, contactless and therefore non-destructive action, short depth of radiation penetration *in vivo* due to the scattering of optical radiation in the tissues, minimal invasiveness, portability and flexibility inherent in the fiber-optic techniques.

Figure 10.7 shows the main species of *neuroplasmonic* whose purpose is to register and visualization of neural signals based on SPR spectroscopy and microscopy.

Despite the fact that the method is based on the use of prisms in neuropasmonics already received a number of important results, they have a number of disadvantages associated with the bulkiness of the prism. Plasmon crystals or gratings promote PPR efficient excitation and eliminate these problems, like a prism. Plasmon crystals represent outstanding advantages, including improved optical signal transmission, field amplification, increased collimation light through subwavelength holes, multimode tunable plasmon resonances, ability to integrate into handheld devices, negative permeability and refraction of light at wavelengths beyond the diffraction limit.

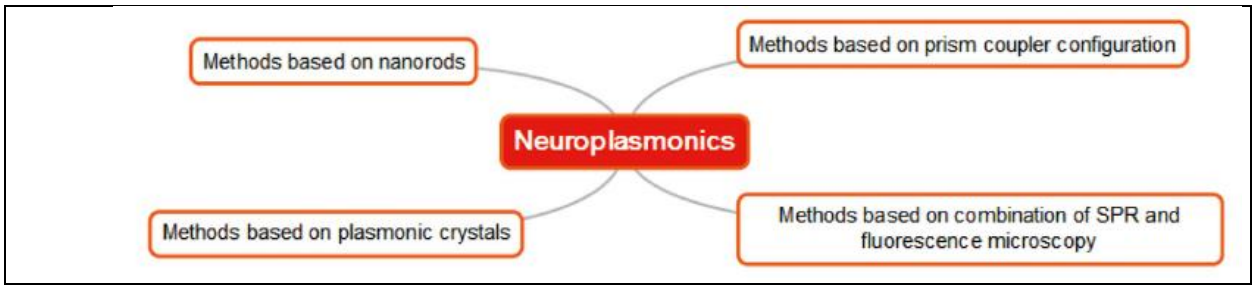



Figure 10.7 - Variations of neuroplasmonics on the implementation of PPR effect

In 2008, Zhang et al. [19] have reported a new non-invasive technique which utilized the shift in the resonance mode of a gold nanoparticle array (plasmonic crystal) to detect the neural cell activity. These plasmonic crystals based on nanoparticles —fabricated by electron beam lithography (EBL) - can be applied in optical *in vivo* imaging and probing local neuron activity on nanometer scale where local membrane ion channels operate. The essence of the experiment (Fig. 10.8) was as follows. Device structure for measuring the response of the "bare plasmonic" template to an external electrostatic field. The SU8 polymer spacer ring defines a circular microchamber filled with air. A modulating potential at 500 Hz was applied between the ITO plates. Second experiment (b): The "plasmon/neuron" device where the perfusion chamber was filled with artificial cerebrospinal fluid (aCSF).

First, they applied an external electric field with similar amplitude of an active neuron across a metal/electrolyte interface. This caused the transient change in the scattered light from the plasmonic crystal based on nanoparticle. The reason was that the external electric field modified the free electron concentration at the metal/electrolyte interface which subsequently led to the change of dielectric constant of metal and the shift in SPP resonance.

Secondly, they have grown dissociated hippocampal neurons onto Au nanoparticle plasmonic array and again corresponding transient signals were recorded when the neurons were stimulated by glutamate, a chemical trigger. Glutamate could change temporarily the neuron activity from single spikes to several consequent spikes called bursts. To complete their study, again in 2009. Zhang et al. [20], have demonstrated that how hippocampal neural spiking activity could be detected in real time by the LSP resonance shift in an embedded Au nanoparticle optical probe.

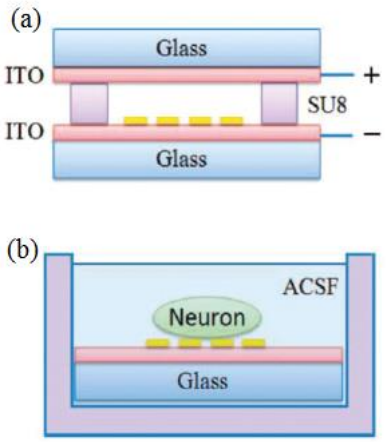



Figure 10.8 – Schematic presentation of two experiments of Zhang in 2009 ([20], Abstract picture)

Their first experiment was based on a device housing the "bare plasmonic" template. The second experiment consisted of "plasmon/neuron" device. In their experiments, they benefited from planar high density and square array of pancake-shaped Au nanoparticles on ITO-coated glass substrate. The biological sample was a tissue from the hippocampus of an E18 rat embryo. Glutamate was injected into the artificial cerebrospinal fluid in order to stimulate the neurons with microsecond duration pulse [21].

This plasmonic crystal was fabricated by EBL. This fabrication method yields spatially controlled nanoparticle arrays with flexibility of the optical spectrum of their LSPs. Furthermore, this team have simulated the mentioned structure by a modal proposed by Mclntyre [22] and used by Lioubimorv et al. [23] in order to describe the changes in the scattering signal in terms of variations in the applied voltage.

In their calculations, they benefited from Drude model for dielectric response of Au and Stern model for describing metal-electrolyte junction. For their bare plasmonic sample, the shift in the resonance wavelength ($\Delta\lambda_{LSP}$) as a result of an applied voltage $V_o$ can be written as

$$\Delta\lambda_{LSP} = -\frac{\varepsilon_0 \omega_P^{*2} \lambda 3_{LSP}}{8\pi c^2 N e d d_{TF} (\varepsilon_\infty + \frac{1-L}{L})} V_0 \tag{10.15}$$

where $\varepsilon_0$, $\omega_P^*$, $\lambda_{LSP}$, $c$, $N$, $d$, $d_{TF}$, $\varepsilon_\infty$ and $L$ are respectively, the electric permittivity of vacuum, gold plasma frequency, LSP resonance wavelength, speed of light in vacuum, electron number density, distance between the plates, Thomas-Fermi screening length, static dielectric constant and the geometrical factor in the polarization direction of the incident electromagnetic wave.

In the presence of hippocampal neural cells,

$\frac{\varepsilon_S}{d_S}(V_0 - V_S)$ and $\frac{1-L}{L}\varepsilon_D$ are substituted for $\frac{V_0}{d}$ and $\frac{1-L}{L}$, respectively.

In brief, this setup can be applied in optically probing local neural activity on submicron scale where local membrane ion channels operate and also, it can be extended to in vivo applications for all-optical recording. Moreover, this technique has the scalability in detecting multiple neuron activities in neural microcircuits according to the geometry of plasmon probe array [20].

## 10.3 Plasmonic substrates for surface-enhanced spectroscopy and medical teragnostics

Progress achieved in nanosized materials technology has renewed interest to surface-enhanced phenomena having history from 1974, when Fleischmann et al. [24] observed an unusual experimental result with the Raman scattering of pyridine molecules on roughened silver electrodes. Of particular interest are the nanoscale noble metals, which have important applications in surface-enhanced Raman scattering (SERS) [25, 26], single-molecule spectroscopy [27, 28], surface-enhanced fluorescence (SEF) [29 -31], radiative decay engineering [32], plasmon-assisted fluorescence, chemical and biological sensoring, and optoelectronic nanodevises [33]. Intensive trends in all these surface phenomena and in the near-field optical microscopy lead to the origin of the emerging field of science, which is now called plasmonics.

In the hierarchy of the nanoplasmonic objects one may select three following types – plasmonic crystals [34], resonance plasmonic particles [35] and plasmonic films [36] (Figure 10.9). Plasmonic film presents the disordered nanotextured metal surface in which surface plasmon resonance (SPP) and/or localized plasmons (LP) could be excited under resonance conditions. Such a substrate has old, preplasmonic terminology - the island vacuum deposited film, just as plasmon resonance particles earlier were named as metal colloids.

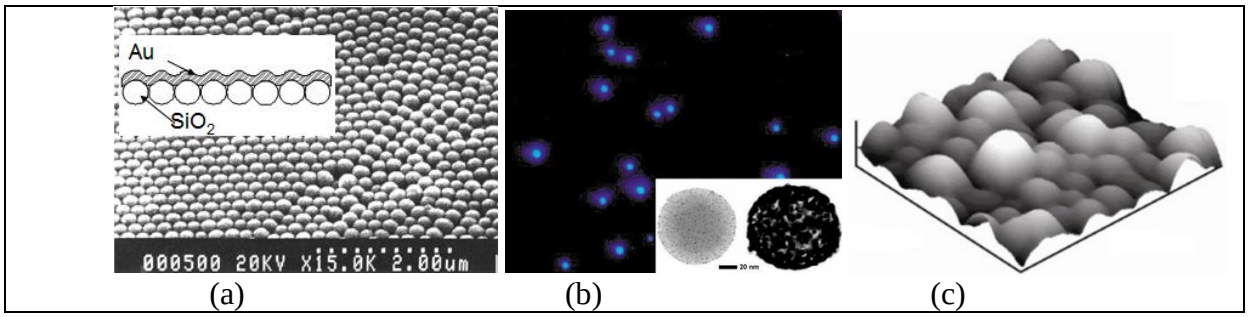



Figure 10.8 – SEM image of gold plasmonic crystal (a), TEM image of silver resonance plasmonic particles (b) and AFM image (900x900 $nm^2$) of gold plasmonic film (c)

The Figure 10.8, *a* shows an example of plasmon crystal formation by vacuum deposition of a thin layer of gold on a colloidal crystal from monodisperse $SiO_2$ globules [13]. Plasmon resonance particles are obtained by reducing the gold or silver salt (Figure 10.8, b) and a new quantum state was presented namely plexcitonic nanoparticles under the plasmonexciton coupling [37].

Silver island vacuum-evaporated films (Fig. 10.8, *c*) or gold colloidal films [38] were likely the most widely used substrates for the enhancement of both Raman and fluorescence. The surface-enhanced fluorescence (SEF) effect, however, has been described as the "weak cousin of the surface-enhanced Raman scattering (SERS) effect" (Moskovits, 1985), owing to its rather tiny enhancement factors. Moreover, fluorescence quenching was considered as the main advantage of surface-enhanced resonance Raman scattering (SERRS) [39]. Nevertheless, SEF in application to analyle molecules with low fluorescence quantum yield gives the excellent result because it was shown [40] that fluorescence enhancements increase as the quantum yield of a free molecule is decreased.

In spite of nonperiodic, nonuniform, and primitive surface morphology, plasmonic films possess some unique properties. The tailoring of its spectral properties allows to selectively excite the SERRS or the enhanced fluorescence of the same biomolecule without changing of light source [41]. These unique properties of plasmon structures are due to the interaction of light with nanoparticles and in the future they will find wide application in various applications of nanophotonics and nanoplasmonics.

One of the most impressive applications of SERS phenomena – intraoperative SERS and image-guided surgery – is presented in the work of S. Nie and coauthors [42].More important

applications of SERS and plasmon particles in medical diagnoctics and therapy (teragnostics) can be found in the review [43].

**References**


1. Zayats A.V. Nano-optics of surface plasmon polaritons / A.V. Zayats, I.I. Smolyaninov, A.A. Maradudin. – Phys. Rep., 408, 2005. – 131–314 p.
2. U. Kreibig, M. Vollmer, Optical Properties of Metal Clusters, Springer, Berlin, 1995.
3. Gersten, J.I. Surface shape resonances in: Surface Enhanced Raman Scattering / J.I. Gersten, R.K. Chang, T.E. Furtak. – Plenum Press, New York, 1982. – 89 p.
4. Jackson, J. D. Classical Electrodynamics / J.D. Jackson. – Wiley, New York, 1962. – 656 p
5. Wokaun, A.A. Radiation damping in surface-enhanced Raman scattering / A.A. Wokaun, J.P. Gordon, P.F. Liao. – Phys. Rev. Lett. 48, 1982, – 957 p.
6. Shchegrov, A.V. Scattering of Surface Plasmon Polaritons by a Circularly Symmetric Surface Defect / A.V. Shchegrov, I.V. Novikov, A.A. Maradudin, Phys. Rev. Lett., 78, 1997. – 4269.
7. Agranovich, V.M. New type of local resonances in thin rough films / V.M. Agranovich,V.E. Kravtsov, T.A. Leskova. – Solid State Commun, 47, 1983. – 925 p.
8. Д. В.Сотников, А. В.Жердев, Б. Б. Дзантиев / Детекция межмолекулярных взаимодействий, основанная на регистрации поверхностного плазмонного резонанса // Успехи биологической химии.-т. 55.- 2015.- с. 391-420.
9. Lakowicz JR. 2004. Radiative decay engineering, 3: surface plasmon-coupled directional emission. Anal Biochem 324:153-169.
10. Malicka J, Gryczynski I, Gryczynski Z, Lakowicz JR DNA hybridization using surface plasmon-coupled emission. Anal Chem. 2003 75(23): 6629–6633.
11. Kitson, S.C. Surface-plasmon energy gaps and photoluminescence / S.C. Kitson, W.L. Barnes, J.R. Sambles. – Phys. Rev. B, v.52, № 15, 1995. – 11441–11445 p.
12. X. Wei et al., Appl. Phys. Lett. 97, 011904 (2010).
13. S.V. Gaponenko, A.A. Gaiduk, O.S. Kulakovich, N.D. Strekal, S.A. Maskevich, O.A. Prokhorov, V.M. Shelekhina / Raman scattering enhancement using crystallographic surface of a colloidal crystal // JETF Letters.–2001.– Vol. 74, № 6. – P.309–311.
14. M. Najiminaini et al., Plasmonics 8, 217 (2013)
15. S.H. Lee et al., Langmuir 25, 13685 (2009).
16. J. Henzie et al.,J.Phys.Chem.B110, 14028 (2006)
17. Y.S. Do et al., Adv. Opt. Mater. 1, 133 (2013)
18. Puscasu et al., J. Appl. Phys. 98, 013531 (2005)
19. J.Q. Zhang, T. Atay, A.V. Nurmikko, Detection of Neural Cell Activity Using Plasmonic Gold Nanoparticles,inConference on Lasers and Electro-Optics/Quantum Electronics and Laser Science Conference and Photonic Applications Systems Technologies (Optical Society of America, 2008).
20. J. Zhang, T. Atay, A.V. Nurmikko, Nano Lett. 9, 519 (2009).
21. S. Venkataramani et al., J. Neurosci. Methods 160, 5 (2007).
22. J.D.E. McIntyre, Surf. Sci. 37, 658 (1973).
23. V. Lioubimov et al., Appl. Opt. 43, 3426 (2004).
24. M. Fleischmann, P. J. Hendra and A. J. McQuillan, (1974). Raman spectra of pyridine adsorbed at a silver electrode, Chem. Phys. Lett 26 ,163–166.
25. Moskovits, Surface-enhanced spectroscopy (1985). Rev. Mod. Phys. 57,783–826.
26. P. L. Stiles, J. A. Dieringer, N. C. Shah, R. P. Van Duyne. (2008). Surface-enhanced Raman spectroscopy. Annu. Rev. Anal. Chem . 1 : 601–626.
27. S. Nie, S.R. Emory, Probing single molecules and single nanoparticles by surface-enhanced Raman scattering. (1997). Science , 275 , 11021–11025.

28. K. Kneipp, Y. Wang, H. Kneipp, L. Perelman, I. Itzkan, R. Dasari, M. Feld (1997). Single molecule detection using surface-enhanced Raman scattering (SERS), Physical Review Letters 78 (9), 1667–1670.
29. K. Sokolov. G. Chumanov, T. Cotton (1998).Enhancement of molecular fl uorescence near the surface of colloidal metal fi lms Anal. Chem . 70, 3898 – 3905.
30. N. Strekal, A. Maskevich, S. Maskevich, J.-C. Jardillier, I. Nabiev (2000). Selective enhance-ment of Raman or fl uorescence spectra of biomolecules using specifi cally annealed thick gold fi lms, Biospectroscopy/ Biopolymers , 57, 325–328.
31. P. J. G. Goulet, Ricardo F. Aroca Surface-enhancement of fl uorescence near noble metal nano-structures, in Topics in Fluorescenec Spectroscopy , 8: Radiative decay engineering ( 2005). C. Geddes and J. Lakowicz eds., Springer Science+Business Media, Inc Inc., new York, 223–247.
32. C.D. Geddes, K. Aslan, I. Gryczynski, J. Malicka, J. Lakowicz (2005). Radiative decay engineering. In: C.D. Geddes, J.R. Lakovicz (eds) Topics in fl uorescence spectroscopy vol 8: Radiative decay engineering. Springer Science+Business Media, Inc, New York, 405–448.
33. S. Chen, R.S. Ingram, M.J. Hostetler, J.J. Pietron, R.W. Murray, T.G. Shaaf, J.T. Khoury, M.M. Alvares, R.L. Whetten, Science 280 (1998) 2098.
34. M. E. Stewart et al. (2006). Quantitative multispectral biosensing and 1D imaging using quasi-3D plasmonic crystals, PNAS 103, 17143–17148.
35. D.A. Shultz (2003). Plasmon resonant particles for biological detection. Curr. Opin. Biotechnol. 14, 13–22.
36. Strekal, O. Kulakovich, V. Askirka, I. Sveklo, S. Maskevich (2008). Features of the Secondary Emission Enhancement Near Plasmonic Gold Film, Plasmonics 4, 1–7.
37. Raman scattering enhancement using crystallographic surface of a colloidal crystal / S.V. Gaponenko, A.A. Gaiduk, O.S. Kulakovich, N.D. Strekal, S.A. Maskevich, O.A. Prokhorov, V.M. Shelekhina // JETF Letters.–2001.– Vol. 74, № 6. – P.309–311.
38. T. Fofang, T.-H. Park, O. Neumann, N. A. Mirin, P. Nordlander, and N. J. Halas, "Plexcitonic nanoparticles: Plasmonexciton coupling in nanoshell j-aggregate complexes," Nano Lett. 8, 3481–3487 (2008).
39. Enhanced Luminescence of CdSe Quantum Dots on Gold Colloids/ O. Kulakovich, N. Strekal, A. Yaroshevich, S.Maskevich, S. Gaponenko, I. Nabiev, U. Woggon, M. Artemyev // Nano Letters. –2002. –Vol. 2, N 12. – P. 1449–1452.
40. R. R. Chance, A. Prock and R. Silbey (1978). Molecular fl uorescence and energy transfer near interfaces. Adv. Chem. Phys . 37, 1–65
41. D. A. Weitz, S. Garoff, J. I. Gersten and A. Nitzan (1983). The enhancement of Raman scattering, resonance Raman scattering, and fl uorescence from molecules adsorbed on a rough silver surface, J. Chem. Phys . 78, 5324–5338
42. Strekal, N. Plasmonic Gold and Silver Films: Selective Enhancement of Chromophore Raman Scattering or Plasmon–Assisted Fluorescence / N. Strekal, S. Maskevich // Reviews in Plasmonics 2010 / ed. C.D. Geddes.– Reviews in Plasmonics. – New York: Springer Science+Business Media, LLC, 2012.–P. 283–302.
43. Mohs, A. M.; Mancini, M. C.; Singhal, S.; Provenzale, J. M.; Leyland-Jones, B.; Wang, M. D.; Nie, S. Hand-Held Spectroscopic Device for in Vivo and Intraoperative Tumor Detection: Contrast Enhancement, Detection Sensitivity, and Tissue Penetration. Anal. Chem. 2010, 82, 9058−9065.
44. Lane L. A., Qian X.mei, Nie S. SERS Nanoparticles in Medicine: From Label-Free Detection to Spectroscopic Tagging/ *Chem. Rev.*, 2015, 115 (19), pp 10489–10529Chem Rev.

## 11. Nanofibers: synthesis, properties and applications

### 11.1 Basics of electrospinning

To understand what electrospinning is, you can look at the mechanism for producing polymer fibers. Conventional fibers of large diameter are obtained by extruding a polymer melt through a spinneret. The resulting elongated melt is dried to form an individual fiber strand. Similarly, electrospinning also involves drawing a molten polymer or polymer solution. However, unlike the conventional spinning method, where is an external mechanical force that pushes the molten polymer through the die, a charge is used in electrospinning that is applied to the liquid to provide a pulling force to the manifold, creating a potential difference. When a sufficiently high voltage is applied, the jet of polymer solution is emitted from the solution. Intermolecular interaction of polymer chains in solution will prevent the decay of the jet. Electrospinning of polymer solutions is based on evaporation of the solvent from the polymer and its solidification to form a polymer fiber.

#### 11.1.2 Electrospinning process

In electrospinning, a high voltage is applied to the polymer liquid, and charges are distributed throughout its volume. When the number of charges in the liquid reaches a critical amount, the jet of liquid breaks out of the solution and forms a Taylor cone. The electroformed jet moves to a lower potential, which in most cases is an grounded electrode. There are many parameters that affect the morphology of the resulting electroformed fiber, making it possible to obtain fibers with "beads" or pores on the surface. It was also possible to electroform the melts of polymers, although polymer solutions are mainly used. Due to this, the parameters of the solvents that affect the electrospinning are of great interest.

Parameters influencing the electrospinning and fibers can be simplified to classify the parameters of polymer solutions and the technological parameters of the electrospinning process, which include the applied voltage, the geometry of the interelectrode space, temperature and humidity. With the understanding of these parameters it is possible to obtain spinning conditions for fiber structures of various shapes and order. It is possible to create nanofibers with different structures by varying the parameters.

#### 11.1.3 Polymer solutions parameters

The properties of the polymer solution have the strongest effect on the resulting fiber during the electrospinning process. Surface tension affects the formation of defects and the length of the fiber. One of the necessary conditions for performing electrospinning is that the solution must be prepared from a polymer with a proper molecular mass and must have a proper viscosity.

The molecular weight of the polymer is related to the long polymer chain, which in turn affects the viscosity of the solution. The increase in the viscosity of the solution is also associated with an increase in the concentration of the polymer. Just like increasing the molecular weight, increasing the concentration will lead to more links between the polymer chains in the solution, which is necessary to maintain the continuity of the jet during electrospinning. At the same time, there is a minimum viscosity required for electrospinning.

For each polymer, there is a minimum viscosity needed to form fibers without spindle-shaped seals. At low viscosity, spindle-shaped seals on the resulting fiber are usually observed. With increasing viscosity, a gradual change is observed in the form of defects from spherical to spindle-shaped, and then to obtaining a defect-free fiber.

The interaction of a jet of solution with its charge determines the distribution of diameters of the resulting fiber. During electrospinning, secondary jets separation from the main jet [1] are possible, which are sufficiently stable to produce fibers of smaller diameter at a certain viscosity. This may explain the different distribution of fiber diameters that is observed in some cases [2, 3,

4]. However, high viscosities can prevent the separation of secondary jets from the main jet, which can contribute to an increase in the diameter of the fibers [5].

Although viscosity plays an important role in the formation of defect-free fibers, it does not determine the concentration at which the fibers are formed by electrospinning.

For polyethylene oxide, it was found that the minimum concentration for forming defect-free fibers was the same, despite a 3.5x increase in viscosity with increasing molecular weight of the polymer, it was increased from 8,000 to 8,000,000. Thus, in this case, it seems that the concentration Greater influence on the formation of defect-free fibers [6].

### 11.1.4 Surface tension

The initiation of electrospinning occurs when a charged solution overcomes its surface tension. However, when the jet is moved to the collecting plate, surface tension can cause defect formation along the jet. Surface tension reduces surface area per fluid unit mass. In the case of a high concentration of free solvent molecules, due to surface tension, these solvent molecules collect and assume a spherical shape. Greater viscosity means that there is a closer interaction between the solvent and the polymer molecules, so when the solution expands under the influence of charges, the solvent molecules will tend to spread through intertwined polymer molecules, while reducing the tendency of the solvent molecules to collect together under the action of surface tension.

A solvent such as ethanol has a low surface tension, so it can be used as an additive to produce defect-free fibers [7]. The addition of surfactants makes it possible to produce more homogeneous fibers. Even when an insoluble surfactant is used as an additive in the form of fine powder, the morphology of the fiber also improves.

### 11.1.5 Conductivity of solution

Electrospinning is the expansion of a solution caused by the repulsion of charges on its surface. As the conductivity of the solution increases, the charge increases, which can be carried by the jet. The conductivity of the solution can be increased by adding ions. Moreover, most drugs and proteins form ions when dissolved in water.

As mentioned before, the defects will be formed if the solution is not fully extended. Adding a small amount of salt or polyelectrolyte to the solution increases the charge-borne charge by promoting the expansion of the solution. As a result, defect-free fibers are obtained, which otherwise would be defective. The increase in the extraction of the solution also contributes to a decrease in the diameter of the fibers [8]. Nevertheless, there is a limit to reducing the diameter of the fiber. With the extension of the solution, the viscoelastic force acting against the Coulomb forces of charges increases [9]. Since the presence of ions increases the conductivity of the solution, the critical stress for electrospinning decreases [10]. Another effect of increasing the number of charges is greater bending instability. As a result, the fiber deposition area increases [11]. This will also contribute to the formation of thinner fibers, since the path of the jet increases.

The size of the ions can influence the morphology of the fibers. The electroformed fibers with the smallest diameter were obtained from the solution with NaCl, and the fibers from the solution with dissolved $KH_2PO_4$ had the largest diameter, and fibers from the solution with $NaH_2PO_4$ had an intermediate diameter. Since sodium and chlorine ions have a smaller atomic radius than the atomic radius of potassium and phosphate ions, they have greater mobility under the external electrostatic field. Incresed elongating force e on the jet caused by more mobile small ions can give possibility to produce fibers with a smaller diameter [8].

To increase the conductivity of the solution while reducing the surface tension, additives such as triethylamine benzyl ammonium chloride are used [12]. They also reduce the fiber diameter. Another way to increase the conductivity of the solution by changing the pH of the solution. In an alkaline medium, the electrospinning of cellulose acetate leads to a significant

decrease in fiber diameter compared to the results obtained in a neutral medium [10]. So cellulose acetate will be subject to deacetylation in an alkaline medium, $OH^-$ ions can have a greater effect on the conductivity and extension of the solution.

However, the interaction between the reagent that is added to improve the conductivity of the solution and the solution itself can also have a significant effect on the fibers produced. When dissolving cellulose acetate in an acidic medium, the electroformed fibers increased in diameter as the conductivity of the solution increased [10]. In some cases, the addition of salt can lead to an increase in the viscosity of the solution. Thus, although the conductivity of the solution increases, the viscoelastic force is stronger than the Coulomb force, because of this, an increase in the diameter of the fibers is observed, instead of decreasing [13].

### 11.1.6 Dielectric permittivity of the solvent

The dielectric constant of the solvent has a significant effect on the electrospinning. As a rule, in a solution with a greater dielectric constant, the formation of spindle-shaped defects and the diameter of the resulting fiber decrease [14]. Solvents such as N, N-dimethylformamide (DMF) are added to the solution to increase its dielectric properties and improve the fiber morphology [15]. The bending instability of the jet also increases with increasing dielectric permittivity. This is manifested in an increase in the area of application of fibers. This can also help reduce the diameter of the fibers due to the increase in the jet path [16].

A solvent with a higher dielectric constant can be added to the solution to increase its electroformability. However, the composition of the mixture will affect the solubility of the polymer, which in turn will affect the morphology of the resulting fiber. When DMF was added to the polystyrene (PS) solution, granules were obtained although the electrospinability of the solution should be better because of the higher dielectric constant of DMF. This may be the result of retraction of the polystyrene molecule due to poor interaction between the polystyrene and the solvent molecules [17].

### 11.1.7 Electrospinning process parameters

Another important parameter that affects the process of electrospinning is various external factors that affect the jet of the solution during electrospinning. The main ones are voltage, feed rate, solution temperature, type of collectors, interelectrode distance. These parameters have a certain influence in the morphology of the fiber, although they are less significant than the parameters of the solutions.

### 11.1.8 Voltage

The most important element in electrospinning is the application of high voltage to the solution. The high voltage induces the necessary charges on the solution and, together with the external electric field, will initiate the electrospinning process when the electrostatic force in the solution overcomes the surface tension of the solution. Typically, a high negative or positive voltage greater than 6 kV can cause droplets to appear at the tip of the needle and distort them in the form of a Taylor cone when a jet occurs [18]. Depending on the speed of the solution, a higher voltage may be required for the Taylor cone to be stable. The Coulomb repulsive force in the jet will stretch the viscoelastic solution. If the applied voltage is higher, the magnitude of the charges will cause the jet to accelerate and a larger amount of solution will be drawn through the tip of the needle. This can lead to a decrease in the size and stability of the Taylor cone [8].

The applied voltage, as well as the resulting electric field, affects the tension and acceleration of the jet, that is, the morphology of the resulting fibers. In most cases, a higher voltage leads to a greater stretching of the solution due to large Coulomb forces in the jet due to a stronger electric field. This reduces the diameter of the fibers [19, 20], and also accelerates the

evaporation of the solvent, which accelerates the drying of the fibers [21]. When a solution with a low viscosity is used, a higher voltage may contribute to the formation of secondary jets during electrospinning. This can reduce the diameter of the fibers [22].

Another factor that can affect the fiber diameter is the flight time of the jet. The greater the value of the flight time of the jet, the more fibers are stretched and lengthened, before hitting the collecting plate. Thus, at a lower voltage, a smaller acceleration of the jet and an electric field, can increase the flight time of the electrospinning jet, which can contribute to the formation of thinner fibers.

In this case, the voltage close to the critical voltage for electrospinning can be advantageous for obtaining thinner fibers [5].

At a higher voltage, it was found that there is a great tendency for the formation of beads [22, 23]. It has also been reported that the shape of beads varies from spindle to spherical with increasing stress [8] Given the increased jet stretching due to higher stress, the number of spindle-shaped defects should be less as reported in some cases [24]. The increased density of spindle-shaped defects due to increased stresses can be the result of increased instability of the jet due to the fact that the Taylor cone retreats to the needle [8, 23]. [25] reports that as the voltage increases, the density of spindle-shaped defects increases, which at a higher voltage are connected, forming a thicker fiber diameter.

High stress affects not only the physical state of the fiber, but also the degree of crystallinity of the polymer fibers. The electrostatic field can cause the ordering of polymer molecules during electrospinning, hence the crystallinity of the fiber increases. However, above a certain voltage, the crystallinity of the fiber is reduced.

With increasing voltage, the acceleration of the fibers also increases. This reduces the flight time of the electrospinning jet. Since the orientation of the polymer molecules will take some time, a shorter flight time means that the fibers will be formed before the polymer molecules have enough time to align themselves. Thus, given the sufficient flight time, the crystallinity of the fiber will improve with a higher voltage [5].

Because electrospinning is caused by charges on the jet, these charges can be influenced by an external electric field, which in turn affects the path of the jet. It is therefore not surprising that attempts have been made to control the morphology of nanofibers through a change in the profile of the electric field between the source of the jet and the collector. This can be achieved by using auxiliary electrodes or by changing the orientation or shape of the manifold. Oriented nanofibres can be obtained by manipulating an electric field.

In electrospinning DC voltage is used, but it is also possible to use alternating current for electrospinning. When the jet moves to the collecting plate, the regular segments of the jet will contain positive or negative charges on it. Since bending instability is the result of Coulomb repulsive forces in the jet, regular segments in the jet of positive or negative stress reduce repulsive forces, thus reducing the bending instability in the jet. Since there is less bending instability and less jet stretching, the emerging fibers have a larger diameter than the fibers that are formed from a constant current of the same voltage.

Another advantage of AC is that there is less tendency to accumulate charges on the fiber after its application. Thus, a thicker layer of electroformed fiber can be assembled especially when a collecting plate is used from the insulator.

### 11.1.9 Spinning solution feed rate

The feed rate determines the volume of the solution available for electrospinning. For a given voltage, there is a corresponding feed rate at which the stability of Taylor cones is maintained. When the feed rate increases, a corresponding increase in fiber diameter or defects occurs.

However, there is a limit to the increase in fiber diameter due to a higher feed rate [26]. If the feed rate corresponds to the entrainment of the solution with the electrospinning jet, then with an increase in the feed rate there must be a corresponding increase in charge. Thus, the corresponding increase in tension of the solution opposes the increase in diameter due to the increased volume

Due to the large volume of the solution, the jet will need more time to dry. As a result, the solvent in the precipitated fiber may not have sufficient time to evaporate during the flight time. Remains of the solvent can cause compounds of fibers, which come into contact, form a web. A lower feed rate is more desirable because the solvent will have more time to evaporate.

### 11.10 Temperature

An increase in the temperature of the solution leads to an increase in the evaporation rate and a decrease in the viscosity of the polymer solution. Polyurethane, electroformable at a higher temperature, forms fibers with a more even distribution of diameters [23]. Perhaps this is due to the lower viscosity of the solution and the greater solubility of the polymer in the solvent, which allows a more even tension of the solution.

At low viscosity, Coulomb forces are able to exert a greater tensile force on the solution, resulting in fibers of smaller diameter [13]. Increased mobility of polymer molecules due to temperature increase also allows Coulomb forces to draw out the solution more. However, in cases where biological substances, such as enzymes and proteins are added to the electrospinning solution, the use of high temperatures can cause the loss of the substance with its functionality.

### 11.11 Effect of Collective Electrode

The electric field between the fiber-forming electrode and the collector for is necessary for initiating the electrospinning process. Thus, in most electrospinning units, the collector is made of a conductive material, such as an aluminum foil, which is grounded so that there is a stable potential difference between the source and the collector. In the case where the non-conductive material is used as a collector, the charges on the electrospinning jet will quickly accumulate on the collector, this will result in fewer fibers being produced [27]. Fibers collected on a non-conductive material usually have a lower packing density than those collected on a conductive surface.

This is caused by an increase in the forces of repulsion of accumulated charges on the collector with a larger layer of accumulated fibers. For conducting collectors, charges on fibers are sprayed, thus allowing more fibers to be attracted to the manifold. As a result, the fibers can be more tightly packed.

For non-conducting collectors, the accumulation of charges can lead to the formation of 3D structures of the fiber due to repulsive forces of similar charges. With a sufficient density of charges on the fiber network, which is formed initially, repulsion of subsequent fibers can lead to the formation of a honeycomb structure [23]. However, even for a conductive collector, when the deposition rate is high and the fiber mesh is thick enough, there will also be a high accumulation of residual charges on the mesh fibers, since polymer nanofibers are generally not conductive. This can lead to dimples on the fibers of the net.

The porosity of the reservoir may have an effect on the placement of fibers. Experiments with porous collectors, such as paper and metal mesh, showed that the fiber mesh collectors had a lower packing density than smooth surfaces, such as metal foil. This can be caused by the diffusion and rate of evaporation of the residual solvent on the collected fibers. In a porous collector, it is possible to quickly evaporate the residual solvent from the fibers due to an increase in the surface area, and smooth surfaces can lead to the accumulation of solvents around the fibers due to the slow evaporation rate.

Due to the capillary spread of moisture and the diffusion of residual solvent in the fibers, the fibers can be combined to give more tightly packed structures. Because Fibers quickly dry on a porous collector, it is more likely that residual charges remain on the fiber, which will repel subsequent fibers. However, on a smooth surface, the residual solvent will induce residual charges to flow into the collector.

**References**


1. Reneker, D. H., Yarin, A. L., Fong, H. and Koombhongse, S. (2000). Bending instability of electrically charged liquid jets of polymer solutions in electrospinning. J. Appl. Phys. 87, pp. 4531-4547.
2. Kim, J. R., Choi, S. W., Jo, S. M., Lee, W. S. and Kim, B. C. (2005). Characterization and Properties of P(VdF-HFP)-Based Fibrous Polymer Electrolyte Membrane Prepared by Electrospinning. J. Electrochem. Soc. 152, pp. A295-A300.
3. Demir, M. M., Yilgor, I., Yilgor, E. and Erman, B. (2002). Electrospinning of polyurethane fibers. Polymer. 43, pp. 3303-3309.
4. Deitzel, J. M., Kleinmeyer, J., Harris, D. and Tan, N. C. B. (2001b). The effect of processing variables on the morphology of electrospun nanofibers and textiles. Polymer. 42, pp. 261-272.
5. Zhao, S. L., Wu, X. H., Wang, L. G. and Huang, Y. (2004). Electrospinning of EthylCyanoethyl Cellulose/Tetrahydrofuran Solutions. J. Appl. Polym. Sci. 91, pp. 242-246.
6. Morozov, V. N., Morozova, T. Y. and Kallenbach N. R. (1998). Atomic force microscopy of structures produced by electrospraying polymer solutions. Int. J.Mass. Spectrom. 178, pp. 143-159.
7. Fong, H. and Reneker, D.H. (1999). Elastomeric nanofibers of styrene-butadiene-styrene triblock copolymer, J. Polym. Sci. Pt. B-Polym. Phys. 37, pp. 3488-3493.
8. Zhong, X. H., Kim, K. S., Fang, D. F., Ran, S. F., Hsiao, B. S. and Chu, B. (2002).Structure and process relationship of Electrospun bioabsorbable nanofibermembranes. Polymer. 43, pp. 4403-4412
9. Choi, J.S., Lee, S.W., Jeong, L., Bae, S.H., Min, B.C., Youk, J.H. and Park, W.H. (2004a). Effect of organosoluble salts on the nanofibrous structure of electrospun poly(3-hydroxybutyrate-co-3-hydroxyvalerate). International Journal of Biological Macromolecules, 34, 4, pp. 249-256
10. Son, W.K., Youk, J.H., Lee, T.S. and Park, W.H. (2004c). Electrospinning of ultrafmecellulose acetate fibers: studies of a new solvent system and deacetylation of ultrafme cellulose acetate fibers, J. Polym. Sci. Pt. B-Polym. Phys. 42, pp. 5-11.
11. Choi, S. S., Lee, S. G., Joo, C. W., Im, S. S. and Kim, S. H. (2004c). Formation of interfiber bonding in electrospun poly(etherimide) nanofiber web. J. Mater. Sci. 39, pp. 1511-1513.
12. Zeng, J., Xu, X., Chen, X., Liang, Q., Bian, X., Yang, L. and ling, X. (2003). Biodegradable electrospun fibers for drug delivery, J. Control. Release, 92, pp. 227-231.
13. Mit-uppatham, C, Nithitanakul, M. and Supaphol, P. (2004). Ultrafme Electrospun Polyarnide-6 Fibers: Effect of Solution Conditions on Morphology and Average Fiber Diameter. Macromol. Chem. Physic. 205, pp. 2327-2338.
14. Son, W. K, Youk, J. H, Lee, T. S. and Park, W. H. (2004a).The effects of solution properties and polyelectrolyte on electrospinning of ultrafme poly(ethylene oxide) fibers. Polymer. 45, pp. 2959-2966.
15. Lee, K.H., Kim, H.Y., Ra, Y.M. and Lee, D.R. (2003b). Characterization of nanostructured poly(e-caprolactone) nonwoven mats via electrospinning, Polymer. 44, pp. 1287-1294.
16. Hsu, C. M. and Shivakumar, S. (2004). N,N-Dimethylformamide Additions to the Solution for the Electrospinning of Poly(e-caprolactone) Nanofibers. Macromol. Mater. Eng. 289, pp. 334-340.

17. Wannatong, L., Sirivat, A. and Supaphol, P. (2004). Effects of solvents on electrospun polymeric fibers: preliminary study on polystyrene. Polym. Int. 53, 1851-1859.
18. Taylor, G. (1964). Disintegration of Water Drops in an Electric Field. Proc. R. Soc. Lond. A. 280, pp. 383-397.
19. Lee, J.S., Choi, K.H., Ghim, H.D., Kim, S.S., Chun, D.H., Kim, H.Y. and Lyoo, W.S. (2004). Role of molecular weight of atactic poly(vinyl alcohol) (PVA) in the structure and properties of PVA nanofabric prepared by electrospinning, J. Appl. Polym. Sci. 93, pp. 1638-1646.
20. Megelski, S., Stephens, J.S., Chase, D.B. and Rabolt, J.F. (2002). Micro- and nanostructured surface morphology on electrospun polymer fibers, Macromolecules. 35, pp. 8456-8466.
21. Pawlowski, K. J., Belvin, H. L., Raney, D. L., Su, J., Harrison, J. S. and Siochi, E. J. (2003). Electrospinning of a micro-air vehicle wing skin. Polymer. 44, pp. 1309-1314.
22. Demir, M. M., Yilgor, I., Yilgor, E. and Erman, B. (2002). Electrospinning of polyurethane fibers. Polymer. 43, pp. 3303-3309.
23. Deitzel, J. M., Kleinmeyer, J., Harris, D. and Tan, N. C. B. (2001b). The effect of processing variables on the morphology of electrospun nanofibers and textiles. Polymer. 42, pp. 261-272
24. Jarusuwannapoom, T., Hongrojjanawiwat, W., Jitjaicham, S., Wannatong, L.,Nithitanakul, M., Pattamaprom, C, Koombhongse, P., Rangkupan, R. and Supaphol, P. (2005). Effect of solvents on electro-spinnability of polystyrene solutions and morphological appearance of resulting electrospun polystyrene fibers. Euro. Polym. J. 41, pp. 409-421.
25. Krishnappa, R. V. N., Desai, K., Sung, C. M. (2003). Morphological study of electrospun polycarbonates as a function of the solvent and processing voltage. J. Mater. Sci. 38, pp. 2357-2365.
26. Rutledge, G. C, Li, Y., Fridrikh, S., Warner, S. B., Kalayci, V. E. and Patra, P. (2000). Electrostatic Spinning and Properties of Ultrafine Fibers, National Textile Center, 2000 Annual Report (M98-D01), National Textile Center, pp. 1-10 Kessick, R. and Tepper, G. (2004). Microscale polymeric helical structures produced by electrospinning. Appl Phys Lett, 84, 23, pp. 4807-4809.

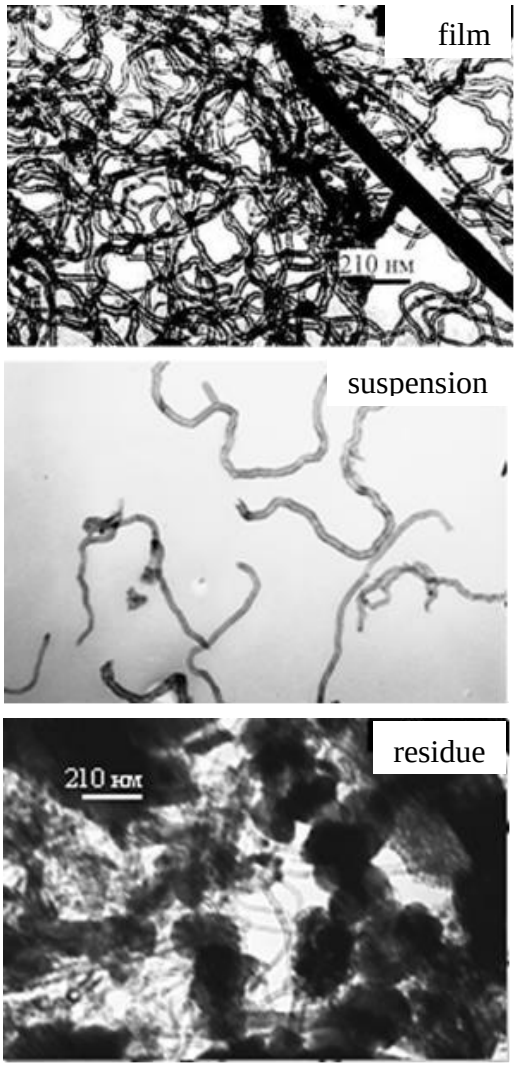
## 12. Elastomeric compositions with carbon nanomaterials
## 12.1 Unfilled elastomeric compositions with carbon nanomaterials

At present, it is an important task to develop composite materials containing nanostructures. This is due to their unique physical and chemical properties, which differ from the properties of macro- and microparticles [1]. However, to achieve positive results from using nanoadditives in the rubber industry, a number of questions need to be addressed. One of these is the correct choice of the type of nanomaterial (in terms of structure and composition, particle size and shape, the presence of active centres, and reinforcing capacity) to achieve a positive effect on a specific elastomer matrix.

Of interest for specialists of the rubber industry are both the processing properties of rubbers and rubber mixes and the service characteristics of vulcanisates [2].

The aim of the present work was to determine the effect of carbon nanomaterials (CNMs) on the plastoelastic properties and vulcanisation kinetics of rubber mixes based on general-purpose and special-purpose rubbers, and also on the physicomechanical properties of vulcanisates based on them. The initial nanomaterial was obtained in a high-voltage discharge plasma. After its complexacid and annealing treatment, the nanomaterial was divided into fractions by ultrasound: "film" fraction - particles of graphite and amorphous carbon and relatively coarse entanglements of tubes and fibres; "suspension" fraction - individual carbon nanotubes; "residue" fraction-entanglements of carbon nanotubes and fibres. Figure 12.1 presents electron micrographs of the different fractions of the carbon nanomaterials.

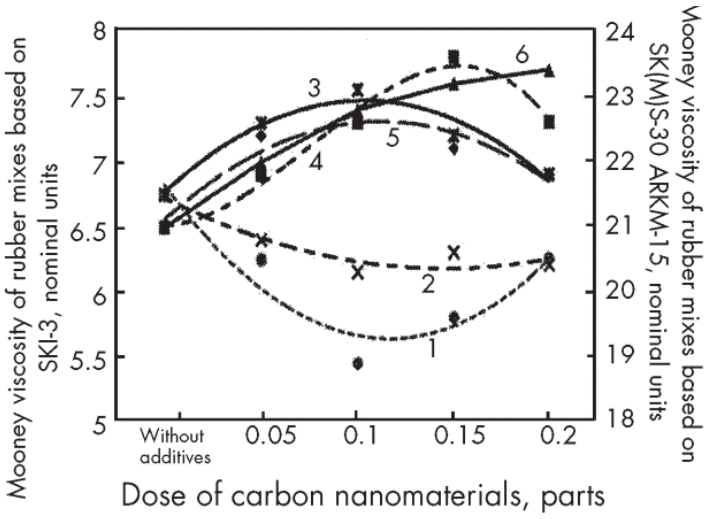
Fig. 12.1 Electron micrographs of the carbon nanomaterials: (a) film; (b) suspension; (c) residue




The investigation was conducted on rubber mixes based on crystallising isoprene rubber SKI-3 and non- crystallising oil-extended butadiene-styrene rubber SKMS-30 ARKM-15, which are general-purpose rubbers, and on rubber mixes based on special-purpose rubbers, namefy butadiene-acrylonitrile rubbers of grades BNKS-18AN and BNKS-40M. The rubber mixes contained CNMs in dosesranging from 0.05 to0.2 phr. Specimens without nanoadditive were used for comparison.

Elastomer composites are extremely complex systems, which makes it difficult to investigate the mechanism of interaction of nanoparticles with rubber macromolecules. Therefore, an investigation was made of model rubber mixes containing no filler, which were prepared on a laboratory mill. To improve the distribution of nanoadditives within the mix, composites of CNMs with zinc white were prepared beforehand (separately for each dose), and these were introduced into the elastomer matrix. The procedure hod been developed and tested at earlier stages of the investigation.

Fig. 12.2 Mooney viscosity of rubber mixes based on general-purpose rubbers: 1 – suspension SKMS; 2 – residue SKMS; 3 – film SKMS; 4 – suspension SKI; 5 – residue SKI; 6 – film SKI

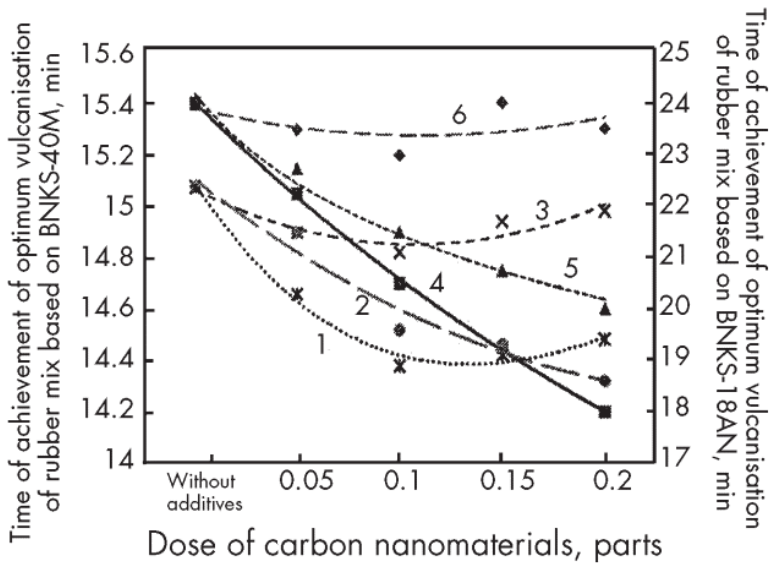



An assessment was made of the change in Mooney viscosity when nanoadditives were introduced, because the viscosity of rubber mixes is associafed with important characteristics of elastomer composites: the glass transition and flow temperatures; processabilify; the physic-mechanical characteristics of the vulcanisates [3]. The Mooney viscosity was determined on an MV2000 viscometer, and tests were conducted by a procedure similar to that recommended by the GOST 10722-76 standard.

It was established that the introduction of CNM fractions in all doses increases negligibly the Mooney viscosity of rubber mixes based on SKl-3 (see Figure 12.2). It must be pointed out that, with change in the structure of the additive, there is a change in the nature of the dependence of viscosity on the dose of CNM fraction. Thus, the maximum viscosity of a rubber mix containing "suspension" fraction is observed at a dose of 0.15 parts, and for "residue" fraction at 0.05 parts. In the case of "film" fraction there is an increase in the viscosity of elastomer composites in the entire range of variation in the amount of nanoadditive. The introduction of the given fraction into rubber mixes based on butadiene-styrene rubber also leads to a certain increase in viscosity. Increase in viscosity in the given cases is due to the formation of physical crosslinks on account of adsorption interaction of rubber mocromolecules with functional groups on the developed surface of the nanoadditives [4]. This adsorption appears most clearly in the case of the "film" and "suspension" fractions, which are more homogeneous in composition. In rubber mixes based on butadiene-styrene rubber there is a reduction in the structure-forming capacity of the additives; this is possibly connected with the oil content in the rubber, which slightfy deactivates the surface of the carbon nanomaterial. Thus, the introduction of fractions consisting of carbon nanotubes, and also accumulations of fibres and nanotubes, into rubber mixes based on an oil-extended rubber leads to a certain reduction in the Mooney viscosity, it seems that to some degree the nanotubes facilitate orientation and promote an increase in the flexibility of macromolecules of the rubber [5].

For mixes based on butadiene-acrylonitrile rubber of different grades, a similar effect of CNM on plastoelastic characteristics can be traced. It was established that the introduction of all CNM fractions leads to a certain reduction in the Mooney viscosity of the elastomer composites. It must be pointed out that the introduction of "him" froction in doses ranging from 0.05 to 0.1 parts leads to a reduction in viscosity, but with subsequent increase in the amount of CNM of the given fraction there is an increase in the viscosity of the rubber mixes. From the data given in Figure 12.3 it can be seen that the "suspension" fraction, which is more homogeneous in composition (structure), reduces the viscosity of the rubber mixes practically linear with increasing dose. The "residue" fraction has no significant effect on the viscosity of rubber mixes based on the given rubbers.

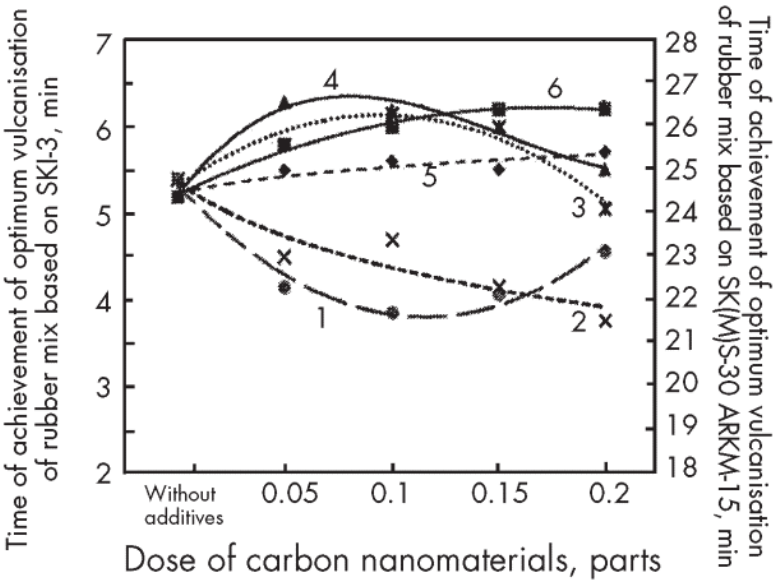
Fig. 12.3 Mooney viscosity of rubber mixes based on Butadiene-acrylonitrile rubbers: 1 – film BNKS-40M; 2 – suspension BNKS-40M; 3 – suspension BNKS-18AN; 4 – film BNKS-18AN; 5 – residue BNKS-40M;




Change in the viscosity of the rubber mixes indicates possible interaction of particles of additives with mocromolecules of the elastomer base. Reduction in the flow resistance of the elastomer composite seems to be due to reduction in the mechanical stresses owing to easier orientation of segments of the macromolecules in the direction of flow and their movement in relation to each other under the action of the nanoadditives. As a result there is an improvement in the processing properties of the rubber mixes, and a reduction in the energy consumption of processing. Less interaction of the "residue" with the polymer base is possibly due to its inhomogeneous composition and structure.

Vulcanisation is a combination of physicochemical processes occurring in the rubber mix, the chief process being the joining (crosslinking) of rubber macromolecules by chemical bonds of different energy and nature in the threedimensional vulcanisation network [6]. In some way, this process is influenced by all ingredients of the rubber mix, and CNMs characterised by large values of the specific surface and by the presence of functional groups can take part in this process, with the formation of both chemical and physical bonds at different stages of the process. Data on the distribution of the time of achievement of the optimum vulcanisation of rubber mixes based on general-purpose rubbers are given in Figure 12.4.

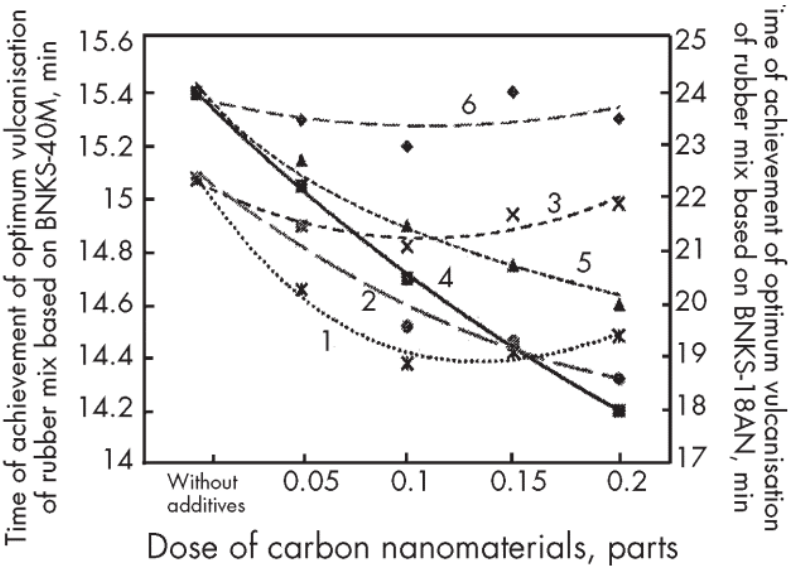
Fig. 12.4 Time of achievement of optimum vulcanisation of rubber mixes based on general-purpose rubbers: 1 – suspension SKMS; 2 – residue SKMS; 3 – film SKMS; 4 – film SKI 5 –residue SKI; 6 – suspension SKI




Study of the vulcanisation kinetics of elastomer composites based on SKl-3 showed that the optimum vulcanisation time increases little. The greatest effect for the "suspension" fraction appears when it is introduced in doses ranging from 0.15 to 0.2 parts. In the case of the introduction of "film* fraction there is an extremal dependence with a maximum point at a content of 0.05 parts. A similar relationship of the effect of the given fraction on the vulcanisation process is observed for rubber SKMS-30ARKM-15. However, the introduction of the "suspension" and "residue" fractions shortens slightfy the time of achievement of the optimum vulcanisation.

Such a nature of the effect of CNMs on the vulcanization kinetics of rubber mixes based on isoprene rubber seems to be due to the ability of nanoadditives to take part in reactionsof decomposition and recombination of radicals of the vulcanising group [7]. The result of interaction is the slowing down of the process of formation of the three-dimensional network. A slightly different nature of the effect of CNMs on elastomer composites based on butadiene-styrene rubber seems to be due to partial deactivation of functional groups on the surface of the "suspension" fraction and to a lesser extent of the "film" fraction by the oil that is contained in the polymer base.

Fig. 12.5 Time of achievement of optimum vulcanisation of rubber mixes based on butadiene-acrylonilrile rubbers: 1 – film BNKS-18AN; 2 – suspension BNKS-18AN; 3 – residue BNKS-18AN; 4 – suspension BNKS-40M; 5 – film BNKS-40M; 6 – residue BNKS-40M




As can be seen from the data presented in Figure 12.5, the nature of the effect of CNMs on the vulcanisation kinetics of rubber mixes based on butadiene-acrylonilrile rubbers of different grades was similar. Thus, the introduction of "film" and "suspension" fractions leads to a reduction in the time of achievement of the optimum degree of vulcanisation with increase in the nanoadditive dose. In the case of the fraction consisting chiefly of carbon nanotubes, by comparison with control specimens, the reduction in time amounted to 17 and 8% for BNKS-18AN and BNKS-40M respectively. When the "residue" fraction is introduced,

no significant change in the vulcanization kinetics is observed. In the given case, most probable is the surface interaction of the nanoadditives not only with the rubber but also with the vulcanising group. The additive particles possibly promoted a more uniform distribution of the components of the vulcanising system within the volume and accordingly a more intensive interaction with the elastomer base.

The behaviour of crosslinked elastomers during deformation is characterised by a predominance of elastic strain over plastic strain. Here, in contrast to other forms of material, the vulcanisate is capable, without failing and without any appreciable residual strains, to withstand an elongation of up to 1000% or more. The physicomechanical properties of the vulcanisates depend to the greatest degree on the rubber, the filler, and the vulcanising system. In the present investigation, use was made of model elastomer composites not containing fillers, and therefore it was urgent to determine the effect of very small additions of CNMs on the structure of the vulcanisation network of the vulcanisates. To assess the interaction of highly dispersed materials with rubber and other ingredients in the process of formation of the three-dimensional network, the nominal tensile strength and the elongation at break were determined, and also the resistance of the specimens to heat ageing.

The causes of change in the properties of the rubbers when they are exposed to temperature are oxidation and the formation of radicals. These processes can be characterised from the degree of change in the physicomechanical properties [8]. Tests of vulcanisates to determine the nominal tensile strength and elongation at break showed that the introduction of CNMs has no significant effect on the physicomechanical properties of the elastomer composites, and they remain at the level reached. Of interest were tests to determine the effect of increased temperatures on the properties of the vulcanisates. Heat ageing was conducted in an oven for 72 h at a temperature of 100°C for the general-purpose rubbers and for 72 h at a temperature of 125°C for the special-purpose rubbers (GOST9.030-74). Table 12.1 gives results concerning change in the elongation at break ($S_\varepsilon$) and in the nominal tensile strength ($S_\sigma$) of vulcanisate specimens after ageing in air.

From the data given in Table 12.1 it can be seen that the introduction of carbon nanomaterials into vulcanisates based on general-purpose rubbers promotes an increase in their heat stability. In vulcanisates based on isoprene rubber, the best combination of properties is possessed by specimens containing CNM of "film" fraction in minimum doses, and of "suspension" fraction in a quantity of 0.15 phr. Elastomer composites based on oil-extended butadiene-styrene rubber are characterised by similar relationships, but the optimum dose of "film" fraction and "suspension* fraction isO.l parts. Here, the greatest effectiveness from introducing carbon nanomaterials into elastomer composites was successfully achieved by using rubber mixes based on polar butadiene-acrylonitrile rubbers. Thus, judging by the change in the physicomechanical characteristics, the temperature resistance increased by more than 25% for specimens based on BNKS-18AN, and 1.2–1.3-fold for specimens based on BNKS-40M.

Table 12.1 Change in elongation at break ($S_\sigma$) and in the normal tensile strength ($S_\varepsilon$)

| Dose of additive (phr) | SKI-3 | | SKMS-30 ARKM-15 | | BNKS-18AN | | BNKS-40M | |
|---|---|---|---|---|---|---|---|---|
| | $S_\sigma$,% | $S_\varepsilon$,% | $S_\sigma$,% | $S_\varepsilon$,% | $S_\sigma$,% | $S_\varepsilon$,% | $S_\sigma$,% | $S_\varepsilon$,% |
| Without additives | 42,9 | -14,6 | -19,1 | -55,6 | -55,0 | -57,7 | -12,9 | -28,9 |
| Film 0,05 | 59,3 | -5,3 | -16,3 | -49,0 | -45,8 | -47,9 | -10,8 | -24,6 |
| Film 0,1 | 58,5 | -7,5 | -13,8 | -43,4 | -36,1 | -36,4 | -8,9 | -18,6 |
| Film 0,15 | 42,5 | -9,3 | -22,4 | -48,5 | -41,2 | -42,5 | -9,9 | -21,9 |
| Film 0,2 | 42,4 | -9,1 | -31,7 | -56,3 | -46,5 | -48,8 | -11,1 | -25,0 |
| Suspension 0,05 | 44,6 | -6,5 | -18,3 | -52,3 | -49,8 | -48,5 | -11,6 | -24,3 |
| Suspension 0,1 | 50,4 | -4,9 | -17,5 | -48,8 | -44,8 | -40,0 | -10,3 | -20,1 |
| Suspension 0,15 | 56,8 | -3,5 | -22,4 | -41,8 | -37,0 | -38,2 | -8,8 | -19,2 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Suspension 0,2 | 49,3 | -1,1 | -26,9 | -35,4 | -40,1 | -36,4 | -9,6 | -18,3 |
| Residue 0,05 | 51,5 | -2,8 | -25,5 | -52,9 | -46,8 | -52,2 | -11,6 | -27,6 |
| Residue 0,1 | 48,9 | -2,7 | -31,8 | -48,8 | -50,0 | -55,6 | -12,3 | -29,4 |
| Residue 0,15 | 49,3 | -2,7 | -25,7 | -53,8 | -46,4 | -54,3 | -11,5 | -28,8 |
| Residue 0,2 | 51,1 | -2,8 | -20,0 | -51,1 | -48,9 | -52,2 | -12,1 | -27,6 |

A possible reason for such an effect of CNMs may be their capacity to act as acceptors of the free radicals formed in the process of ageing of the vulcanisates, and thereby the slowing down of the process itself [9]. Bearing in mind that nanoparticles can interact with components of the vulcanising system, it can be assumed that they promote a reduction in polysulphidity and the formation of stronger chemical bonds in the three-dimensional network of the vulcanisates [10].

Thus, the use of CNM as a modifying additive is most expedient in rubber mixes based on butadiene-acrylonitrile rubbers, as it leads to an improvement in the processing properties of the elastomer composites.

**REFERENCES**


1. P.A. Vityaz' and K.A. Solntsev, Technology of Structural Nanostructural Materials and Coatings. Belaruskaya Navuka, Minsk, 282 pp. (2011).
2. Yu.F. Shutilin, Reference Book on the Properties and Application of Elastomers. Monograph. Voronezh State Technological Academy, Voronezh. 871 pp. (2003).
3. F.A. Makhlis and D.L. Fedyukin, Technological Handbook on Rubber. Khimiya, Moscow, 400 pp. 11989|.
4. I.M. Tsypkina and A.P. Voznyakovskii, Kauch. I Rezina, (1): 10 (2003).
5. V.M Goncharov and D.V. Ershov, Kauch. I Rezina, (1): 16 (2007).
6. A.E. Kornev etal.. Technology of Elastomer Composites. College Textbook. Eksim, Moscow, 288 pp. (2000).
7. Z.V. Priss and M.S. Fel'dshman, Kauch. I Rezina, (12):21 (1977).
8. D.L. Fedyukin and F.A. Makhlis, Mechanical and Processing Properties of Rubbers. Khimiya, Moscow. 240 pp. (1985).
9. A.P. Voznyakovskii etal., Kauch. Rezina, (1):6 (1998).
10. A.I. Lyamkin et al., Kauch. i Rezina, (5): 16 (2005).

## 13. Paintwork materials modified with carbon nanoparticles

Nanotechnologies and their introduction into technological processes is one of the preferred direction of the many countries development. Nanotechnological development and related branch of science and engineering connecting with creation, investigation and using objects with nanosize elements will lead to profound changes in many spheres of human activity – in material science, power-engineering, electronic engineering, information science, machine industry, medicine, agricultural industry, ecology [1].

Obviously, it is connected with possibility of modification with nanomaterials of polymer matrixes differented by their chemical structure and application.

In fact, nanoparticles can add innovation properties to usual polymer materials (antibacterial, self-cleanable, electroconductive, optical and etc). The modification with nanosize particles is interesting due to the presence of such fine fillers in polymeric matrix composite, which is leading to polymer structural change because of macromolecules orientation in border layer. As reducing particle size filler leads to area per unit volume increasing that is why the whole polymer pass into border layer condition and composite assumes another properties in comparison with unfilled polymer materials.

This chapter is devoted to study the influence of nanosize additives on properties of paintwork composite materials and coatings on their base.

### 13.1 Nanotechnologies in paintwork materials

For a long time nanomaterials are using in paintwork compositions. First of all this is natural materials, such as talc, mica, asbestos, products on the base of clay (bentonite), which can be used as fillers and functional additives. Soluble glass, partially hydrolyzed ethyl orthosilicate (ethyl silicate) and microgel binders are used as film-forming material [1]. Now nanotechnologies give to the paintwork industry a wide range of new possibilities and play a great role in development of modern paints and varnishes. Deriving special nanostructures and using nanoparticles in paints and coatings allows to producers create product with significantly improve or new properties. For example:

– self-cleanable coating (with «lotus effect»);
– antibacterial paints;
– mar-proof coatings;
– photocatalytic paints;
– transparent coatings with improved protection from UV rays;
– coatings able to screen magneto-electric emission;

In the long term, It is planned to create such unique materials as:

– self-recovering coatings;
– inhibitive primer without содержащие chrome;
– electroconductive paints;
– coating with changeable reflecting ability;
– heat-insulating paintwork materials;
– paints for solarcell array [2].

### 13.2 Types, properties and produce methods of nanomaterials.

Now the produce methods and properties of nanomaterials consisting of building block, which sizes are ten or one nanometers, have been investigating intensively. A scale of this value fits to proteins molecule or smaller but bigger than sizes and masses of simple organic molecules. In area of such size the rules, which are working for solid body and are described with band structure model do not function. However, rules and phenomena that are typical for molecule are not applicable.

For another thing, nanomaterials due to their sizes have comparatively high quantity of surface atoms possessing special properties, which are not same as solid body properties. It make possible to obtain the coatings keeping such properties as flexibility and high hardness, to produce materials, which can be stable to the majority of corrosion environment and can have a high hydrophobic properties. At the same time, it is possible to control distribution homogeneousness of organic and nonorganic particles on the molecule level.

Nonorganic components of such materials are nanopowder of metals, silicon dioxide, barium sulphate, aluminium oxide, zirconium. Polymer powders can be used as fillers distributing into film binder. Obtaining in the last ten years nanomaterials on the carbon base such as fullerene, which consist of 60 ($C_{60}$) and also 70, 84, 96 carbon atoms, explosion nanocarbon, nanodiamond, nanotube are of great interest [3–7].

Fullerene in distinction from other carbon form soluble in organic solvent and form true solution. Coatings with new properties, for example, more lightweight can be produce on the base of fullerene substituting fillers with particles size 50 – 100 mkm.

### 13.3. Explosion carbon

Explosion carbon is a product obtaining by explosive synthase method. The substance consist of small diamond particles with roentgen size 4 – 5 nm and spherical graphite particles with the same size. Usually explosion carbon has 30 – 80% diamond particles. Diamond and graphite parts in explosion carbon connect with each other in some way, as it is impossible to divide them by natural approach even with intensive hypersonic affect and gravitational field. Explosion carbon is used in industry in oiling-reconstructive composition for combustion engine showing unusual properties in tribotechnical systems. It can be used in other application fields.

Explosion nanodiamand is a product making from explosion carbon by selective oxidation of graphite part. Now it is known that explosion nanodiamand intensively interact with environment showing various and not always comprehensible properties including biological systems by the reason of big quantities of surface atoms. Explosion nanodiamand is used as an additive in reinforced electrodeposited coatings and polishing agent, semiconducting and optical material, primers [8].

Nanotubes [9, 10] is a material consisting of total carbon tubes with diameter into several atomic diameter (as multilayer as monolayer) and length until millimeter unit. Nanofibers, which are characterized some dissimilarity in length and diameter are similar in properties with nanotubes. Nanotubes can be produced by several methods such as catalytical growth in hydrogen atmosphere with methane addition or carbon oxide disproportionation under the pressure. Nanotubes and nanofibers find application in electrochemistry and carbon composite materials production.

A traditional carbon material carbon black using in paintwork industry has in their composition nanoscale structures. However overall level of investigation in this area is insufficient in spite of large production volume and numerous applications.

Investigations connected with development of new type of coatings on the base of nanocarbon composite materials are carried out intensively. It is supposed that such coatings have a low capillary tension and hydrophobic properties. It was determined that they have increased strength properties because Young's modulus along the nanotube axis exceeds the same modulus of monocrystal diamante. It is known that electrodeposited coating using combined metal deposition and nanodiamant particles are exceed in several times traditional electrodeposited coatings (ones).

Such fillers can be used in combination with acrylic, polyurethane, ethoxyline and other form binders in the form of organic solvent and water dispersion.

The nanoparticle production method are different from usual technological production processes of dispersed system by technical approach. The methods are divided on to groups:

– Dispergating basing on size reduction, disintegration («top-down»),

– Condensation with help of which nanostructures are obtained from smallest initial components be most exact from atoms and molecules («bottom-up»). [1]

One of the great problem fine-dispersed particles introduction in polymer matrix is connect with flocculating phenomenon and forming large coagulative aggregates with sizes 6 – 12 mkm. It is possible to elude such undesirable processes by special additives [11, 12] selection and regulation of introduction process.

## 13.4 The influence of the introduction method of nanoparticles in solvent borne coating materials

The series of experiments were carried out in Belarusian State Technological University. The results show influence of the methods of introduction carbon nanoparticles on the properties solvent borne alkyl and epoxy primers, physical and mechanical (hardness, adhesion, impact strength) and protective properties (resistance to static action of water and sodium chloride) of the coatings on their basis.

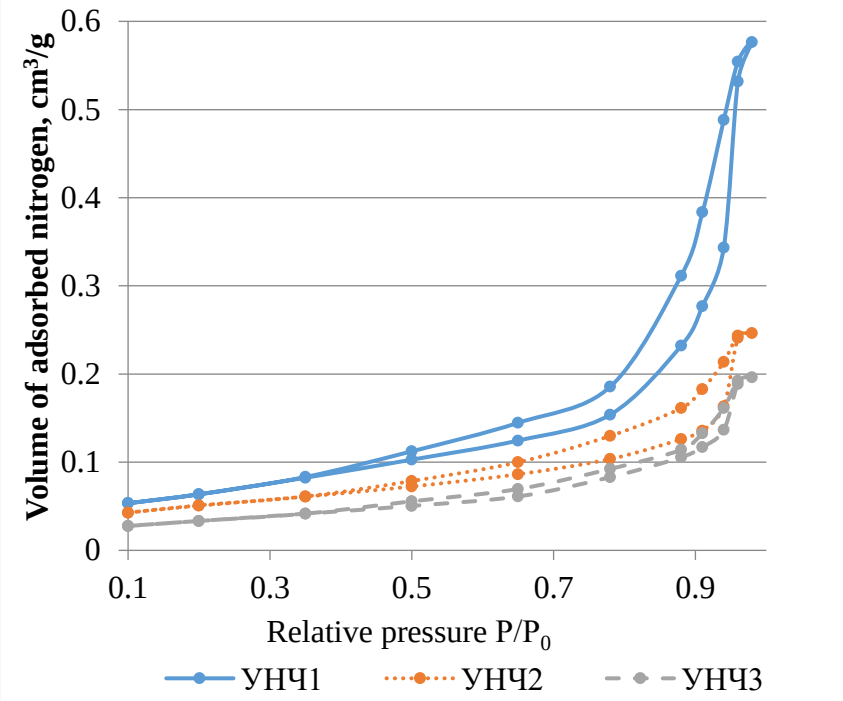
The objects of investigation were carbon nanoparticles obtained of following condensate methods:

– CNP1 – carbon nanoparticles obtained catalytic synthesis in the fluidized bed;

– CNP 2 – unactivated carbon nanoparticles obtained in in plasma of high-voltage discharge;

– CNP 3 – activated by oxygen-containing groups carbon nanoparticles obtained in plasma of high-voltage discharge.

The surface area of investigated CNP was estimated by low-temperature method BET (Brunauer, Emmett, Teller) on the instrument NOVA 2200 [13].




| Index name | CNP 1 | CNP 2 | CNP 3 |
|---|---|---|---|
| Surface area, $m^2$/ g | 234 | 174 | 121 |

Figure 13.1 – The determination of CNP surface area

The biggest surface area was determined for CNP1 obtained catalytic synthesis in the fluidized bed, the smallest – for CNP3. This parameter describes the value of nanotubes surface area, which will deal the amount of possible absorbed molecules of film-forming agent, solvents and other functional additives, so influence on coating structure changes.

*The CNP introduction method.*

CNP were added in the pigmented coating materials by the previous dispersing during 15 minutes in solvent and condensation agent in box of ultrasonic bath - Bandeline Sonorex (operating frequency 40 kHz). Then CNP were mixed with coating material by milling during 20 minutes in a laboratory dissolver DISPERMAT®CA (rate of mixing about of 2000 Rev/min) [14, 15].

CNP were added at primer in a form of suspensions in a solvent or condensation agent. CNP were inputted in the amounts of 0,005, 0,010% 0,100% of the mass of paint material taking into account the nonvolatile content.

*Testing methods.* Properties of coating materials were determined according to standard procedures. The degree of grinding was determined according to GOST 6589 (ISO 1524), mass fraction of solids GOST 17537 (ISO 3251). Coating materials were applied using a pneumatic on standard substrates: glass plates of special purpose (size 100×100×5 mm (ISO 1514)); steel plates with a size of 70×150 mm and thickness of 0.8–1.0 mm. The coating formation was carried out under natural conditions at a temperature of (20 ± 2)°C. The physical and mechanical properties of coatings were determined after 2 days after application: hardness by pendulum device type TML, the Kennings pendulum (GOST 5233, ISO 1522), impact strength (GOST 4765, the diameter of the striker 8 mm), the adhesion method of lattice cuts on a 6-point scale (GOST 15140, ISO 2409). The resistance to static action of corrosion environments at a temperature of (20 ± 2)°C was investigated after 10 days according to GOST 9.403 (ISO 2812). Tests were carried out to visible changes of appearance of the investigated coatings.

*The investigation of influence of the introduction method of CNP on properties alkyl primers based coatings.* CNP were added into the alkyl primer, which represents as a suspension of pigments and fillers in alkyd-styrene film-forming material with adding functional additives and solvents. Table 13.1 presents the properties of the unmodified alkyd primer and physical-mechanical, and protective properties of coatings based on it.

Table13.1 – The properties of alkyd primer and coatings based on it

| The name of parameter | The value of parameter |
|---|---|
| Color of coating | Grey |
| Degree of grinding, mcm, not more | 30 |
| Mass fraction of nonvolatile substances, %, not less | 53,0 |
| Hardness by pendulum device type TML (Kennings), not less<br>– rel. un.<br>– sec | <br>0,38<br>95 |
| he adhesion method of lattice cuts, grade, no more | 0 |
| Impact strength, cm, not less | 60 |
| Resistance to static action, days, not less than<br>– water<br>– 0,5% NaCl solution<br>– 3% NaCl solution | <br>4<br>4<br>4 |

The primer is characterized by good physical-mechanical and protective properties. CNP were added at alkyd primer by three ways according the table 13.2.

Table 13.2 – The methods of CNP introduction into alkyd primer

| Designation of method | Dispersion phase | Stage of introduction into coating material |
|---|---|---|
| Method 1A | o-xylene | Before application |
| Method 2A | polar solvents | |
| Method 3A | polar solvent | On the stage of dispersing during primer obtaining |

*Method 1A.* The properties of alkyd primer coatings modified before application with sustention CNP in o-xylene (method 1A) are shown in the table 13.3.

## 14. Plasma-chemical synthesis of nanocomposite polymer coatings

It is found that electron-beam evaporation of sulfur and zinc powder mixture has a number of kinetic features. The melting of sulfur is observed at the first step of electron beam influence on the mixture of the powders in the crucible after some breakdown time and it is accompanied by its intense transition into the gas phase (Fig. 14.1, stage I). The intense growth of deposition speed is stated at this stage (Fig. 14.1, stage II). Further (Fig. 14.1, stage III), when boiling is finished in the area adjacent to the crucible, we observe intense pink glow. At this stage, according to the results of the X-ray structure analysis of the crucible material, there is no unbound sulfur. The deposition of the coating on the substrate is made with persistent glow. At stage III, the deposition speed considerably decreases, stabilizes and on the average it is around 0.12 nm/s.

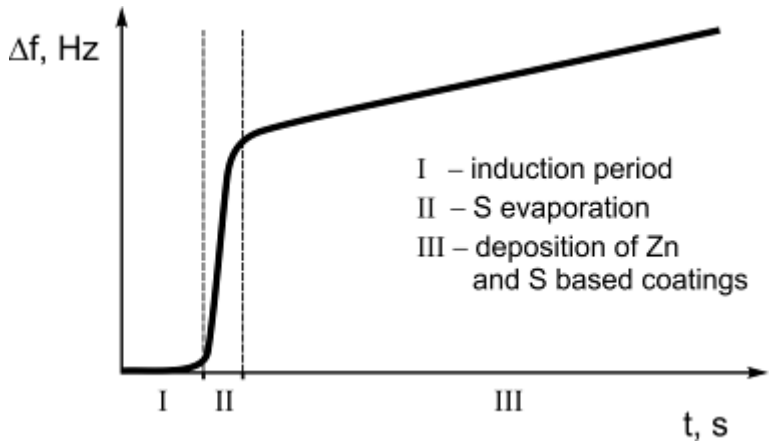
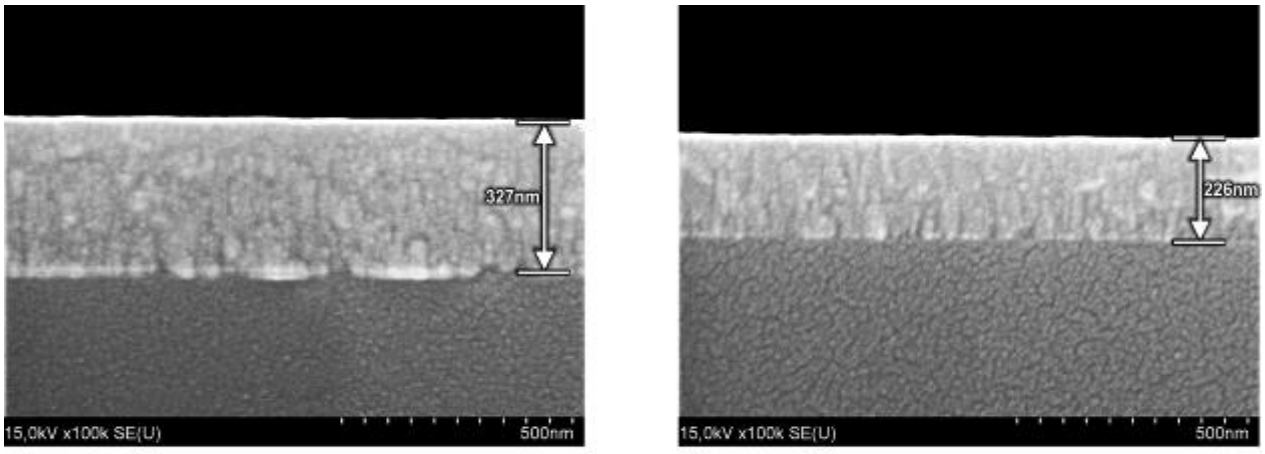



Fig. 14.1. Evaporation kinetics of sulfur and zinc powders mixture

The electron heating of powders mixture with subsequent melting of sulfur can initiate the formation of ZnS chemical compound. The energy released in chemical interaction of zinc and sulfur, is sufficient to convert generating ZnS compound into the gas phase with subsequent formation of the coating. Thus, in our opinion, the formation of zinc sulfide doesn't happen on the substrate, as for the majority of vacuum deposition methods [5, 67], but during the formation of the gas phase.

It is established that the introduced method allows forming quite continuous, densely packed and pinhole-free coatings which have the grain texture in the boundary layers near the substrate (Fig. 14.2).




(a) (b)

Fig. 14.2. Cross-section SEM images of ZnS coatings, thickness of 330 nm (a) and 230 nm (b)

The analysis of absorption spectra of the formed coatings shows that the substrate temperature does not have any considerable influence on the value of the band gap (Fig. 14.3), the value of which is determined by the chemical composition, grains size that form the coating [5, 11, 68]. For the coatings with 330 and 230 nm thickness, the value of *Eg* is 3.58 and 3.54 eV respectively. It should be mentioned that when the coatings are formed by thermal evaporation of zinc sulfide in vacuum, or by chemical bath deposition technique using an aqueous solution and other methods, we observe, as a rule, the high influence of thermal treatment on the width of the forbidden area. It can be explained by the formation of the defect structure and its change at heating, inhomogeneous distribution of chemical elements along the thickness of the layer [5, 69, 70].

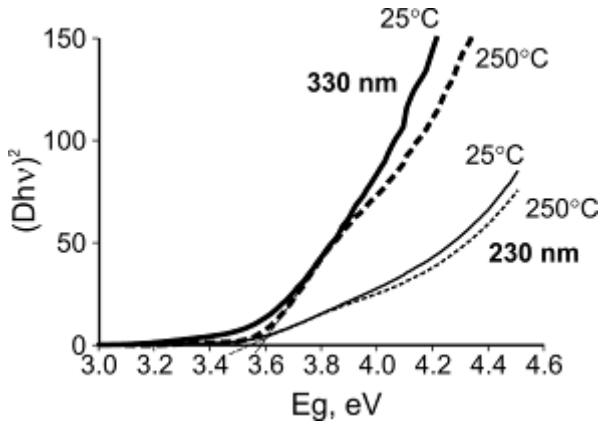



Fig. 14.3. $(\alpha h\nu)^2$ versus photon energy *hν* plots of different thickness coatings formed at the temperature of 25 and 250 °C

The absence of the founded significant influence of the substrate temperature on *Eg* is, apparently, a consequence of the formation of zinc sulfide particles out of it. Otherwise the diffusion processes defined by the substrate temperature would lead to a significant change in the structure of the deposited layer, and consequently in *Eg* value. Moreover, the non-typical fact experimentally established by us, proving the increase in *Eg* band gap width observed with the growth in coating thickness, doesn't match in some way the results previously obtained by the authors of ZnS vacuum coating [5, 67, 71]. In our case, it is due to the primary growth of ZnS particles in zone of gas phase formation, and not on the substrate. In the initial growth stages of ZnS coating, due to the influence of the substrate, a boundary layer with more defects is formed with a clearly defined grain texture (Fig. 14.2) and, accordingly, with a smaller *Eg* band gap width. As the coating thickness increases, the contribution to *Eg* value of this boundary layer decreases.

The distribution of chemical elements along the thickness of the layer depending on the substrate temperature determined by AES is showed in Fig. 14.4. At the deposition on the low temperaturesubstrate, the character of chemical elements distribution along the thickness of the deposited layer is almost homogeneous and corresponds to ZnS composition (atomic proportion of sulfur and zinc is 1:1). The temperature of the substrate rises up to 250 °C, as shown in Fig. 14.4b, and it leads to some changes of stoichiometry what can be explained by more intensive desorption of sulfur atoms.

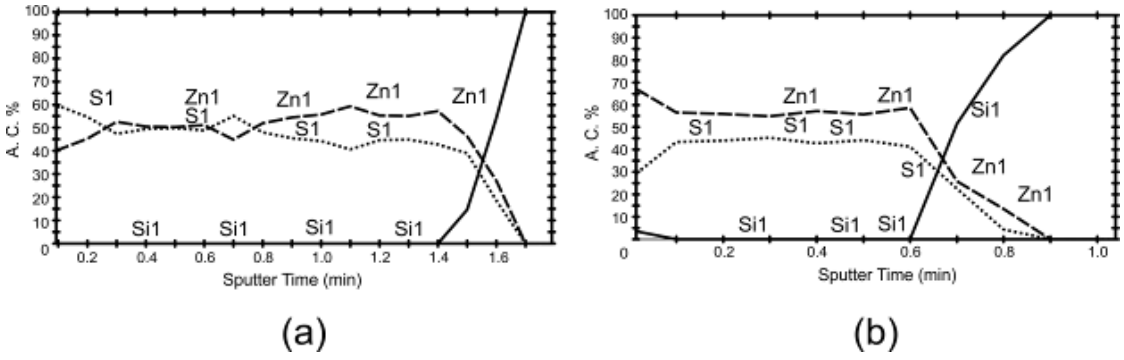



Fig. 14.4. Distribution of zinc and sulfur along the coating thickness formed at 25 °C (a), 250 °C (b) temperature of the substrate

At the same time, the temperature of the substrate and the width of the coating have noticeable influence on their morphology (Fig. 14.5).

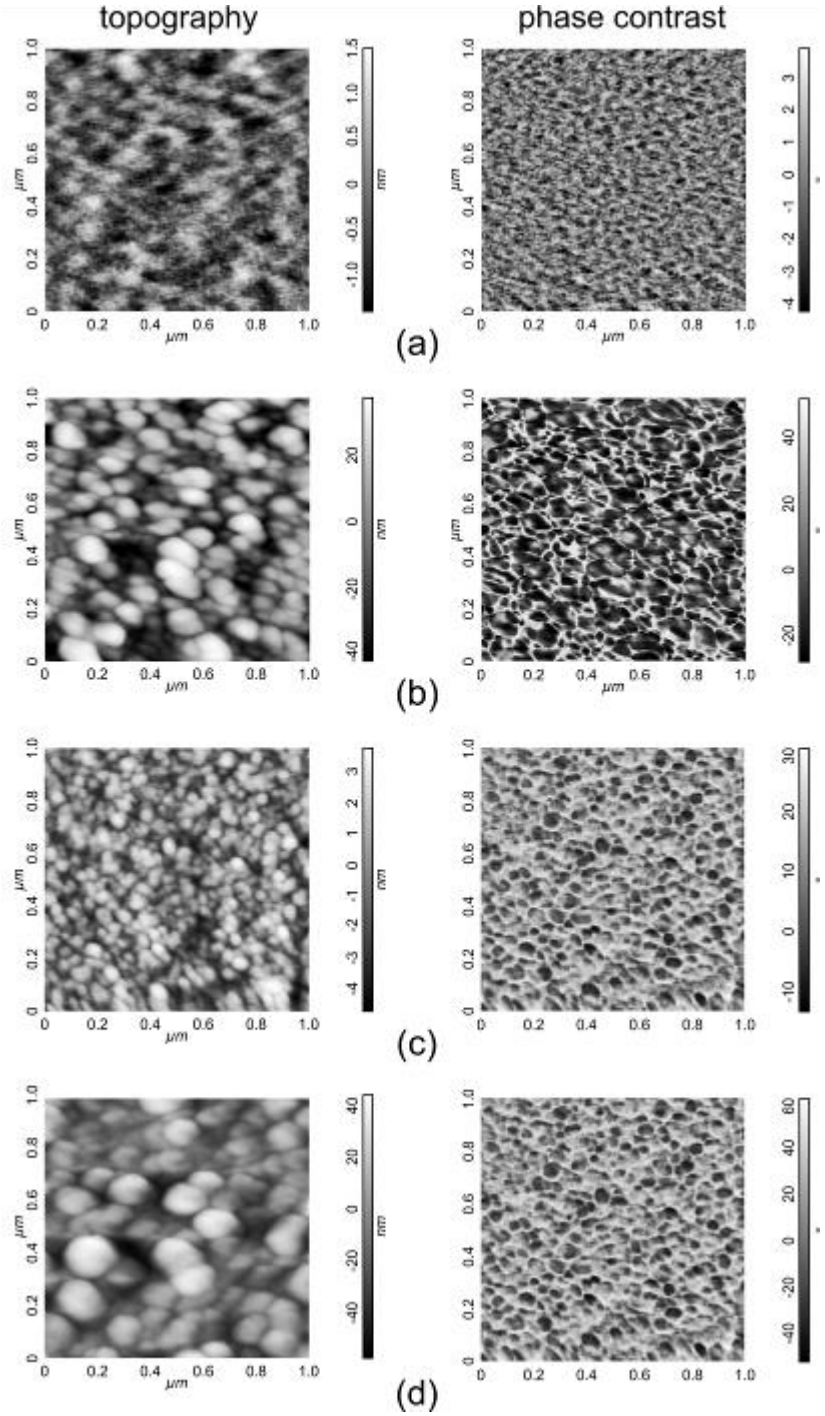



Fig. 14.5. AFM topography (first column) and phase contrast (second column) of a ZnS coatings: a) $T_{substrate}$ =25°C, 330 nm, b) $T_{substrate}$ =250°C, 330 nm, c) $T_{substrate}$ =25°C, 230 nm, d) $T_{substrate}$ =250°C, 230 nm. Scan size 1 μm × 1 μm.

The coating with 330 nm width formed at 25◦C temperature of the substrate is characterized by very flat surface: the level difference is not more than 3 nm. To estimate any discrete surface compositions forming the coating is impossible. The heating of the substrate during the deposition process initiates the formation of large grains with average size more than 70 nm. The height of crystals grains is more than 60 nm in such layers. The coating of 230 nm thickness formed at the temperature of 25◦C as well as the layer of 330 nm thickness have atomic flat surface (the grain size is not more than 8 nm). But in comparison with the coating of 330 nm

thickness, the surface of such layers has structural formations with the average size of more than 30 nm on the substrate surface.

The comparison of AFM data and the band gap width of ZnS coatings with different thicknesses allow stating the fact known from the papers of other authors [71] about the increase of crystals in size and correspondingly the decrease of band gap width in value. The comparison of these data at the different temperatures indicates the mutual compensating effect of ZnS particles growth, on the one hand, and some changes in chemical composition (Fig. 14.4), on the other hand, as well as the formation of less defective structure [5, 71]. It should be noted that the objective of the work was not to study the interactions of the coatings properties with their structure and composition. It was necessary to show the possibility to form ZnS coating via low-energy electron flow.

The intense evaporation of sulfur by electron beam is a highly undesirable process: the performance of high-vacuum pumping system decreases; the use of shutters is needed. The search of technological methods reducing the duration of the II period dispersing of the sulfur and zinc powders mixture is a prerequisite for the practical realization of the method.

The possible reactions in evaporation area allow us to suppose that the use of dispersed copper as an additional component of the mechanical mixture of zinc and sulfur is aimed to influence in active manner the selectivity of evaporation, chemical composition of the deposited coating, and, as the result, their properties, in particular, the value of the band gap. This effect is due to the fact that the heating of copper with sulfur in vacuum leads to the formation of copper sulfide, which, compared with other sulfides, is difficult to be converted into the gas phase by low energy electron flow [72]. According to the studies, the electron beam evaporation of three-compound mixture is indeed accompanied by a decrease induration of II stage on the kinetic curve. This can be explained by chemical interaction of sulfur with copper and the formation of copper sulfide.

With the method of X-ray spectral microanalysis (EDAX) it is established that the injection of copper into the crucible leads to the noticeable rise of sulfur content in the coating in comparison with the coating formed without copper. This can be explained by the fact that with the presence of copper in the crucible, the part of sulfur passes to the bound state that prevents its intensive transition into the gas phase at stage II. Therefore, at the beginning of stage III of the dispersing process, more sulfur contains in the target. The copper content increased in the target can increase the proportion of bound sulfur in the form of copper sulfide which is not able to pass into the gas phase. This can be confirmed by the fact that the increasing of copper content > 20% results in a drastic decreasing of coating deposition rate almost to zero. Therefore, the concentration of sulfur in the coating is maximal at considerably low content of copper in the mixture (Table 14.1).

Table 14.1 The chemical composition and values of *E*g of the formed coatings

| Copper content in the mixture of zinc and sulfur, % | Content of elements in the coating, at.% | | $E_g$ |
|---|---|---|---|
| | Zn | S | |
| 5 | 37.74 | 62.26 | 3.60 |
| 7 | 37.90 | 62.10 | |
| 10 | 40.77 | 59.23 | 3.54 |
| 15 | 46.63 | 53.37 | |

The increase of copper fraction in the target is accompanied by a decrease in *Eg* value, what may be the effect of the formation of more defective structure (Fig. 14.6).

Thus, the introduction of copper to the target is not an effective technique to reduce the duration of stage II of electron beam dispersing process, as it reduces the deposition rate, increases

unbound sulfur fraction in the flow of zinc sulfide particles, which leads to the formation of defects in thin layer structure.

During making and carrying out of experiments, an interesting phenomenon was observed. After being vacuumized ($10^{-3}$Pa) for 30-40 min, the mechanical mixture of copper and sulfur powders (molar fraction of compounds is 1:1) changes color from yellow-brown to blue-black. The powder, formed after vacuumization, was studied using XRD analysis (Fig. 14.7). The registered reflections of the X-ray picture definitely indicate the presence of copper sulfide [73–75].

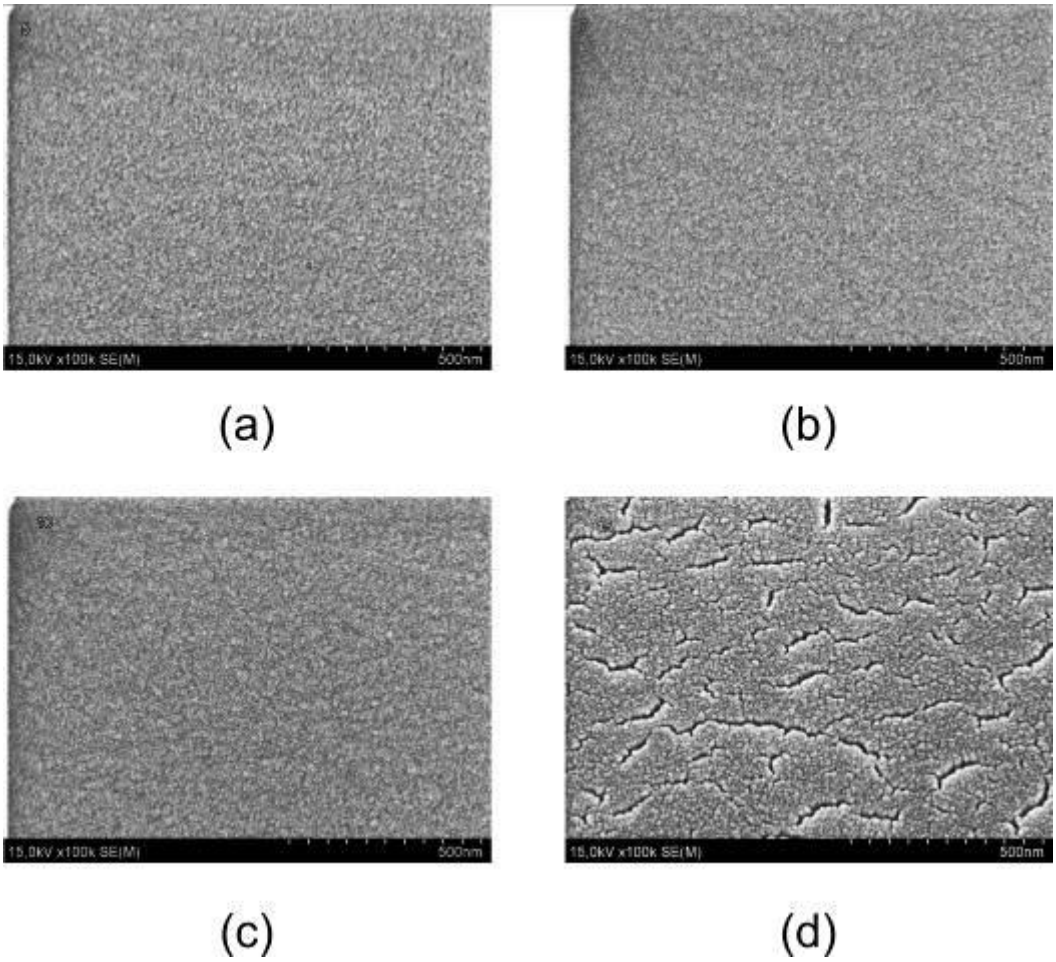



Fig. 14.6. SEM images of the coatings formed with 5% (a), 7% (b), 10% (c) and 15% (d) copper in the crucible

The use of mechanical mixture of zinc powder and the powder obtained by vacuuming of sulfur and copper powders mixture, allows eliminating the unwanted sulfur intensive evaporation by electron beam (stage II). At the same time, it does not affect the rate of coatings deposition, as in the case of dispersed copper use. The flow of electrons initiates the reactions of zinc sulfide formation easily passing into the gas phase due to the energy released during the reaction. The adding of the copper powder to the sulfur copper and zinc powders mixture similar to the use of zinc and copper sulfide powders mixture allows depositing the coatings in which the copper is absent (Fig. 14.8). The copper remains in the target in the form of sulfur compound.

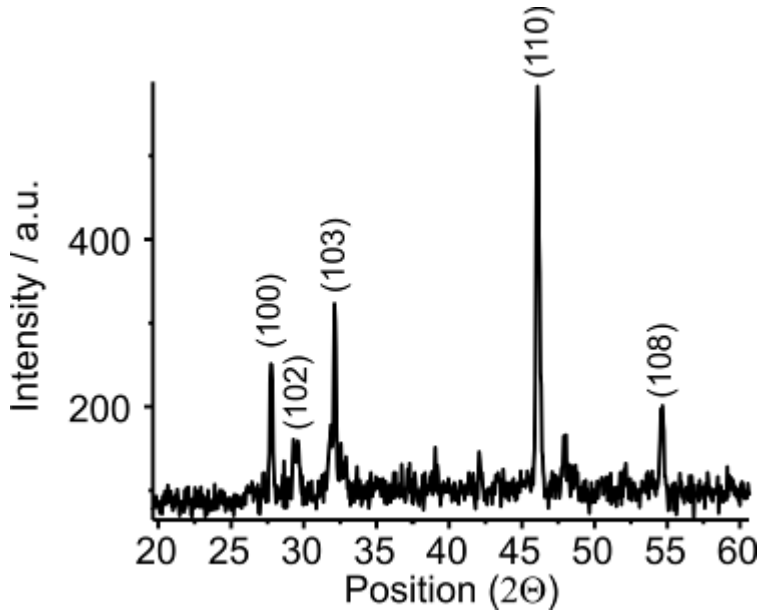



Fig. 14.7. XRD pattern of the mechanical mixture of copper and sulfur powders after vacuumizing

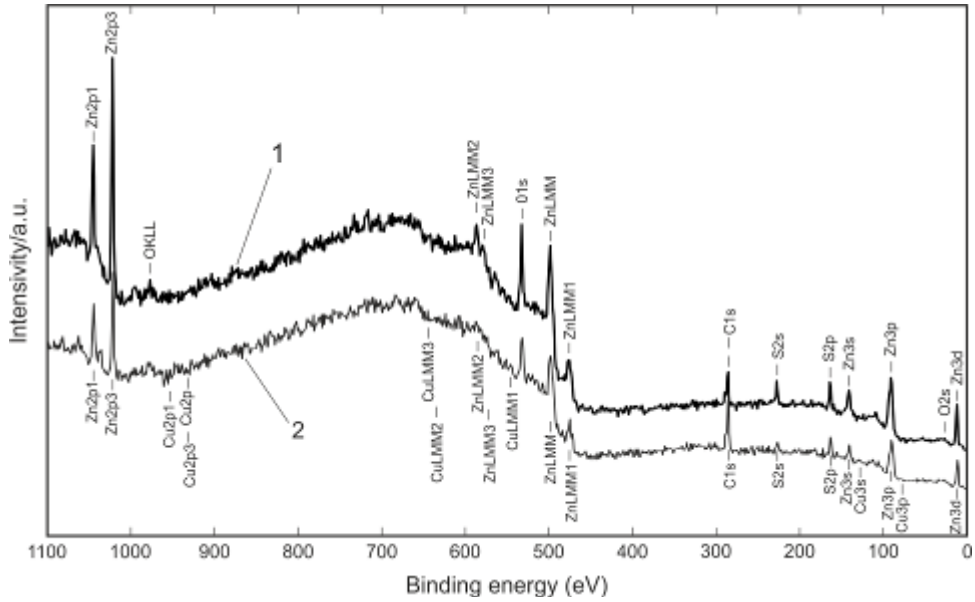



Fig. 14.8. XPS spectra of the coating formed by the evaporation of the mixture of zinc and sulfur powders (1), the mixture of $Cu_xS_y$ powders and zinc (2)

The TEM data show that when copper sulfide is used, as well as in the case of the formation of coatings with zinc and sulfur powders, the coatings form crystals, the size of which considerably varies (Fig. 14.9). The average size of ZnS crystals, formed with zinc and sulfur powders, is not more than 25 nm. This result confirms the correctness of the conclusions made in UV–vis analysis and AFM research results. When the powders of copper sulfide and zinc are used as a target, the size of crystalline grains is considerably bigger than the size of crystalline grains which form the coating deposited with zinc and sulfur powders. And their shape is more expressive.

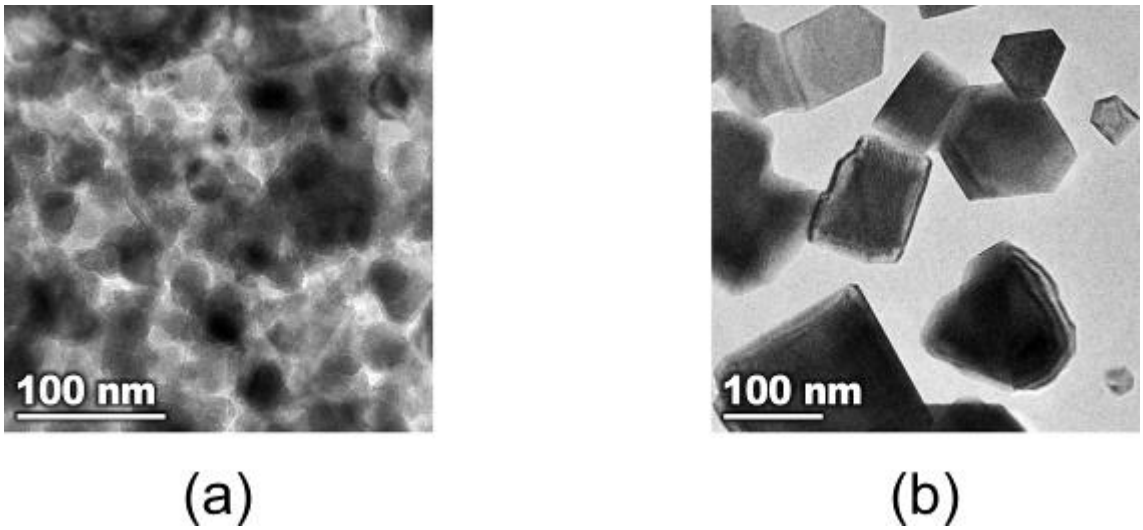



Fig. 14.9. TEM images of the coating formed with Zn and S powders (a) and $Cu_xS_y$ powders and Zn (b) used as targets

## 14.1. Structural properties of zinc oxide coatings doped with copper

The process of electron beam dispersion of targets based on nitrates may be conventionally divided into 3 periods; (I) no visible changes in the readout of the QCM sensor on the background of intense gas release in the vacuum chamber, (II) beginning of the coating deposition process and (III) intense condensation of the thin layer from the gas phase.

The impact of the low-energy electron beam on the zinc nitrate powder is accompanied by its rapid melting in crystallization water at initial stage. The deposition is virtually absent, despite the high pressure in a vacuum chamber, which is above $10^{-1}$ Pa (Fig.14.10, a, stage I). Thus, when heated, the salt intensively loses water molecules that pass into the gas phase and are removed with the vacuum pumping system.

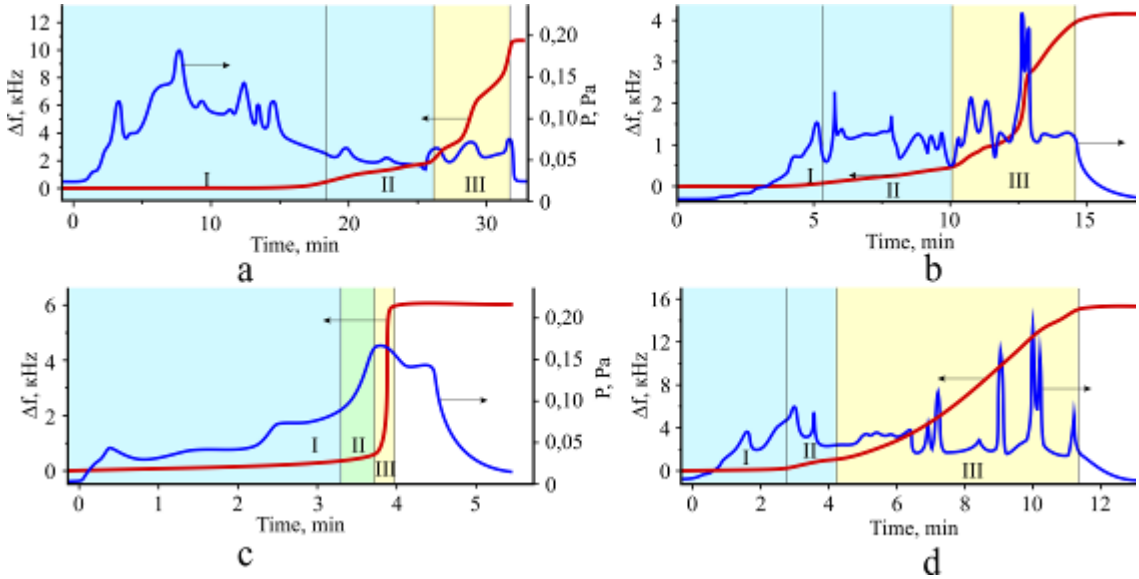



Fig. 14.10. Kinetic dependence of the low-energy electron-beam dispersion of Zn $(NO_3)_2 \cdot 6H_2O$ (a), Zn + $Zn(NO_3)_2 \cdot 6H_2O$ (1:1) (b), $Cu(NO_3)_2 \cdot 3H_2O$ (c), Zn + $Zn(NO_3)_2 \cdot 6H2O$ + $Cu(NO_3)_2 \cdot 3H_2O$ (1:1) (d) targets

It has been observed that the electron beam causes to increase in viscosity of melt and its color changes from transparent to light-yellow. At this dispersion stage, the beginning of the

coating deposition process is registered with a noticeable pressure reduction in a vacuum chamber (Fig. 14.10, a, stage II).

Fig.14.11 shows the IR spectra of coating obtained with substrate temperature of 25 °C. The double absorption band in the frequency ranges from 1250 to 1500 $cm^{-1}$ is an evidence of ($-NO_3$) groups stretching vibrations [76]. Thus, the period of intensive salt melting in its crystallization of water is followed by a period of its transition into the gas phase. Thus, the effect of electron flow on salt decomposition is actively realized with formation of an oxide (change in target color from transparent to light yellow). The comparative analysis of the IR spectra of the coatings and the zinc nitrate hexahydrate powder indicates the coincidence in their chemical composition (Fig. 14.11).

At the final stage, an intense electrical discharge appears in the target area. At that moment, the QCM detects an intensive growth in deposition rate (Fig. 14.10, a, stage III). A pale yellow powder was formed and the gas phase generation was stopped in certain period of time. The pressure in a vacuum chamber restores to its initial value prior to deposition. According to the X-ray analysis data, the light-yellow powder generated in the target by the electron beam impact is zinc oxide. The IR-spectrum of the thin layer formed at stage III corresponding to IR-spectra of zinc salt, which indicates the presence of a significant proportion of salt in the gas phase formed at this stage.

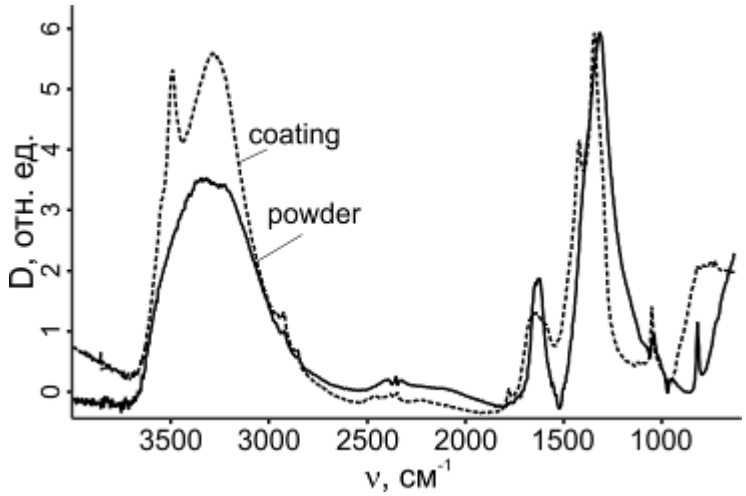



Fig. 14.11. IR spectra of coating obtained with substrate temperature of 25 °C and initial powder

The significant changes in kinetics of volatile substances are observed when zinc dust is used for coating (Fig. 14.10, b). The use of the mixture of zinc dust and its salt as a target makes it possible to exclude melting of salt in its crystallizing water virtually completely, as well as to decrease the duration of period I significantly. It should be noted that almost immediately after the electron flow impact on the target, an intensive plasma glow is detected near the evaporation zone.

In the vacuum chamber, the electron-beam exposure on copper nitrate is followed by a relatively monotonous increase in pressure and by coating deposition at high speed (Fig. 14.10, c). Under the influence of low-energy electron flow, salt melted quickly without noticeable changes in QCM (Fig. 14.10, c, stage I). However, the length of this period is significantly less as compared to the same period for the powder of zinc nitrate. The color of melt in the target was turning brown, which indicate the partial decomposition of copper nitrate in it. At the same time, in the process of evaporation of copper nitrate the plasma discharge was not registered.

In contrast, the dispersion process of the ternary mixture (Zn + $Zn(NO_3)_2 \cdot 6H_2O$ + $Cu(NO_3)_2 \cdot 3H_2O$) was characterized by the appearance of plasma discharges. Thus, the occurrence of plasma discharges is caused first of all by processes of zinc nitrate decomposition. The ternary mixture dispersion process is characterized by the highest decrease in the length of period I (Fig. 14.10, d).

One should emphasize that a plasma discharge occurs against dazzling white luminescence of particular target portions. This may indicate their high heating temperature. The low-energy electron-beam evaporator used in this work is designed exclusively for dispersion of polymeric materials. It is neither capable to melt the metal nor even heat it to any significant temperature.

The kinetics of dispersing composite targets is determined by chemical interactions between their components. The chemical reactions take place not only by the electron beam impact, but also at the stage of targets making. Thus, in the process of mixing powders of zinc nitrate and zinc dust, the intensive mixture heating was observed by release of brown gas. The reaction may be written as follows:

$Zn(NO_3)_2 \cdot 6H_2O + 5Zn = Zn(NO_2)_2 + 2ZnO + 3Zn$ (residue) $+ 6H_2O\uparrow + Q$

where Q is moderate heating. As a side reaction, the reaction of nitrogen (IV) oxide (brown gas) should be noted:

$Zn(NO_3)_2 \cdot 6H_2O + Zn = 2ZnO + 2NO_2\uparrow + 6H_2O\uparrow + Q$

Thus, the electron beam initiates the following basic interaction:

$Zn(NO_2)_2 + 2ZnO + 3Zn = 6ZnO + N_2\uparrow + Q$ (intensive heating)

The plasma discharge occurs in a nitrogen atmosphere. An indirect confirmation of that is the red-pink color of the plasma glow.

It is well known that the reaction of thermal destruction of anhydrous copper nitrate may be written as follows:

$2Cu(NO_3)_2 = 2CuO + 4NO_2 + O_2$.

When using hydrate, the detachment of water molecules precedes the destruction of salt molecules and their transition into the gas phase. The water detachment process realizes at the stage of the monotonic increase in pressure without any noticeable changes in QCM (Fig. 14.10, stage I & II).

All the above chemical interactions may be realized in three-component mixtures under the impact of electron flow.

Chemical analysis of the target material after dispersion showed the presence of copper and zinc in oxidized form.

The results of TEM analysis of the coating formed by $Zn(NO_3)_2 \cdot 6H_2O$ electron-beam dispersion at stage III are shown in Fig. 14.12. The coating was formed at the final stage of period III. This was done in order to limit the impact of zinc nitrate on the results of TEM investigations.

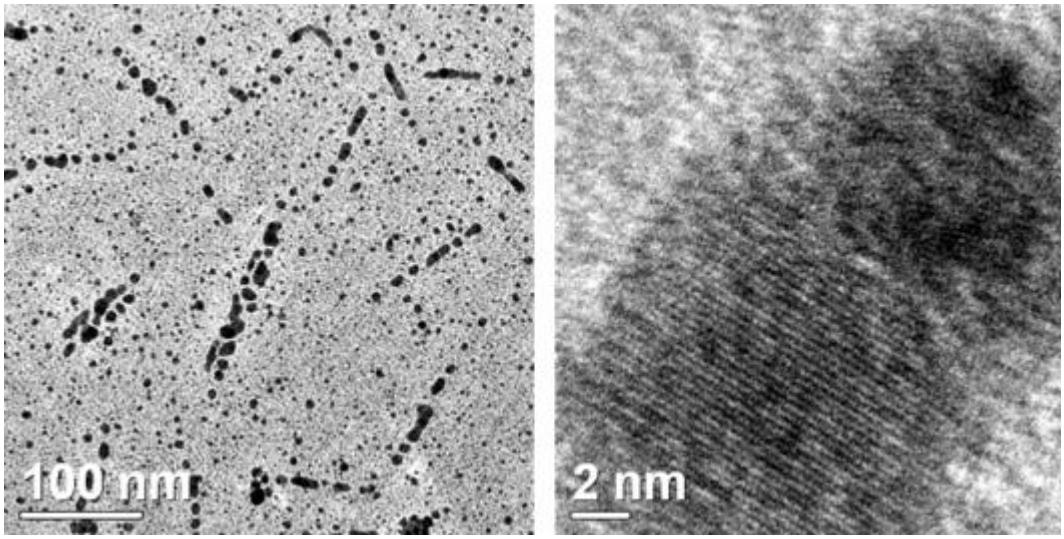



Fig. 14.12. TEM images of the coating formed by low-energy electron beam dispersion of $Zn(NO_3)_2 \cdot 6H_2O$

Fig. 14.13 shows thick coating formed by using binary and ternary targets due to merging of smaller nanoparticles.

One should note that the transition of zinc oxide nanoparticles formed in the target onto the coating is possible in several ways: through intensive gas emission during salt decomposition; during the sublimation process. On the basis of numerous studies, one could argue that it is unlikely to capture nanoparticles of zinc oxide formed in the target by evaporating salt followed by their transfer on the coating.

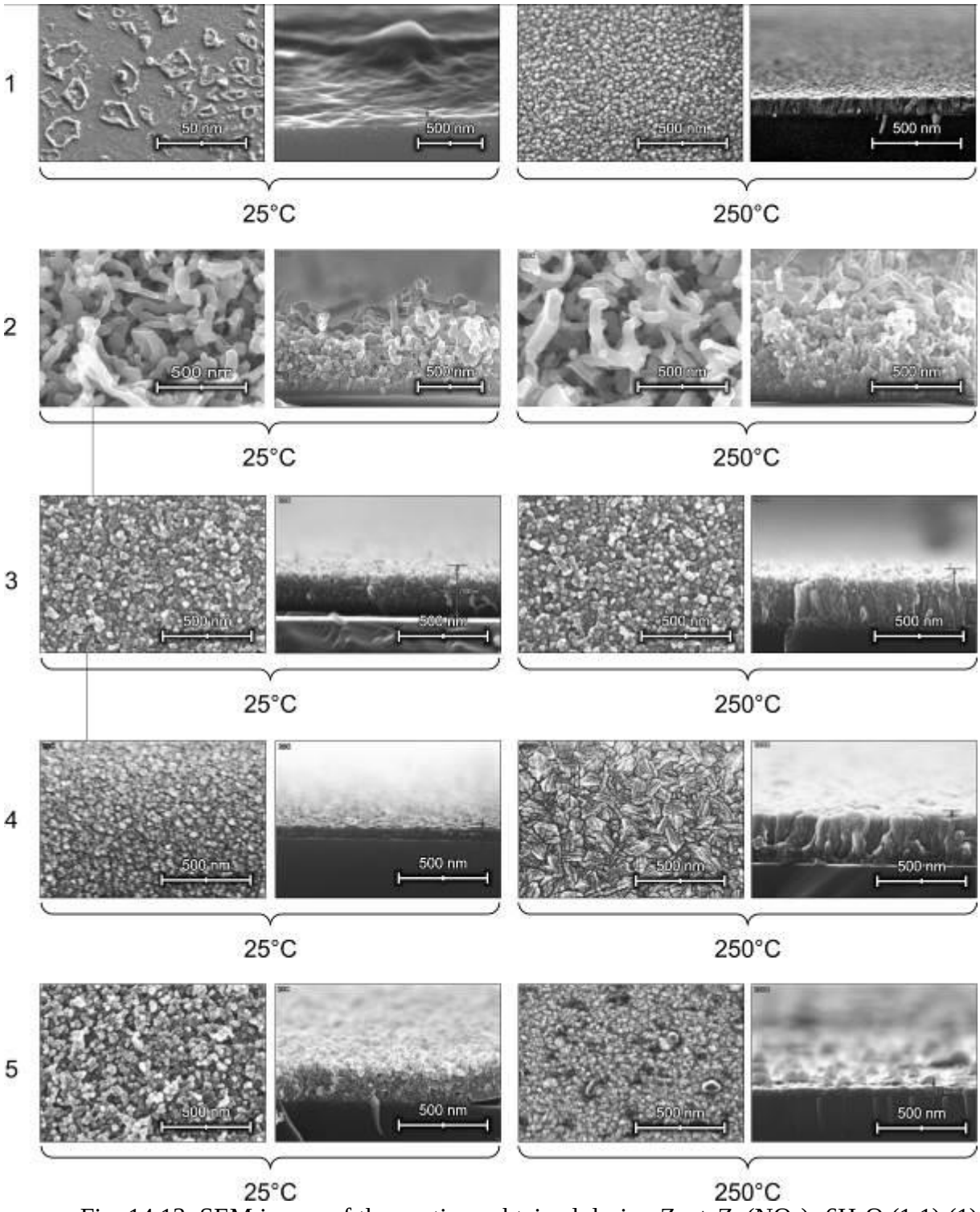



Fig. 14.13. SEM image of the coatings obtained during Zn + $Zn(NO_3)_2 \cdot 6H_2O$ (1:1) (1), (Zn + $Zn(NO_3)_2 \cdot 6H_2O$) + $Cu(NO_3)_2 \cdot 3H_2O$ evaporation with 1:0.25 (2), 1:0.5 (3), 1:1 (4), 1:2 (5) mixture ratio. Substrate temperature 25°C and 250°C

Thermal salt decomposition accompanied by intense local gas emission is the main cause of zinc oxide nanoparticles appearance on the coating (Fig. 14.12) during zinc nitrate dispersion. In case of using mixtures of zinc dust and salt, the transition of zinc nanoparticles onto the coating is caused by realization of the second noted way. The sublimation temperature of ZnO is 1800 °C and it is impossible to reach it with the help of electron-beam evaporator used in the studies as well as at the expense of the reaction energy decomposition of nitrate (the endothermic reaction $2Zn(NO_3)_2 = 2ZnO + 4NO_2\uparrow + O_2\uparrow$). Apart from noted dazzling white luminescence, the high temperature in the electron beam zone is indicated by the presence of metallic zinc nanorods in the upper layer of the coating based on ($Zn + Zn(NO_3)_2 \cdot 6H_2O$) + ($Cu(NO_3)_2 \cdot 3H_2O$) 1:0.25 (Fig. 14.13, (2)). Zinc melting temperature is known to exceed 400 °C. Thus, the transition of zinc oxide particles into the gas phase in addition to intensive gas emission is caused by high temperature. It is a consequence of exothermic reactions occurring between the target components.

The concentration of copper nitrate in the target and temperature of the substrate significantly influence the coating morphology. As shown in Fig. 14.13, the deposition of the coating at 250 °C has been significantly higher speed as compared to deposition on the substrate heated to 25 °C. Such abnormal change in deposition rate is explained by the effect of salt decomposition processes. According to this process the interaction of salt molecules with the surface heated up to 250 °C led the decomposition of molecules and the obtained zinc oxide molecules are involved into the formation of coating. Whereas during decomposition on the surface with 25 °C temperature, the oxide coating is formed only by zinc oxide particles formed at salt decomposition directly in the zone of electron flow impact on the target (thermal decomposition) as well as in its surface area, where intense plasma discharges are detected (photochemical decomposition).

It should be also noted that the coatings deposited on the high-temperature substrate have an evident columnar structure, which is appropriate at the primary growth of layers as a result of capture of the particles from the gas phase.

High content of original metal salt in the thin layer is the reason of high surface heterogeneity of the coating, its inflation on particular parts (Fig. 14.13 (1)). The defects are associated with the salt ability to intensively absorb moisture from the atmosphere.

Numerous researches showed high repeatability of the structures and the properties of thin layers based on $Zn + Zn(NO_3)_2 \cdot 6H_2O$ (1:1) and $Zn + Zn(NO_3)_2 \cdot 6H_2O + Cu(NO_3)_2 \cdot 3H_2O$ (1:1) formed on heated coatings. In such layers, using the method of Auger spectroscopy, homogenous distribution of oxygen and zinc is registered with high frequency throughout the coating thickness (Fig. 14.14).

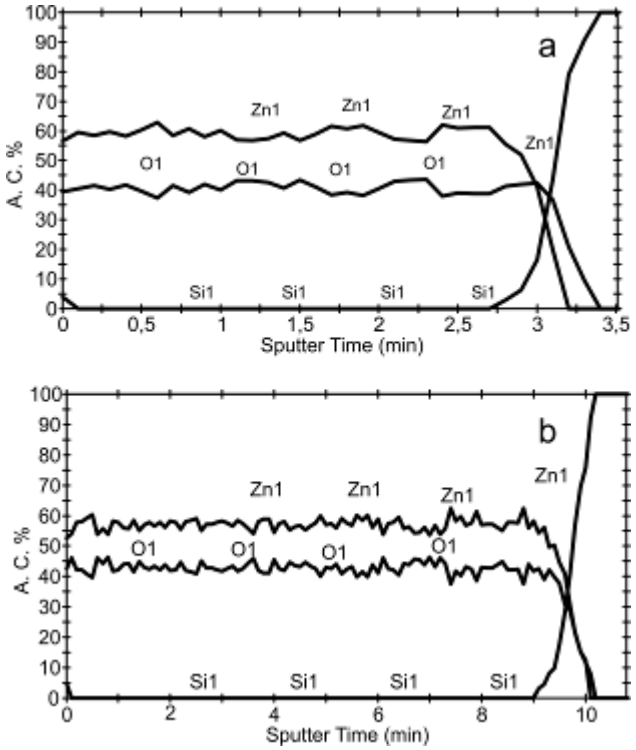



Fig. 14.14. Depth distribution of zinc and oxygen of the coating obtained by Zn – $Zn(NO_3)_2 \cdot 6H_2O$ (1:1) (a), Zn + $Zn(NO_3)_2 \cdot 6H_2O$ + $Cu(NO_3)_2 \cdot 3H_2O$ (1:1) (b) evaporation. The substrate temperature equals 250°C

At the same time, the results of researches indicate a high content of reduced zinc in the coating (approximately 18%) at the evaporation of Zn + $Zn(NO_3)_2 \cdot 6H_2O$ (1:1) mixture and less than 15% at the evaporation of Zn + $Zn(NO_3)_2 \cdot 6H_2O$ + $Cu(NO_3)_2 \cdot 3H_2O$ (1:1) mixture). The injection of copper nitrate into the evaporation zone allows slightly reducing the surplus of zinc concentration over oxygen. This can be explained by higher reactive activity of zinc as compared to copper. Zinc oxidation is possible at the process of copper recovery from its salt in direct contact of zinc with copper nitrate.

The measurements of the ratio of Zn and Cu content averaged through the layer thickness, which were made using the EDMA method, shows that the change in the target composition is an effective method to regulate the degree of doping of zinc oxide coatings with copper (Table 14.2).

Table 14.2 Zn and Cu ratio and the band gap values of the coatings deposited on the substrate at 250°C

| Components ratio in the target (Zn + $Zn(NO_3)_2 \cdot 6H_2O$)): ($Cu(NO_3)_2 \cdot 3H_2O$) | Content, at.% | | $E_g$ |
|---|---|---|---|
| | Zn /(Cu+ Zn) | Cu /(Cu+ Zn) | |
| 1:0 | 100 | 0 | 3.28/3.17* |
| 1:0.25 | 99.44 | 0.56 | - |
| 1:0.5 | 94.10 | 5.90 | 3.33 |
| 1:1 | 91.79 | 8.21 | 3.28/3.13* |
| 1:2 | 61.34 | 38.66 | 2.67 |

* – the coating is deposited at the temperature of the substrate 25°C and it is annealed at 250°C for 30 minutes

The X-ray diffraction data indicate on active participation of copper in the processes of structure formation (Fig. 14.15). The increase of copper content in thin layer, the (101), (100) then (002) peaks typical to crystalline ZnO are sequentially disappearing on the X-ray diffraction pattern.

According to Drmosh et al [20], the small amounts of copper in the thin layer (Cu at. % < 5.9) causes the *c*-axis orientation of ZnO crystallites perpendicular to the substrate, as confirmed by the data of scanning electron microscopy (Fig. 14.13). The higher copper content in the thin layer prevents the orientation growth of ZnO crystallites. It is known that the presence of three expressed peaks between 30° and 40° 2Θ indicates the polycrystalline hexagonal wurtzite structure. The disappearance of those peaks is possibly conditioned by the effect of copper on amorphization of the thin layer structure [77]. At Cu at. % > 5.9 the thin layer of ZnO formed is amorphous. The effect of copper is shown in increased sizes of the crystalline areas, reduced internal stresses and *c* parameter of the crystal lattice (Table 14.3). A similar effect of copper had already been registered in a number of researchers [20, 77, 78]. The decrease in the interplanar spacing (002) indicates the increase in $V_o$ oxygen vacancies [78].

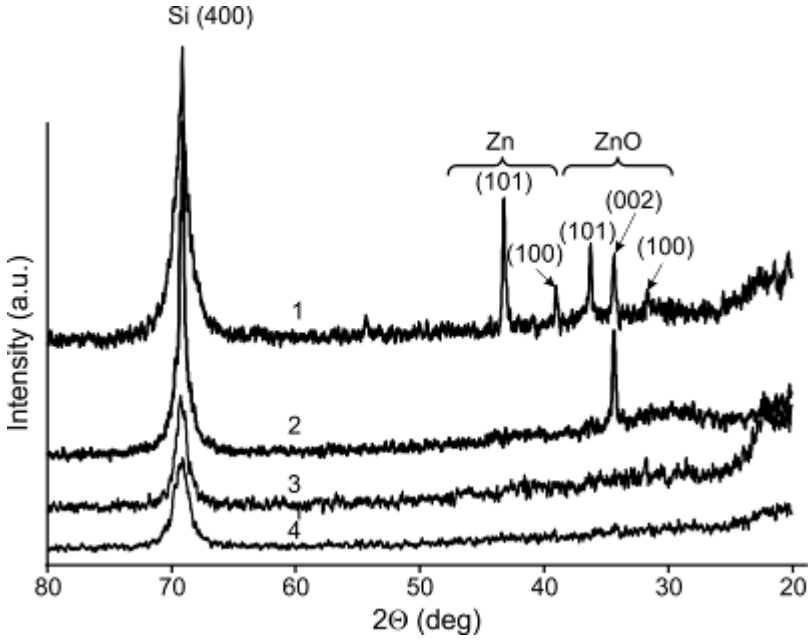



Fig. 14.15. X-ray patterns of the coatings obtained at (Zn + $Zn(NO_3)_2 \cdot 6H_2O$) + $Cu(NO_3)_2 \cdot 3H_2O$ evaporation with 1:0.25 (1); 1:0.5 (2), 1:1 (3), 1:2 (4) components ratio in the mixture

Table 14.3 Parameters of the (002) diffraction peak of ZnO thin copper containing coatings formed on Si substrates

| Components ratio | 2$\theta$ (°) | FWMH (°) | *c* (nm) | d (nm) | D (nm) | σ (GPa) |
|---|---|---|---|---|---|---|
| 1:0,25 | 34.35 | 0.2332 | 0.5217 | 0.1365 | 35.66 | 0.54 |
| 1:0,5 | 34.37 | 0.1992 | 0.5214 | 0.1364 | 41.70 | 0.40 |

The chemical composition of coatings was studied by decomposition of the Zn2p3/2, O1s and Cu2p3/2 bands on constituent separate peaks (Fig. 14.16). The peaks at 933.7 eV and 932.5 eV are correlated with CuO and $Cu_2O$ and Cu, respectively [79]. The binding energies of $Cu^0$ and $Cu^+$ are very similar in the Cu2p3/2. The peaks near at 1021.4 eV are correlated with ZnO ($Zn^{2+}$ chemical state), the peak 530.1 eV corresponds to $O^{2-}$ ions, which is typical for normal ZnO wurtzite structure, whereas the higher energy peak of 531.7 eV can be determined by different binding types, e.g., by adsorbed hydroxyl groups [80]. The X-ray photoelectron spectroscopy data

indicate the increase in the number of zinc atoms bound with oxygen atoms when copper concentration increases in the bulk of the coating. It is indicated by a shift in Zn2p3/2 towards smaller energies [80].

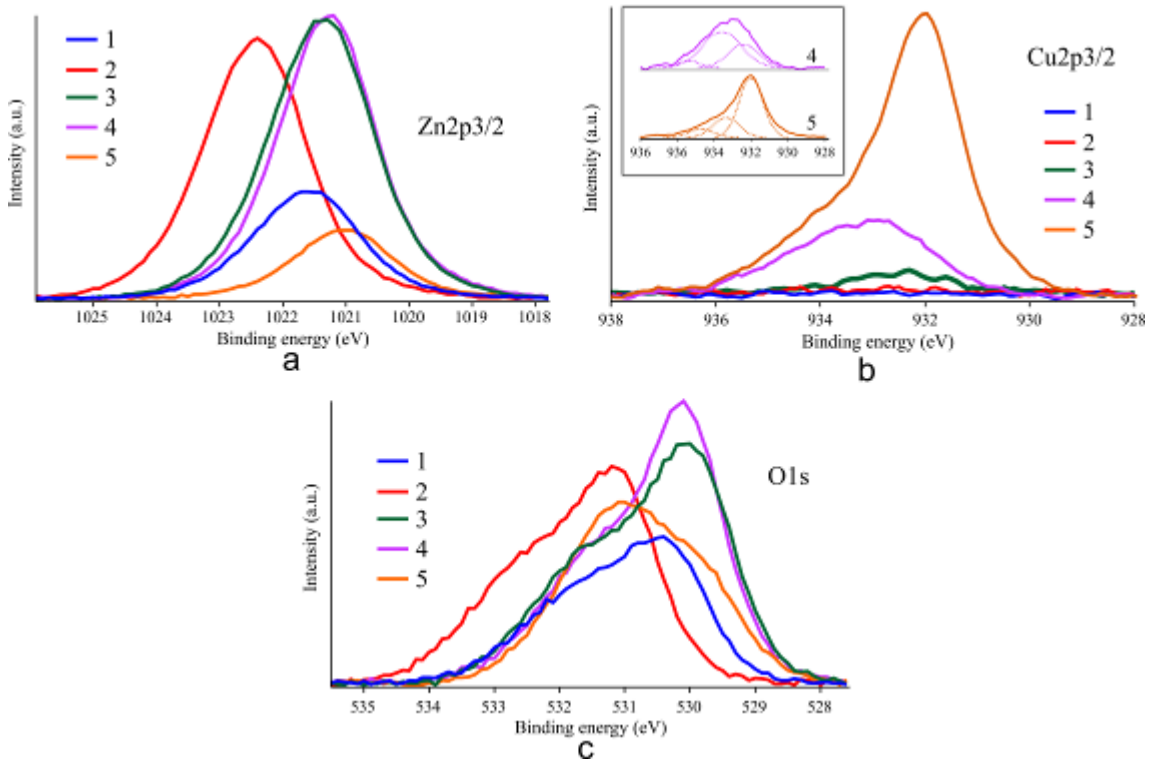



Fig. 14.16. XPS spectra of the coatings on Si substrates (250°C) formed at Zn – $Zn(NO_3)_2 \cdot 6H_2O$ (1:1) (1), (Zn + $Zn(NO_3)_2 \cdot 6H_2O$) + $Cu(NO_3)_2 \cdot 3H_2O$ evaporation with 1:0.25 (2); 1:0.5 (3), 1:1 (4), 1:2 (5) components ratio in the mixture

For the coatings with Cu at. % < 5.9 throughout the thin layer, copper is mainly in the oxidized state (CuO, $Cu_2O$). The coating formed by using the target with 1:2 components ratio is characterized by a high content of metal copper throughout the thin layer. The results obtained confirm the reduction of copper by zinc from its salt with subsequent decomposition of zinc nitrate to oxide at high temperature of the substrate. The lack of expressed peaks of copper oxide on X-ray patterns [81, 82] indicates the amorphous structure of oxides formed.

The chemical composition and structural properties of coating also have an influence on the value of energy band gap ($E_g$), which changes from 2.6 to 3.32 eV in the coatings (Table 14.2). As already noted, the injection of copper into the coating causes an increase in zinc oxidation by reducing the density of oxygen vacancies. At the same time, the crystallinity reduces, ultimately lead to increase in $E_g$ value. It is known that the band gap values differ significantly for CuO and $Cu_2O$. For bulk $Cu_2O$, the value equals to 2.17 eV, with 1.2 eV for CuO [83, 84]. Therefore, the presence of copper in the form of oxide in ZnO layer might cause the decrease in $E_g$. These features apparently cause nonmonotonic changes in $E_g$ with increasing concentration of copper. The energy band gap increases with low value of Cu in ZnO film due to structural change material, however, at higher concentration of Cu *Eg* decreases due to presence of copper oxide in more and more quantity.

## 14.2 Conclusion

We propose a new method to form ZnS coating by exposing mechanical mixture of sulfur and zinc powders to low-energy electron flow in vacuum. This method does not require additional annealing of the formed layer, and it can be used to apply coatings on the finished products of different nature, including polymers ones. The formation of zinc sulfide particles does not happen on the substrate, but out of it. The formed layers are characterized by chemical and structural

homogeneity in thickness. The substrate temperature influences insignificantly the morphology of the deposited layers, but it does not have significant influence on the value of band gap width. It was founded that the mechanical mixture of copper and sulfur powders in vacuum initiates a chemical interaction between sulfur and copper forming copper sulfide powder. The exposure of the mechanical mixture of zinc and copper sulfide powder to low-energy electron flow in vacuum is accompanied by the deposition of ZnS coating, the composition of which does not include copper. The influence of copper is shown up in increasing in size of the crystals forming ZnS coating.

It has been concluded that the injection of zinc dust into the evaporation zone increases the stability of the evaporation process and contributes fast deposition of film. The transition of zinc oxide particles formed in the target is caused by high temperature in the dispersion zone due to exothermic reactions occurring between the target components as well as intensive gas release during salt decomposition. The increase in temperature of the substrate up to 250 °C contributes to the increase in the growth rate and the deposition of structurally homogenous coatings with homogenous distribution of zinc and oxygen throughout the thickness. The injection of copper nitrate into the evaporation zone allows to increase the degree of zinc oxidation. It has been found that the doping of zinc oxide coatings with copper is due to disorienting effect, which causes the decrease in internal stresses.

The injection of copper into the coating has a nonmonotonic effect on the bandgap value, which varies from 2.6 to 3.33 eV. At low values of copper in the coating, an increase in $E_g$ has been found due to structural changes in the matrix layer of ZnO, whereas at high copper concentration there is a decrease in $E_g$ due to the presence of copper oxide in the coating as an independent phase.

**References**

[1] V.G. Korsakov, M.M. Sychov, V.V. Bakhmetyev, Synthesis and properties ofnano disperse semiconductors A2B6 and nanophosphors, Condensed Matterand Interphases 14 (2012) 41–52.

[2] X. Fang, T. Zhai, U.K. Gautamb, L. Li, L. Wu, Y. Bando, D. Golberg, ZnS nanostructures: from synthesis to applications, Progress in Materials Science 56 (2011)175–287.

[3] Yu.N. Bobrenko, N.V. Yaroshenko, G.I. Sheremetova, T.V. Semikina, B.S. Atdaev, Ultraviolet photosensors based on ZnS thin films, Tekhnologiya i Konstruirovanie v Elektronnoi Apparature 5 (2009) 29–31.

[4] P.N. Krylov, E.A. Romanov, I.V. Fedotova, Thermal annealing influence on optical properties of nanocrystalline zinc sulfide films, Semiconductors 45 (2011)1571–1575.

[5] Y.P. Venkata Subbaiah, P. Prathap, K.T. Ramakrishna Reddy, Structural, electrical and optical properties of ZnS films deposited by close-spaced evaporation, Applied Surface Science 5 (2006) 2409–2415.

[6] E.V. Karaksina, T.A. Gracheva, D.N. Shevarenkov, Structural defects in CVD ZnS, Inorganic Materials 46 (2010) 6–11.

[7] A.N. Georgobiani, A.N. Gruzintsev, Effect of surface morphology on manganese ion luminescence kinetics in ZnS:$Mn^{2+}$ films, Inorganic Materials 47 (2011)107–111.

[8] N. Kamoun Allouche, T. Ben Nasr, N. Turki Kamoun, C. Guasch, Synthesis and properties of chemical bath deposited ZnS multilayer films, Materials Chemistry and Physics 123 (2010) 620–624.

[9] N. Fathy, M. Ichimura, Photoelectrical properties of ZnS thin films deposited from aqueous solution using pulsed electrochemical deposition, Solar Energy Materials and Solar Cells 87 (1–4) (2005) 747–756.

[10] W. Tang, D.C. Cameron, Electroluminescent zinc sulphide devices produced by sol–gel processing, Thin Solid Films 280 (1–2) (1996) 221–226.

[11] M. Öztas¸, M. Bedir, S¸ Ocak, R. Güler Yyldırım, The role of growth parameters on structural, morphology and optical properties of sprayed ZnS thin films, Journal of Materials Science: Materials in Electronics 18 (5) (2007) 505–512.

[12] V.L. Gayou, B. Salazar-Hernandez, M.E. Constantino, E. Rosendo Andres, T. Dıaz, R. Delgado Macuil, M. Rojas Lopez, Structural studies of ZnS thin films grown on GaAs by RF magnetron sputtering, Vacuum 84 (2010) 1191–1194.

[13] A. Kronenberger, A. Polity, D. M. Hofmann, B. K. Meyer, Structural, electrical, and optical properties of hydrogen-doped ZnO films, Phys. Rev. B 86 (2012) 115334.

[14] S. Eustis, D. C. Meier, M. R. Beversluis, B. Nikoobakht, Analysis of copper incorporation into zinc oxide nanowires, ACS Nano. 2 (2008) 368–376.

[15] O. Lupan, T. Pauporté, B. Viana, P. Aschehoug, Electrodeposition of Cu-doped ZnO nanowire arrays and heterojunction formation with p-GaN for color tunable light emitting diode applications, Electrochimica Acta. 56 (2011) 10543–10549.

[16] N.L. Tarwal, K.V. Gurav, S.H. Mujawar, S.B. Sadale, K.W. Nam, W.R. Bae, A.V. Moholkar, J.H. Kim, P.S. Patil, J.H. Jang Photoluminescence and photoelectrochemical properties of the spray deposited copper doped zinc oxide thin films, Ceramics International, 40 (2014) 7669–7677.

[17] H. Czternastek, ZnO thin films prepared by high pressure magnetron sputtering, Opto-Electron. Rev. 12(1) (2004) 49–52.

[18] Z. Yan, Z. T. Song, W. L. Liu, Q. Wan, F. M. Zhang, S. L. Feng, Optical and electrical properties of p-type zinc oxide thin films synthesized by ion beam assisted deposition, Thin Solid Films. 492 (2005) 203–206.

[19] B. L. Zhu, J. Wang, S. J. Zhu, J. Wu, R. Wu, D. W. Zeng, C. S. Xie, Influence of hydrogen introduction on structure and properties of ZnO thin films during sputtering and post-annealing, Thin Solid Films. 519 (2011) 3809–3815.

[20] Q. A. Drmosh, S. G. Rao, Z. H. Yamani, M. A. Gondal, Crystalline nanostructured Cu doped ZnO thin films grown at room temperature by pulsed laser deposition technique and their characterization, Applied Surface Science 270 (2013) 104–108.

[21] S. A. Studenikin, N. Golego, M. Cocivera, Optical and electrical properties of undoped ZnO films grown by spray pyrolysis of zinc nitrate solution, J. Appl. Phys., 83, (1998) 2104–2111.

[22] A.V. Ragachev, M.A. Yarmolenko, A.A. Rogachev, D.L. Gorbachev, Bing Zhou, Chemical composition, morphology and optical properties of zinc sulfide coatings deposited by low-energy electron beam evaporation, Appl. Surf. Sci. 303 (2014) 23–29.

[23] C.D. Geddes, K. Aslan, I. Gryczynski, J. Malicka, J.R. Lakowicz, Noble-metal surfaces for metal enhanced fluorescence, Rev. Fluoresc. 2004 (2004) 365–401.

[24] C.L. Haynes, C.R. Yonzon, X. Zhang, Surface-enhanced Raman sensors: earlyhistory and the development of sensors for quantitative biowarfare agent and glucose detection, J. Raman Spectrosc. 36 (2005) 471–484.

[25] R. Dong, B. Tian, C. Zeng, T. Li, T. Wang, J. Zhang, Ecofriendly synthesis and photocatalytic activity of uniform cubic Ag@AgCl plasmonic photocatalyst, J.Phys. Chem. C 117 (2013) 213–220.

[26] V. Zaporojtchenko, R. Podschun, U. Schurmann, A. Kulkarni, F. Faupel, Physico-chemical and antimicrobial properties of co-sputtered Ag–Au/PTFE nanocomposite coatings, Nanotechnology 17 (2006) 4904–4908.

[27] D.A. Boyd, L. Greengard, M. Brongersma, M.Y. El-Naggar, D.G. Goodwin, Plasmon-assisted chemical vapor deposition, Nano Lett. 6 (2006) 2592–2597.

[28] H. Biederman, Plasma Polymer Films, Imp. Col. Press, London, 2004.

[29] P. Gritsenko, A.M. Krasovsky, Thin-film deposition of polymers by vacuum degradation, Chem. Rev. 103 (2003) 3607–3649.

[30] A.V. Rahachou, A.A. Rogachev, M.A. Yarmolenko, Jiang Xiao-Hong, Liu ZhuBo, Molecular structure and optical properties of PTFE-based nanocomposite polymer-metal coatings, Appl. Surf. Sci. 258 (2012) 1976–1980.

[31] A. Heilmann, J. Werner, In situ observation of microstructural changes of embedded silver particles, Thin Solid Films 317 (1998) 21–26.

[32] A. Fridman, Plasma Chemistry, Cambridge University Press, 2008.

[33] A.A. Rogachev, A.V. Rogachev, M.A. Yarmolenko, A.I. Egorov, Morphology of polytetrafluoroethylene coatings formed from the active gas phase and its thermal modification, Russ. J. Appl. Chem. 79 (2006) 1207–1209.

[34] Jörg Friedrich, Mechanisms of plasma polymerization – reviewed from a chemical point of view, Plasma Process. and Polym. 8 (2011) 783–802.

[35] R. Barbey, L. Lavanant, D. Paripovic, N. Schuwer, C. Sugnaux, S. Tugulu, H.-A.Klok, Polymer brushes via surface-initiated controlled radical polymerization: synthesis, characterization, properties, and applications, Chem. Rev. 109 (2009)5437–5527.

[36] G. Borkow, J. Gabbay Copper as a biocidal tool, Curr. Med. Chem. 12 (18) (2005) 2163–2175.

[37] V.B. Preethi Sudha, Sheeba Ganesan, G.P. Pazhani, T. Ramamurthy, G.B. Nair, Padma Venkatasubramanian, Storing Drinking-water in Copper pots Kills Contaminating Diarrhoeagenic Bacteria, J. Health Popul. Nutr. 30 (1) (2012) 17–21.

[38] G. Borkow, J. Gabbay, Putting copper into action: copper-impregnated products with potent biocidal activities, FASEB J. 18 (14) (2004) 1728–1733.

[39] Shikha Jain, Niharika Nagar, Vijay Devra, Synthesis and characterization of highly efficient copper nanoparticles and their catalytic application in oxidative kinetic study, Adv. Appl. Sci. Res. 6 (6) (2015) 171–180.

[40] G.V. Ramesh, S. Porel, T.P. Radhakrishnan, Polymer thin films embedded with in situ grown metal nanoparticles, Chem. Soc. Rev. 38 (2009) 2646–2656.

[41] Chun-Chin Chen, Franklin Chau-Nan Hong, Structure and properties of diamond-like carbon nanocomposite films containing copper nanoparticles, Appl. Surf. Sci. 242 (2005) 261–269.

[42] Wei Zhang, Yi-He Zhang, Jun-Hui Ji, Jun Zhao, Qing Yan, Paul K. Chu, Antimicrobial properties of copper plasma-modified polyethylene, Polymer, 47 (2006) 7441–7445.

[43] Zhiguang Guo, Weimin Liu, Bao-Lian Su, Superhydrophobic surfaces: From natural to biomimetic to functional, J. Colloid Interface Sci. 353 (2011) 335–355.

[44] H. J. Ensikat, P. Ditsche-Kuru, C. Neinhuis, W. Barthlott, Superhydrophobicity in perfection: the outstanding properties of the lotus leaf, Beilstein J. Nanotechnol. 2 (2011) 152–161.

[45] P. Zhang, F. Y. Lv, A review of the recent advances in superhydrophobic surfaces and the emerging energy-related applications, Energy. 82 (2015) 1068–1087.

[46] A. C. Lima, J. F. Mano, Micro/nano-structured superhydrophobic surfaces in the biomedical field: part II: applications overview, Nanomedicine (Lond.) 10(2) (2015) 271–297.

[47] H. Jeong, J. Heo, B. Son, D. Choi, T. H. Park, M. Chang, J. Hong, Intrinsic Hydrophobic Cairnlike Multilayer Films for Antibacterial Effect with Enhanced Durability, ACS Appl. Mater. Interfaces. 7(47) (2015) 26117–26123.

[48] U. Cengiz, H. Y. Erbil, Superhydrophobic perfluoropolymer surfaces having heterogeneousroughness created by dip-coating from solutions containinga nonsolvent, Appl. Surf. Sci. 292 (2014) 591–597.

[49] S. Heinonen, E. Huttunen-Saarivirta, J.-P. Nikkanen, M. Raulio, O. Priha, J. Laakso, E. Storgards, E. Levanena, Antibacterial properties and chemical stability of superhydrophobic silver-containing surface produced by sol–gel route, Colloids and Surfaces A: Physicochem. Eng. Aspects. 453 (2014) 149–161.

[50] R. Taurino, E. Fabbri, D. Pospiech, A. Synytska, M. Messoria, Preparation of scratch resistant superhydrophobic hybrid coatings by sol–gel process, Prog. Org. Coat. 77 (2014) 1635–1641.

[51] L. Ovaskainen, S. Chigome, N. A. Birkin, S. M. Howdle, N. Torto, L. Wagberg, C. Turner, Superhydrophobic polymeric coatings produced by rapid expansionof supercritical solutions combined with electrostatic deposition (RESS-ED), J. of Supercritical Fluids. 95 (2014) 610–617.

[52] B.-Y. Jeong, E.-H. Jung, J.-H. Kim, Fabrication of superhydrophobic niobium pentoxide thin films by anodization, Appl. Surf. Sci. 307 (2014) 28–32.

[53] J. Liang, K. Liu, D. Wang, H. Li, P. Li, S. Li, S. Su, S. Xu, Y. Luo, Facile fabrication of superhydrophilic/superhydrophobic surface ontitanium substrate by single-step anodization and fluorination, Appl. Surf. Sci. 338 (2015) 126–136.

[54] S. Rezaei, I. Manoucheri, R. Moradian, B. Pourabbas, One-step chemical vapor deposition and modification of silica nanoparticles at the lowest possible temperature and superhydrophobic surface fabrication, Chem. Eng. J. 252 (2014) 11–16.

[55] F. Henry, F. Renaux, S. Coppee, R. Lazzaroni, N. Vandencasteele, F. Reniers, R. Snyders, Synthesis of superhydrophobic PTFE-like thin films by self-nanostructuration in a hybrid plasma process, Surface Science 606 (2012) 1825–1829.

[56] R. Jafari, R. Menini, M. Farzaneh, Superhydrophobic and icephobic surfaces prepared by RF-sputtered polytetrafluoroethylene coatings, Appl. Surf. Sci. 259 (2012) 719–725.

[57] H.-K. Kim, Y.-S. Cho, Fabrication of a superhydrophobic surface via spraying with polystyrene and multi-walled carbon nanotubes, Colloids Surf. A: Physicochem. Eng. Asp. 465 (2015) 77–86.

[58] D. Gong, J. Long, P. Fan, D. Jiang, H. Zhang, M. Zhong, Thermal stability of micro–nano structures and superhydrophobicityof polytetrafluoroethylene films formed by hot embossing via apicosecond laser ablated template, Appl. Surf. Sci. 331 (2015) 437–443.

[59] Z. Duan, Z. Zhao, D. Luo, M. Zhao, G. Zhao, A facial approach combining photosensitive sol–gel with self-assembly method to fabricate superhydrophobic $TiO_2$ films with patterned surface structure, Appl. Surf. Sci. 360 (2016) 1030–1035.

[60] Z. Li, Hierarchical ZnO films with microplate/nanohole structures induced by precursor concentration and colloidal templates, their superhydrophobicity, and enhanced photocatalytic performance, J. Mater. Res. 29(1) 2014 115 – 122.

[61] A. A. Rogachev, M. A. Yarmolenko, A. V. Rogachou, D. V. Tapalski, X. Liu, D. L. Gorbachev, Morphology and structure of antibacterial nanocomposite organic-polymer and metal-polymer coatings deposited from active gas phase, RSC Adv. 3 (2013) 11226–11233.

[62] Crystalline Olefin Polymers. Raff, R A V & Doak, K W (Eds). Published by Interscience Publishers, John Wiley, 1965.

[63] J. Dechant, R. Danz, V. Kimmer and R. Schmolke, Infrakrasnaya spektroskopiya polimerov (Infrared Spectroscopy of Polymers). p. 472, Moscow, 1976.

[64] Bijoy Barman, K.C. Sarma, Luminescence properties of ZnS quantum dots embedded in polymer matrix, Chalcogenide Letters, 8 (2011) 171–176.

[65] B. Bhattacharjee, D. Ganguli, K. Iakoubovskii, A. Stesmans, S. Chaudhuri, Synthesis and characterization of sol–gel derived $ZnS:Mn^{2+}$ nanocrystallites embedded in a silica matrix, Bull. Mater. Sci., 25 (2002) 175–180.

[66] H.Q. Bian, S.Y. Ma, F.M. Li, H.B. Zhu, Influence of ZnO buffer layer on microstructure and Raman scattering of ZnO:Ag film on Si substrate, Super lattices and Microstructures 58 (2013) 171–177.

[67] E. Ritter, Optical film materials and their applications, Applied Optics 15 (10) (1976) 2318–2327.

[68] R.F. Khairutdinov, Chemistry of semiconductor nanoparticles, Russian Chemical Reviews 67 (1998) 125–139.

[69] S.W. Shin, S.R. Kang, J.H. Yun, A.V. Moholkar, J.-H. Moon, J.Y. Lee, J.H. Kim, Effect of different annealing conditions on the properties of chemically deposited ZnS thin films on ITO coated glass substrates, Solar Energy Materials & Solar Cells95 (2011) 856–863.

[70] F. Göde, E. Güneri, A. Kariper, C. Ulutas, F. Kirmizigül, C. Gümüs, Influence of annealing temperature on the structural, optical and electrical properties of amorphous zinc sulfide thin films, Journal of Physics 326 (2011) 012020,http://dx.doi.org/10.1088/1742-6596/326/1/012020.

[71] P. Prathap, N. Revathi, Y.P. Venkata Subbaiah, K.T. Ramakrishna Reddy, Thickness effect on the microstructure, morphology and optoelectronic properties of ZnS films, Journal of Physics: Condensed Matter 20 (2008)035205.

[72] I.L. Knunyants (Ed.), Chemical Encyclopedia, vol. 3, Great Russian Entsikl., Moscow, 1992.

[73] R. Blachnik, A. Müller, The formation of Cu2S from the elements: I. Copper use din form of powders, Thermochimica Acta 361 (2000) 31–52.

[74] A. Phuruangrat, T. Thongtem, S. Thongtem, Characterization of copper sulfidehexananoplates, and nanoparticles synthesized by a sonochemical method, Chalcogenide Letters 8 (2011) 291–295.

[75] Y. Zhao, H. Pan, Y. Lou, X. Qiu, J. Zhu, C. Burda, Plasmonic $Cu_{2-x}S$ nanocrystals: optical and structural properties of copper-deficient copper(I) sulfides, Journal of the American Chemical Society 131 (12) (2009) 4253–4261.

[76] L.J. Bellamy, The Infra-red Spectra of Complex Molecules. Molecular spectra, Methuen, Methuen & Co. Ltd., London, 1954, pp. 323.

[77] J.-H. Bang, H.-S. Uhm, Won Kim, J.-S. Park, Effects of additive gases and plasma post-treatment on electrical properties and optical transmittance of ZnO thin films, Thin Solid Films 519 (2010) 1568–1572.

[78] R. Chauhan, A. Kumar, R. Pal Chaudhary, Synthesis and Characterization of Copper doped ZnO nanoparticles, J. Chem. Pharm. Res. 2 (2010) 178–183.

[79] J. Ghijsen, L.H. Tjeng, J. van Elp, H. Eskes, J. Westerink, G.A. Sawatzky, M.T. Czyzyk, Electronic structure of $Cu_2O$ and CuO, Phys. Rev. B. 38 (16) (1988) 11322–11330.

[80] J. He, B. Tan, Y. Su, S. Yang, Q. Wei, XPS Analysis of ZnO thin Films Obtained by Pulsed Laser Deposition, Adv. Mat. Res. 383-390 (2012) 6293–6296.

[81] H. L. Zhang, G. Y. Zhao, L. Z. Xu, Preparation of the photosensitive copper complex and CuO film pattern, Appl. Surf. Sci. 274 (2013) 397–400.

[82] N. Ekthammathat, T. Thongtem, S. Thongtem, Antimicrobial activities of CuO films deposited on Cu foils by solution chemistry, Appl. Surf. Sci. 277 (2013) 211–217.

[83] D. P. Singh, N. R. Neti, A. S. K. Sinha, O. N. Srivastava, Growth of Different Nanostructures of $Cu_2O$ (Nanothreads, Nanowires, and Nanocubes) by Simple Electrolysis Based Oxidation of Copper, J. Phys. Chem. C. (2007) 1638–1645.

[84] E. Darezereshki, F. Bakhtiari, A novel technique to synthesis of tenorite (CuO) nanoparticles from low concentration $CuSO_4$ solution, J. Min. Metall. Sect. B-Metall. 47 (2011) 73–78.

## 15. Carbon coatings doped with metals
### 15.1. Introduction

Diamond-like carbon (a-C) gain ever-growing popularity in scientific and industrial communities on the dependence of the excellent properties like high hardness, low friction coefficient, high wear resistance, better chemical inertness, biocompatibility and optical transparency in IR region [1-3]. Various deposition techniques have been used to fabricate a-C films, such as plasma enhanced chemical vapor deposition (PECVD) [4], magnetron sputtering [5], pulsed laser deposition (PLD) [6], and cathode arc evaporation (CAE) [7]. Compared with other methods, CAE is a low temperature deposition method with a better ion ratio and higher ion energy. By CAE method, films with strong sp3 bonding are easily produced. During deposition, ion bombardment on the substrates improves adhesion between the film and the substrate [8].Preparation of a-C films using CAE technique, the power supply is operated in direct-current or pulsed mode. Bias voltage is an important parameter for direct-current cathode arc evaporation technique. The effect of bias voltage on the structure and properties of a-C films was reported by many researchers [9-10]. Pulse frequency is crucial influencing factors for pulsed cathodic arc method. However, negligible efforts related to influences of pulse frequency on structure and properties of a-C films were found in the literature. Variation of pulse frequencies can modify the energy and quantity of depositing ions, which further effects on the growth of films. Therefore, understanding the interrelation of pulse frequencies and the growth of a-C films is of cardinal significance to optimize the microstructure and associated mechanical properties. Hard a-C films have good prospects for solving various technical problems due to the high mechanical, tribological and optical properties [11]. However, practical industrial applications of a-C films were still restricted because of the high intrinsic compressive stress. In recent years, the incorporation of metallic elements into the a-C films provides an effective way to improve the properties of the films, such as Ti doping, Cu doping, W doping, Cr doping and etc. [12-15]. Incorporating metals in a-C films considerably influence their microstructure and lead to significant changes in electrical conductivity, biocompatibility, mechanical, optical and tribological properties. Compared with Ti-doped a-C (a-C:Ti) films containing TiC nanometer grains, copper atoms are not chemically bound with carbon in Copper-doped a-C films (a-C:Ti) [12]. In addition, some researchers also proposed a mechanism for the effect of incorporating metal on the tribological properties of the a-C films, including the relationship between variation of metals-doping and the graphitization of films [16]. It signifies that doping metal (Ti or Cu) is a promising technology improving the structure and mechanical properties of a-C films. Unfortunately, many of the researchers used to focus their attention on a-C films doped with small concentration of Ti or Cu (up to 10 wt. %) [17, 18], but neglected the researching of morphology and mechanical properties of high alloyed a-C films (metal concentration 10-40 wt. %) [19]. Diamond-like amorphous carbon (DLC) films are composed of a $sp^2$ hybrid graphite phase and a $sp^3$ hybrid diamond phase. The graphite phase exists in the highly crosslinked network structure of diamond phase, which improves the a-C film by higher hardness, resistivity, thermal conductivity and chemical inertness. Therefore, the a-C films show great potential in applied mechanics, acoustics, microchip manufacture etc. [20, 21]. E.g., some a-C films doped with metals such as W, Au, Pt etc. reveal excellent performance regarding to adhesion, corrosion resistance and residual stress [29].

When choosing an ultimate element, Cu can be regarded as an ideal alternative for Pt and Au, due to its low cost, high stability and excellent electrical conductivity [30, 31, 32].

However, Cu in a-C films cannot form carbide, but only a weak chemical bond between carbon and copper, which, in turn, affects the relative content of the graphite phase and the diamond phase. As it is known, the relative content of the phases influences the film properties, therefore, special methods of using Cu can be applied to get the necessary performance of a film [33, 34, 35]

a-C films are composed by amorphous carbons of a-C or a-C:H with a high proportion of $sp^3$ bond. a-C films are characterized by their high hardness, low friction coefficient, chemical

inertness, optical transparency and superior electrical resistance, and therefore have widespread applications in fabricating mechanical, acoustic, electronic and optical devices and in the preparation of magnetic media protection coatings [36, 37, 38]. It is worthwhile to mention that one important feature for a-C films is its facile manipulation of their physical properties, ranging from the mechanical properties to the photoconductivity, by simply controlling the relative ratio of the $sp^3$ to $sp^2$ carbons in films. However, the shortcomings of a-C films, involving the large internal stress, low adhesion and thermal stability have hindered their practical applications in engineering. For example, conventional a-C films exhibit limited adhesion to many metallic substrates and high internal compressive stress (up to 10 GPa), as a result, leading to peeling-off of the films from the substrate when the film thickness is more than 1μm. Consequently the development of a-C films with improved adhesive to metallic substrate has become one major topic in engine applications [39]. To reduce internal stress, wear and friction, and to improve adhesion between film and substrate, one of the most efficient methods is metal doped a-C film deposited by introducing metal elements into the system [40, 41]. Unfortunately, the effects of metals implanted on the structure and further on the properties of a-C films are very complex, as the functions of each metal element varies. For example, some metals (e.g. Ag, Cu, etc.) do not chemically bond with carbon and only arranged in the form of nanoparticles within the diamond-like carbon matrices. The other species (e.g. Ti, W, etc.) can be chemically bonded with carbon atoms which results in the formation of the composite film with metal carbide phases. Although the introduction of metallic species can reduce the internal stress of a-C films, which on the other hand will unavoidably change the films' hardness. To our knowledge, no report has been given on the effects of binary metal doped a-C composite films on the structures and properties of a-C films with the aim to improve its surface properties. Therefore the investigation on the relationship between the overall physical properties, in particular the internal stress and the hardness of the films, and the molecular structure of the films is of great value.

## 15.2. Experimental

A vacuum deposition device was used to obtain the doped and multilayer composite a-C films, as shown in Figure. 15.1. This device contains two DC cathode arc evaporators with magnetic separation of plasma-ion source for sputtering non-magnetic metal target (such as Cr, Ti, Al, Cu and etc.), pulsed cathode arc evaporator for synthesis of a-C films, a gas-ion source for heating and cleaning of substrates.

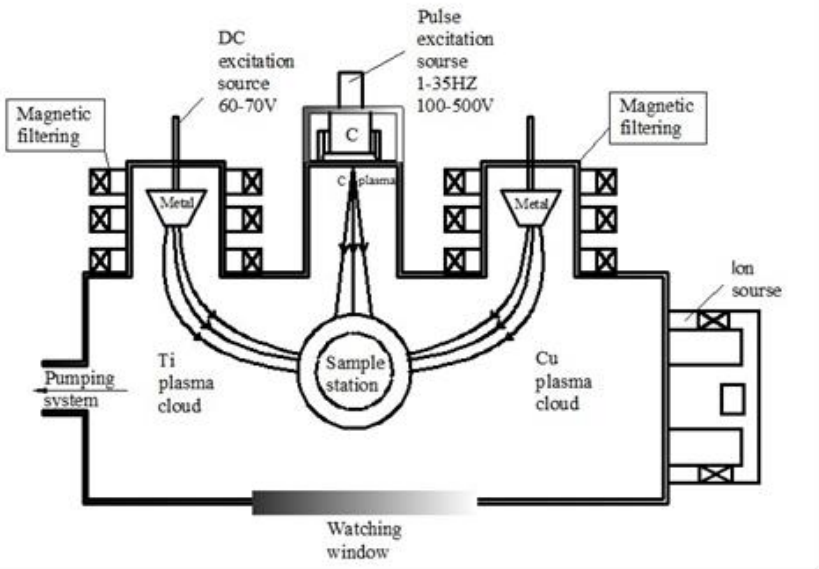



Fig. 15.1. Schematic diagram of deposition device with double-excitation cathode arc source.

A metallic Cu and Ti targets of 99.99% purity were used as DC evaporation cathodes. A graphite target of 99.5% purity was used as evaporation cathode of pulsed arc to obtain a-C films. The values of voltage, impulse frequency and number of impulses were changed from 250 to 350 V, 1-20 Hz and 5000, respectively. Silicon (100) substrate was ultrasonically cleaned for 20 min in acetone, ethanol and deionized water. Crystalline n-Si (100) wafers were used as substrate for the measurement of residual stress in films. At a base pressure of $6\times10^{-4}$ Pa, the substrate was etched by ion sputtering source for 15 min to clean the oxide layer on surface. $Ar^+$ energy and ion current density were 0,4 keV and ~25 A/m$^2$, respectively. The substrate was cooled till room temperature (above 20 °C) after sputtering clean. The speed of the sample holder with substrate was at 2 min$^{-1}$.It was prepared some type of samples of carbon based films. a-C films with different contents of Ti and Cu alloyed were researched, respectively, the content of doped metal in the films was from 10 to 40 wt.%. Doping of titanium or copper in a-C films was carried out by a separate electric arc evaporator. The change the content of metals in the coating was achieved by varying the arc current. All samples were annealed in vacuum with the pressure of $2\times10^{-3}$ Pa at 350 °C for 60 min.T he film thickness was measured by a step device (Ambios Technology XP-2) and Stylus Profiler (Dektak XT, Bruker) The Dektak XT is a semi-automated stylus profiler that can be used to measure step height with better than 5 Å repeatability, surface roughness, as well as 3D surface mapping. On the surface of substrate a line was drawn before putting it in the chamber, after depositing films, then wipe off the line with anhydrous ethanol for thickness measurement.

*Raman spectroscopy and chemical composition.* The microstructure of the films was analyzed using a Renishaw inVia Raman spectrometer. Raman spectra were excited with the 514.5 nm emitting lines from an Ar-ion laser at an incident power of 20 mW. The distribution pattern of elements was evaluated by scanning electron microscopy (SEM) using TESCAN Vega II LSH equipped with Oxford «INCA Energy 350» Instruments Energy Dispersion Spectroscopy (EDS) facilities.

*The microhardness test.* The microhardness of a-C films was measured by AFFRI DM-8 Vickers sclerometer, with the use of diamond indenter with the angle of 172°30′ between the long edge of rhombic pyramid. The indentation load and duration time were 245.1 mN and 10 s, respectively

*AFM test*. Surface morphology and roughness ($R_{MS}$)of films was observed by atomic force microscopy (AFM) using a multimode scanning microscope Solver-PRO P47 (NT-MDT) in tapping mode, with a scanning scope of 5×5 μm and a scan rate of 1.0 μm/s..

*Residual test*. According to Stoney formula, total internal stress (σ) of the films was determined by the change in curvature radius of the Si substrate before and after the deposition using double-crystal X-ray device. Both the radius of curvature R and the film thickness $t_f$ are also measured by the stylus profiler[83].

*Wear and friction.* The tribological property was studied by frictional wear testing that was performed using a ball-disc friction method. The coefficient of friction (CoF) of the films was measured by a tribotester (UMT2-EC, Bruker) with a ball made of $Si_3N_4$ (diameter 5/32″). The test was conducted with a load of 0.5 N at a sliding speed of 0.1 m/s in ambient air under the relative humidity 42÷45% and temperature 22 °C.

## 15.3. Results

### 15.3.1. Morphology and mechanical properties of high alloyed carbon films prepared by pulsed cathodic discharge

As shown in Fig. 15.2 the content of metals increase with the growing of discharge current to 80 A. On continuous increase of arc current, the concentration of metals in a-C films significantly decreases. It can be explained that the droplet phase was generated in the intensive mode, which is then separated from the flow by a magnetic separator. All Raman spectra are characterized by a broad almost symmetric peak in the range of 1100-1800 cm$^{-1}$ centered at

approximately 1560 $cm^{-1}$, which represents the typical characteristic of hydrogenated amorphous a-C films. This confirms that the films are amorphous[13, 19, 29]. The peak in all spectra can be fitted as G band (~1560 $cm^{-1}$) and D band (~1350 $cm^{-1}$) by two Gaussian curves, revealing the existence of mixed $sp^2$ and $sp^3$ bonding in a-C films.

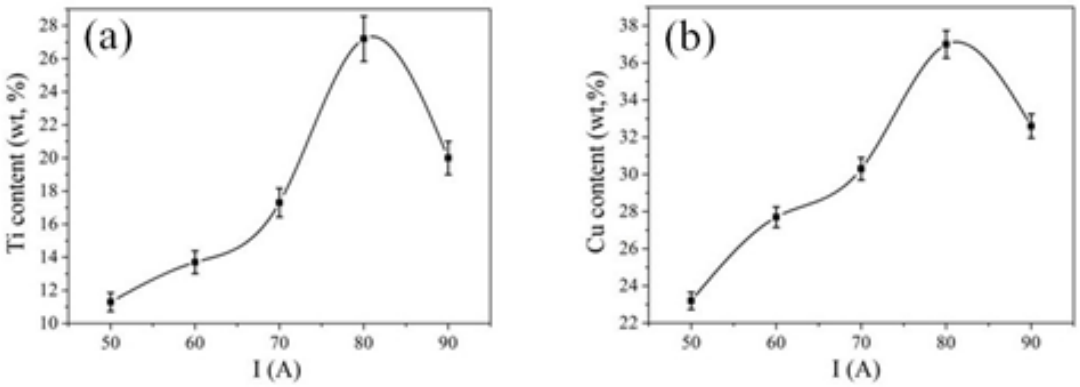



Fig. 15.2**.** The concentration of metal (Ti, Cu) in the alloyed a-C films deposited at different arc current: (a) Ti-doped a-C films; (b) Cu-doped a-C films.

The Raman spectra of synthesized a-C:Ti films with varied Ti content were recorded as shown in Fig. 15.3.

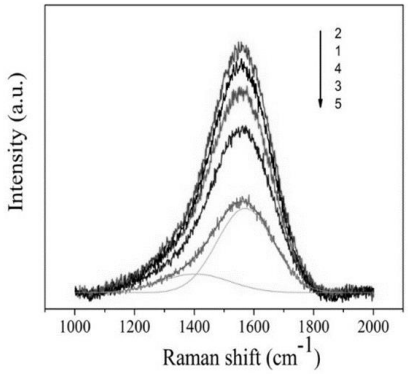



Fig. 15.3 Raman spectra for a-C:Ti with different Ti concentration. (The numeration of curves is as in Table 15.2)

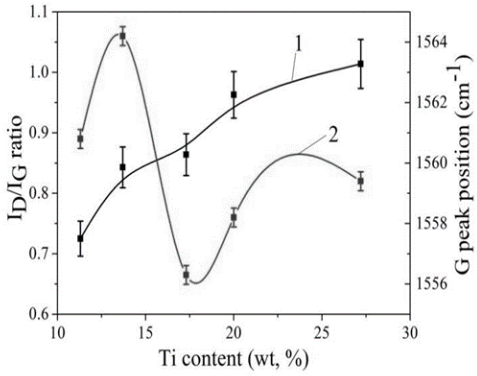



Fig. 15.4 (1)–The G-peak position, (2)–$I_D/I_G$ ratio of the a-C:Ti films with different Ti concentration.

The variations of $I_D/I_G$ ratio and the position of the G-peak of the films as a function of the Ti content are shown in Fig. 15.4. As it is shown at the Fig. 15.4, that with increasing Ti content of the films, the ratio $I_D/I_G$ is increased. This effect implied that with higher Ti contents, the number of $sp^3$ bonds in the a-C component decreased, and the $sp^2$ bonds in the graphite phase dominated the structure. This may be due to the formation of TiC damage the carbon three-dimensional network in the a-C films. The position variation of the corresponding G peak also shows that a metal Ti doping induced variation in the film structure. Changing the position of the corresponding G peak also shows that the structure variation is induced by Ti doping in the films.

The Raman spectra of synthesized -doped a-C:Cu films with varied Cu content were recorded as shown in Fig. 15.5. The variations of $I_D/I_G$ ratio and the position of the G-peak of the films as a function of the Cu content are shown in Fig. 15.6.

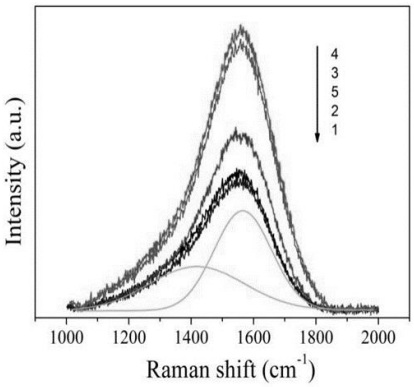



Fig. 15.5. Raman spectra for a-C:Cu with different Cu concentration. (The numeration of curves is as in Table 15.2).

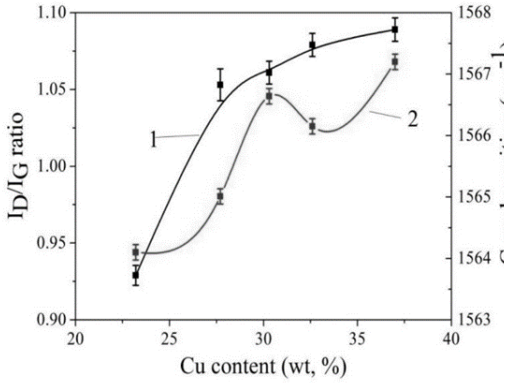



Fig. 15.6 (1) –The G-peak position, (2) –$I_D/I_G$ ratio of the a-C:Cu films with different Cu concentration.

It is seen that the $I_D/I_G$ ratio increase monotonically with increasing Cu content, basically consistent with the $I_D/I_G$ variation of Ti-alloyed a-C films. The results indicate that the proportion of $sp^2/sp^3$ bonds of films increases. However, their G peak position changes are inconsistent. This may be due to Ti with C to form a hard TiC phase, while the Cu atoms and C atoms do not react, in the form of Cu nanoparticles distributed in the a-C film.

The morphological properties of Ti-doped a-C films and Cu-doped doped a-C films (before annealing) are shown in the Table 15.2.

Table 15.2. Morphological parameters of a-C:Ti and a-C:Cu films.

| Sample | № | Arc current A | Metal content wt, % | $R_{ms}$ μm | Thickness μm | Wear rate $10^{-15}m^3/(N\times m)$ |
|---|---|---|---|---|---|---|
| a-C:Ti | 1 | 50 | 11.3 | 11.8 | 0.41 | 0.56 |
| | 2 | 60 | 13.7 | 11.1 | 0.48 | 0.65 |
| | 3 | 70 | 17.3 | 11.2 | 0.51 | 0.72 |
| | 4 | 80 | 27.2 | 4.3 | 0.53 | 0.91 |
| | 5 | 90 | 20.0 | 16.0 | 0.47 | 0.77 |
| a-C:Cu | 1 | 50 | 23.2 | 8.7 | 0.36 | 0.13 |
| | 2 | 60 | 27.7 | 11.1 | 0.45 | 0.14 |
| | 3 | 70 | 30.3 | 4.2 | 0.44 | 0.13 |
| | 4 | 80 | 37.0 | 9.0 | 0.50 | 0.16 |
| | 5 | 90 | 32.6 | 19.1 | 0.53 | 0.21 |

As the Ti concentration increases from 11.3 to 17.3 wt. %, the roughness of Ti alloyed a-C films remains unchanged. With further increase of the Ti content, the roughness of the films decreases rapidly. For Cu alloyed a-C films, the roughness of the Cu alloyed a-C films is minimized, when the Cu concentration is 70 A. From the above made analysis, under certain

process conditions, metal (Ti or Cu) can reduce the roughness of the a-C films [91]. When the arc current continues to increase from 80 A, the roughness of Ti or Cu alloyed a-C films increases sharply, this may be due to the large metal particles are produced. Although the magnetic separator has a filtering effect, part of the large particles are still deposited on the film surface.

Fig. 15.7 shows AFM images of surface morphology of the metal-doped (Ti or Cu) a-C films before and after annealing. It can be seen that surface morphology of Ti or Cu alloyed a-C films contains some particles with a large size and shows an uneven particle distribution on the surface, which is largely similar to the amorphous a-C films. It is believed that the large particles are consisted of $sp^3$ hybridization phase of carbon atoms and metallic particle agglomeration. From Table 15.2, Cu-doped a-C films shows the lower wear rate. From the Fig. 15.7e, Cu-doped a-C films show bigger aggregated particles in comparison to Ti-doped a-C films. This result implied that with higher Cu content, evaporated Cu particles aggregate and easily form droplet phases. The droplet phases will act as a solid lubricant during the friction, thus providing the films with high tribological properties. After annealing, particle size and density of the Cu-doped films increase significantly. This is apparently due to the occurrence of the surface segregation, which is formed over the equilibrium structure with the minimum level of internal stress. From the Fig. 15.12a and b, Ti alloyed a-C film surface is not flat, the size of 0.4-0.5 μm particles are uniformly distributed on the surface. After annealing, the surface of Ti alloyed carbon films becomes smoother; particle size and density on the surface significantly reduce. These surface morphology changes of the films are the result of intensive diffusion process, accompanied by the formation of metal carbide.

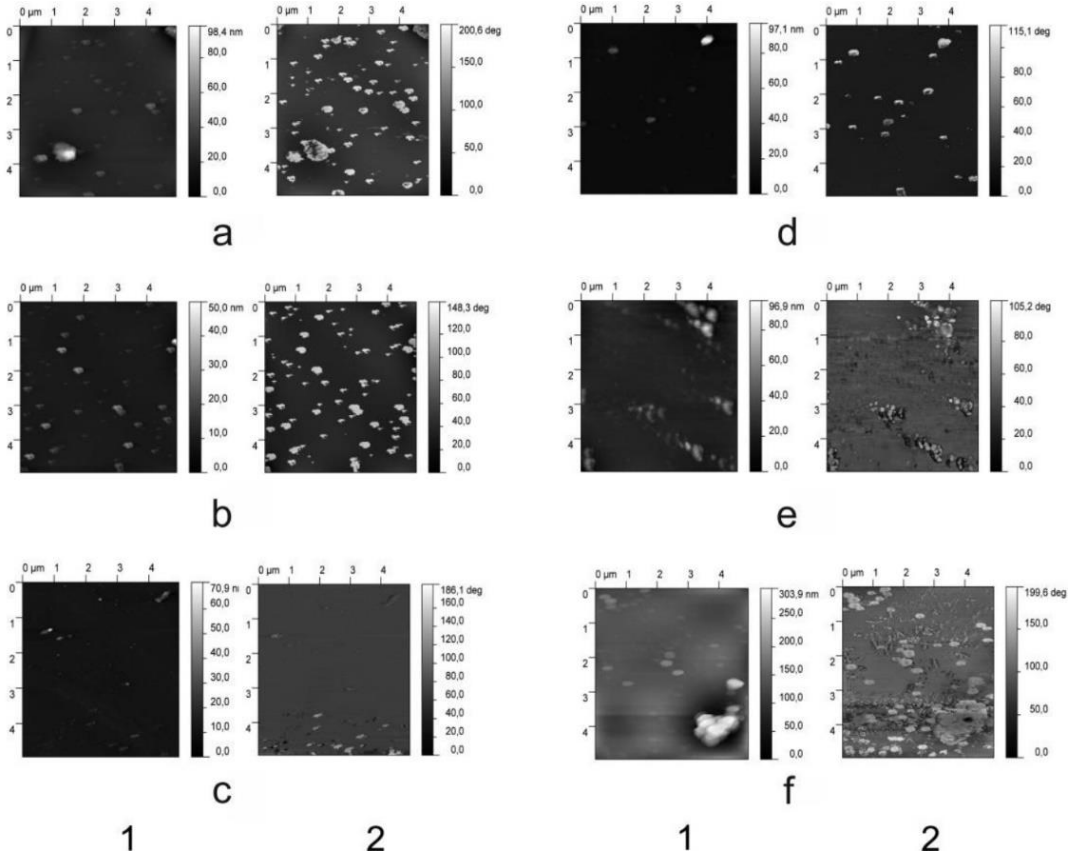



Fig. 15.7. AFM images of Ti-doped and Cu-doped a-C films before and after annealing at 350 °C in vacuum: (a, c) Ti concentration 13.7 %; (b, d) Ti concentration 27.2 %; (e, f) Cu concentration 27.7 %. 1-morphology; 2-phase contrast.

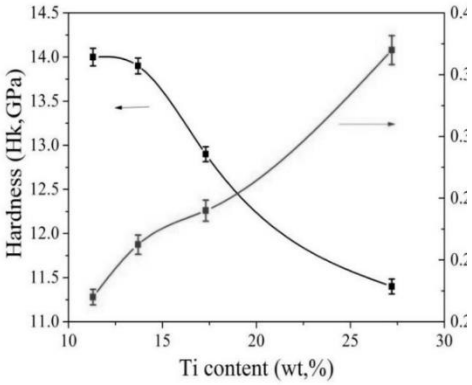



Fig. 15.8. Hardness and friction coefficient as a function of the Ti content

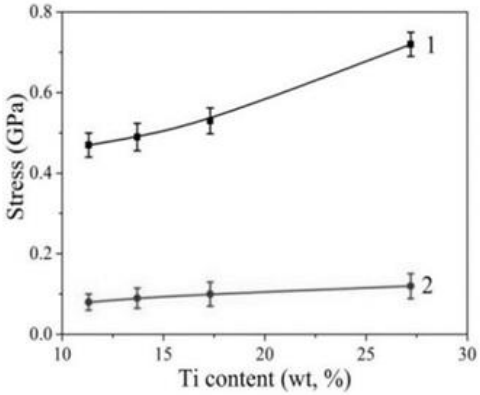



Fig. 15.9. Internal stress as a function of the Ti content. 1 – before annealing, 2 – after annealing

Fig. 15.8 displays the hardness and friction coefficient of the Ti-doped a-C films as a function of the Ti concentration. With the increase of Ti content, the hardness of the films sharply decline. This is due to the shielding effect of Ti atoms and the interaction between Ti atoms and C atoms, restraining the growth of the hard carbon particles. Studies have shown that the variation of Ti content directly affects tribological properties of the films [92]. The Fig.14 shows that with increasing Ti concentration, the friction coefficient of the films increased significantly. The titanium-doped carbon films have a higher friction coefficient, leading to the intensive wear on the surface, as shown in Table 15.2.

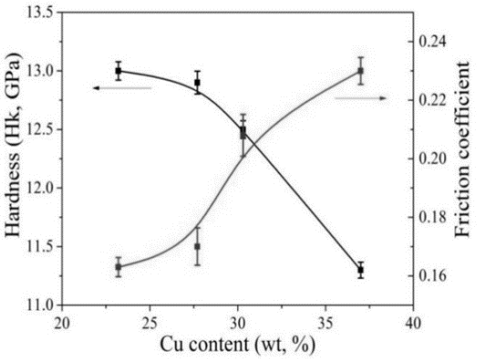



Fig. 15.10. Hardness and friction coefficient as a function of the Cu content for a-C:Cu.

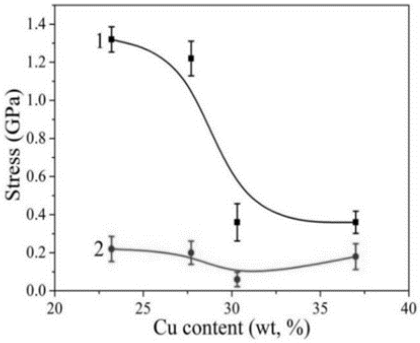



Fig. 15.11. Internal stress as a function of the Cu content. 1 – before annealing, 2 – after annealing

Fig. 15.9 shows the internal stress of the films as a function of the Ti concentration. Before annealing, the Ti content of the films increases to 17.3 %, the internal stress of Cu alloyed a-C films has a slightly rise. When Ti content continues to increase, the internal stress of the film increased significantly. After annealing, the internal stress of the films maintains substantially constant as the Ti content increases. Comparing the Fig.14, it is found that internal stress of the films is reduced more than 5 times after the annealing treatment. Because the internal stress of films is caused by the structure incompatibility. After annealing, the incompatible metastable structure transition to quasi-equilibrium state structure, leading to the intrinsic stress decrease [78].

Fig. 15.10 displays the hardness and friction coefficient of the a-C:Cu films as a function of the Cu concentration. It can be seen, that the mechanical and tribological properties of a-C films doping copper have almost monotonic dependence of the Cu content. With the increase of the Cu content, the hardness of the films decreases and the friction coefficient increases. The comparison

of the Fig. 15.12 and 15 shows that the hardness of a-C:Cu films decreases more quickly than Ti-doped a-C films. It may be due to the formation and growth of carbide in Ti-containing a-C films, slowing the hardness of films decrease [94]. From Table 15.2 and Fig. 15.10, Cu-doped a-C films have a sufficiently low friction coefficient and wear rate. This is due to the Cu atoms in the form of nanoparticles are presented in the a-C films and have played a role of the balls in the rubbing process, so the films have good tribological properties [95]. Fig. 15.11 shows the internal stress of the films as a function of the Cu concentration. Before annealing, with the increase of copper concentration, the stress of Cu-alloyed a-C films significantly decreases. On the reason that copper atoms have a full d orbital and are not chemically bound with carbon but they arrange themselves in the form of nanoparticles within the a-C matrix. This could significantly effect on the release of the stress, but the hardness will have greater loss [92, 94]. After annealing, the internal stress of the films sharply decreases, and with the increase of the Cu content, the internal stress of the films substantially does not change.

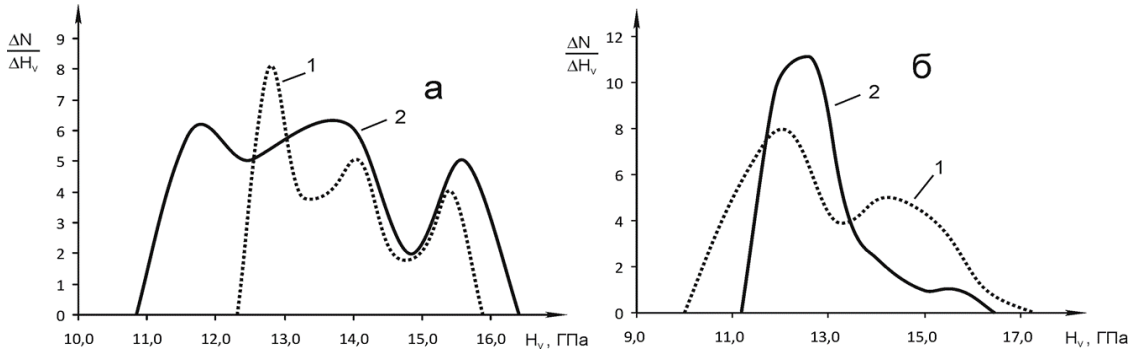



Fig. 15.12. The distribution of the values of microhardness surface areas a-C:Ti (a), a-C:Cu (b) with an arc current 50 A (1) and 70A (2)

Fig. 15.12a shows the distribution curve of measurement number on the hardness values of Ti alloyed a-C films. It can be seen, that three peaks are presented in the graph. Three peaks can be related with the hardness values of monocomponent a-C (maximum hardness $H_k$ = 15.5±0.5 GPa), titanium carbide ($H_k$ = 13.5±0.5 GPa) and highly dispersed droplet phase titanium ($H_k$=12.5±1.2 GPa). It can be concluded that Ti alloyed a-C films contain three phases: metallic phase titanium, diamond-like carbon and titanium carbide [69].

Fig. 15.12b shows the distribution curve of measurement number on the hardness values of Cu alloyed a-C coating. The distribution curve of Cu-doped a-C films contain two peaks, corresponding respectively to the different components of the films: superfine metallic phase and a modified carbon layer. On increasing the concentration above 30% copper layer is determined by microhardness mechanical properties of the metallic phase.

The researches have shown that the hardness of a-C films is generally 10÷30 GPa and the internal stress is 0.5÷7 GPa [75, 82]. Based on the above analysis available, the hardness of high Ti alloyed a-C films is 11÷14 GPa and the internal stress is 0.4÷0.7 GPa; the hardness of high Cu alloyed a-C films is 11÷13 GPa and the internal stress is 0.3÷1.3 GPa. In comparison with the pure a-C films, high Ti or Cu alloyed a-C films have high hardness and low internal stress. Thus, to some extent, the metal doping could optimize the mechanical properties of the a-C films.

## 15.4. Conclusions

a-C films have been deposited on Si substrates using pulsed cathodic arc plasma technique. The effects of the copper doping on the optical properties of deposited amorphous carbon films were investigated. It was found out that within the range of Cu concentration from 0% to 29.7 wt.% the optical band gap is decreasing with the increase of copper concentration. The optical properties of a-C:Cu films are related to their electronic band structure. AFM data have shown that the morphology of the a-C:Cu films depends on the concentration of copper content in the films,

moreover, RMS increases with increasing the concentration of doping element. Raman and XPS spectroscopy revealed the effect of copper doping on the carbon atom phase composition of C:Cu films.

**References**


[1] K. Bewilogu, D. Hofmann History of diamond-like carbon films — From first experiments to worldwide applications Surf. Coat. Technol Vol. 242 (2014), p.214.

[2] R.A. Ismail, W.K. Hamoudi, K.K. Saleh: Effect of rapid thermal annealing on the characteristics of amorphous carbon/n-type crystalline silicon heterojunction solar cells Mat. Sci. Semicon. Proc Vol. 21 (2014), p.194.

[3] J. Miksovsky, A. Vossa, R. Kozarovad, T. Kocourek, P. Pisarik, G. Ceccone, W. Kulisch, M. Jelinek, M.D. Apostolova, J.P. Reithmaier, C. Popov: Cell adhesion and growth on ultrananocrystalline diamond and diamond-like carbon films after different surface modifications Appl. Surf. Sci Vol. 297 (2014), p.95.

[4] D. Caschera, F. Federici, S. Kaciulis: Deposition of Ti-containing diamond-like carbon (DLC) films by PECVD technique Mater. Sci. Eng Vol. C 27 (2007), p.1328.

[5] Guojia Ma, Shuili Gong: A study of structure and properties of Ti-doped DLC film by reactive magnetron sputtering with ion implantation Appl. Surf. Sci Vol. 258 (2012), p.3045.

[6] D.L. Pappas, K.L. Saenger, J. Bruley: Pulsed laser deposition of diamond - like carbon films Journal of applied physics Vol. 71 (1992), p.5675.

[7] M.C. Chiu, W.P. Hsieh, W.Y. Hob, D.Y. Wang, F.S. Shieu: Thermal stability of Cr-doped diamond-like carbon films synthesized by cathodic arc evaporationThin Solid Films Vol. 476 (2005), p.258.

[8] K. Weng, Y. Chen, T. Lin, D. Wang: Metal-doped diamond-like carbon films synthesized by filter-arc depositionThin Solid Films Vol. 515 (2006), p.1053.

[9] P. Tsai, K. Chen: Evaluation of microstructures and mechanical properties of diamond like carbon films deposited by filtered cathodic arc plasmaThin Solid Films Vol. 516 (2008), p.5440.

[10] C. Chang, J. Jao, T. Chang, W. Ho, D. Wang: Influences of bias voltage on properties of TiAl-doped DLC coatings synthesized by cathodic arc evaporationDiamond Relat. Mater Vol. 14 (2005), p.2127.

[11] M. Moseler, P. Gumbsch, C. Casiraghi, A.C. Ferrari, J. Robertson, The ultrasmoothness of diamond-lime carbonScience, 2005; 309, 1545-1547.

[12] J. Robertson, Diamond-like amorphous carbon. Mater. Sci. Eng. R 37 129-281 (2002).

[13] Davis, C.A., Amaratunga, G.A.J., Knowles, K.M.: Growth mechanism and cross-sectional structure of tetrahedral amorphous carbon thin films. Phys. Rev. Lett. 80, 3280 (1998).

[14] N. Dwivedi, S. Kumar, H.K. Malik, C. Sreekumar, S. Dayal, C.M.S. Rauthan, O.S. Panwar, J. Phys. Chem. Solids. 73, 308-316. (2012)

[15] L. Yang, A. Neville, A. Brown, P. Ransom, A. Morin, Tribol. Int. 70, 26-33. (2014)

[16] C.W. Zou, H.J. Wang, L. Feng, S.W. Xue, Appl. Surf. Sci. 286, 137-141, (2013).

[17] Mednikarov B., Spasov G., Babeva Tz., Pirov J., Sahatchieva M., Popova C., Kulisch W., J. Optoelectronics Advanced Materials, 2005, 7, 1407 – 1413

[18] Selvakumar N., Barshilia H. C., Solar Energ. Mat. Sol. Cells, 2012, 98. 1–23.

[19] Bewilogu K., Hofmann D., Surf. & Coat. Technol. 2014, 242, 214–225.

[20] Zhou B., Jiang X., Rogachev A.V., Shen R., Adv. Mater Res, 2013. 629, 25 – 31.

[21] Jansson U., Lewin E., Thin Solid Films, 2013, 536, 1–24.

[22] Siklitsty V. I., Yastrebov S. G., Lodygin A. B., Chaos, Solitons, Fractals, 1999, 09. No 12.

## 16. Sol-gel synthesis of functional materials

### 16.1. The formation of optical materials using sol-gel processes

The first experiments on receiving "vitreous material" from colloidal gels, according to the reviews of X. Dislich [1], L. Hench and J. Vest [2], were carried out by M. Ebelman and T. Graham in 1864. However the extremely long time (about a year) was required to prevent transformation of volume silica gel into thin powder and therefore the offered method hasn't drawn attention of technologists.

In 1971 H. Dislich [3] received the first disc-shaped sample with the diameter of several centimeters from uniform borosilicate glass by hot pressing of the gel powder synthesized from alkoxide of silicon and organic compounds of boron, which has been recognized then as one of considerable achievements of sol-gel technology . Sumio Sakka together with colleagues from Kyoto University in Japan aimed at synthesizing the materials in the form of glass blocks, fibers and plates, using the initial solution consisting of tetraethylorthosilicate (TEOS), ethanol, water and hydrochloric acid [4]. He called this method a direct alkoxide sol-gel process since formation of material of future glass was combined with process of gelation of colloidal system in a molding form, and vitreous dioxide of silicon was synthesized by heat treatment of dried up silica gel. This is how the term "a formed gel-glass" appeared. At the same time difficulties in the of synthesis of dry gels of the considerable size without their cracking on pieces, the considerable periods of drying, bad reproducibility and availability of alternative ways of production of glass have not allowed this option of sol-gel process to come to the level of developments and implementation at that time.

Long-term experience of researches has shown that the direct alkoxide method of synthesizing the volume-formed glassware is limitedly suitable for the technical application in view of low economic efficiency, the poor reproducibility and ecological danger because of need of burning off the large numbers of the organic additives that regulate drying, and applications of methyl compounds of silicon acid.

The "colloid" sol-gel method of vitreous materials synthesis which allows to form coarse-pored monoliths of the aggregated ultrafine particles of silicon dioxide which are in a condition of sol or transferred to solid and fluid system from silica powders [5-7] has a number of restrictions. There are difficulty of receiving the homogeneous gels, impossibility of doping of vitreous materials, tendency to crystallization, high temperatures of processing in a void-free state.

The known way of synthesis of vitreous materials by hydrolysis of silicates of alkali metals was also not widely used for formation of large preparations in view of complexity of processes of scavenging of alkali metals and problems of destruction of ultraporous xerogels when drying. However it is successfully applied during the creation of the transparent matrixes of small size to introduce of the photoactive organic dyes and optical composites of the organic-inorganic nature.

Efforts of researchers and developers of sol-gel processes were aimed at the search for the ways of modifying the alcoxide sol-gel process which would allow to overcome the faults and to realize the advantages of "traditional" options. The sceme of "straight lines" options of sol-gel processes of synthesis of vitreous materials is given in fig. 16.1.

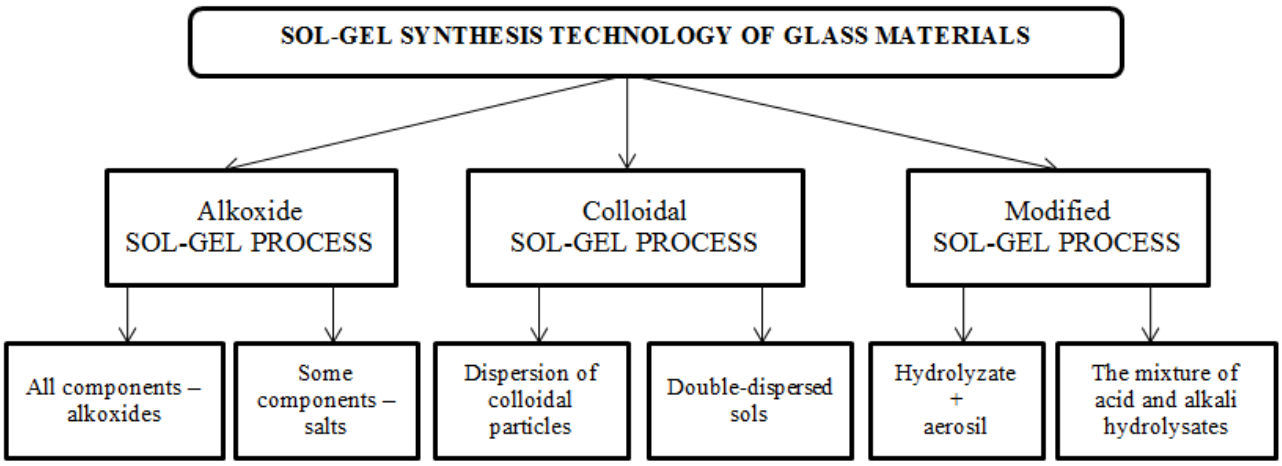



Fig.16.1. Scheme variants of "direct" sol-gel synthesis process of glass materials

Thus, physical and chemical bases of alcoxid sol-gel processes had been rather deeply developed by the beginning of the nineteenth years of the last century. The patterns of evolution of colloidal silica system on the basis of organic compounds of silicon were experimentally determined, and the most optimum system for practical use has been recognized the $Si(OC_2H_5)_4$ – $C_2H_5OH$ – $H_2O$ with use of HCl as the hydrolysis process catalyst.

### 16.1.1. Alcoxid sol-gel process: the modern condition of the question

#### 16.1.1.1. Basic concepts and definitions

The term **silica** is used as a short, convenient designation of dioxide of silicon in all its crystal, amorphous, hydrated and the hydroxylated forms.

**Sol** is a two-phase system consisting of the colloid particles distributed in liquid.

The **colloidal particles** are solid particles with a diameter of 1 ÷ 100 nm.

**Silica sol** – silica colloidal system consisting of Si02 particles of colloidal size in a liquid medium.

**Gel** – an interconnected three-dimensional solid mesh with pores of submicron size, filled with a liquid (hydrogel) or gas (xerogel).

If the liquid in the pores is water, we use the term akvagel, and in the case of filling alcohol – **alkogel**.

The generic term gel is often applied to various types of ultraporous gel structures, and low density gel obtained by hypercritical drying is called airgel. The gel is considered to be dry when physically adsorbed water is completely removed from his body. This occurs at temperatures ranging from 100 to 180 °C.

The dried gel still contains a large amount (6.6 to 2 nm) of chemisorbed hydroxyl groups (OH) on the surface of the pores. Heat treatment of gels in the range 500-800 ° C leads to a desorption of hydroxyls from the pore surface; such gels are called stabilized gels.

Heat treatment of the gels at elevated temperatures (800-1200 °C) significantly reduces the number of pores and their interconnection due to the mechanism of viscous flow. This process is called sintering or compaction. The porous gel transforms into a dense monolithic glass or when all the pores disappear.

**Microamorfous silica** includes sols, gels, powders and gel-glass, which consist mainly of primary particles of a size of less than one micron and a specific surface area greater than 3 $m^2/g$.

Mikroamorfous silica is divided into two classes:

1. Pyrogenic silica (Aerosil) is an amorphous anhydrous silica particles obtained at high temperature by oxidation of silicon compounds such as SiCl4, SiH4, SiHCl3 and others at vapor state or in an oxygen-native carbohydrate flame.

2. Hydroxylated microamorfous silica is formed on the embryos generated from solutions of $Si(OH)_4$ monomer, formed by hydrolysis of organic silicon compounds (alcoxid sol-gel process) or silicic acid salts. Tetraedrica meshes of the silica lining during the condensation of $Si(OH)_4$ monomer are amorphous and have a spherical shape.Thus mikroamorfous hydroxylated silica is obtained from the solution either in the form of spherical amorphous particles (sol) or aggregates into a gel or coagulates at the sudden change in pH or temperature forming a powder. Initially formed particles may contain some uncondensed Si-OH (**silanol group**) groups and, depending on the growth conditions, OH groups either remain on the surface, or are distributed over the particle diameter by certain law in the amorphous network of the silica.

According to the classification G. Scherer [8] the pores in the xerogels obtained by the sol-gel method, depending on the size, are divided into three groups: **micropores** (<1 nm), **mesopores** are pores or transient pores (1-30 nm) and **macropores** (> 30 nm).

The sols formed from particles of fumed silica and hydroxylated amorphous silica obtained by hydrolysis of alkoxide compounds of silicon with subsequent polycondensation are called mixed or hybrid sols. And the transition of sol into the gel by destabilizing (when the temperature or pH of the system is changed) is called gelation. Gelation occurs in the forms (molds) sol is poured into after the destabilization. During the maturation period (aging) of gel, which occurs for quite a long time (from several hours to several days), the process of hardening of the gel network through polycondensation reactions and the growth of necks in the structure of silica oxide framework and extruding of intermicellar liquid (syneresis) continues.

Drying is a complex process of the removal of liquid from the gel body by the mechanisms of flow and evaporation, the gel shrinks and significantly increases the mechanical strength.

The **sol-gel process** – a brief outline of the evolution of the colloidal silica system in the phase transition sol → gel → monolithic solid, whereby the vitreous material (porous, monolithic or composite) is formed.

**Sol-gel technology** is the technology for the preparation of porous or solid chemical products, in particular, on the basis of microamorphous silica glass, xerogels, powders, films or fibers.

Xerogel or glass or composite that keeps the shape of the injection mold in the process of the evolution from the hydrogel to glass perform and undergoing uniform shrinkage along the three coordinate axes is formed by a direct sol-gel process of molded vitreous material.

### 16.1.1.2 Physico-chemical basis of the alcoxide sol-gel processes

The "classic" version of alkoxide sol-gel process proposed by S.Sakka in 1985 [4], is the formation of sol from TEOS by its hydrolysis and polycondensation of the resulting silicic acid monomers under the conditions of acid catalysis, gelation drying of the gel, and finally sintering it to a state of transparent glass material (quartz glass and a gel). The stages of the "classical" Sol-gel process contain a number of chemical and physical processes that form the basic technological scheme:

- mixing of initial components;
- hydrolysis of TEOS and polycondensation;
- casting of sols in shapes;
- gelation;
- the maturation of gels;
- drying of gels;
- bonding to the glassy state.

A simplified block diagram of alkoxide sol-gel process is shown in Fig. 16.2.

Physics-chemical regularities of the evolution of colloid-silica system at the phase transition from sol state to a transparent monolithic solids has been the subject of numerous studies, with the greatest number of works devoted to the study of system $(C_2H_5O)_4Si$ - $C_2H_5OH$ - $H_2O$ - HCl as the most suitable for practical use [10, 2, 21-24].

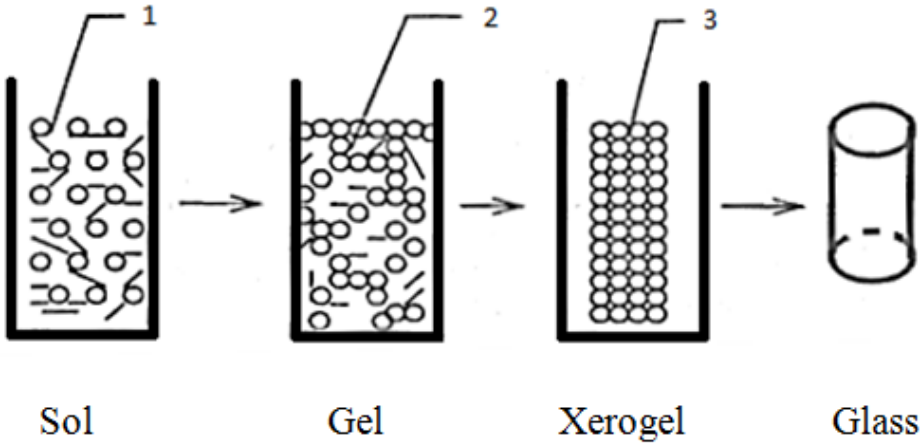
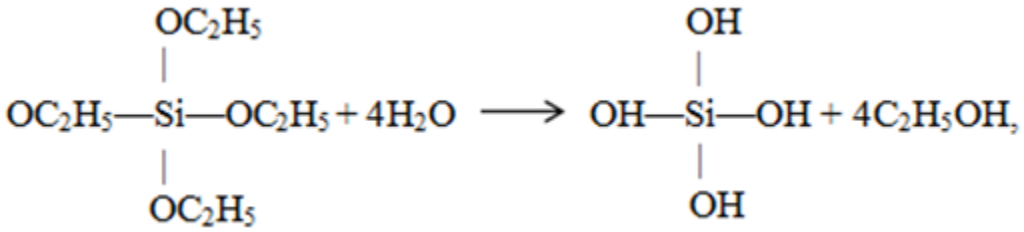



Fig. 16.2. Structural scheme of the sol-gel process: 1 - the solid phase particles of $SiO_2$; 2 - liquid phase ($H_2O$, $C_2H_5OH$); 3 - gas phase.

### 16.2. The hydrolysis and polycondensation

4 molecules of water are required in accordance with the chemical reaction for the complete hydrolysis of TEOS molecules:



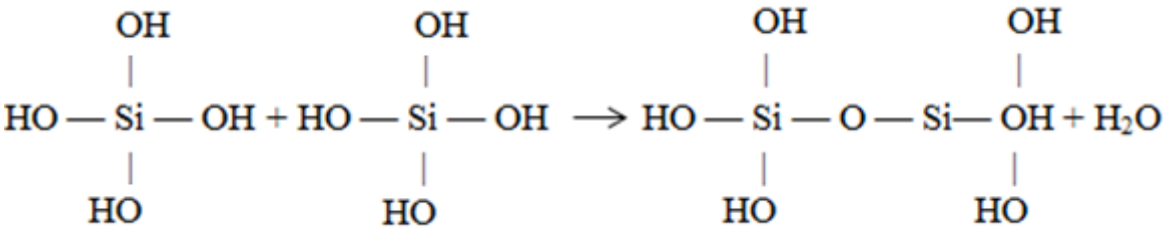
however, it is well known [9] that the process of hydrolysis with $(C_2H_5O)_4Si$ takes place in four stages:

$$\begin{aligned}
&Si(OEt)_4 + H_2O \xrightarrow{[k_0]} SiOH(OEt)_3 + EtOH,\\
&SiOH(OEt)_3 + H_2O \xrightarrow{[k_1]} Si(OH)_2(OEt)_2 + EtOH,\\
&Si(OH)_2(OEt)_2 + H_2O \xrightarrow{[k_2]} Si(OH)_2(OEt) + EtOH,\\
&Si(OEt)_3(OEt) + H_2O \xrightarrow{[k_3]} Si(OH)_4 + EtOH
\end{aligned}$$

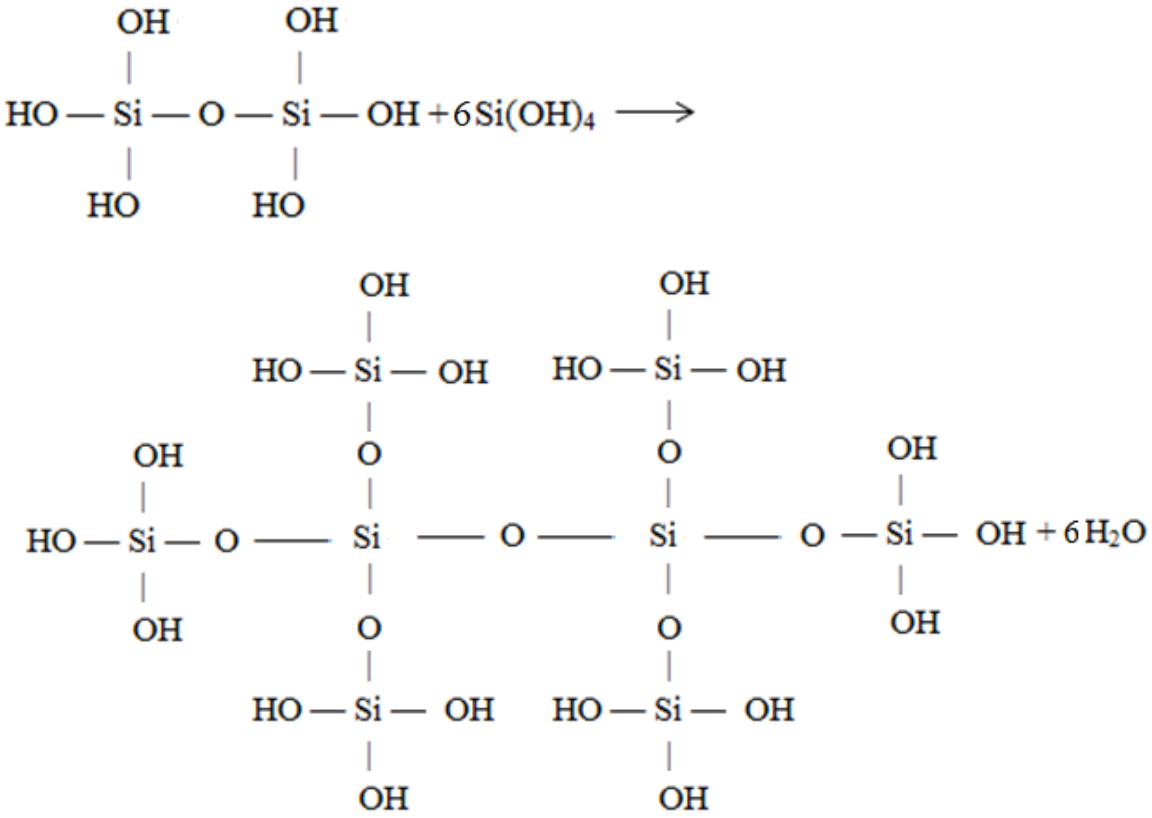
rate of hydrolysis is influenced by many factors, in particular, the presence of a catalyst, the temperature, pH, concentration of the solvent etc.

The total hydrolysis reaction (16.1) is a chemical reaction of the first order.

Hydrated silica tetrahedron formed as a result of hydrolysis interact in a condensation reaction forming =Si – O – Si= connection:

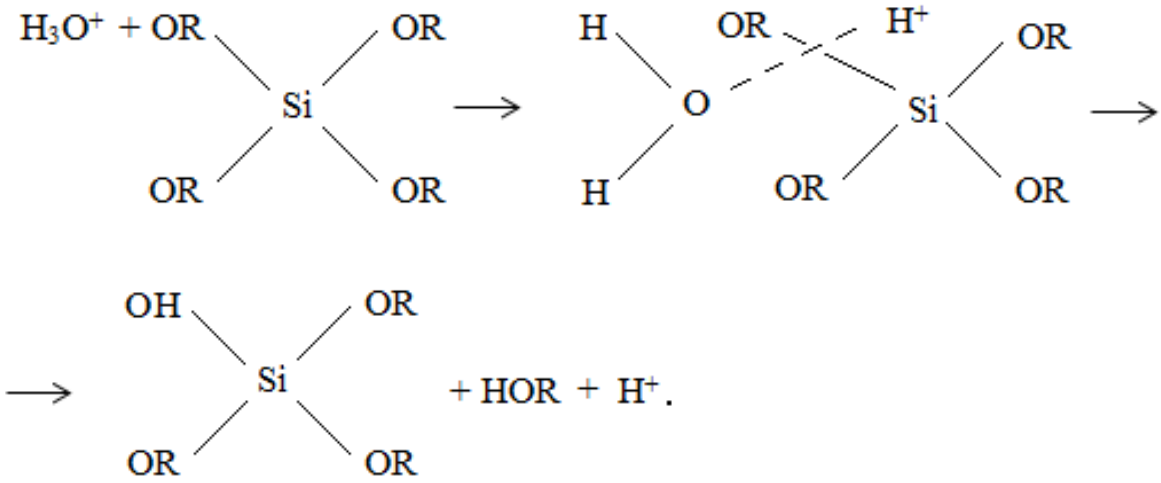


The connection between several tetrahedron is described as a chemical reaction of polycondensation, for example:



Water molecules formed by reactions of condensation and polycondensation, in turn, are involved in the hydrolysis of TEOS and the intermediate silanols. Thus, the reaction of hydrolysis and polycondensation occur simultaneously in the solution of $(C_2H_5O)_4Si$ - $C_2H_5OH$ - $H_2O$ which complicates the formation of polymeric structures greatly . The reaction rate of the polycondensation is also affected by many technological factors, the main ones are: the concentration of water and solvent, pH, type of catalyst.

The hydrolysis reaction of TEOS follows the mechanism of electrophilic substitution in acid catalysis:



The mechanism of the reaction of polycondensation of the hydrolyzed products can be represented by the following scheme [9]:

$$\equiv Si - OH + S \leftrightarrow \equiv Si - O^- + H^+S$$

$$\equiv Si - OH + \equiv Si - O^- \leftrightarrow [\equiv Si - O \cdots Si \equiv]^-$$

$$\downarrow$$

$$\equiv Si - O - Si \equiv + OH^-$$

$$OH^- + H^+S \rightarrow H_2O + S,$$

where S is a mixture of ethanol and water.

The researches of the kinetics of hydrolysis and condensation of TEOS process in the presence of large quantities of water, when using HCl as a catalyst and solvent – ethanol, shows that the TEOS hydrolysis reaction rate is significantly greater than the rate of polycondensation reaction products of hydrolysis.

In addition, reverse reactions can undergo in the system of TEOS – ethanol – water – acid, namely the reaction of esterification and depolymerization by cleavage of the linkages Si – O – Si by the action of both alcohol and water:

Consideration of the dependency of speed of hydrolysis, and depolymerization polikondensatsii on pH (Fig. 16.3) [9] shows that at low pH (acid hydrolysis), the rate of hydrolysis is the higher, and at polycondensation and depolymerization is least.

It was found experimentally that spherical structures consisting of primary particles with a diameter of 1-2 nm are formed in these conditions, especially at high dilutions of the system , which aglomerate in the secondary hydrated particles with lateral sizes of 4-6 nm [2].

Thus, it was stated, in conditions of high dilution polycondensation process similar to process, passing of silica polymerization in the water solution, resulting in formation of colloidal particles with smooth surface, but generated particles vastly saturated hydroxyls, which concentration increases from the center to the surface.

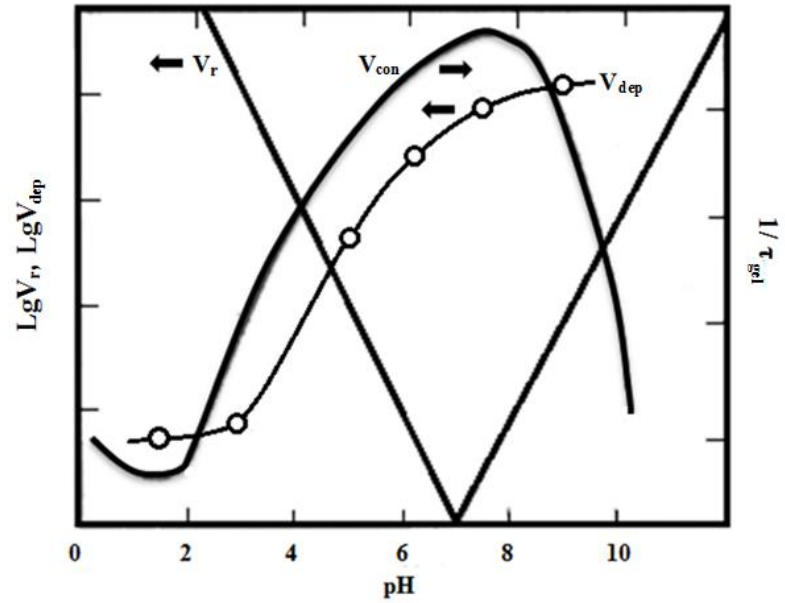



Fig. 16.3. Impact of pH on the rate of hydrolysis processes, condensation and depolymerization in the silica-colloidal system [9]: $V_r$ – rate of hydrolysis reaction; $V_{con}$ – rate of condensation reaction; $V_{dep}$ – rate of depolymerization reaction; $\tau_{gel}$ – time of gelation.

TEOS gydrolysis is more often carried in the environment of ethanol, ‘cause it good solvent for TEOS, while the last one water immiscible. Besides, the presence of ethanol in reaction environment allows moving reaction of hydrolysis to the right, promoting more complete reaction.

On the fig.16.4 is the diagram, showing the regions of TEOS and water’s immiscibility and allowing us to choose necessary concentrations of $Si(OC_2H_5)_4$ in ethanol [10]. This diagram is

convenient for practical using and give us opportunity to determine the equivalent amount (in %) of $SiO_2$ in the triple system $Si(OC_2H_5)_4 – H_2O – C_2H_5OH$.

In our diagram we can see that for the fully mutual solubility of system and if necessary carrying of TEOS hydrolysis with formation 3D volume silica particles, equivalent concentration $SiO_2$ in the system can't exceed 5-6 weight %.

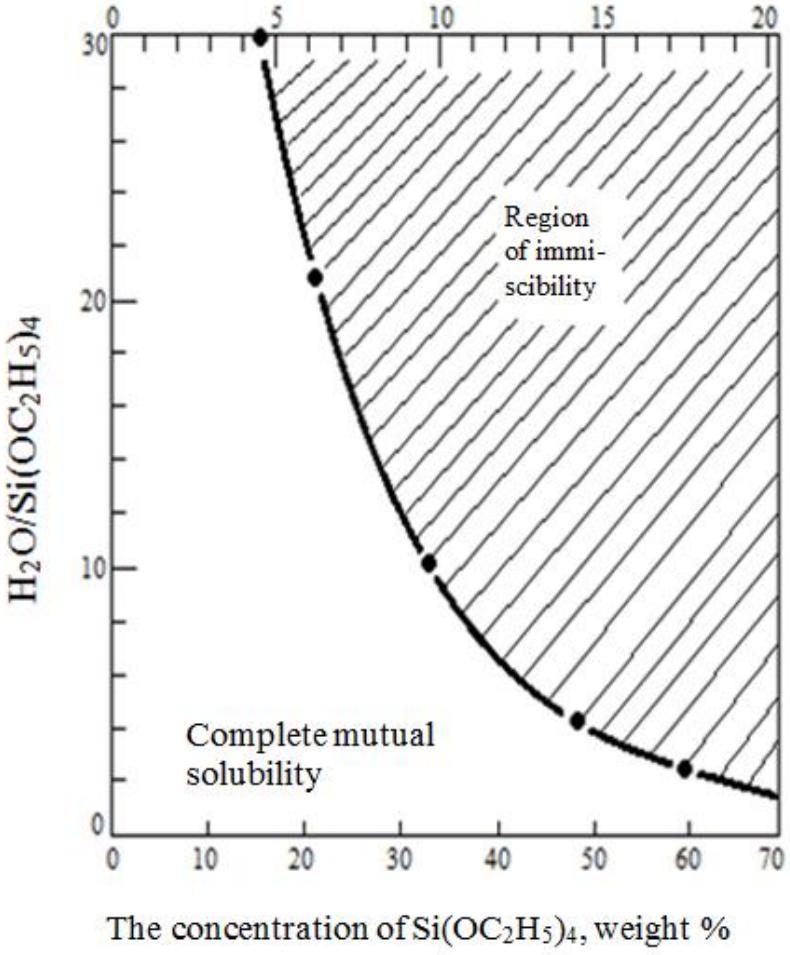



Fig. 16.4. - Mutual solubility as the ratio function of $H_2O/Si(OC_2H_5)_4$ and concentration of $Si(OC_2H_5)_4$ in the triple system – $Si(OC_2H_5)_4 – H_2O – C_2H_5OH$.

### 16.2.1 Gelation

Sol-gel transition represents multi-step hierarchical process, spreading from the atomic-molecular to macroscopic level.

Gelation represents synthesis of inorganic structure grid in the liquid – colloidal solution with room temperature or a bit higher, and generating solid body (a two- or three-phase gel) has an amorphous structure grid unlike solids resulting in crystallization of solutions and melts.

In the hydroalcoholic silica-colloidal solutions sol-gel transition generally set experimentally to viscosity changing over time. Gelation time ($\tau_{gel}$) linear increasing with the ratio of **ethanol: TEOS** , but exist best ratio **water : TEOS**, providing minimal $\tau_{gel}$ and value of this ratio increasing with ethanol content (from 4 to 6) [11].

By consideration behavior of colloidal silica particles in the liquid phase it's necessary to take fully into account pH impact on the gelation, which generally shown on the fig. 16.5 [9]. Curve ABC characterizes silica particles behavior when pH changes and without salts in this system.

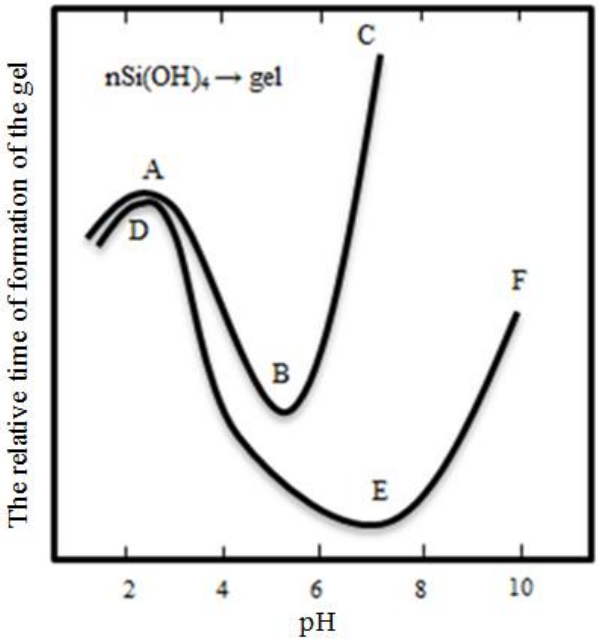



Fig 16.5 pH impact on the silica gelation without (curve ABC) and with (curve DEF) salts.

The stability of sols with maximum lifetime with best gelation period we can see about pH=1.5-3, while the minimum sols stability and fast flowing gelation process observed when pH=5-6. Higher than pH 7 gel doesn't form 'cause in this silica particles charged and only their growth occurs. Curve DEF shows the behavior of silica when in the system electrolytes with concentration higher than 0.2-0.3M are present.

When administered in stabilized (pH=1.5-3) sol hydroxide-ions and increasing pH (sol destabilization) process of agglomeration colloidal particles sharply accelerated. Clutch of silica particles occurs through formation of connections Si – O – Si. When silica particles collides with each other all neutral groups S—OH on their surfaces and ionize groups SiO mutually condensed with the connections formation Si – O – Si.

### 16.2.2 Maturation and drying

Important stage in the sol-gel process is maturation of gel – syneresis. It is accompanied by a gel structure seal and crowding out of solvent that amplified during drying with increasing of temperature. Due to capillary pressure and gel shrinkage in him arise significant internal tensions that lead to formation of fracture. Value of internal tensions can be reduced by using a chemical supplements controls process of drying or by using a drying in the conditions that weaken surface tension (hypercritical conditions) or in a result of formation bimodal pore structure with micro- and mesopores, formatting by addition ultrafine filler particles.

The main reason of gel destruction during the drying is capillary pressure ΔP proportional to energy specific surface area on the border liquid – solid and inversely proportional to radius pore r [12]:

$$\Delta P = \frac{2\gamma * cos\theta}{r}$$

where the θ – contact angle on the border liquid – solid – air.

According to classical concept theory of drying porous bodies there are three stages of drying:

**Stage 1 –** period constant speed of evaporation when decrease gel volume equally to liquid volume that lost by drying. Wherein gel grid deforms by a large capillary forces that cause gel shrinkage. That behavior is typical for gels with average pore size more than 20 nm.

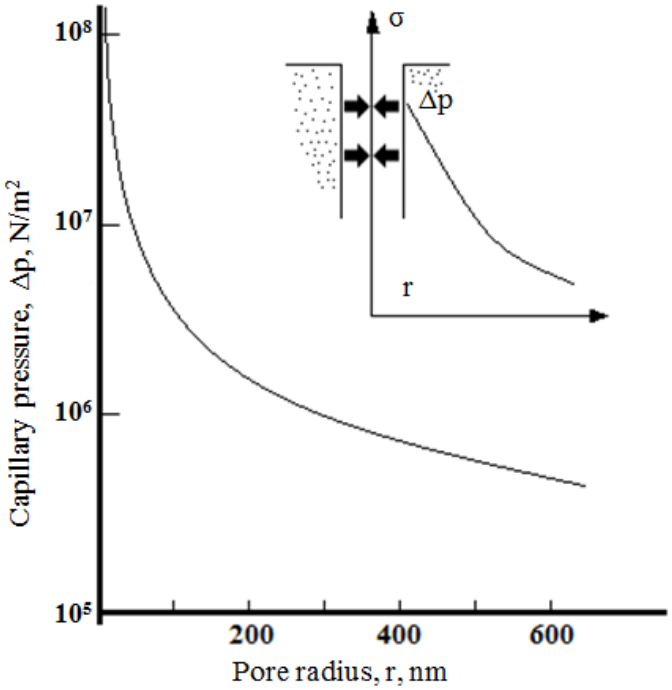



Fig.16.6 – capillary pressure as the function of pore radius; top part of figure – stress distribution in gel structure with pressure from center pore.

**Stage 2** begins when the "critical point" reached i.e. gel shrinkage is finished and strength of silicon-oxygen gel frame becomes sufficiently large in order to resist further shrinkage. In this "critical point" contact angle approaching to zero and liquid meniscus radius equated with pore radius. This condition creates the biggest capillary pressure and pores begin to be cleaned from the liquid. Since the speed of evaporation on stage 2 is decreases this period called "first period of speed drop"

**Stage 3 –** Third stage of drying occurs when the liquid surface of the pores is removed by vapor diffusion of their surface. Thus there is still greater retardation of liquid removal speed before reaching the equilibrium state (the "second speed drop period").

It should be noted that according to studies [13, 14] kinetics drying of gels formed from silicon alkoxides and having pores less than 20 nm, at the initial stage is substantially different from the classical one. It was found that at the first stage of drying of such gels are the evaporation rate of the fluid is not constant but decreases over time, and this dependence is associated with the initial pore size in the gel, as well as the drying conditions. Gels during drying in the first step is a significant reduction in pore size in gels and hence the vapor pressure decreases.

### 16.2.3 Sintering

At the final stage of the sol-gel process - sintering to solid state monolithic glassy body - the formation of defects related to expansion and cracking in the case of gels hydroxyls porosity less critical, as well as by recrystallization of the components of the gel matrix. Control physic-chemical processes taking place during sintering, possibly as stage the final stage by adjusting temperature conditions, the composition of the gas phase and in the initial stage by the introduction of fillers, modifying the three-dimensional porous structure of the silicon skeleton [9].

It is known that gels when heated seal takes place in several stages, depending on the temperature. In the range of 500-700 ° C. sintering is carried out by diffusion processes caused by condensation of silicon hydroxide chemical reaction. This step is followed by reduction of weight samples to T occurs with elimination of water as steam:

$\equiv Si - OH + HO - Si\equiv \rightarrow \equiv Si - O - Si\equiv + H_2O\uparrow$

The activation energy of sintering at this stage is determined by the activation energy and the dehydration process is about 22-24 kJ / mol. Shrinkage of the sample is due to the coalescence of the gel particles by thermal polycondensation, and depends on the composition of the gas atmosphere, the initial water content and texture of the gel.

At higher temperatures, in the range 750-1000 ° C sintering is carried out through two mechanisms: a diffuse and viscous flow, and a greater contribution to the process of making a seal mechanism of viscous flow and to describe this stage of the model of Frenkel:

$$\frac{\Delta L}{L} = \frac{\delta}{2r\eta}\tau$$

where δ - the density of the substance; r - radius of the particles; η - viscosity; T - interfacial energy.

Assuming that η = Aexp (Q / RT), heating rate V = dT / dτ, a σ and A - constant value, the equation (1,5) can be transformed in this way:

$$\frac{\Delta L}{L} = \left(\frac{\sigma R T^2}{2rAVQ}\right) \cdot exp\left(\frac{-Q}{RT}\right)$$

At temperatures above 1000 ° C sintering mechanism becomes purely viscous, and if up to 900 ° C the gel particles do not change at T> 1000 ° C they are growth.

In the work [15] was detected and experimentally confirmed special behavior hydrated xerogels synthesized by alkoxides hydrolysis with using acid catalyst.

On the curve of linear shrinkage (fig. 16.7A) detected three plots with different tilt angles that interpreted in the following way. On the first stage of compaction polycondensation is prevailing reaction and forms triple gel grid (150-550 ºC) at the same time the compaction of silica-oxygen carcass take place, with shrinkage that leads to restricting primary particles into the secondary agglomerates with higher coordination.

Silica-oxygen basis of xerogel consist of particles which density decreases from center to peripheral part. Researches for small-angle dispersion of X-ray radiation confirms formation of loose polymeric highly hydrated particles, especially in conditions of high concentration ethanol that prevents the formation of dense colloids and growth of isthmus between particles of silica.

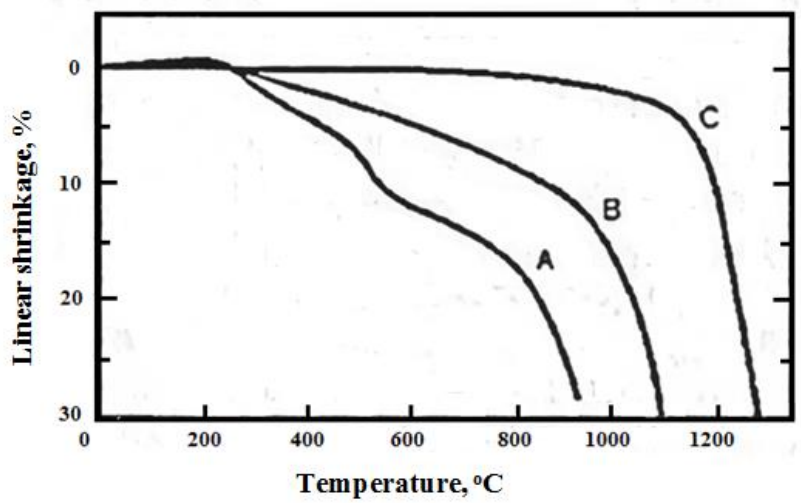



Fig. 16.7. Dependence of linear gel shrinkage on processing temperature; A – alkoxides gel (acid catalysis); B - alkoxides gel (alkaline catalysis); C – colloidal gels.

At the stage 2, structural relaxation occurs with an increase in the basis density of the gel (nanolevel). The basis density of the gel in stage 2 is approximately 0.78 relative to the basis density of gels formed from aquasols [15]. Curve C in fig. 16.7 well confirms the fact that up to

temperatures of approximately 1100 °C in aquagels there is no structure relaxation, and the mechanism of viscous flow activated at temperatures above 1100 °C, leads to the formation of high-density quartz glass at 1200 °C.

Thus, it has now been unequivocally established that for gels obtained from alkoxides with small and medium amounts of water for hydrolysis, the structure of the silica-oxygen scaffold is mainly formed as a product of a polycondensation reaction starting in solution and continues during gelling and conversion gel → glass. In solution, these reactions lead to the formation of a metastable, hydroxylated disordered network, which under high temperatures undergoes strong structural relaxation at the nano- and micro-dimensional level. Depending on the rate of heating there are several options for transforming this metastable grid into a glassy state (glass transition temperature $T_g$) i.e. a significant structural relaxation of the silica-oxygen basis occurs, which can lead to the destruction of the gel and at high rates of temperature rise the structural relaxation slows down, which can lead to the formation of a large-porous disordered structure with smaller shrinkage at the macrolevel (curve b).

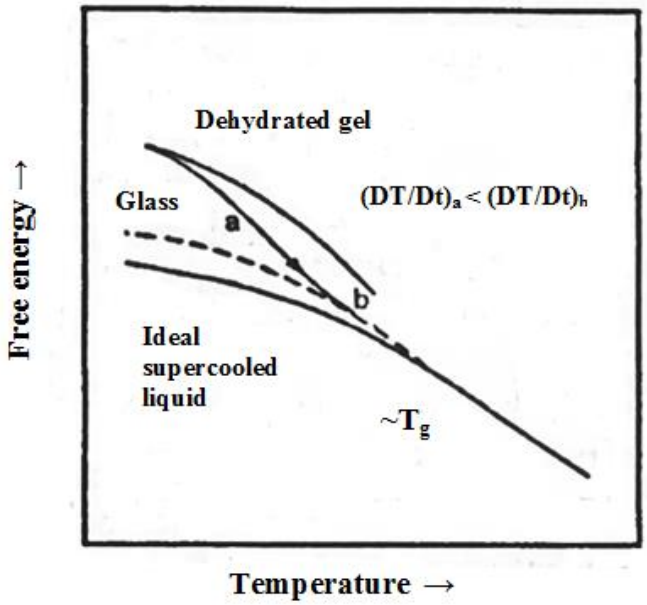



Fig. 16.8. Transformation metastable structure of gel into the glassy state.

## 16.3. Synthesis features of monolithic materials in a modified sol-gel process

### 16.3.1 Synthesis of bulk-molded optical materials

To overcome the disadvantages of the "classical" alkoxide sol-gel synthesis process of glass materials [4] was proposed several options for its modification (fig.16.9).

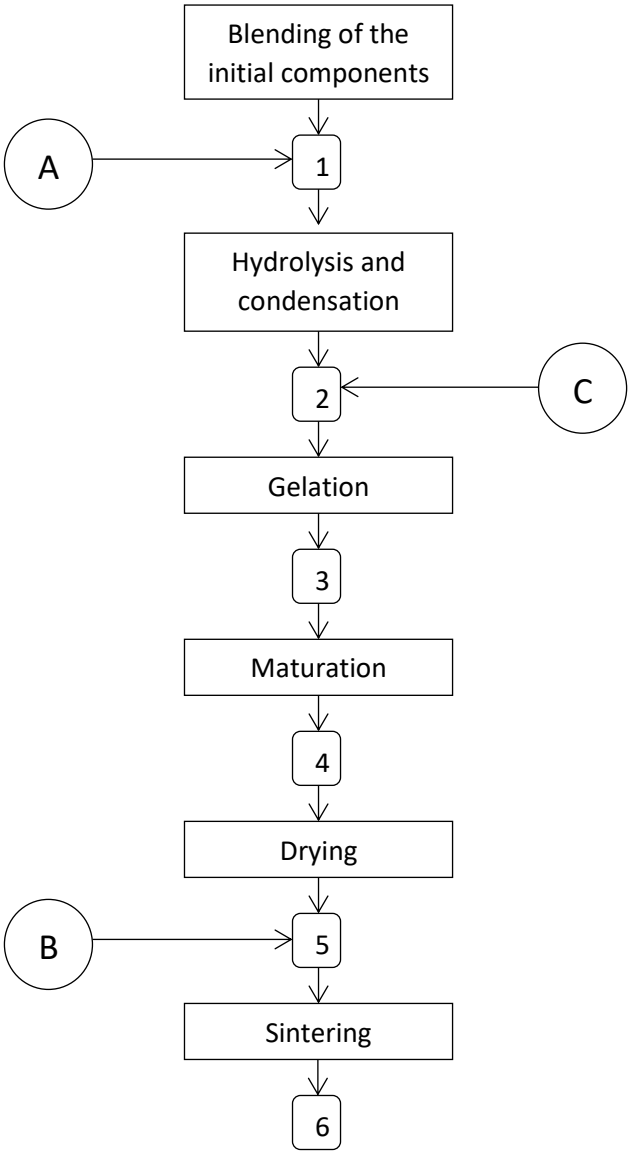



Figure 16.9. Modification variants of the of the "classical" alkoxide sol-gel process: A – the addition of CARD; B – "Combined" sol-gel process; C – addition of silica fillers: 1 – solution; 2 – sol; 3 – alkogel; 4 – ripe gel; 5 – xerogel; 6 – glassy material

Basic modification of the "classical" alkoxide sol-gel synthesis process of glass materials:

A. The sol-gel process with chemical additives regulating drying (CARD).These additives are added to the initial mixture to form interconnected pores in the xerogel homogeneous system, to improve drying conditions of gels, producing porous and solid bulk glass materials of sufficient size without cracking and foaming during heat treatment. It reagents having low values of surface tension and does not adversely affect the processes of hydrolysis, polycondensation and maturation. Among the most common are: formamide [17, 18] and dimethyl-formamide [19], glycerin [20], ethylene [21] and oxalic acid [16].

However, using CARD in the alkoxide sol-gel process was unable to obtain larger glass, moreover, the drying processes gels with specific pore surface reaches 800 $m^2/g$, are very long (more than one month) and organic materials administered in colloidal systems in significant amounts are difficult to remove and contaminate the formed glassy material.

B. A method of a dual processing (combined sol-gel process) has been proposed A. Rabinovich and others [22] and was developed in corporation AT&T (USA) [23, 24] in order to obtain high-purity quartz preforms and support tubes for drawing optical fibers. This method involves the formation of the silica gel with the specified parameters of the silicon alkoxide by hydrolysis and polycondensation of an acid, followed by heat treatment and mechanical dispersion

of the active powder to condition $SiO_2$. The fine powder scattered on fractions, used for the preparation of the water-based slurry.

Glassy articles according to this process are formed by centrifugal casting methods similar to existing technologies in the ceramic technologies, followed by drying and sintering macroporous preforms at elevated temperatures.

The disadvantages of this method are: the technological complexity of the process; the necessity to introduce mechanical crushing operations; high sintering temperature (1700 °C) to a state of porous monoliths optical transparent quartz glass.

C. Adding to colloidal system formed from silicon alkoxides, spherical particles of $SiO_2$, obtained by alkaline hydrolysis of the silica gel dispersed [9,25]. This variant enables significantly increase the solids content of solspurposefully form mesoporous structure of the gel and xerogel, get strong silica gel with the desired size and complex shapes. By adding silica fillers in alkoxide-formed sols, mixed colloids with a wide particle size (polydisperse) were obtained. They have specific structural and rheological properties. Addition of aerosil and dry gels into the sol necessitates the use of special activation methods – ultrasonic dispersion and stabilization.

Quartz glass with zero-porous was synthesized by using a method of filler adding in alkoxide-forming sols obtained by acid hydrolysis TEOS with the following ultrasonic homogenization of mixed sol [21]. As a solid filter, aerosol A-175 was introduced into the solutions at a molar ratio to TEOS from 0 to 1,5 the pH of the sols was adjusted to 5 units and the gelling and drying process was carried out at room temperature in closed glass and Teflon vessels. The drying was controlled by changing the rate of evaporation from the surface of the gel and the shrinkage of the gels after drying was 10-20% of the volume of the initial solutions, depending on the content of the aerosol in them.

In the work it was established that when the aerosol is introduced, the yield of the final product (glassy $SiO_2$) reaches 25%, i.e. increases more than 2 times in comparison with the alkoxides-forming gel while maintaining its monolithicity, and the glass formation occurs already at 800 $^oC$, as evidenced by the identity of the IR-specter of heat-treated samples and quartz glass. However at a temperature of about 1100 $^oC$, crystallization of the glass is observed with the release of α-cristobalite, which is facilitated by the slow heating and prolonged isothermal exposures. The resulting glass samples had small dimensions (up to 20 mm in diameter), and the mass content of OH-groups could not be reduced less than 8 - $10^{-2}$% , which causes foaming and glass bubbling when heated above 1150 $^oC$.

These shortcomings were overcome in works [22, 26], in which the process of formation of a monolithic quartz glass having a low content of OH-groups (<5 ppm) is describe. The sols were prepared by mixing TEOS, water, hydrochloric acid and aerosol in a molar ratio 1: 22: 0.01: 1.25, followed by gelation at pH 3,7 – 4,5. The gels were dried at a temperature of 60 $^oC$ after 5 days of maturation in the syneresis fluid. Typical physical parameters of xerogels were as follows: pore specific surface area of about 440 $m^2/g$; The pole volume is 0,87 $cm^3/g$; The average pore radius is 40 A. The analysis with the help of SEM of the cleaved surface of the gel showed the formation of a layered type structure. Transparent quartz glass was obtained by sintering the gel at a temperature of about 1500 $^oC$ in a helium atmosphere, but a serious problem, as the authors of [26] pointed out, is the cracking on the layer that is observed during compaction of gels during heat treatment. To reduce the probability of cracking and to maximize the removal of OH-groups from the pore surface, a rather complex scheme of dehydroxylation was developed, including the stages of heat treatment in vacuum, in atmosphere ($CCl_4$, $N_2$) and dechlorination in an atmosphere of dried and purified oxygen (Figure 16.10).

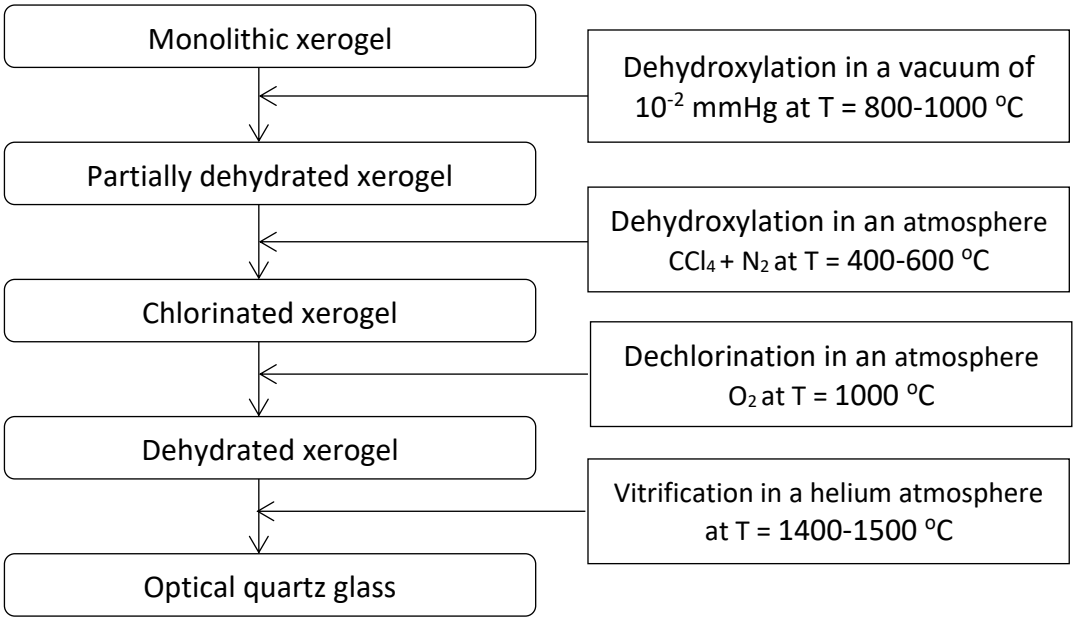



Fig. 16.10. Scheme dehydroxylation of xerogels in the production of optical quartz glass with the use of fillters.

Various variants of modifying alkoxides sols by the addition of ultradisperse silica particles have been studied by other authors, where spherical powders $SiO_2$ synthesizes by alkaline hydrolysis of TEOS [27] or aerosils obtained by plasma hydrolysis of $SiCl_4$ were used as filters [28].

Thus, it seems that the most effective process for the synthesis of monolithic quartz glass of sufficient size and complex forms is the alkoxide sol-gel process modified in the part of the formation of mixed sols by the introduction of ultradispersed fillers, which leads to an increase in the yield of vitreous $SiO_2$, a decrease in the probability of cracking and foaming of gels in the course of heat treatment, and also allows to obtain alloyed glassy materials.

**REFERENCES**

1. DISLICH H. *Glassy and Crystalline Systems from Gels: Chemical Basis and Technical Application* // J.Non-Cryst.Sol. – 1983. – Vol. 57. – P. 371-388.

2. HENCH L., WEST J. *The Sol-Gel Process* // Chem.Rev. –1990. – Vol. 90. – P. 33-72.

3. DISLICH H. A *New Route to Multicomponent Oxyde Glasses* // Angew. Chem. Int. Ed. Engl. –1971. – Vol. 10. – P. 363-370.

4. SAKKA S. *Sol-Gel Synthesis of Glasses: Present and Future* // Am. Ceram. Soc. Bull. – 1985. – Vol. 64. – P. 1463-1466.

5 RABINOVICH E.M., JOHNSON D.W., McCHESNEY J.B. e.a. *Preparation of High-Silica Glasses from Colloidal Gels: I. Preparation for Sintering and Properties of Sintered Glasses* // J.Amer.Ceram.Soc. – 1983. – Vol. 60. – №10. – P. 683-688.

6. CLASEN R. *Preparation of High-Purity Silica Glasses by Sintering of Colloidal Particles* // Glastech.Ber. – 1987. – Vol. 60. – № 4. – P.125-132.

7. CLASEN R. *Preparation and Sintering of High-Density Green Bodies to High-Purity Silica Glasses* // J.Non-Cryst.Sol. – 1987. – Vol. 89. –P. 335-344.

8. SCHERER G.W. *Recent Progress in Drying of Gels* // J.Non-Cryst.Sol. – 1991. – Vol. 147-148. – P. 363-374.

9. BRINKER C.J. *Hydrolysis and Condensation of Silicates: Effect on Structure* // J.Non-Cryst.Sol. – 1988. – Vol. 100. – P. 31-50.

10. YOLDAS B.E. *Modifications of Molecular Size and Structure During the Hydrolytic Policondensation of Metal Alkoxides* // Desing. of New Materials, Ed. D.L.Cocke, Clearfold A., Plenum Press. – 1987. – №4 . – P. 13-37 .

11. KINOUCHI FILHO O., AEGERTER M.A. *Rheology of the Gelation Process of Silica Gel* // J.Non-Cryst.Sol. – 1988 . – Vol. 105. – P. 191-197.

12. ZARZYCKI J. *Monolithic Xero- and Aerogels for Gel-Glass Process* // Ultrastruct. Process Ceram., Glasses and Composites., New York e.a., 1984. – P. 27-42.

13. SCHERER G.W. *Recent Progress in Drying of Gels* // J.Non-Cryst.Sol. –1991. – Vol. 147-148. – P. 363-374.

14. SCHERER G.W. *Stress and Fracture During Drying of Gels* // J.Non-Cryst.Sol. – 1999. – Vol. 121. – P. 104-109.

15. BRINKER C.J., ROTH E.P., TALLANT D.R. e.a. Relationships between Sol to Gel and Gel to Glass Conversions: Structure of Gels During Densification // Sci. Ceram. Chem. Process.: Proc. 2ND Int. Conf. Ultrastruct. Process. Ceram., Glasses and Compos., Palm Coast, Fla, Feb.25 – March 1, 1985, N.Y. e.a. – 1986 . – P.37-51.

16. HENCH L.L. *Use of Drying Control Chemical Additives (DCCAs) in controlling Sol-Gel Processing* // Sci. Ceram. Chem. Process.: Proc. 2nd Int. Conf. Ultrastruct. Process. Ceram., Glasses and Compos., Palm Coast, Fla, Feb. 25- March 1, 1985 , N.Y. e.a. – 1986. – P.52-64.

17. ORCEL G., HENCH L., ARTAKI I. e.a. *Effect of Formamide Additive on the Chemistry of Silica Sol-Gel. II. Gel Structure* // J.Non-Cryst.Sol. – 1988. – Vol.105. – P.223-231.

18. KAJI H. , NAKANISHI K. , SOGA N. *Formation of Porous Gel Morphology by Phase Separation in Gelling Alkoxy-Derived Silica. Affining between Silica Polymers and Solvent* // J.Non-Cryst.Sol. – 1995. – Vol. 181. – P. 16-26.

19. SAKKA S., KOZUKA H, ADACHI T. *Stability of Solutions, Gels and Glasses in the Sol-Gel Glass Synthesis* //J.Non-Cryst.Sol. – 1988. – Vol. 102. – P. 263-268.

20. HENCH L.L. *Use of Drying Control Chemical Additives (DCCAs) in Producing Gel Monolith* // Glass. Curs. Issues. Proc. NATO Adv. Study. Inst. Tenerife., Apr. 2-13, 1984, Dordrecht e.a. – 1985. – P. 259-262.

21. LUO S., TIAN K. Low Temperature Synthesis of Monolithic Silica Glass from the System $Si(OC_2H_5)_4 – H_2O – HCl – HOCH_2CH_2OH$ by the Sol-Gel Method // J.Non-Cryst.Sol. – 1988. – Vol. 100. – P. 254-262.

22. RABINOVICH E.M., McCHESNEY J.B., JOHNSON D.W.Jr. *Gel-Derived Glasses for Optical Fibers Prepared from Alkoxides and Fumed Silica* // Sci. Ceram. Chem. Process.: Proc.2nd Int. Conf. Ultrastruct. Process. Ceram., Glasses and Compos., Palm Coast, Fla, Feb. 25 - March 1, 1985, N.Y. e.a. – 1986. – P. 208-216.

23. McCHESNEY J.B., JOHNSON D.W., FLEMING D.A. e.a. Hybridezed Sol-Gel Process for Optical Fibers // Electron.Lett. – 1987. – Vol. 23. – № 19. – P. 1005-1007.

24. McCHESNEY J.B., JOHNSON D.W.Jr. *Large Silica Bodies by Sol-Gel for Production of Optical Fibers* // Abstr. Of Int. Conf. Optical Fiber Communication (OFC'97), 3-10 May. – 1997. – Tu B3.

25. SHIBATA S., KITAGAVA T., HORIGUSHI H. *Fabrication of Fluorine-Doped Silica Glasses by the Sol-Gel Method* // J.Non-Cryst.Sol. – 1988. – Vol. 100. – P. 269-273.

26. KUNDU D., De G., KARMAKAR B. Monolithic Silica Glass from Hybrid Silica Gels with Layered Structure // Abstr. of 6th Int. Workshop on Glasses and Ceramics from Gels, Seville (SPAIN), Oct. 6-11. – 1991. – P. 1309.

27. RUVOLO E.C.Jr., BELLINETTI H.L., AEGERTER M.A . Silica Glass: Preparation and Characterisation //J.Non-Cryst.Sol. – 1990. – Vol. 121. – P. 244-249.

28. ZHONGCHENG W., RENJIE T., QING S., MEI E C. Synthesising the F-Doped Silica Tube Using the Sol-Gel Process //Proc. of XVII Int. Congress on Glass, Vol. 4., Beijing. – 1995. – P. 544-547.

## 17. Physics of Carbon Low-dimensional Systems and Device Structures

Production and research of low-dimensional systems in their physical, chemical and biological aspects, as well as innovations emerging from them are a concern of the whole world of engineering [1–8]. The article provides a brief overview of the theoretical and experimental studies of physical properties of carbon low-dimensional systems, as well as elements of the devices on their basis. Results recently obtained in the Physics Department of BSU and associated working hypotheses are presented (see also [9–15]). In conclusion the outlook of work for some low-dimensional systems are discussed.

First, let us explain terminology used in the text and illustrations (see for example, [16, 17]): allotropy, nanotechnology, low-dimensional system, nanomaterials and nanoindustry.

*Allotropy* is a phenomenon in which substances of the same chemical elements in one of the state of matter (solid, liquid, gaseous or plasma) exist as two or more modifications of the structure with different properties. The term allotropy applies only to substances of the same atoms (e.g., carbon atoms), regardless of the state of matter of these substances, in contrast to the term "polymorphism" which is used only in relation to a solid state of matter of substances regardless of whether the substance is simple (composed of atoms of the same chemical element) or complex (consisting of atoms of different chemical elements). The terms "allotropy" and "polymorphism" are identical for simple solids (e.g., white and red phosphorus, gray and white tin, graphite and diamond).

*Nanotechnology* is a set of methods of production, diagnostics and application of objects with a characteristic size of about 0.1 nm up to 100 nm. By varying the size, shape and relative positioning of nanoobjects, even while maintaining their chemical composition, it is possible to control their properties (e.g., melting point, solubility or transparency). The following methods are used to produce objects in nanotechnology: assembling objects by selective transfer of single atoms from the probe (probe tip in a scanning probe microscope) on the surface of the studied object and creation of "atomic patterns" — information-carrying medium (i.e., in a broad sense an art of building [18]), selective removal of atoms from multicomponent compounds by ion beams [19], as well as electron beams [20] and finally the traditional cut-off of excess material (from the "bulk" to the final product), for example by ablation — the removal of a surface layer of a solid when exposed to laser radiation and/or plasma flux [21–23].

*Low-dimensional systems* is a consolidated system of many particles, the length of which along at least one spatial direction (in the Cartesian coordinate system) is comparable to the one of the parameters of the dimension of length, which characterize the states and/or processes in the system (see, e.g., [24]). To such parameters we may attribute the mean free path of the particles (between scattering, recombination or decay events), the average wavelength of the particles (after de Broglie), and others. The movement of delocalized (tentatively "free") particles due to quantum-size effects (in one, two or three directions) leads to a qualitative restructuring of their energy as a function of (quasi)momentum and spin (intrinsic magnetic moment of the particle) in the quantization direction. (The concept of quasimomentum is a generalization of the concept of momentum for the case of motion of a particle in space with potential energy periodically dependent on the coordinate [25].) If the energy spectrum quantization of (quasi)particles in the system takes place in one direction only, the system is called two-dimensional, if in two directions only — one-dimensional, and if in all three — zero-dimensional. Low-dimensional systems can be *two-dimensional* (graphene surface, interface or stacking fault of atomic planes in the crystal), *one-dimensional* (carbine — a chain of carbon atoms, carbon nanotube, quantum thread or dislocation in crystal), and *zero-dimensional* (fullerene, atomic cluster, quantum dot and pore in a continuous medium (matrix) or a point defect in the crystal lattice). Low-dimensional systems can be *fractal* (with fractional dimensionality) [26]. To the stationary low-dimensional systems belongs a semiconductor super-lattice consisting of crystalline layers of different chemical

composition, crystallographic orientation or structure, whose period in one direction is less than the mean free path of electrons and/or holes. To the non-stationary low-dimensional systems belong solitons, the domains of the electric field in the Gunn diodes, "threads" of current and light in semiconductors and others. The moving elements of low-dimensional systems in the condensed state are quasiparticles [photons, conduction electrons, holes (electron vacancies in the covalent chemical bond) trions, magnons, etc.]. The excited states of the system of particles can be represented as a set of quasiparticles, only if the interaction of the quasiparticles is weak, as in a dilute gas.

Materials consisting of low-dimensional systems are called *nanostructured* (*nanostructural*) materials (or briefly, *nanomaterials*). The main structural elements of the nanomaterials are atomic clusters, granules, pores of nanometer size (conventionally between 1 and 100 nm) in solid matrix, fibers, tubes, layers, etc. Structural elements (low-dimensional systems) add distinctions to nanomaterials if compared to the uniform materials of the same chemical composition.

*Nanoindustry* is a branch of industry based on the application of nanotechnologies. It allows to create nanomaterials (on the basis of the low-dimensional systems), as well as elements of tools, devices, mechanisms and machines with the characteristic size of the functional components conventionally from 1 to 100 nm. Nanoindustry involves the complex of technical processes, metrology services, as well as institutions, organizations and enterprises engaged in the transfer of knowledge, skills and know how (innovations) from researchers to engineers, from engineers to manufacturers and then carries out the commercialization of science-intensive products.

## 17.1. Allotropes of carbon.

Agglomeration of carbon atoms in consolidated system can take part with formation several stable (under certain conditions) allotropic forms (see, e.g., [27–30]; Fig. 1). For carbon it is typical to change type of hybridization of valence orbitals (single-electron wave functions of atoms) in the process of formation of a C–C-bonds and bonds with the atoms of other chemical elements (primarily hydrogen, iron, silicon, nitrogen and boron). [Changing the valence of atoms depending on their environment allows, in particular, "reinforce" three-dimensional (bulk) materials with fractal systems [31–33].] For example, graphane — graphene passivated on both sides with hydrogen is a two-dimensional counterpart of the diamond lattice with $sp^3$-hybridized carbon atoms [34]. (At room temperature graphane is a dielectric, unlike graphene which is a conductor of electricity.)

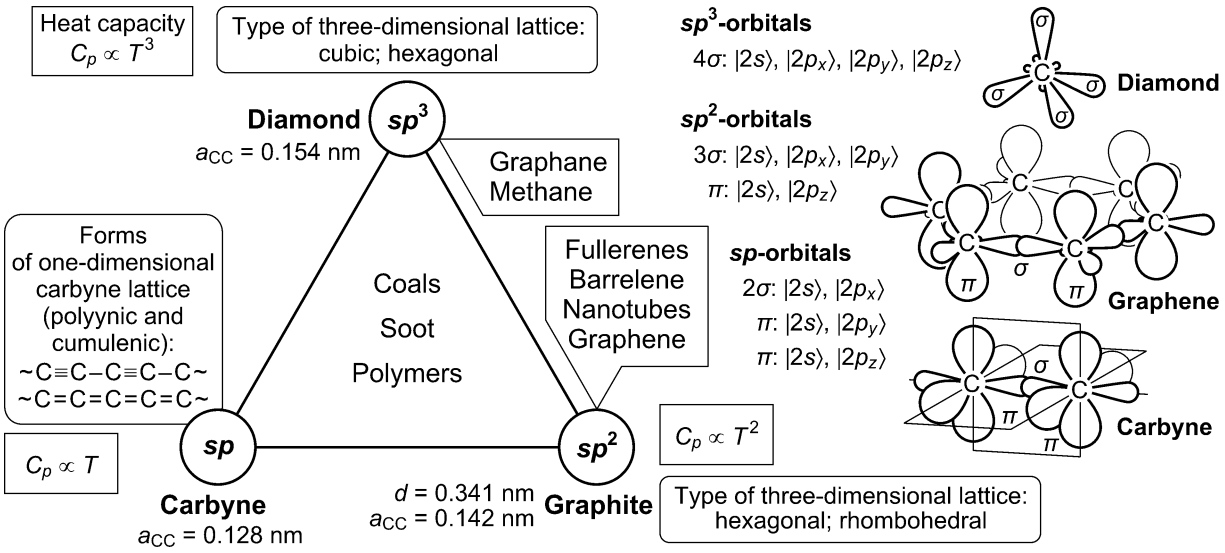



Fig. 1. Allotropes of carbon associated with changes in hybridization valence orbitals (see also [27–30]). Here $a_{CC}$ is the distance between the nearest atoms, $d$ is the distance between atomic

planes in the graphite, σ and π are the hybrid orbitals of carbon atoms, $C_p$ is the thermal capacity at constant pressure, $T$ is the temperature; ∝ is a symbol of proportionality. (Graphane and methane are compounds composed of carbon and hydrogen.)

### 17.2. Zero-dimensional systems.

For the fullerenes (see, e.g., [13, 35]), as well as for the close to them types of low-dimensional systems the approaches for their classification were developed [36] and a number of practical results in the synthesis technology were obtained (see Fig. 2a).

In the work [37] using the method of molecular orbital the totally symmetric vibrations of fullerene $C_{60}$ were studied. Comparison of the calculations with the data on Raman scattering in fullerites [38] shows the possibility of the Fermi resonance (interaction between modes of vibration of the atoms due to their anharmonicity) for two types of totally symmetric vibrations [39] of the $C_{60}$ cage molecule.

A model [13, 40, 41] of the fullerene $C_{20}$ formation and of its subsequent growth in higher fullerenes and carbon nanotubes by merging $C_{10}$ rings and $C_{10}^{*}$ stars (predicted by authors isomer in the form of five-pointed star consisting of ten carbon atoms) present in the carbon plasma was proposed (Fig. 2b).

The modeling of configurational transitions in atomic clusters is a challenging task [42]. In particular, this is due to the fact that low-dimensional systems are characterized by both spontaneous and induced point symmetry breaking (see, e.g., [43, 44]). For example, in the work [45] the analogue of the dynamic Jahn–Teller effect in the dication carbododecahedron $C_{20}^{2+}$ due to the repulsion of uncompensated positive charge on the carbon atoms (the Coulomb distortion) was discovered by calculations. The $D_3$-symmetry if the ground state of the $C_{20}^{2+}$ was found and its infrared vibrational spectrum was calculated. The dynamics of the fission of the charged $C_{60}$ accompanied with the emission of $C_{20}^{2+}$ were studied [46] by the quantum-chemical calculations.

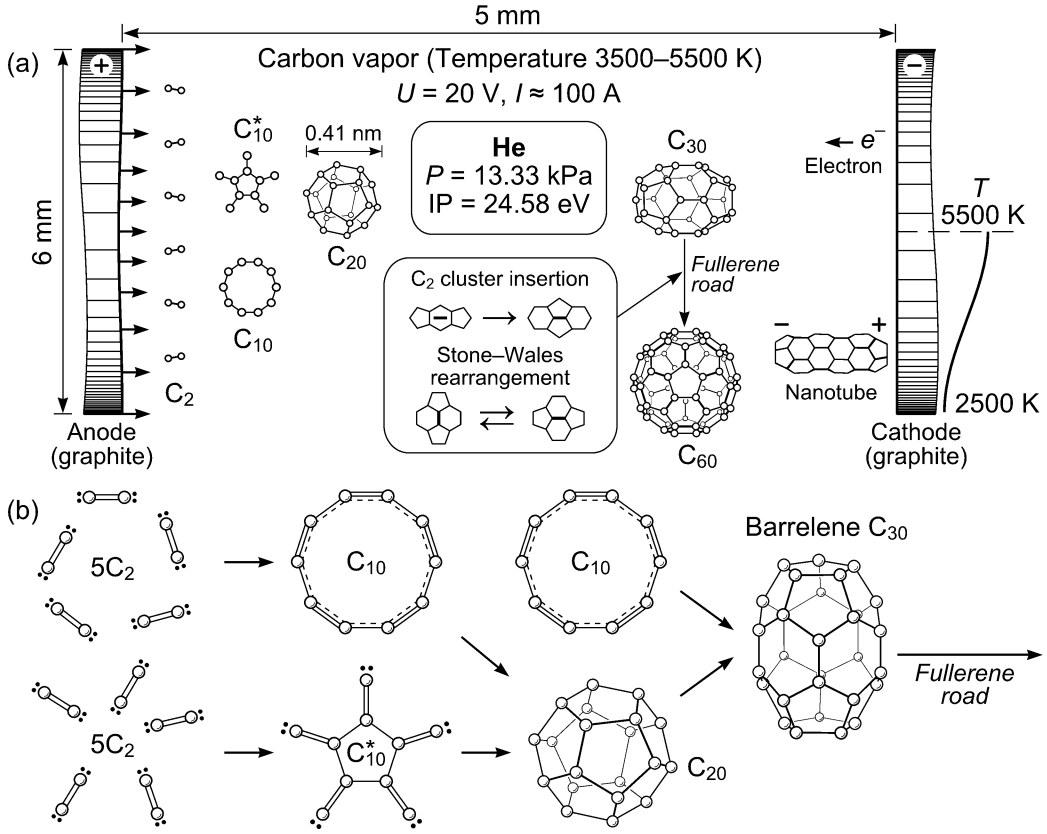



Fig. 2. Scheme (after [13]) of the fullerene formation in the carbon plasma (a) and the possible pathway of the creation of the molecules $C_{20}$ and $C_{30}$ (b). Here are shown the typical size of graphite electrodes in the form of discs, the voltage *U* and the current *I*, plasma temperature *T* and distribution of *T* at the surface of the electrodes, partial pressure *P* of helium and ionization potential IP of He atoms

In the work [47] structural and energy characteristics of the smallest magnetic endofullerene Fe@$C_{20}$ were calculated by the methods of quantum chemistry. It was stated that the ground state of Fe@$C_{20}$ is a septet, and the magnetic moment of the Fe@$C_{20}$ was estimated to be 8 Bohr magneton.

Here we note that the schemes of application of carbon fullerenes in hot-electron magnetic tunnel transistor [48] and elements of quantum computers [49] were proposed.

The application of fullerenes in the industry is just in its beginning. For example, in the work [50] it was found that small concentrations of a mixture of fullerenes $C_{60}$–$C_{70}$ (up to 0.255 wt.%) in the graphite feed material with Ni–Mn as metal catalysts reduce the activation energy of the phase transition from graphite to diamond by 1.6 times (diamond synthesis was carried out in the temperature range of 1600–1800 K at a pressure of 5 GPa).

### 17.3. One-dimensional systems.

Carbon nanotubes (see, e.g., [51, 52]) are objects of both fundamental research and applications. Single-walled nanotubes are represented by two indices $(m, n)$, which determine their chiral vector: $\mathbf{C}(m, n) = m\mathbf{a}_1 + n\mathbf{a}_2$, where *m* and *n* are natural numbers including zero, $\mathbf{a}_1$ and $\mathbf{a}_2$ are the basis vectors of hexagonal graphene lattice. When we roll the graphene plane into the nanotube the beginning and the end of chiral vector are "glued" together. Carbon nanotubes of the *zigzag* type have indices $(m,0)$ and of the *armchair* type — $(m,m)$.

In the work [53] the single-walled carbon nanotubes filled with potassium atoms in which every 10 carbon atoms capture one electron of doping atoms were theoretically analyzed. It is assumed that the positive charge is distributed inside the volume of the nanotube and the negative — on its surface. In the single-electron approximation, the dependence of the Fermi quasimomentum of conduction electrons inside the nanotube on their concentration and tube radius for an arbitrary number of sub-bands of the transverse motion was found. On the basis of this approach the graduated dependence of the conductivity of nanotube metallic subsystem on its radius was calculated, which subsequently found support in [54]. This feature of the doped nanotubes can be used in the design of logic elements for computers [8].

The characteristics of a carbon $(8, 8)$ nanotube encapsulating a single endofullerene $Fe@C_{20}$ were studied [47, 55]. A magnetic nanorelay scheme based on the two carbon nanotubes filled with magnetic endofullerenes was developed. Nanorelay is closed as a result of bending of nanotubes under the magnetic force. The performance of the nanorelays based on $(8, 8)$ and $(21, 21)$ nanotubes fully filled with $Fe@C_{20}$ was calculated.

In the work [56] a model of the electronic structure of the nanoobject (a stack with a stacking fault) made of the lead phthalocyanine (PbPc) macromolecules was proposed. The calculation by molecular orbitals method shows the transfer of electric charge of $-0.6e$, where $e$ is the elementary charge, from Pb atoms to the macrocyclic ring Pc ($C_{32}H_{16}N_8$) of the molecule. The model explains observed [57, 58] effect of an electric switch in the PbPc films as a "turning inside out" [59] of PbPc molecules in a stack.

Dimerization of the C–C bonds in single-walled carbon nanotubes predicted and quantitative estimates of the Peierls distortions obtained [60]. In the free electron approximation [61] the model of the electronic structure of single-walled nanotubes taking into account the finiteness of the carbon skeleton thickness was developed. Within the tight-binding model for the electronic spectrum of single-walled carbon *armchair* nanotubes it is shown [62] that at the temperature of liquid helium the rate of energy transfer from the lattice conduction electrons (to carbon nanotubes frame) is determined by the interaction of $\pi$-electrons with the twiston mode of the phonon spectrum. The calculation shows the prospects of creating the bolometer on the basis of those "metal" carbon nanotubes. (The calculated time constant for such a bolometer is about a microsecond.)

A new way to control the movement of nanoelectromechanical systems based on carbon nanotubes proposed in [63]. Namely, as a result of chemical adsorption of H and F atoms on the open ends of a single-walled nanotube it gets an electric dipole moment. This dipole can be set in motion by means of a nonuniform electric field.

DC electrical resistance of a semiconductor quantum wire in dielectric solid matrix due to gas interactions of delocalized along the wire electrons with longitudinal acoustic (LA) phonons of matrix was calculated in the work [64]. Currently, a nanocomputer with a bit logic based on arrays of nanowires lattices where each lattice site is a programmable transistor are being developed [65].

In the work [66] the ballistic motion of a single conduction electron in a quasi-one-dimentional wire located in the dielectric environment with distributed inductance was analyzed. It was shown the possibility of the existence of inductive soliton (or *inducton*) in the wire and its parameters were estimated. It was found that the current pulse of inducton is compressed in the direction of motion with an increase in the inductance of the environment. We should note that in the work [67] was proposed the concept of the electrostatic inducton as the electron moving in a two-dimensional semiconductor quantum well parallel to the metal plane situated nearby.

Using method of molecular orbitals in the tight binding approximation [68] atomic and band structure of the single-walled carbon *zigzag* type nanotube (6,0) at its axial relative elongation $\varepsilon$ were calculated. It was shown that at the ground state the nanotube has a Kekule structure. At

relative elongation $\varepsilon_t \approx 9\%$ of the nanotube the calculations show a structural phase transition leading to the quinoid structure. This phase transition is accompanied by the transition of the nanotube from the narrow gap to the moderate gap semiconductor that can be applied in a mechanical deformation sensor (see Fig. 3). The predicted transition is also relevant in connection with the indicated [69] possibility of using structural instability of one-dimensional systems as the physical basis for the operating of devices of molecular electronics.

Interaction and DC tunneling conductivity between coaxial ends of carbon nanotubes located opposite to each other were studied on the example of two (11, 11) nanotubes with open ends passivated by hydrogen atoms [70]. A force sensor scheme based on the interaction between the ends of coaxial nanotubes was proposed, and the relationship between the tunnel conductivity and the measured force for the studied nanotubes were established. Estimates of the operational characteristics of the device as a magnetic field sensor were given based on the measurements of the magnetic force acting on the coaxial nanotubes filled with magnetic endofullerenes $(Ho_3N)@C_{80}$, which have according to [71] magnetic moment $21\mu_B$.

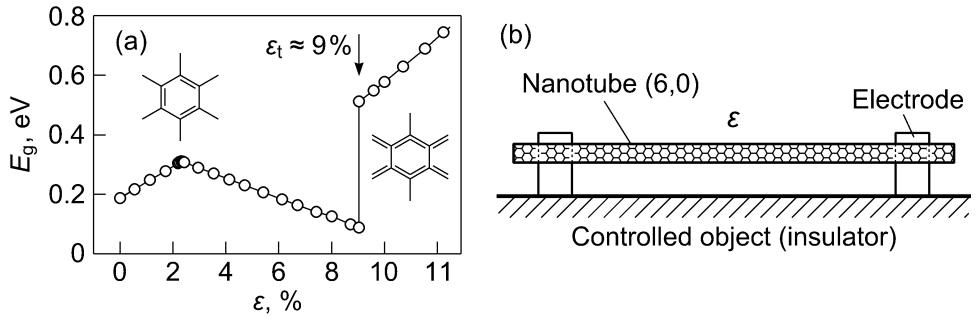



Fig. 3. (a) The phase transition from Kekule to quinoid structure under the tension (relative elongation ε) of the single-walled carbon (6, 0) nanotube and the scheme of the strain sensor (b) on its basis [68]. Here $E_g$ is the band gap of the nanotube

In the works [72–74] the DC electrical conductivity of fibers and layers of carbon nanotubes, as well as "epoxy resin / carbon nanotubes" composites in the temperature range 1.8–300 K were studied. Higher conductivity of composite "epoxy resin + multi-walled nanotubes" in comparison with composites "epoxy resin + single-walled nanotubes" with the same weight content of the filler (carbon nanotubes) is associated with a smaller number of contact barriers to electron transitions between individual nanotubes. It was shown [73, 74] that DC electrical conductivity of the carbon nanotube array at cryogenic temperatures is increased under the influence of electromagnetic radiation (in the frequency range 0.5–7.3 THz) due to bolometric heating and changes in the rate of hopping migration of electrons under the radiation. In the work [75] it is established that the kinetic inductance of the individual nanotubes contributes to the low-frequency impedance of the fibers composed of the single-walled carbon nanotubes (in the range of cryogenic temperatures and at a constant electric bias).

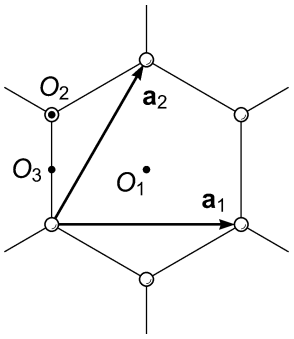
### 17.4. Two-dimensional systems.

In the works [76, 77] a new calculating algorithm for construction of carbon funnels from graphene was proposed. (Factorization of the hexagonal lattice by a shift to the chiral vector $\mathbf{C}(m,n) = m\mathbf{a}_1 + n\mathbf{a}_2$ gives the tube $(m,n)$; see, e.g., [78]. Factorization of the hexagonal lattice by a rotation from the symmetry group of the lattice provides a conical surface — a funnel; see Fig. 4.)

To classify monolayer funnels it is sufficient to consider the of conjugacy classes of the group of rotations and their action on the lattice (see, e.g., [39]). This gives only 8 funnel types defined by rotations around the three types of points $O_1$, $O_2$ and $O_3$. For each of these apexes, there are several angles values. All transformations for all possible funnel types are shown in Fig. 4.

We should note here that the graphene funnel may serve as concentrators of solar radiation in the nonimaging optics [79], as well as may find applications in nanolithography [80, 81].

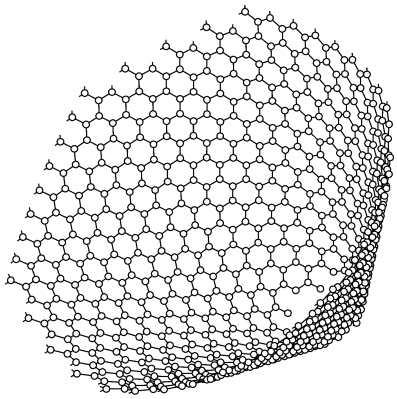
Modeling of locally symmetric carbon surfaces with defect of structure (eg. graphene with a compact convex, consisting of several pentagons, hexagons and heptagons) was carried out [77, 82] by the method of block-regular structures. A model of the cantilever made of the carbon surface was proposed [83]. It should be noted that the method is based on the experimentally known stable local atomic configurations (or blocks).




(a)

| **Funnel apex** | **Angle of rotation** |
|---|---|
| $O_1$, center of regular hexagon | $\pi/3$, $2\pi/3$, $\pi$, $4\pi/3$, $5\pi/3$ |
| $O_2$, vertex of regular hexagon | $2\pi/3$, $4\pi/3$ |
| $O_3$, center of edge of regular hexagon | $\pi$ |

(b)



(c)

Fig. 4. (a) Fragment of the graphene lattice (with points of the type $O_1$, $O_2$ and $O_3$), $\mathbf{a}_1$, $\mathbf{a}_2$ are the basis vectors. (b) Rotations around points $O_1$, $O_2$ or $O_3$, are possible in the graphene group symmetry (within the accuracy of a shift by the vectors $\mathbf{a}_1$ and $\mathbf{a}_2$). (c) Monolayer carbon funnel determined according to [76] by the rotation with the center at the $O_1$ point by the angle $\pi$

In [84] by the quantum-chemical method of molecular orbitals were calculated *zigzag* type (*n*zGNR) graphene nanoribbons, where *n* is a number of *zigzag* chains. Calculations of the electronic energy band structure shows that a narrow nanoribbon (4zGNR) is a semiconductor in both antiferromagnetic (AFM) and ferromagnetic (FM) states. Wide nanoribbon (10zGNR) in the AFM state is a semiconductor (bandgap $\approx$0.1 eV), while in FM state nanoribbon is a half-metal (electrical conductor for one spin orientation).

A new concept of electromechanical nanodynamometer based on the relative shift of the layers of a bilayer graphene was proposed in the works [85, 86]. In the proposed device the external force $\mathbf{F}_{ext}$ acting on one of the layers of graphene causes a relative shift of this layer and the corresponding change in electrical conductivity between the layers (Fig. 5). The interaction force between layers $\mathbf{F}_{int}$ returns the top layer of graphene to the initial position after removal of the external action.

Recently, were proposed the tunneling diode based on double-layer graphene with and without dielectric spacer [87] and multi-layer graphene membrane based memory cell [88].

In the work [89] the possibility of creating alternating current electric generator on the basis of field emission of electrons from graphene nanoribbons was analyzed. Note that in contrast to the known schemes (see, e.g., [90]), where field emission occurs from the butt end of the nanowire

(bunch of carbon nanotubes), we propose a more simple and production-friendly construction in the form of a double-clamped graphene nanoribbon cathode placed above the flat anode surface was proposed.

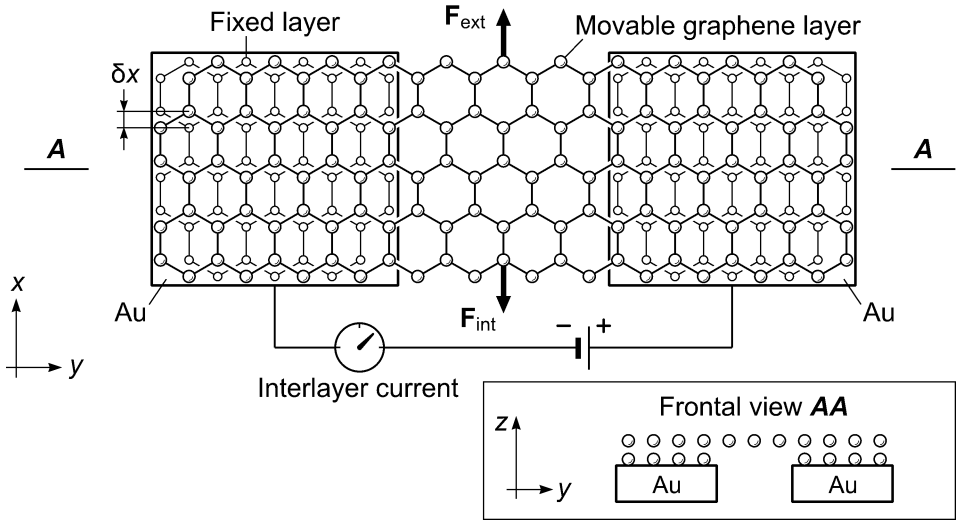



Fig. 5. Scheme of the nanodynamometer based on graphene bilayer [85]. Mobile upper layer is in mechanical and electrical contact with two stationary lower graphene layers placed on two gold electrodes

### 17.5. Nanostructured materials.

One of the types of disordered solids are nanocomposites made of low-dimensional systems of different chemical composition. The model of two-dimensional disordered system and method of calculation of its macroscopic DC conductivity were proposed [91]. The calculations using the model show correspondence with the experiment [92] on conductivity of partially graphitized regions created on the surface of the polycrystalline diamond films by ion implantation.

Using electron spin resonance (ESR) the nanostructuring of the natural diamond crystal grains due to ionizing irradiation from nuclear reactor was found [93], as well as magnetism of nanoporous diamond films [94] (see also nanostructuring and (quasi)ferromagnetism of silicon crystals [95] implanted with high fluences of high-energy ions of noble gases).

By means of ESR studies of natural diamond crystals it was found (see review [15]) that as a result of their high-energy implantation with various ions (kinetic energy of the ion more than 1 MeV/nucleon) the quasi-one-dimensional regular stable structures have been formed along the direction of irradiation (ion tracks).

In the works [96–99] the diamond monocrystals (with weight up to 0.5 carat) synthesized in a “split-sphere” type apparatus and polycrystalline diamond films obtained by CVD-method (chemical vapor deposition) in the “doping” hydrocarbon nitrogen atmosphere were studied. For the first time, it was observed the inversion of ESR signal from nitrogen atoms in synthetic diamond monocrystals and polycrystalline diamond films at room temperature with increasing microwave electromagnetic radiation power (up to 70 mW in $H_{102}$-resonator). Improving the structural perfection of films (based on the data of Raman scattering in films) leads to a decrease of the microwave power at which the signal is inverted (the first derivative of the resonant absorption of microwave radiation on the external magnetic field strength). Inversion of the in-phase ESR signal (magnetic field modulation frequency of 100 kHz) is associated with the bistability of the nitrogen atom (P1-center) coordination in diamond matrix [100].

In the work [101] the stationary states of a single nitrogen atom replacing carbon atom in the diamond crystal lattice site were calculated by the method of molecular orbitals. (N atom together with the four nearest C atoms is called a P1-center or nitrogen in *C*-form.) The possibility of making a maser based on the P1-centers operating at room temperature is discussed. (The scheme of the maser based on organic molecular crystal is presented in [102].)

An abnormal increase in the low-frequency permittivity of synthetic diamond single crystals as a result of boron atoms doping (electrically polarizable hydrogen-like acceptors substituting carbon atoms in the crystal lattice, and being in the neutral charge state) was found at room temperature [103].

A mapping of the Raman scattering [104] on an oblique thin section in natural diamonds implanted with high-energy ions (ion kinetic energy > 1 MeV/nucleon) was realized and a projected range of ions was identified by a degree of disordering of the crystal lattice due to implantation.

By the ESR method at room temperature a state of spin glass of radicals (irradiation-induced defects with uncompensated spin magnetic moments of the electron) in irradiated with fast reactor neutrons polycrystalline CVD diamond and solitary centers of their magnetic ordering in the form of ferrons were revealed [105, 106]. Looking forward, these results can be applied to increase the information capacity of condensed matter, in which the greatest possible amount of information recorded is determined by the number of minima of the adiabatic electron term [107].

Note that the ESR signal inversion (in-phase with the magnetic field modulation) of the coal samples in the lump form with increase in their size along the magnetic component of the microwave field in microwave spectrometer $H_{102}$-resonator was observed in [108]. In this work within the concept of spin nanomechanics [109] the possibility of stimulation of magnetoplasticity in rocks is discussed [110].

Finally, we should note that the diamonds with point defects (centers) of crystalline structure are promising for the creation of injection lasers operating on intracenter radiative transitions of electrons and/or holes, for which a zero-phonon lines in the spectrum of photoluminescence are observed [111, 112]. For example, a diamond *p*–*i*–*n*-diode was fabricated [113] and for a direct bias at room temperature the emission of single photons from single electrically neutral NV-centers in the *i*-region was detected. In diamond crystals NV-centers are photostable solid-state sources of single photons at room temperature [114].

### 17.6. Prospects

In order to create, research and practically apply low-dimensional condensed systems in their physical, chemical and biological aspects (for functional elements of devices in a size range from 0.1 μm to 1 nm) it is necessary to develop the quantum theory of states and processes in such systems. The issue is that the basis for the description of the electric, magnetic, optical, acoustic and thermal phenomena in the functional elements based on three-dimensional crystals size of which is greater than or of the order of magnitude of 0.1 μm is a band theory of condensed systems. (The history of the electronic band structure, which describes the state of the electrons, photons and phonons, as well as the processes involving them in crystalline insulators, semiconductors and metals is presented in [115, 116].) However, when the size of the elements and structures is from 0.1 μm to 10 nm the electronic band theory can not be applied for a low-dimensional systems due to violations of the translational symmetry (see, e.g., [117]). By contrast, the quantum theory of solitary atomic clusters, which is the basis for the description of objects of atomic range (0.1–1 nm), also can not be used to adequately describe the conditions and processes in low-dimensional systems. Therefore, it is necessary to develop the local theory, which consider states and processes

in individual atoms (molecules, clusters), taking into account their agglomeration in low-dimensional system. The local (atomic-molecular) theory does not overstate the role of translational symmetry in the formation of the properties of matter: the structure itself and the properties of matter are the result of the electronic structure of its constituent atoms [118].

It seems that the development of methods for creating and diagnostics, as well as the expansion of the applications of low-dimensional systems (and nanomaterials) in electronics, photonics, acoustics and spintronics of the new generation will happen traditionally: "from understanding the situation to the practice of calculation and experiment". Some areas of research on this topic may be noted:

1) Development of methods of mathematical modeling and physical-chemical principles of the molecular architecture of low-dimensional quantum systems and device structures based on them. The solution of this problem will allow to realize the "low tonnage" technologies not by the method of "straight stacking" but rather purposefully and economically.

2) Development of quantum theory of ionization equilibrium and hopping charge transfer in low-dimensional systems for the purposes of hydrogen and solar energetics. The solution of this problem will allow creating new materials for photovoltaic cells and "non-poisoned" electrodes for water photolysis.

3) Integrating magnetism into semiconductor micro- and nanoelectronics (see, e.g., [119]). Engineering development of magnetic carbon low-dimensional systems on silicon wafers will allow to extend their use in spintronics.

4) Development of methods for the formation of low-dimensional systems (fibers, rolls and ribbons) due to the interaction of compression plasma fluxes and intense laser radiation with the surface of synthetic diamond crystals for the purpose of creation elements of technical devices on their basis.

5) Establishing the dependence of mechanical strength of the diamond crystals and silicon carbide on the Fermi level (chemical potential) position due to its intrinsic defects of crystal structure and impurity atoms taking into account their electrical and magnetic activity. This will allow to simulate the destruction processes of those materials in the devices of power electronics and optics.

6) Development of the concept of spin nanomechanics of coals of a different metamorphic grade. This probably may allow to predict (and prevent) sudden outbursts of coal in mines during its production.

It is clear that in scientific researches and applications of the results we should rely on the experiment, theory and practice, due to which the physics and technology of low-dimensional carbon systems originated and developing.

## References

[1] M.C. Roco, Ch.A. Mirkin, M.C. Hersam. *Nanotechnology Research Directions for Societal Needs in 2020: Retrospective and Outlook* (Dordrecht, Springer, 2011).

[2] M.A. Ratner, D. Ratner. *Nanotechnology: A Gentle Introduction to the Next Big Idea* (Upper Saddle River, NJ, Prentice Hall PTR, 2003).

[3] *Nanotechnologies in Electronics*, ed. by Yu.A. Chaplygin (Moscow, Tekhnosfera, 2005) [in Russian].

[4] R.A. Andrievskii. *Fundamentals of Nanomaterials Science. Possibilities and Problems* (Moscow, Binom, 2012) [in Russian].

[5] *Nanotechnologies in Semiconductor Electronics*, ed. by A.L. Aseev (Novosibirsk, SB RAS, 2007) [in Russian].

[6] N.G. Rambidi. *Nanotechnologies and Molecular Computers* (Moscow, Fizmatlit, 2007) [in Russian].

[7] N. Kobayashi. *Introduction to Nanotechnology* (Binom, Moscow, 2008) [in Russian].

[8] *Nanoelectronics and Information Technology: Advanced Electronic Materials and Novel Devices*, ed. by R. Waser (Weinheim, Wiley, 2012).

[9] N.A. Poklonski. *Proc. of Int. Winter School on Physics of Semiconductors 2012, St.-Petersburg–Zelenogorsk, 24–27 Feb. 2012* (St.-Petersburg, Ioffe Physical-Technical Institute, 2012) pp. 7–10 [in Russian].

[10] N.A. Poklonski. Proc. of Int. Winter School on Physics of Semiconductors 2010, St.-Petersburg–Zelenogorsk, 25 Feb. – 1 March 2010 (St.-Petersburg, Ioffe Physical-Technical Institute, 2010) pp. 48–52 [in Russian].

[11] N.M. Lapchuk, V.B. Odzhaev, N.A. Poklonski, D.V. Sviridov. BSU Bull.: Phys. Math. Inform., **No. 1**, 4–14 (2009) [in Russian].

[12] N.A. Poklonski, A.K. Fedotov. BSU Bull.: Phys. Math. Inform., **No. 3**, 49–55 (2006) [in Russian].

[13] N.A. Poklonski, E.F. Kislyakov, O.N. Bubel', S.A. Vyrko. Bulletin of the Foundation for Fundamental Research, **No. 3**, 18–36 (2005) [in Russian].

[14] V.A. Dorosinets, A.G. Zakharov, N.M. Lapchuk, M.G. Lukashevich, N.A. Poklonski, A.R. Chelyadinskii, V.Yu. Yavid, A.M. Yanchenko. *Selected Scientific Works of Belarusian State University. Vol. 4. Physics* (BSU, Minsk, 2001) pp. 413–434 [in Russian].

[15] D.P. Ertchak, V.G. Efimov, V.F. Stelmakh. J. Appl. Spectrosc., **64** (4), 433–460 (1997).

[16] *Encyclopedia for Pupils and Students*. In 12 Vols. *Vol. 2. Physics. Mathematics*, ed. by N.A. Poklonski (Minsk, P. Brouka Belarusian Encyclopedia, 2010) [in Russian].

[17] *Encyclopedia for Pupils and Students.* In 12 Vols. *Vol. 4. World of Engineering*, ed. by N.A. Poklonski (Minsk, P. Brouka Belarusian Encyclopedia, 2010) [in Russian].

[18] A. Bejan, S. Lorente. J. Appl. Phys., **113**, 151301 (2013).

[19] B.A. Gurovich, D.I. Dolgii, E.A. Kuleshova, E.P. Velikhov, E.D. Ol'shanskii, A.G. Domantovskii, B.A. Aronzon, E.Z. Meilikhov. Phys. Usp., **44** (1), 95–105 (2001).

[20] S.T. Skowron, I.V. Lebedeva, A.M. Popov, E. Bichoutskaia. Nanoscale, **5** (15), 6677–6692 (2013).

[21] V.I. Konov. Laser Photonics Rev., **6** (6), 739–766 (2012).

[22] G.N. Makarov. Phys. Usp., **56** (7), 643–682 (2013).

[23] M.R. Baklanov, J.-F. de Marneffe, D. Shamiryan, A.M. Urbanowicz, H. Shi, T.V. Rakhimova, H. Huang, P.S. Ho. J. Appl. Phys., **113**, 041101 (2013).

[24] M.S. Dresselhaus. Mater. Res. Lett., **2** (1), 1–9 (2014).

[25] R. Peierls. Sov. Phys. Usp., **34** (9), 817–817 (1991).

[26] M.J. Gazalé. *Gnomon: From Pharaohs to Fractals* (Princeton, Princeton University Press, 1999).

[27] V.V. Brazhkin. Phys. Usp., **52** (4), 369–376 (2009).

[28] *Carbon Investigation — Progress and Problems*, ed. T.A. Sladkova (Moscow, Nauka, 2007) [in Russian].

[29] M. Inagaki, L.R. Radovic. Carbon, **40** (12), 2279–2282 (2002).

[30] R.B. Heimann, S.E. Evsyukov, Y. Koga. Carbon, **35** (10-11), 1654–1658 (1997).

[31] G.A. Domrachev, A.I. Lazarev, B.S. Kaverin, A.N. Egorochkin, A.M. Ob"edkov, E.G. Domracheva, L.G. Domracheva, G.V. Markin, E. Huipe Nava, A.A. Sorokin, O.N. Suvorova, V.L. Karnatsevich, A.I. Kirillov, A.A. Zakurazhnov. Phys. Solid State, **46** (10), 1969–1983 (2004).

[32] A.D. Pogrebnjak, A.G. Ponomarev, A.P. Shpak, Yu.A. Kunitskii. Phys. Usp., **55** (3), 270–300 (2012).

[33] R.A. Andrievski. Phys. Usp., **56** (7), 261–268 (2013).

[34] D.C. Elias, R.R. Nair, T.M.G. Mohiuddin, S.V. Morozov, P. Blake, M.P. Halsall, A.C. Ferrari, D.W. Boukhvalov, M.I. Katsnelson, A.K. Geim, K.S. Novoselov. Science, **323** (5914), 610–613 (2009).

[35] G.N. Churilov, N.V. Bulina, A.S. Fedorov. *Fullerene Synthesis and Theory of Formation*, ed. by V. F. Shabanov; Kirensky Institute of Physics SB RAS (Novosibirsk, SB RAS Publ., 2007) [in Russian].

[36] P.W. Fowler, D.E. Manolopoulos. *An Atlas of Fullerenes* (Oxford, Clarendon Press, 1995).

[37] O.N. Bubel', S.A. Vyrko, E.F. Kislyakov, N.A. Poklonski. JETP Lett., **71** (12), 508–510 (2000).

[38] P.J. Horoyski, M.L.W. Thewalt, T.R. Anthony. Phys. Rev. B, **54** (2), 920–929 (1996).

[39] N.A. Poklonski, A.T. Vlassov, S.A. Vyrko. *Finite Symmetry Groups. Fundamentals and Applications* (Minsk, P. Brouka Belarusian Encyclopedia, 2011) [in Russian].

[40] N.A. Poklonski. *Proc. of Int. Winter School on Physics of Semiconductors 2003, St.-Petersburg–Zelenogorsk, 28 Feb. – 3 March 2003* (St.-Petersburg, Ioffe Physical-Technical Institute, 2003) pp. 75–77 [in Russian].

[41] N.A. Poklonski, S.V. Ratkevich, S.A. Vyrko. J. Phys. Chem. A, **119** (34), 9133–9139 (2015).

[42] R.S. Berry, B.M. Smirnov. Phys. Usp., **56** (10), 973–998 (2013).

[43] D.V. Shirkov. Phys. Usp., **52** (6), 549–557 (2009).

[44] M.L. Perrin, R. Frisenda, M. Koole, J.S. Seldenthuis, J.A. Celis Gil, H. Valkenier, J.C. Hummelen, N. Renaud, F.C. Grozema, J.M. Thijssen, D. Dulić, H.S.J. van der Zant. Nat. Nanotechnol., **9** (10), 830–834 (2014).

[45] N.A. Poklonskii, E.F. Kislyakov, O.N. Bubel', S.A. Vyrko. J. Appl. Spectrosc., **69** (3), 323–327 (2002).

[46] N.A. Poklonski, E.F. Kislyakov, O.N. Bubel', S.A. Vyrko. Proc. SPIE, **5509**, 179–186.

[47] N.A. Poklonski, E.F. Kislyakov, S.A. Vyrko, N.N. Hieu, O.N. Bubel', A.I. Siahlo, I.V. Lebedeva, A.A. Knizhnik, A.M. Popov, Yu.E. Lozovik. J. Nanophotonics, **4**, 041675 (2010).

[48] M. Gobbi, A. Bedoya-Pinto, F. Golmar, R. Llopis, F. Casanova, L.E. Hueso. Appl. Phys. Lett., **101**, 102404 (2012).

[49] S.C. Benjamin, A. Ardavan, G.A.D. Briggs, D.A. Britz, D. Gunlycke, J. Jefferson, M.A.G. Jones, D.F. Leigh, B.W. Lovett, A.N. Khlobystov, S.A. Lyon, J.J.L. Morton, K. Porfyrakis, M.R. Sambrook, A.M. Tyryshkin. J. Phys.: Condens. Matter, **18** (21), S867–S883 (2006).

[50] Ф.М. Шахов, С.В. Кидалов. ФТТ, **56** (8), 1571–1574 (2014).

[51] P.N. Dyachkov. *Electronic Properties and Application of Nanotubes* (Moscow, BINOM. Laboratoriya znanii, 2015) [in Russian].

[52] M.V. Kharlamova. Phys. Usp., **56** (11), 1047–1073 (2013).

[53] N.A. Poklonskii, E.F. Kislyakov, G.G. Fedoruk, S.A. Vyrko. Phys. Solid State, **42** (10), 1966–1971 (2000).

[54] S. Suzuki, F. Maeda, Y. Watanabe, T. Ogino. Phys. Rev. B, **67**, 115418 (2003).

[55] N.A. Poklonski, E.F. Kislyakov, S.A. Vyrko, N.N. Hieu, O.N. Bubel', A.I. Siahlo, I.V. Lebedeva, A.A. Knizhnik, A.M. Popov, Yu.E. Lozovik. SPIE Newsroom., 19 Nov. 2010 (3 pp.) [doi:10.1117/2.1201010.003091].

[56] N.A. Poklonski, E.F. Kislyakov, D.I. Sagaidak, A.I. Siaglo, G.G. Fedoruk. Tech. Phys. Lett., **27** (5), 180–182 (2001).

[57] C. Hamann, H.-J. Höhne, F. Kersten, M. Müller, F. Przyborowski, M. Starke. Phys. Status Solidi B, **50** (1), K189–K192 (1978).

[58] Th. Frauenheim, C. Hamann, M. Müller. Phys. Status Solidi A, **86** (2), 735–747 (1984).

[59] A.S. Nizovtsev, S.G. Kozlova. J. Phys. Chem. A, **117** (2), 481–488 (2013).

[60] N.A. Poklonski, E.F. Kislyakov, O.N. Bubel', S.A. Vyrko. In: *Carbon Nanostructures* (Minsk, A.V. Luikov HMTI NASB, 2006) pp. 265–269 [in Russian].

[61] N.A. Poklonski, E.F. Kislyakov, S.L. Podenok. *Physics, Chemistry and Application of Nanostructures: Reviews and Short Notes to Nanomeeting-2003, Minsk, 20–23 May 2003*, ed. by V.E. Borisenko, S.V. Gaponenko, V.S. Gurin (Singapore, World Scientific, 2003) pp. 186–189.

[62] N.A. Poklonski, E.F. Kislyakov, L. Kuzmin, M. Tarasov, E.E.B. Campbell. *Physics, Chemistry and Application of Nanostructures: Reviews and Short Notes to Nanomeeting-2005, Minsk, 24–27 May 2005*, ed. by V.E. Borisenko, S.V. Gaponenko, V.S. Gurin (World Scientific, Singapore, 2005) pp. 235–239.

[63] O.V. Ershova, Yu.E. Lozovik, A.M. Popov, O.N. Bubel', N.A. Poklonskii, E.F. Kislyakov. Phys. Solid State, **49** (10), 2010–2014 (2007).

[64] N.A. Poklonskii, E.F. Kislyakov, S.A. Vyrko. Semiconductors, **37** (6), 710–712 (2003).

[65] J. Yao, H. Yan, S. Das, J.F. Klemic, J.C. Ellenbogen, C.M. Lieber. PNAS, **111** (7), 2431–2435 (2014).

[66] N.A. Poklonski, A.T. Vlassov, S.A. Vyrko, E.F. Kislyakov, S.V. Ratkevich, A.I. Siahlo. *Physics, Chemistry and Applications of Nanostructures. Reviews and Short Notes: Proc. of the Int. Conf. Nanomeeting-2013, Minsk, 28–31 May 2013*, ed. by V.E. Borisenko et al. (World Scientific, Singapore, 2013) pp. 36–39.

[67] S. Bednarek, B. Szafran. Phys. Rev. B, **73**, 155318 (2006).

[68] N.A. Poklonski, S.V. Ratkevich, S.A. Vyrko, E.F. Kislyakov, O.N. Bubel', A.M. Popov, Yu.E. Lozovik, N.N. Hieu, N.A. Viet. Chem. Phys. Lett., **545**, 71–77 (2012).

[69] Yu.H. Krieger. J. Struct. Chem., **40** (4), 594–619 (1999).

[70] A.M. Popov, I.V. Lebedeva, A.A. Knizhnik, Yu.E. Lozovik, N.A. Poklonski, A.I. Siahlo, S.A. Vyrko, S.V. Ratkevich. Comput. Mater. Sci., **92**, 84–91 (2014).

[71] M. Wolf, K.-H. Müller, Yu. Skourski, D. Eckert, P. Georgi, M. Krause, L. Dunsch. Angew. Chem. Int. Ed., **44** (21), 3306–3309 (2005).

[72] P. Kuzhir, V. Ksenevich, A. Paddubskaya, T. Veselova, D. Bychanok, A. Pliyushch, A. Nemilentsau, M. Shuba, S. Maksimenko, L. Coderoni, F. Micciulla, I. Sacco, G. Rinaldi, S. Bellucci. Nanosci. Nanotechnol. Lett., **3** (6), 889–894 (2011).

[73] V. Ksenevich, J. Galibert, V. Samuilov. In: *Carbon Nanotubes*, ed. by J.M. Marulanda (Rijeka, Groatia, InTech, 2010) pp. 123–145.

[74] V.K. Ksenevich, V.B. Odzaev, Z. Martunas, D. Seliuta, G. Valusis, J. Galibert, A.A. Melnikov, A.D. Wieck, D. Novitski, M.E. Kozlov, V.A. Samuilov. J. Appl. Phys., **104**, 073724 (2008).

[75] V.K. Ksenevich, N.I. Gorbachuk, N.A. Poklonski, V.A. Samuilov, M.E. Kozlov, A.D. Wieck. Fullerenes, Nanotubes, and Carbon Nanostructures, **20** (4-7), 434–438 (2012).

[76] A.T. Vlassov, N.A. Poklonski, S.A. Vyrko. *Proc. of Int. Conf. Actual Problems of Solid State Physics (SSP-2009), Minsk, 20–23 Oct. 2009* (Minsk, Varaksin, 2009) Vol. 3, pp. 53–55 [in Russian].

[77] N.A. Poklonski, S.A. Vyrko, A.T. Vlassov. *Contributed Papers of VI Int. Conf. Plasma physics and plasma technology (PPPT-6), Minsk, Sept. 28 – Oct. 2, 2009* (Minsk, Polyfact, 2009) Vol. II, pp. 740–743.

[78] D.V. Kolesnikov, V.A. Osipov. Phys. Part. Nuclei, **40** (4), 502–524 (2009).

[79] R. Winston, J.C. Miñano, P. Benítez, W.T. Welford. *Nonimaging Optics* (Amsterdam: Elsevier Academic Press, 2005).

[80] R.P. Seisyan. Tech. Phys., **50** (5), 535–545 (2005).

[81] R.P. Seisyan. *Proc. of Int. Winter School on Physics of Semiconductors 2006, St.-Petersburg–Zelenogorsk, 24–27 Feb. 2006* (St.-Petersburg, Ioffe Physical-Technical Institute, 2012) pp. 7–10 [in Russian].

[82] A.T. Vlassov, N.A. Poklonski, S.A. Vyrko. *Proc. of Int. Conf. Actual Problems of Solid State Physics (SSP-2011), Minsk, 18–21 Oct. 2011* (Minsk, Varaksin, 2011) Vol. 2, pp. 303–305 [in Russian].

[83] N.A. Poklonski, A.T. Vlassov, S.A. Vyrko. *Proc. of Int. Conf. Shell and Membrane Theories in Mechanics and Biology: from Macro- to Nanoscale Structures, Minsk, 16–20 Sept. 2013*, ed. G.I. Mikhasev, H. Altenbach (Minsk, BSU, 2013) pp. 95–98 [in Russian].

[84] N.A. Poklonski, E.F. Kislyakov, S.A. Vyrko, O.N. Bubel', S.V. Ratkevich. J. Nanophotonics, **6**, 061712 (2012).

[85] N.A. Poklonski, A.I. Siahlo, S.A. Vyrko, A.M. Popov, Yu.E. Lozovik, I.V. Lebedeva, A.A. Knizhnik. J. Comput. Theor. Nanosci., **10** (1), 141–146 (2013).

[86] A.M. Popov, I.V. Lebedeva, A.A. Knizhnik, Yu.E. Lozovik, B.V. Potapkin, N.A. Poklonski, A.I. Siahlo, S.A. Vyrko. J. Chem. Phys., **139**, 154705 (2013).

[87] I.V. Lebedeva, A.M. Popov, A.A. Knizhnik, Yu.E. Lozovik, N.A. Poklonski, A.I. Siahlo, S.A. Vyrko, S.V. Ratkevich. Comput. Mater. Sci., **109**, 240–247 (2015).

[88] A.I. Siahlo, A.M. Popov, N.A. Poklonski, Yu.E. Lozovik, S.A. Vyrko, S.V. Ratkevich. Physica E, **84**, 348–353 (2016).

[89] N.A. Poklonski, E.F. Kislyakov, S.A. Vyrko. *Proc. Int. Sci. Conf. Shell and Membrane Theories in Mechanics and Biology: from Macro- to Nanoscale Structures, Minsk, Sept. 16–20, 2013*, ed. by G.I. Mikhasev, H. Altenbach (Minsk, BSU, 2013) pp. 105–108.

[90] V.I. Kleshch, A.N. Obraztsov, E.D. Obraztsova. JETP Lett., **90** (6), 464–468 (2009).

[91] N.A. Poklonski, A.A. Kocherzhenko, A.I. Benediktovitch, V.V. Mitsianok, A.M. Zaitsev. Phys. Status Solidi B, **243** (6), 1212–1218 (2006).

[92] I.A. Dobrinets, A.M. Zaitsev, N.A. Poklonski, A.D. Wieck. *Proc. of 8th Int. Symp. on Diamond Materials, Paris, France, April 27 – May 2, 2003* (Electrochemical Society, 2003) Abs#522.

[93] N.A. Poklonski, T.M. Lapchuk, N.I. Gorbachuk, V.A. Nikolaenko, I.V. Bachuchin. Semiconductors, **39** (8), 894–897 (2005).

[94] N.A. Poklonski, A.V. Khomich, N.M. Lapchuk, V.G. Baev, M. Sambuu. *Proc. of Russian Conf. Physical Problems of Hydrogen Energetics, St.-Petersburg, 21–23 Nov. 2005* (St.-Petersburg, Ioffe Physical-Technical Institute, 2005) pp. 20–23 [in Russian].

[95] S.V. Adashkevich, N.M. Lapchuk, V.F. Stel'makh, G.G. Fedoruk, E.N. Shumskaya. JETP Lett., **84** (10), 547–550 (2006).

[96] N.A. Poklonski, N.M. Lapchuk, T.M. Lapchuk. JETP Lett., **80** (12), 748–751 (2004).

[97] N.A. Poklonski. Tech. Phys. Lett., **32** (4), 309–311 (2006).

[98] N.A. Poklonskii, T.M. Lapchuk, V.G. Baev, G.A. Gusakov J. Appl. Spectrosc., **73** (1), 5–9 (2006).

[99] N.A. Poklonski, N.M. Lapchuk, A.V. Khomich, F.-X. Lu, W.-Zh. Tang, V.G. Ralchenko, I.I. Vlasov, M.V. Chukichev, M. Sambuu. Chin. Phys. Lett., **24** (7), 2088–2090 (2007).

[100] N.A. Poklonski, E.F. Kislyakov, O.N. Bubel', S.A. Vyrko. In: *Carbon Nanostructures* (Minsk, A.V. Luikov HMTI NASB, 2006) pp. 260–264 [in Russian].

[101] N.A. Poklonski, E.F. Kislyakov, O.N. Bubel', S.A. Vyrko. *Physics, Chemistry and Applications of Nanostructures. Reviews and Short Notes to Nanomeeting-2011: Proc. of the Int. Conf., Minsk, 24–27 May 2011*, ed. by V.E. Borisenko et al. (World Scientific, Singapore, 2011) P. 110–113.

[102] M. Oxborrow, J.D. Breeze, N.M. Alford. Nature, **488** (7411), 353–356 (2012).

[103] O.N. Poklonskaya, BSU Bull.: Phys. Math. Inform., **No. 3**, 38–44 (2013) [in Russian].

[104] O.N. Poklonskaya, A.A. Khomich. J. Appl. Spectrosc., **80** (5), 715–720 (2013).

[105] O.N. Poklonskaya. BSU Bull.: Phys. Math. Inform., **No. 2**, 60–65 (2013) [in Russian].

[106] O.N. Poklonskaya, Doklady NASB, **57** (5), 49–54 (2013) [in Russian].

[107] M.D. Bal'makov. Phys. Usp., **42** (11), 1167–1173 (1999).

[108] N.A. Poklonski, S.A. Vyrko, O.N. Poklonskaya, N.M. Lapchuk, S. Munkhtsetseg. J. Appl. Spectrosc., **80** (3), 366–371 (2013).

[109] V.I. Alshits, E.V. Darinskaya, M.V. Koldaeva, R.K. Kotowski, E.A. Petrzhik, P. Tronczyk. Phys. Usp., **60** (3), 305–318 (2017).

[110] A.L. Buchachenko. Phys. Usp., **57** (1), 92–98 (2014).

[111] A.M. Zaitsev. *Optical Properties of Diamond. A Data Handbook* (Springer, Berlin, 2001).

[112] I.A. Dobrinets, V.G. Vins, A.M. Zaitsev. *HPHT-Treated Diamonds: Diamonds Forever* (Springer, Berlin, 2013).

[113] A. Lohrmann, S. Pezzagna, I. Dobrinets, P. Spinicelli, V. Jacques, J.-F. Roch, J. Meijer, A.M. Zaitsev. Appl. Phys. Lett., **99**, 251106 (2011).

[114] I. Aharonovich, S. Castelletto, D.A. Simpson, C.-H. Su, A.D. Greentree, S. Prawer. Rep. Prog. Phys., **74**, 076501 (2011).

[115] A.H. Wilson. Proc. R. Soc. London Ser. A, **371** (1744), 39–48 (1980).

[116] R. Peierls. Contemp. Phys., **21** (1), 3–17 (1980).

[117] V.K. Malinovskii. Phys. Solid State, **41** (5), 725–728 (1999).

[118] V.Ya. Shevchenko. *Nanostructured materials: Belarus–Russia–Ukraine. NANO-2014: Plenary reports of IV Int. Conf., Minsk, 7–10 Oct. 2014,* ed. P.A. Vityaz' et al. (Minsk, Belaruskaya navuka, 2015) pp. 56–67 [in Russian].

[119] Ch. Seife. Science, **309** (5731), 82 (2005).

## 18. Microsensors and nanosensors
## 18.1. Introduction

The trend of electronic miniaturization is an evolution which aims to produce ever smaller electronic circuits. When considering computers, the first generation computers (1945 - 1955) were constructed using vacuum tubes. Using semiconductor transistors, the second generation computers (1955 – 1965) appeared. Especially the rise of Integrated Circuits and the Very Large Scale Integrated Circuits decreased the size of the third generation computers (1965 – 1980) and the fourth generation computers (from 1980). The ongoing miniaturization is strongly related with the Moore's law which states that the number of transistors integrated in a processor doubles every 18 months. Due to the ongoing miniaturization and the ever decreasing unit cost, a fifth generation computers appeared which are almost "invisible". This fifth generation computers are embedded in watches, bank cards, mobile phones, …

Notice however, this miniaturization does not only appear in electronic products since it also appears in mechanical and optical applications. Very often, the miniaturization of electronic, mechanical and optical components is combined in one single device. Micro-Electro-Mechanical Systems or MEMS combine miniaturized mechanical elements and micro-electronics. MEMS are also called Micro Systems Technology (MST) or Micromachines.

MEMS generally contain a number of microsensors, a number of micro actuators and integrated microelectronics. A sensor is a device which converts a nonelectrical physical quantity (pressure, velocity, acceleration, light intensity, temperature, …) into an electrical quantity (voltage, current, capacity, …). A microsensor is a very small sensor having sizes which can even be smaller than $1\ \mu m$.

There exists a very broad range of microsensors. In the present text, we will restrict our attention to a limited number of microsensors used in MEMS. More precisely, we will study the behaviour of a number of accelerometers (from Analog Devices and from Hitachi).

An actuator is a device which converts a source of energy (electrical energy, pneumatic pressure, hydraulic fluid pressure, thermal energy, …) into a motion. A micro actuator is a very small actuator having sizes in the micrometer range. When considering such micro actuators, a distinction can be made between electrostatic, electromagnetic and piezoelectric actuators.

In general, the components of a MEMS have sizes between $0.1\ \mu m$ and $100\ \mu m$. Related with these micro-dimensions, the entire MEMS devices generally range from $20\ \mu m$ to $1000\ \mu m = 1\ mm$. Due to the ever ongoing miniaturization, even NEMS (Nano-Electro Mechanical Systems) appeared based on modern nanotechnology.

### 18.2: Accelerometers

An accelerometer is used to measure an acceleration. Several types of accelerometers are available like capacitieve accelerometers, piezoresistive accelerometers, piezoelectric accelerometers, thermal accelerometers. When considering MEMS technology, capacitive accelerometers are frequently used.

#### 18.2.1: The use of the ADXL05

When considering MEMS technology, capacitive accelerometers are frequently used. For instance the ADXL05 accelerometer of Analog Devices is a capacitive accelerometer which also contains signal conditioning circuitry. The ADXL05 is able to measure positive and negative accelerations to a maximum level of $\pm\ 5g$.

Figure 17.1 visualises the ADXL05 acceleration sensor at rest. The differential capacitor sensor consists of fixed outer plates and center plates attached to the beam which moves due to an

applied acceleration (in reality, a sufficiently large number of plates are used). Two capacitances $CS1$ and $CS2$ are obtained forming a capacitive divider. In Figure 1, $CS1$ and $CS2$ equal each other.

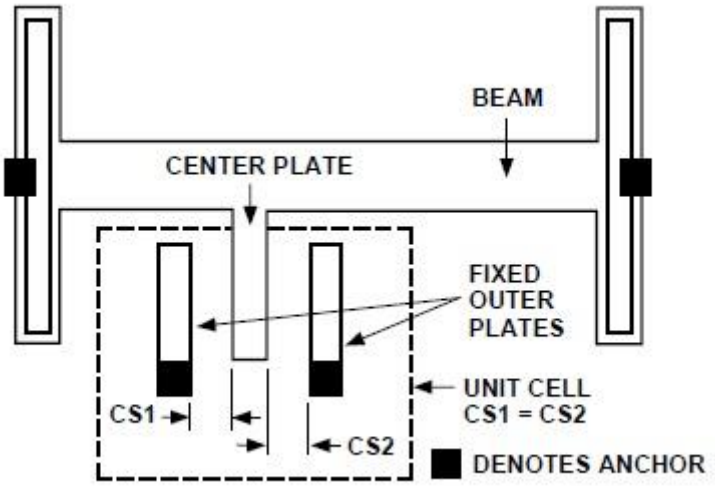



Figure 17.1: Diagram of the ADXL05 acceleration sensor (source: Analog Devices)

Figure 17.2 visualises the situation where an acceleration is applied. Due to this acceleration, the beam (and the center plate) moves closer to one of the fixed outer plates. In Figure 17.2, this implies an increase of $CS2$ and a decrease of $CS1$. The larger the acceleration, the larger the increase of $CS2$ and the larger the decrease of $CS1$.

The capacitive divider containing $CS1$ and $CS2$ is driven differentially by a $1\ MHz$ square wave as visualised in Figure 17.3 (there is a phase difference of $180°$ between the $1\ MHz$ voltages applied to the lower and the upper outer fixed plates). In case $CS1$ and $CS2$ equal each other, no AC voltage is obtained at the beam (only a DC voltage of $1.8\ V$ is obtained). In case $CS1$ and $CS2$ do not equal each other, a $1\ MHz$ signal is obtained at the beam. The larger the acceleration, the larger the difference between $CS1$ and $CS2$, the larger the amplitude of the obtained $1\ MHz$ signal.

A detailed analysis of Figure 17.3 is outside the scope of this text, but using a demodulator the $1\ MHz$ voltage at the beam of the sensor is converted into a DC signal. The output of the preamplifier equals $1.8V\ \pm 200\ mV/g$. The ADXL05 measures in a range of $\pm 5g$ implying a voltage $V_{PR}$ ranging from $0.8\ V$ to $2.8\ V$.

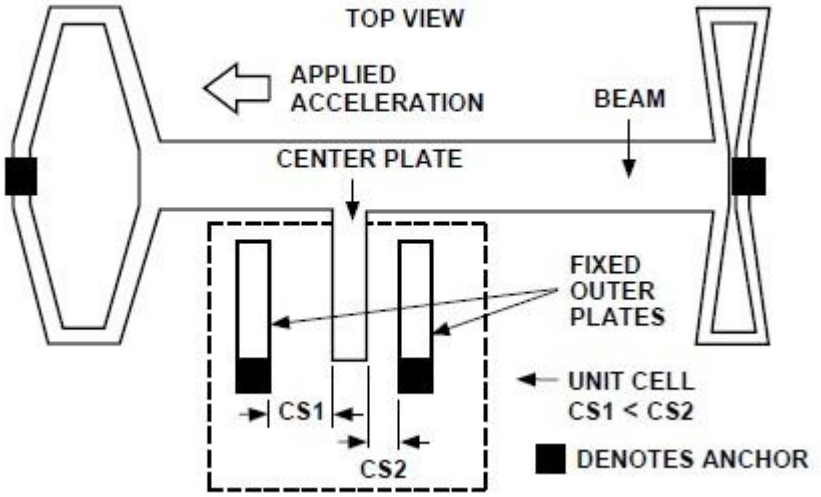


Figure 17.2: Diagram of the ADXL05 acceleration sensor (source: Analog Devices)

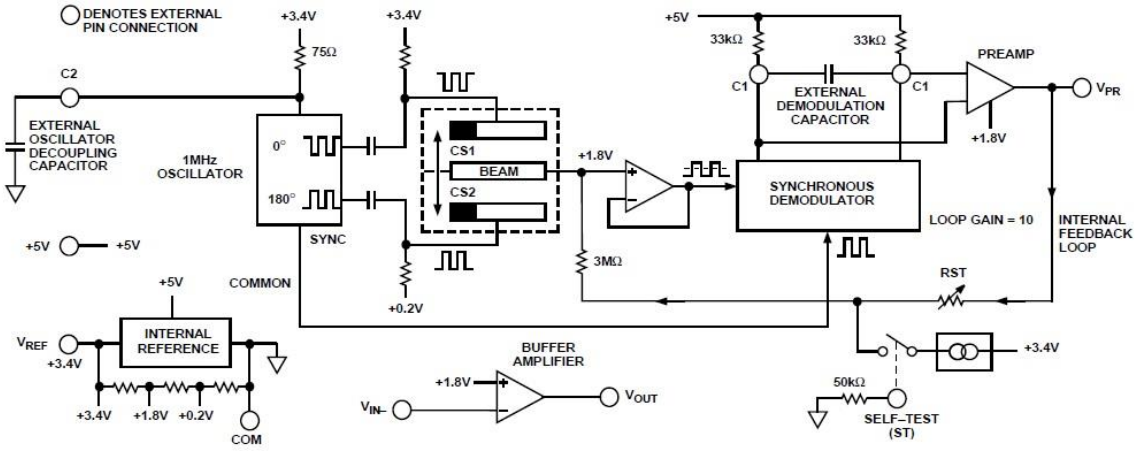



Figure 17.3: Functional block diagram of the ADXL05 (source: Analog Devices)

As visualised in Figure 17.4, the preamplifier is followed by a buffer amplifier and the user chooses three resistors $R_1$, $R_2$ and $R_3$. The output of the buffer equals

$$V_{OUT} = \left(\frac{R3}{R1}(1.8V - V_{PR})\right) + \left(\frac{R3}{R2} + 1\right)1.8V.$$

By an appropriate choice of $R_1$, $R_2$ and $R_3$, the range of the output voltage is determined. In case of Figure 17.4 with $R1 = 49.9\,k\Omega$, $R2 = 274\,k\Omega$ and $R3 = 100\,k\Omega$ accelerations in the range $\pm 5g$ imply an output voltage of $2.5V$ in case of an acceleration of $0g$ in combination with a $\pm 2.0\,V$ full-scale swing (corresponding with an output sensitivity of $400\,mV/g$). More precisely, the output ranges from $0.5V$ to $4.5V$. Such an analog DC voltage can be converted into a digital value using an ADC (Analog to Digital Converter).

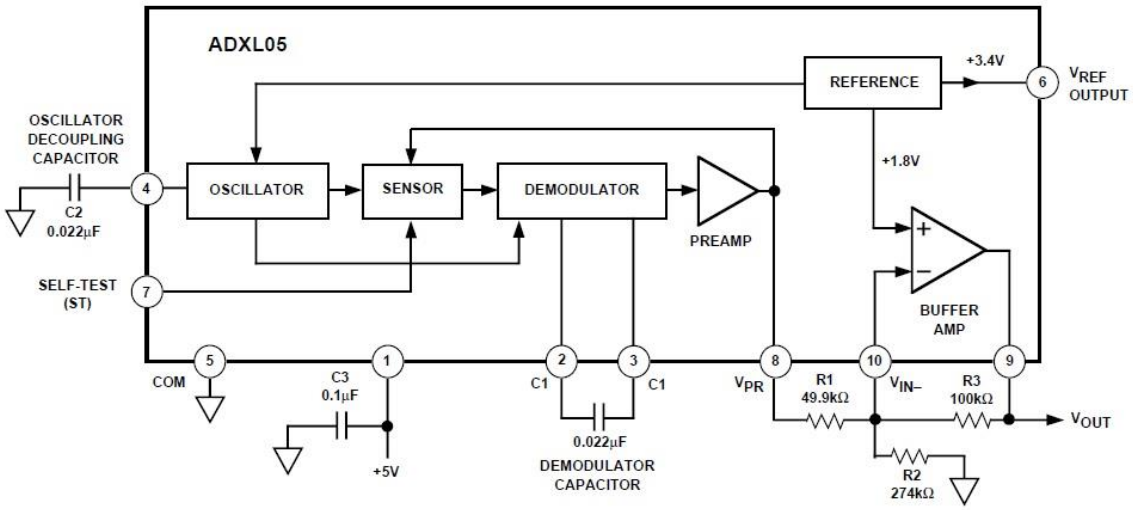



Figure 17.4: ADXL05 application (source: Analog Devices)

### 18.2.2: The use of the ADXL05, ADXL202, ADXL210, ADXL 150 and ADXL 250

The ADXL05 is a single axis MEMS accelerometer. As visualised in Figure 17.5, the ADXL05 only measures the acceleration in one single direction i.e. along one single sensitive axis (called the X axis). Perpendicular to this X axis, there is a transverse Y axis and a transverse Z axis.

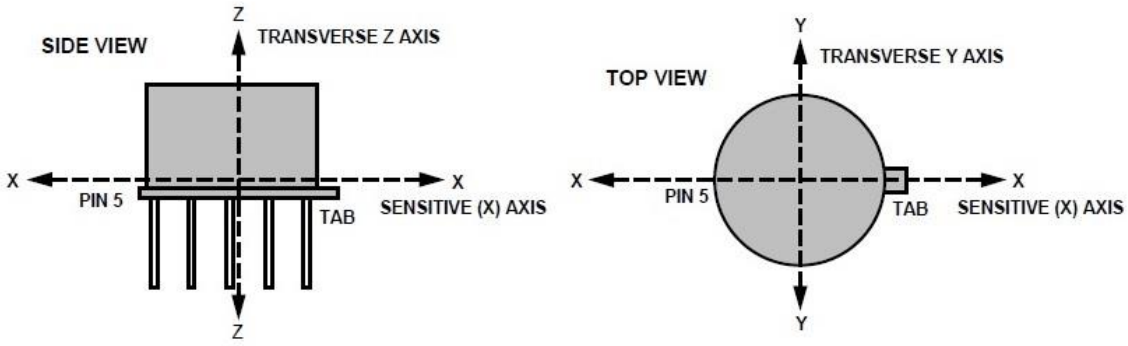



Figure 17.5: Sensitive X axis and transverse Y and Z axes (source: Analog Devices)

Not only the ADXL05 but also the ADXL150 is a single axis accelerometer. Notice however there also exist dual axis and three axis accelerometers. For instance the ADXL250 is a dual axis accelerometer measuring the acceleration in the X and the Y directions as visualised in Figure 17.6 (also the ADXL202 and the ADXL210 are dual axis accelerometers of Analog Devices). For instance the ADXL345 is a three axis accelerometer measuring the acceleration in the X, Y and Z directions.

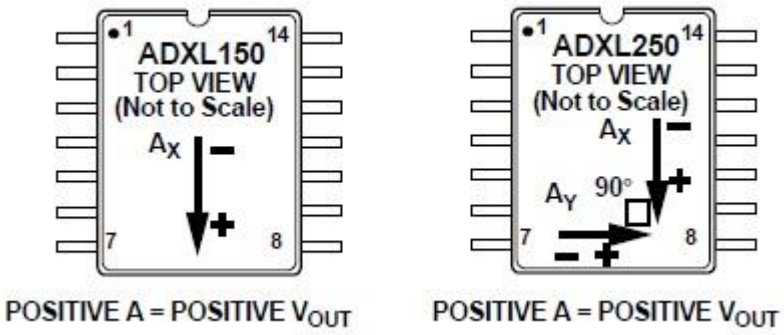



Figure 17.6: Single and dual axis accelerometers (source: Analog Devices)

### 17.3: Practical applications of accelerometers
### 17.3.1: A broad range of applications

Accelerometers are used in a very broad range of applications, only a limited number of examples are mentioned here. Accelerometers can be used to measure vibrations on machines. For instance when measuring the vibrations of industrial machinery, information can be obtained concerning the bearings. By detecting the wear of the bearings, these bearings can be replaced before they are really broken which avoids an unplanned standstill of the installation.

When considering offshore wind turbines (especially when considering farshore wind farms), it takes a lot of effort to perform maintenance and repairs due the large distance between the shoreline and the wind turbine farm. Due to this reason, it is important to limit the number of

interventions as much as possible. Due to bad weather conditions, it can also be difficult or even impossible to reach an offshore wind turbine to perform reparations. This can imply a long standstill of the broken installation which accounts for a large loss of electrical energy production. Accelerometers can be used for condition monitoring. For instance, by measuring the vibrations of the drive train, information can be obtained concerning several parts of this drive train. This allows to perform maintenance and repairs just in time before a standstill of the installation occurs avoiding loss of electrical energy production. Possibly, useless maintenance and useless replacements of parts can be avoided.

Accelerometers are also used in cars. For instance an airbag contains accelerometers. In case of a car accident, a very fast decrease of the speed will be measured and this information can be used to activate the airbag to protect the passengers. Accelerometers are also used in an unmanned vehicle to track its position. When considering an underwater vehicle where no GPS signals are available, accelerometers and gyroscopes can be used to calculate the orientation and the position of the vehicle. By integrating the angular velocity provided by a rate-gyroscope, the orientation of the vehicle is known. By a double integration of the accelerations provided by an accelerometer, the velocity and the position is known.

Accelerometers are also used in consumer electronics. For instance a free-fall sensor detects when a device has been dropped and is falling down. This information can be used to park the hard disk in order to prevent damage when reaching the ground. Accelerometers are also used when sending a parcel by post. The accelerations are measured during the transport and by logging the data, a sudden fall of the parcel will be logged. In case the parcel is broken when arriving, the logged data allows to detect when the fall occured and who is responsible for the damage.

Accelerometers which are able to measure static accelerations like the Earth's gravity are also used to measure the tilt of electronic devices. Figure 17.7 visualises an application of the ADXL05 accelerometer. By mounting the MEMS acceleration sensor in an electronic device, its tilt with respect to the gravitational force can be measured.

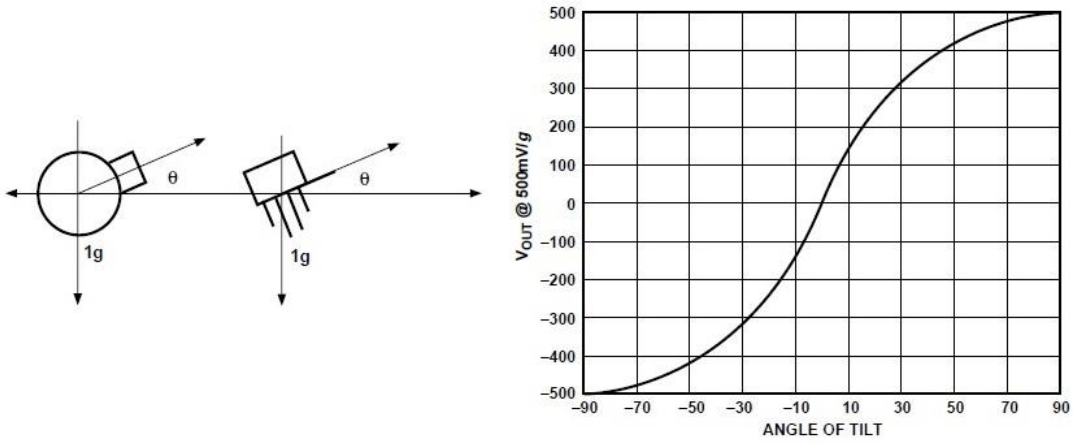



Figure 17.7: Measuring the tilt of an electronic device (source: Analog Devices)

In case the tilt angle $\theta$ equals zero, the accelerometer is mounted in order to have the sensitive X axis perpendicular with the gravitational force i.e. parallel to the Earth's surface. In case $\theta = 0$, the gravitional force is not detected giving a zero acceleration output. As the tilt angle $\theta$ increases, the component of the gravitational force along the sensitive X axis increases which will be measured.

Suppose the accelerometer has an output voltage $a$ in case no acceleration has been detected. Suppose the accelerometer has an accelerometer scale factor $b$ (expressed in $V/g$). The output voltage is given by

$$V_{OUT} = b.sin(\theta).1g + a$$

This expression allows to calculate the tilt angle $\theta$ as (which is visualised in Figure 17.7)

$$\theta = arcsin\left( \left[ \frac{V_{OUT} - a}{b \,.\, 1g} \right] \right).$$

As a particular application, MEMS accelerometers (in combination with gyroscopes) are used to detect the orientation of touchscreens.

### 18.3.2: Medical applications

Accelerometers are also used in a broad range of medical applications including sport and leisure. Accelerometers are used in pedometers. Using such a portable electronic device, it is possible to count the steps taken by someone by interpreting the motions of the hands, the hips or the feed. These pedometers are useful when coaching sportsmen but also to encourage a large range of people to walk more and improve their health.

When considering elderly people, acceleration sensors can be used for fall detection. Fall incidents for elderly people can cause serious injuries such as hip fracture or head traumas. Due to these fall incidents there is an increased morbidity. By detecting the fall, by triggering an alarm and by providing medical assistance the consequences of such a fall are seriously reduced. Based on two approachable papers "Reliability Study of the Hitachi H34C Accelerometer in Wireless Body Area Networks for Fall Detection" and "Fall Detection Using Multi-sensory Accelerometer Sensor Network" the use of MEMS accelerometers has been studied for fall detection.

#### 18.3.2.1: Assignment 1

Study the paper "Reliability Study of the Hitachi H34C Accelerometer in Wireless Body Area Networks for Fall Detection" and answer the questions below.

1) Give an overview of the structure of the paper.

2) Why is it useful to realise fall detection of elderly people.

3) Which acceleration sensor has been used in the paper? Did the authors use a single axial, a dual axial or a 3-axis accelerometer?

4) How does the accelerometer send its measurement data to the outer world?

5) Describe the static method used to test the sensitivity of the acceleration sensor?

6) Descibe the dynamic test method using a vibration table.

7) How is the acceleration calculated in the X direction, in the Y direction and in the Z direction? How do the authors calculate the magnitude of the resulting acceleration.

8) In the paper, a falling human body has been simulated by a falling wooden stick. What accelerations have been measured by the accelerometer mounted on the wooden stick. What accelerations have been obained in the X direction, in the Y direction and in the Z direction? Give an interpretation of the evolution of the magnitude of the resulting acceleration.

#### 18.3.2.2: Assignment 2

Study the paper "Fall Detection Using Multi-sensory Accelerometer Sensor Network" and answer the questions below.

1) Give an overview of the structure of the paper.

2) How many acceleration sensors have been used in the paper and where are these sensors mounted? Did the authors use single axial, dual axial or 3-axis accelerometers?

3) How do the authors calculate the magnitude of the resulting acceleration?

4) The authors use the Kaiser-Teager energy operator. What expression is used to realise this Kaiser-Teager operator.

5) Study the acceleration measurements of patient 1 in Figure 17.4. How do the resulting magnitudes of the accelerations look like. What effect is obtained by using the Kaiser-Teager energy operator?

## 18.4: Nano Electro Mechanical Systems

In the previous paragraph, we mainly studied Micro Electro Mechanical Systems (MEMS) having dimensions ranging from approximately $0.1\,\mu m$ to $100\,\mu m$. Due to the ever ongoing miniaturization, even NEMS appeared based on nanotechnology where sizes from $1\,nm$ to $100\,nm$ appear.

Nanotechnology is actually engineering at a molecular scale i.e. the lower limit is set by the size of atoms. Indeed, carbon-carbon bond lenths and the spacing between these atoms in a molecule are approximatley in the range from $0.1\,nm$ to $0.15\,nm$. The upper limit for the nonatechnology is often set close to the lower limit of the microtechnology. Nanotechnology includes the dimensions where quantum mechanical effects appear which do not appear when considering microtechnology.

The world of nanotechnology is quite broad and includes among others nanoelectronics, nanosensors, nanorobotics and the already mentioned NEMS. When using nanotechnology in electronic components nanoelectronics appears. Nanoelectronics is a technology which is really different from the traditional transistor based electronics. Related with Moore's law which states that the number of transistors integrated in a processor doubles every 18 months, the sizes of the transistors decreased to approximately $25\,nm$ which requires the development of new methods and materials to construct these devices.

Also when considering memory storage, traditional electronic memories mainly rely on the use of transistors. However, using reconfigurable interconnections between vertical and horizontal wiring arrays allows to obtain ultra high density memories (for instance leading to nano-RAM). Hewlett-Packard proposed the production of a memristor (a memory resistor i.e. a resistor having memory) allowing to replace the existing Flash memories.

Nanotechnology, more precisely nanoelectronics, also aims to produce diplays based on carbon nanotubes allowing to obtain displays with very low energy consumption. The rise of nanowires and other nanostructured materials could also lead to the production of more efficient solar cells than the traditional semiconductor based solar cells.

When considering nanosensors, especially medical applications arise. In order to improve the medical diagnoses, it would be advantageous to measure the concentrations of biomolecules in real time. Detecting toxic chemicals and measure their concentrations in the envirmonment is also a challenge. Nanorobotics also have a promising future when considering nanomedicine when trying to detect and destroy cancer cells. Moreover, possibilities arise concerning drug delivery to the patient or when realising surgical instrumentation. Nanorobots are also involved in the realisation of nanoscale 3D printing.

### References

W. Catteeuw, H. Hallez, J. Boydens, Reliability Study of the Hitachi H34C Accelerometer in Wireless Body Area Networks for Fall Detection, Annual Journal of Electronics, Volume 7, paper presented at the International Scientific Conference Electronics – ET2013, Sozopol, Bulgaria, 18 – 20 September 2014 (pp. 50-53), Technical University of Sofia.

J. Fraden, Handbook of Modern Sensors: Physics, Designs, and Applications, Springer, London, 2010.

H. Hallez, P. Alcala, J. Boydens, Fall Detection Using Multi-sensory Accelerometer Sensor Network, Annual Journal of Electronics, Volume 8, paper presented at the International Scientific Conference Electronics – ET2014, Sozopol, Bulgaria, 11 – 13 September 2014 (pp. 51-53), Technical University of Sofia.

L. Janssens, Sensors, Group T: International Engineering School Leuven, 2015.

J. Peuteman, J. Daem, P. De Witte, F. Loret, P. Staelens D. Boone, Unmanned underwater, ground and aerial vehicles: different applications, with technical similarities, International Symposium on "Light Weight Unmanned Vehicle Systems and Subsystems", March 11-13, 2009, Ostend, Belgium (10 pages).

A.S. Tanenbaum, Structured Computer Organization, Prentice Hall - Pearson, Upper Saddle River, New Jersey, 2006.

Datasheets

ADXL05 $\pm 1g$ to $\pm 5g$ Single Chip Accelerator with Signal Conditioning, Analog Devices.

ADXL150/ADXL250 $\pm 5g$ to $\pm 50g$, Low Noise, Low Power, Single/Dual Axis iMEMS Accelerometers, Analog Devices.

ADXL202/ADXL210 Low Cost $\pm 2g$ to $\pm 10g$ Dual Axis iMEMS Accelerometers with digital output, Analog Devices.

ADXL345 3-Axis, $\pm 2g$, $\pm 4g$, $\pm 8g$, $\pm 16g$ Digital Accelerometer, Analog Devices.