I R A M P
4(1), June 2013, pp. 1-22

© International Science Press
ISSN: 2229-3159

REVIEW ARTICLE

# Some Aspects of Nanophysics

**V. P. GAVRILENKO**[*, **]

[*]*Center for Surface and Vacuum Research, 40 Novatorov St., Moscow 119421, Russia*
[**]*National Research University 'Moscow Institute of Physics and Technology', 9 Institutsky per. Dolgoprudny, Moscow Region, 141700, Russia*

Translated from Russian by Nikolay Kryukov

**ABSTRACT:** In this review we discuss the following aspects of nanoscale physics: (1) Low-dimensional objects and materials; (2) Ballistic transport of electrons in low-dimensional structures and conductance quantization; (3) Laser physics; (4) Nanooptics.

## 1. INTRODUCTION

On the twenty-ninth of December, 1959, at the annual meeting of the American Physical Society at the California Institute of Technology future Nobel laureate Richard Feynman gave the lecture called "There's Plenty of Room at the Bottom". In his lecture, he said: "I would like to describe a field, in which little has been done, but in which an enormous amount can be done in principle. This field is not quite the same as the others in that it will not tell us much of fundamental physics (in the sense of, "What are the strange particles?") but it is more like solid-state physics in the sense that it might tell us much of great interest about strange phenomena that occur in complex situations. Furthermore, a point that is most important is that it would have an enormous number of technical applications. What I want to talk about is the problem of manipulating and controlling things on a small scale… It is a staggeringly small world that is below. In the year 2000, when they look back at this age, they will wonder why it was not until the year 1960 that anybody began seriously to move in this direction".

From a historical point of view, it is believed that starting from this lecture by Feynman, nanotechnology was first recognized as a feasible area of activity.

In accordance with the definition of nanotechnology, given in the Technical Committee "Nanotechnology" of the International Organization for Standardization (ISO), nanotechnology is:

1. Understanding and control of matter and processes at the nanoscale, typically, but not exclusively, below 100 nanometres in one or more dimensions where the onset of size- dependent phenomena usually enables novel applications,
2. Utilizing the properties of nanoscale materials that differ from the properties of individual atoms, molecules, and bulk matter, to create improved materials, devices, and systems that exploit these new properties.

The definition of nanotechnology is broad enough. Under this definition, nanoscale physics can be seen as part of nanotechnology, within which the study of physical phenomena on the size scale from about 1 nm to 100 nm is carried out; it can also be seen as the development of relevant objects, materials and devices, which have the above-mentioned physical phenomena at the base of their functioning.

It should be noted that in the transition from bulk materials to objects with dimensions in the nanometer range, the properties of these objects can significantly change. This is due firstly to the fact that in the nanometer scale, quantum properties begin to appear. Second, nanomaterials have a large specific surface area. The properties of atoms on the surface are different from those of the same atoms inside the material. For example, high catalytic activity of certain nanomaterials is related to this.

The present review is devoted to the examination of certain aspects of nanoscale physics.

## 2 . LOW-DIMENSIONAL OBJECTS AND MATERIALS

With continuous decrease in the size of the sample of large (macroscopic) values, such as meters or centimeters, to very small values, the properties of the sample initially remain essentially unchanged, then slowly begin to change, but at a size less than or of the order of 100 nm can change drastically. In many cases, the most important role at small sizes is played by the so-called quantum size effects. One of the most important conditions for which there are quantum size effects, is the requirement that the sample size (at least in one dimension) is less than or of the order of the de Broglie wavelength for electrons $\lambda$.

The de Broglie wavelength for an electron moving in a crystal with a velocity $v$ has the value

$$\lambda = \frac{h}{m^* v} = \frac{h}{\sqrt{2m^* E_{kin}}}$$

where $m^*$ is the effective mass of the electron, and $E_{\mathrm{kin}}$ is the kinetic energy of the electron.

Let us consider, as an example, free electrons in semiconductor crystals. At room temperature $E_{\mathrm{kin}} \approx kT = 0.026$ eV the effective mass $m^*$ of the electron in different semiconductors varies widely. For instance, for GaAs $m^* \approx 0.068m_0$, for ZnO $m^* \approx 0.28m_0$, forZnSe $m^* \approx 0.17m_0$, where $m_0$ is the rest mass of free electrons. For these semiconductors the de Broglie wavelength (at room temperature) is $\lambda \approx 29$ nm (for GaAs), $\lambda \approx 14$ nm (for ZnO), $\lambda \approx 18$ nm (for ZnSe).

In metals at room temperature free electrons have energies near the Fermi energy (a few eV). The mass of free electrons in metals is close to $m_0$. Therefore in metals, the de Broglie wavelength of the electrons is less than or of the order of 1 nm, which is considerably less than in semiconductors. Therefore, quantum size effects are easier to implement technologically in semiconductors.

For the observation of quantum size effects it is also important that the mean free path of electrons should be greater than the geometric dimensions of the object under consideration, and the object boundaries should be sufficiently smooth.

The simplest kinds of low-dimensional objects are quantum wells, quantum wires and quantum dots.

The quantum well is a two-dimensional (2D) object. It is a thin layer of a semiconductor crystal with a thickness $d$ of the order of the de Broglie wavelength. The system of electrons in a crystal is a two-dimensional electron gas. The movement of electrons in this layer is limited in the direction $x$, in which the thickness $d$ is measured, and is not limited along the two other directions. Moving in the direction $x$, the electron is not able to leave the layer as its work function (it is about several eV) considerably exceeds the thermal motion energy (0.026 eV at room temperature). Therefore, the motion of an electron in the $x$-direction can be approximately regarded as a one-dimensional motion in an infinitely deep square potential well of width $d$.

The quantum wire is a one-dimensional (1D) object. The motion of electrons is limited along two axes ($x$ and $y$), and is not limited to axis $z$. Electron motion in the plane ($x$, $y$) can be approximately regarded as a movement in a two-dimensional infinitely deep potential well.

The quantum dot is a zero-dimensional (0D) object. The movement of electrons is restricted in three dimensions. The motion of an electron in the quantum well can be approximately regarded as a movement in three-dimensional infinitely deep potential well.

The most important quantum-mechanical characteristics of crystal semiconductor structures are, firstly, the possible (allowed) values of the electron energy and the dependence of these energies on the quasi-momentum, and, secondly, the dependence of the electron density of states on the electron energy $g(E)$. The density of states $g(E)$ is defined as the number of quantum states of electrons per unit volume (for 3D systems), area (for 2D systems) or length (for 1D systems).

The quantum-mechanical characteristics of low-dimensional systems (2D, 1D, 0D objects) differ significantly from those of 3D systems.

For 3D systems (unlimited crystals), the electrons can move freely in any direction. Near the bottom of the conduction band (for a 3D semiconducting crystal) electron energy $E$ depends on the crystal momentum quadratically:

$$E = \frac{\hbar^2}{2m^*}\left(k_x^2 + k_y^2 + k_z^2\right) \tag{1}$$

where $k_x$, $k_y$ and $k_z$ are components of the quasi-wave vector $\mathbf{k}$ of the electron along the axes $x$, $y$ and $z$, respectively. The vector $\mathbf{k}$ is related to the electron quasi-momentum $\mathbf{p}$ by the relation $\mathbf{k} = \mathbf{p}/\hbar$. In (1) it is assumed for simplicity that the electron effective mass $m^*$ is the same for motion in all directions.

For 3D system for the electronic density of states, the following relation holds:

$$g(E) = \frac{2^{1/2}(m^*)^{3/2}}{\pi^2\hbar^3}E^{1/2} \tag{2}$$

where the electron energy $E$ is measured from the bottom of the conduction band. Thus, for a 3D system, density of states is close to zero in the vicinity of the bottom of the conduction band. The dependence $g(E)$ (2) is shown in Fig. 1a.

For the quantum well (2D electron gas) electron motion along the x-axis is quantized. The energy of this motion is determined by the formula

$$E_{n_x} = \frac{\pi^2\hbar^2 n_x^2}{2m^* d_x^2}, n_x = 1, 2, 3, ... \tag{3}$$

The total energy of the electron can be represented as

$$E = E_{n_x} + \frac{\hbar^2 k_y^2}{2m^*} + \frac{\hbar^2 k_z^2}{2m^*}$$

In accordance with (3), for the quantum well the energy levels depend on the quantum number $n_x$ and on the well width $d_x$. In the case of the quantum well, the dependence of the electron density of states (per unit area) on energy $g(E)$ is stepwise. This dependence is shown by the solid line in Fig. 1b. A sudden change in the function $g(E)$ occurs at $E_1$, $E_2$, $E_3$, ..., defined by (3).

For the quantum wire (1D electron gas) electron motion along the two axes ($x$ and $y$) is quantized. The total energy of the electron may be represented as

$$E = E_{n_x, n_y} + \frac{\hbar^2 k_z^2}{2m^*} \tag{4}$$

where $E_{nx,\, ny}$ is the energy of the quantum motion of electrons:

$$E_{n_x, n_y} = \frac{\pi^2\hbar^2}{2m^*}\left(\frac{n_x^2}{d_x^2} + \frac{n_y^2}{d_y^2}\right), n_x, n_y = 1, 2, 3, ... \tag{5}$$

For the quantum wire dependence of the density of states (per unit length) on energy $g(E)$ has a number of sharp peaks corresponding to energies $E_{nx,\,ny}$ (4). Figure 1c shows the typical dependence of the density of states on the energy for the quantum wire.

For the quantum dot (0D electron gas) electron motion is quantized along all three axes (axes $x$, $y$ and $z$). In the case of the quantum dot, electron energy can be represented as

$$E_{n_x,n_y,n_z} = \frac{\pi^2\hbar^2}{2m^*}\left(\frac{n_x^2}{d_x^2}+\frac{n_y^2}{d_y^2}+\frac{n_z^2}{d_z^2}\right), n_x, n_y, n_z = 1,2,3,... \tag{6}$$

The electron energy spectrum of the quantum dot is discrete. The density of electron states of a quantum dot is a set of delta functions, as shown in Figure 1d. Quantum dots are often referred to as artificial atoms. The position of the energy levels of quantum dots can be changed by varying the size of the quantum dot, defined by the values $d_x$, $d_y$ and $d_z$.

Functioning of many nanoelectronics devices is determined by the above-mentioned peculiarities of the energy spectra and the electron density of states for low-dimensional nano-objects.

The above changes of the energy spectrum occur for low-dimensional semiconductor crystals, including the spectrum of the electrons in the conduction band and the spectrum of the holes in the valence band. In addition, in low-dimensional semiconductor systems the band gap increases compared to bulk semiconductors. This leads, in particular, to blue-shift in the optical absorption spectrum of low-dimensional systems. As an example, Fig. 2 shows the dependence of optical absorption spectra of CdSe quantum dots on their radius [1]. The increase in the band gap is due to the confinement effect of electrons and holes that arise at small sizes of low-dimensional objects. Often, in the calculation of the energy spectrum of the quantum dots the electrons and holes in the quantum dot are considered to be under the influence of a centrally symmetric confinement potential $V_0$, which is constant for distances greater than the radius of the quantum dot and vanishes inside the dot. Figure 3 (see [2]) shows a comparison of experimental

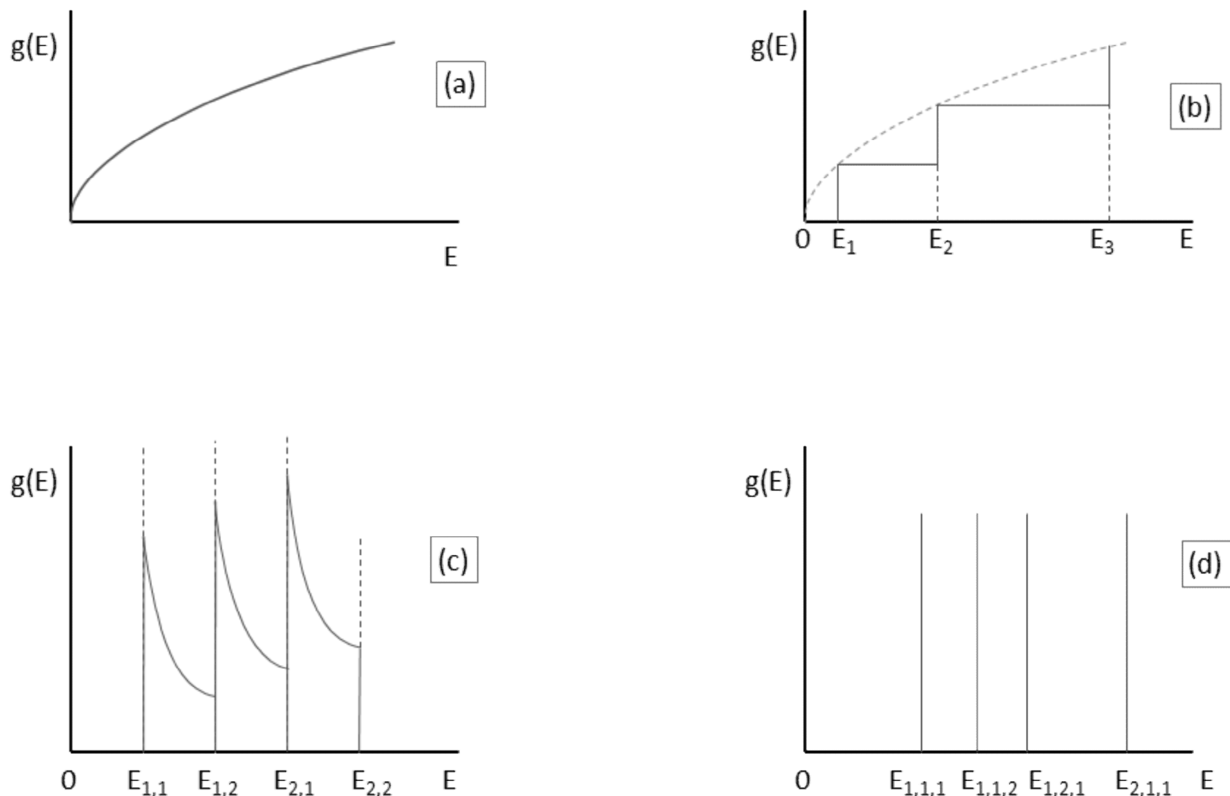


**Figure 1: The dependence of the density of states on energy. (a) 3D electron gas in an infinite crystal. (b) 2D electron gas (quantum well). (c) 1D electron gas (quantum wire). (d) 0D electron gas (quantum dot).**

and theoretically calculated values of the band gap of two colloidal quantum dots (CdSe and CdTe) depending on their radius. The calculations assumed that the confinement potential $V_0$ is 400 meV in all cases. In the calculations in [2] the choice of the value of the specified confinement potential $V_0$ was determined by the band gap of the matrix $E_g$, in which the quantum dots were placed ($V_0 \approx 0.08E_g$).

In order to create low-dimensional objects and structures, two approaches are used. In the first approach, based on special nanotechnology techniques, geometrical structures are formed with the desired shape and size. The second approach is based on the fact that using an electric field, one can manage the geometry of the isolated areas in semiconductors, characterized by a certain type of charge (electrons or holes) and certain concentration. The electric field may be the internal field in the semiconductor and also the electric field caused by voltage application from the outside to the system containing a semiconductor. Structures containing semiconductors in the second approach may be, for example, the structures metal-oxide-semiconductor, metal-semiconductor, and also heterostructures.

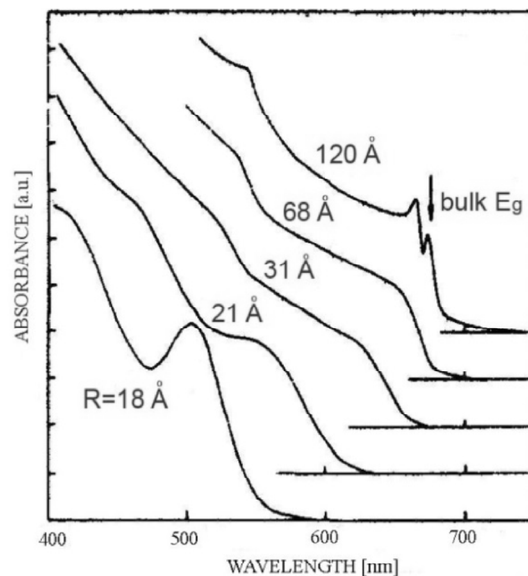


**Figure 2: Optical absorption spectra of CdSe with various radii at 2 K. The arrow indicates the bulk CdSe. After [1].**

*Superlattices.* Superlattices are crystalline structures in which, besides the periodic potential of the crystal lattice, there is another periodic potential, the period of which is significantly larger than the lattice constant, and that period is at the nanoscale. The repetition period of the layers in the superlattice is in the range of several nanometers to tens of nanometers. It is important to note that the semiconductor crystal lattice constants are a few tenths of a nanometer (for example, for silicon, germanium and gallium arsenide lattice constants are in the range of 0.5 to 0.6 nm). There are two main types of semiconductor superlattices: doped superlattices and compositional superlattices. Doped superlattices consist of a periodic sequence of p-type and n-type layers of the same semiconductor. Acceptor atoms in p-type layers bond the electrons which are supplied from n-type layers. As a result, an additional periodic potential is created by alternating charges of ionized donors and acceptors. Compositional superlattices are periodic heterostructures of alternating layers of different composition and a different band gap, but with similar values of the lattice constant. As a result, an additional periodic potential is created by means of periodic variation of the band gap. The concept of artificial superlattices was proposed in [3]*. At present, one of the main methods of creating superlattices is the method of molecular-beam epitaxy (MBE).

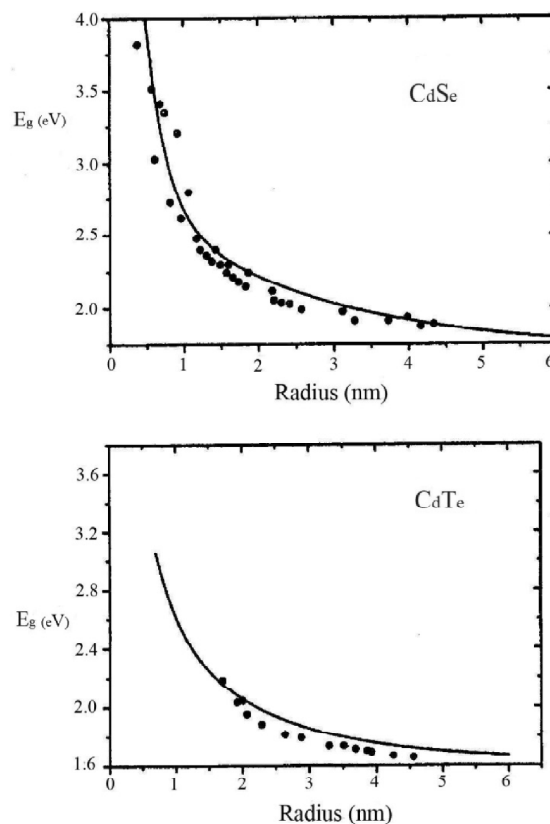


**Figure 3: Comparison between experimental and calculated effective bandgap energies for semiconductor quantum dots (CdSe and CdTe) as functions of the dot radius. The experimental data are plotted with symbols and the theoretical predictions with solid curves. The confinement potential is V0 = 400 meV in all case. After [2].**

*One of the authors of this work, Leo Esaki, is one of the laureates of the Nobel Prize in Physics in 1973, which he was awarded together with Ivar Giaever "for their experimental discoveries regarding tunneling phenomena in semiconductors and superconductors, respectively" (half of the Nobel Prize). The other half of the Nobel Prize in Physics in 1973 was given to Brian David Josephson "for his theoretical predictions of the properties of a supercurrent through a tunnel barrier, in particular those phenomena which are known as the Josephson effects".

Additional periodic superlattice potential significantly alters the original band structure of semiconductors. The superlattice can be regarded as a set of periodically spaced quantum wells. The band structure of the superlattice can be found on the basis of solutions of the Schrödinger equation with the additional periodic potential. Since the potential of the superlattice is periodic, energy levels of the quantum wells are broken down into bands. This is analogous to breaking down of the energy levels of electrons of the outer shells of the atoms during crystal formation. Since the period of the superlattice is much greater than the lattice constant, the resulting superlattice energy bands (minibands) represent a finer breaking down of the energy bands of initial semiconductor crystals. The above consideration suggests that the period of superlattices is substantially less than the electron mean free path. The energy spectrum of superlattices can be easily rebuilt by changing the thickness of the layers, which is easily accomplished by the MBE method. Fig. 4 shows the splitting of the conduction and valence bands in the composite superlattice into minibands.

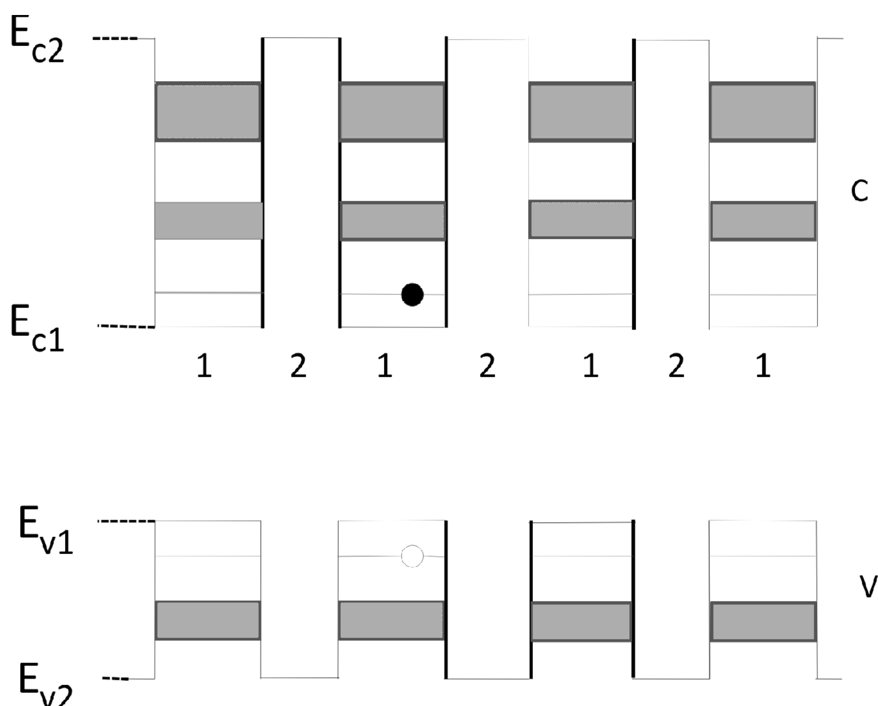


**Figure 4: Energy diagram of the composite superlattice. C is the conduction band, V – the valence band.**

Superlattices exhibit a number of interesting physical effects, which are reviewed, for example, in [4]. One such effect is the appearance of negative differential conductivity [3]. Such negative differential conductivity is due to the fact that electrons accelerated in a miniband perpendicularly to the layers of the superlattice can have a negative effective mass. We would also like to note the study of Bloch oscillations of electrons [5], infrared spectroscopy of electron states [6], study of the electron transport in electric fields [7], study of the nonlinear dynamics of carriers in superlattices [8]. Superlattices are used in a number of semiconductor devices such as lasers (see Section 4 of this article), light-emitting diodes, photodetectors, transistors with negative differential resistance.

## Graphene

Graphene is a truly two-dimensional crystal which consists of a single layer of carbon atoms assembled in a hexagonal lattice. In this lattice, $2s$, $2p_x$, and $2p_y$ orbitals of a carbon atom are hybridized so that each carbon atom is connected to three neighboring carbon atoms by strong $sp^2$- (or σ-) bonds. The remaining $p_z$- (or π-) orbital determines the low-energy electronic structure of graphene, including the valence band and the conduction band. The variety of physical and chemical properties of graphene is due to its crystalline structure and the presence of π-electrons in the atoms

that constitute graphene. Production and study of the first samples of graphene was carried out in 2004 [9][*]. In the first experimental study of the properties of graphene a simple but effective mechanical exfoliation method was used for extracting thin layers of graphite from a graphite crystal with Scotch tape. Then these layers were transferred to a silicon substrate.

Graphene has a number of unique electronic properties, including very high electron mobility at room temperature. It can be as high as about 200 000 cm$^2$ V$^{-1}$ s$^{-1}$. For comparison, electron mobility in pure crystalline silicon is two orders of magnitude lower, at about 1400 cm$^2$ V$^{-1}$ s$^{-1}$. The presence of high mobility makes graphene a promising material for the creation of high-frequency transistors.

In the structure of the energy bands of single-layer graphene there is no band gap, and at the points of contact of the conduction and valence bands, the energy spectrum of electrons and holes for single-layer graphene is a linear function of the wave vector (in conventional semiconductors and dielectrics, the energy of electrons and holes near the bottom of the conduction band and the top of the valence band depend on the wave vector quadratically). Ultrarelativistic particles, photons and neutrinos have a similar energy spectrum. We can say that the effective mass of electrons and holes in graphene in the vicinity of the points of contact of the conduction band and valence band is zero. However, despite the similarity of the electrons and holes in graphene to photons and other massless particles, they also have significant differences. First, the electrons and holes in graphene are charged particles. Second, the electrons and holes in graphene are fermions.

The unique electronic properties of graphene can be seen in optics as well. In particular, using graphene, it is relatively easy to "see" the fine-structure constant $\alpha$. For this, it is necessary to compare the intensity of the light passing through the aperture covered by graphene to the intensity of the light transmitted through the uncovered aperture. For visible light, the coefficient of transmission of light for single-layer graphene is given by a simple formula: $T \approx 1 - \pi\alpha \approx 0.977$. The first experimental results for graphene's optical transmittance $T$, including both single-layer graphene, and few-layer graphene, were obtained in [10].

Since 2005, a large amount of work has been done to study the quantum Hall effect (QHE) in graphene. The first works in which the QHE was observed in graphene were [11, 12]. It turned out that the QHE in graphene can be observed even under room temperatures (at high magnetic fields). For graphene, the Hall conductance is

$$R_{xy}^{-1} = \pm 4\left(n + \frac{1}{2}\right)\frac{e^2}{h} \tag{7}$$

where $n$ is a non-negative integer and ± stands for electrons and holes, respectively. Although the QHE was observed in many 2D systems[*], the QHE in graphene is very different from "normal" QHE, because equation (7) contains the term 1/2. Such difference is related to the properties of graphene, including the relativistic nature of the motion of charged particles in graphene.

One of the main areas of research of graphene is related to using graphene in electronics. In the future, graphene could supplement or partially replace silicon devices in micro/nanoelectronics. However, there is the following problem. Using graphene directly when creating a field-effect transistor is not possible due to lack of band gap. In other words, it is not possible to make a significant difference in the resistance under any applied voltages to the gate - that is, one cannot set two conditions that are suitable for binary logic: conductive and non-conductive. One possible way of solving this problem is to use the effect of quantum confinement of a charge in graphene nanoribbons [14, 15]. For such graphene nanoribbons, the band gap already exists. Another way is chemical modification of

---

[*]Two co-authors of this work, Andre Geim and Konstantin Novoselov, were awarded the 2010 Nobel Prize in Physics for "groundbreaking experiments regarding the two-dimensional material graphene".

[*]Note that the first quantum Hall effect was observed in [13] in a two-dimensional electron gas, realized with a silicon metal-oxide-semiconductor field-effect transistor. In this work, it was found that the Hall conductance has fixed values $R_{xy}^{-1} = ne^2/h$, where $n$ is an integer ($n$ = 1,2,3, ...). The experiments were carried out at low temperatures (typically 4.2 K) in order to suppress disturbing scattering processes originating from electron-phonon interactions. Explanation of the quantum Hall effect is based on the account of quantization of electron motion in the plane of the two-dimensional system (Landau quantization), if a strong magnetic field is applied perpendicular to the plane. At present, on the basis of the quantum Hall effect the standard unit of electrical resistance is realized. For his discovery, Klaus von Klitzing was awarded the 1985 Nobel Prize in Physics.

graphene. Such a modification can be made, using, e.g., reaction with hydrogen atoms in the hydrogen plasma (see, for example, [16]). This may form graphane – a two-dimensional material, in which one carbon atom is bonded with one hydrogen atom and three carbon atoms. The third possible way is to use a two-layer graphene instead of a single-layer graphene [17, 18]. For a two-layer graphene, dependence of electron energy (near the bottom of the conduction band) and holes (near the top of the valence band) on the wave vector is quadratic (in contrast with single-layer graphene). However, for bilayer graphene the bandgap is also absent, as the bottom of the conduction band and top of the valence band are in contact with each other. However, if an electric field is applied perpendicularly to the plane of the bilayer graphene, the shape of the conduction band (near the bottom) and the valence band (near the top) will change in such way that the contact zone disappears, and the band gap appears.

Another application of graphene is proposed in [19] and is based on the use of graphene as a very sensitive sensor for detection of molecules of chemical substances that are attached to the surface of the graphene film. In this paper it is shown that if graphene has micrometer dimensions, it is possible to detect a single molecule. The principle of operation of this sensor is based on the fact that different molecules on the graphene surface can act as donors and acceptors, which in turn leads to a change in the resistance of graphene. The high sensitivity of graphene to the appearance of individual molecules on its surface is due to a very low level of intrinsic fluctuations caused by the thermal motion of charges and defects in graphene.

One of the main methods for the identification of graphene is Raman spectroscopy. The method of Raman spectroscopy is a rapid and non-destructive method allowing to distinguish single-layer graphene from graphite, as well as two-and three-layer graphene [20, 21]. Figure 5 shows a typical Raman spectrum of graphite and single-layer graphene. A characteristic feature of these spectra is the presence of two peaks. One peak is a G peak. It corresponds to the Raman shift at ~1580 $cm^{-1}$. This peak characterizes oscillations of $sp^2$ bonds in the carbon atoms. Another peak is a 2D peak at ~2700 $cm^{-1}$. 2D peak is the second order of the D peak. In turn, D peak is due, probably, to the breathing modes of $sp^2$ atoms in rings (for example, see [22]). Because of the selection rules, the D peak appears in Raman spectra of graphene and graphite only at significant number of defects. The absence of D peak in Fig. 5 indicates that under the experimental conditions for graphene, as well as for graphite, a significant number of defects was absent. In the spectrum of single-layer graphene there is a sharp 2D symmetrical peak. When we are dealing with graphite, 2D peak is considerably asymmetric. With the increase in the number of graphene layers, the Raman shift of 2D peak monotonically increases [20]. Further information on the number of graphene layers can be obtained using the intensity ratio $I_G/I_{2D}$ [23].

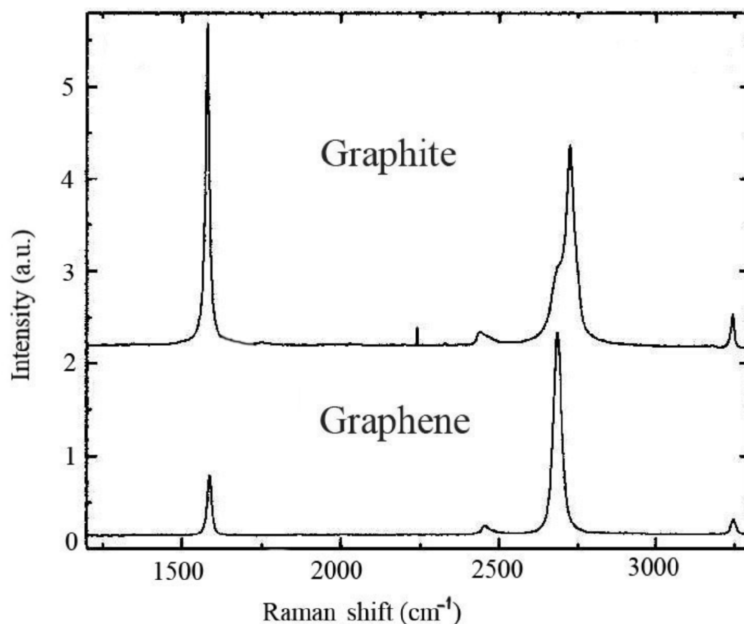


**Figure 5: Comparison of Raman spectra at 514 nm excitation for bulk graphite and graphene. After [20].**

## Carbon Nanotubes

Carbon nanotubes were discovered in the study of products which are formed by electric arc discharge with carbon electrodes (in a helium atmosphere) [24]. The author of [24] carried out a study of reaction product formed at the cathode using electron microscopy. Electron microscopy showed that on the cathode carbon needle-like tubes with a diameter of 4 to 30 nm appear. This was the first observation of carbon nanotubes.

A single-wall carbon nanotube is a ribbon of single-layer graphene folded into a cylinder. Graphene can be rolled into a cylinder in many ways, which leads to a large variety of types of carbon nanotubes. Nanotubes can be described by their chiral vector:

$$\vec{C} = n\overrightarrow{a_1} + m\overrightarrow{a_2}$$

where $\overrightarrow{a_1}$ and $\overrightarrow{a_2}$ are the basis vectors of the graphene lattice, and $n$ and $m$ are integers (see Fig. 6). We draw two lines $A_1$ and $A_2$ perpendicular to vector $\vec{C}$ through the beginning of $\vec{C}$ and the end of $\vec{C}$, and cut the ribbon from the graphene layer along these lines. If we roll the ribbon into a cylinder so that lines $A_1$ and $A_2$ are aligned, we will obtain a structural model of a carbon nanotube. The length of the circumference of the cross section of the nanotube equals the magnitude of the vector $\vec{C}$. The diameter of a single-wall carbon nanotube is determined by the integers $n$ and $m$:

$$d = \frac{d_0}{\pi}\sqrt{3(n^2 + m^2 + nm)}$$

where $d_0 = 0.142$ nm is the distance between adjacent carbon atoms in the graphene plane.

Single-wall carbon nanotubes (SWCNTs) can possess both metallic and semiconducting electronic properties. SWCNTs are metal tubes if the difference $n - m$ is a multiple of three. In all other cases SWCNTs are semiconducting tubes. For metallic SWCNTs the bandgap is zero (see Fig. 7b, 7c), and for semiconducting SWCNTs it is a finite value (see Fig. 7a). The typical values of the band gap of semiconducting SWCNTs are in the range of tenths of eV to about 1 eV.

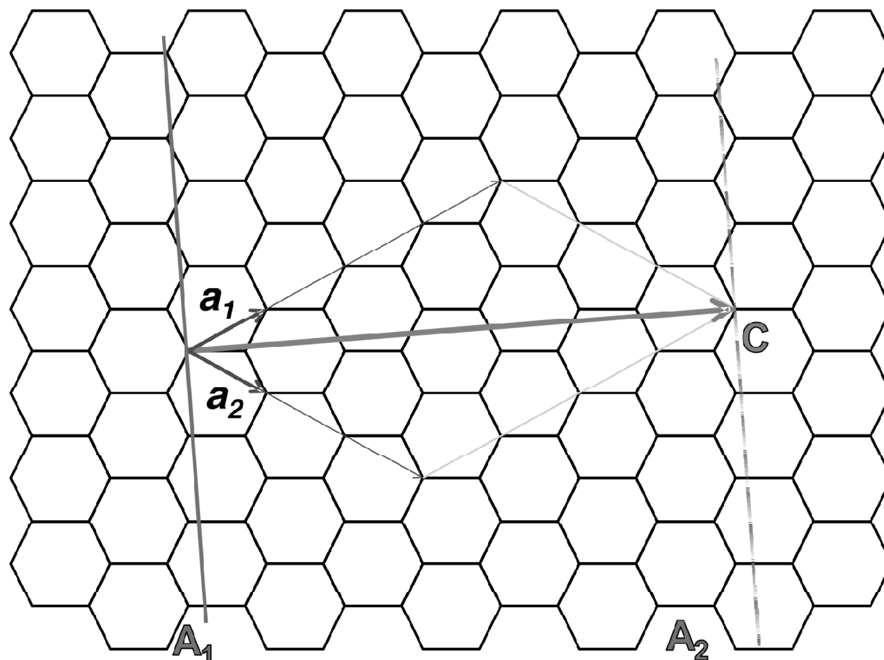


**Figure 6: Construction of a structure model of a single-wall carbon nanotube using a graphene sheet. Vectors $\vec{a}_1$ and $\vec{a}_2$ form basis vectors for the graphene lattice. Here the chiral indices (n,m) are: n = 4, m = 3.**

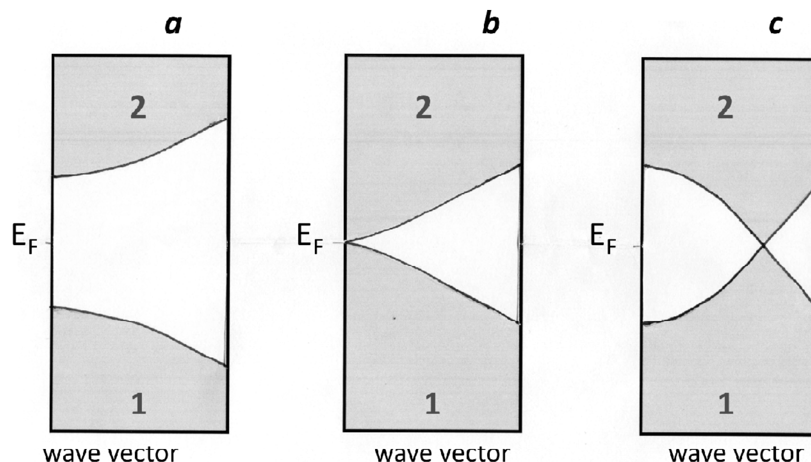


**Figure 7: Electronic structure of single-wall carbon nanotubes (SWCNTs). (a) Semicomducting SWCNT. (b), (c) Metallic SWCNTs; the chiral indices (n, m) are as follows: for Fig. 7b n is a multiple of 3, m=0; for Fig. 7c n=m. The Fermi level $E_F$ separates populated (1) and vacant (2) states.**

A typical dependence of the electron density of states on the electron energy for semiconducting and metallic SWCNTs is shown in Fig. 8 (cf. [25]). Similar dependences were observed experimentally using tunneling spectroscopy (see, e.g., [26, 27, 28]). In Fig. 8, the symbols $S_{11}$, $S_{22}$ show low-energy band gap transitions in semiconducting SWCNTs, and the symbol $M_{11}$ shows a low-energy band gap transition in metallic SWCNTs. The energy gaps, corresponding to these transitions, $E_{11}{}^{S}$, $E_{22}{}^{S}$ and $E_{11}{}^{M}$, are approximately expressed by simple equations (see [29])

$$E_{11}^{S} \approx \frac{A}{d}, E_{22}^{S} \approx \frac{2A}{d}, E_{11}^{M} \approx \frac{3A}{d} \tag{8}$$

where $d$ is the diameter of SWCNT, and the value of $A$ does not depend on the diameter $d$ or on chiral indices $n$ and $m$. The peaks corresponding to transitions $S_{11}$, $S_{22}$, $M_{11}$ are experimentally observed in the absorption spectra of SWCNTs (see, e.g. , [30, 31, 32]). Since the SWCNT product, obtained by synthesis of SWCNTs, generally contains the mixture of a large number of semiconducting and metallic SWCNTs, the experimental data on the intensities of lines $S_{11}$, $S_{22}$, $M_{11}$ in the absorption spectrum of SWCNTs provide information on the relative quantity of the semiconducting and metallic SWCNTs in the SWCNT samples [30, 32].

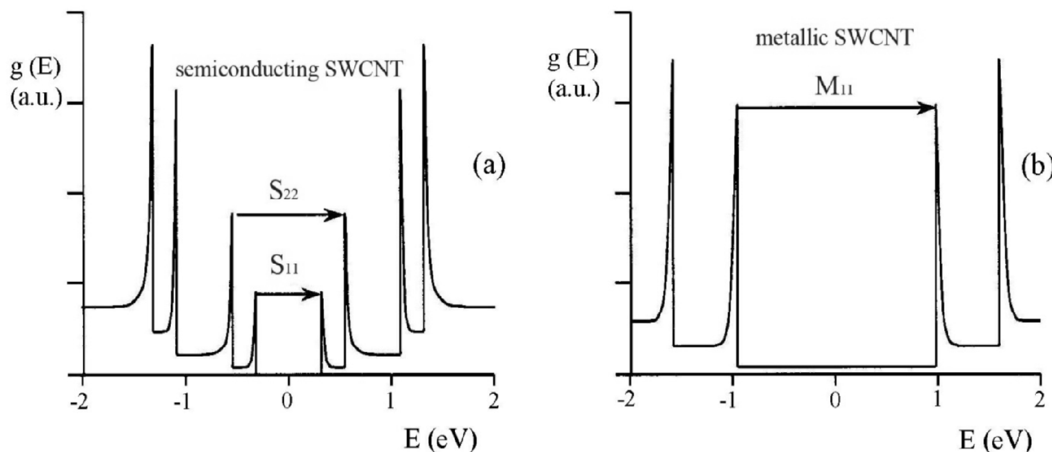


**Figure 8: Density of electronic states of semiconducting (a) and metallic (b) single-wall carbon nanotubes (SWCNTs). Absorption spectra of SWCNTs are formed as a result of interband transitions, including the transitions $S_{11}$ and $S_{22}$ (the first and second interband transitions of semiconducting SWCNTs, respectively) and the first interband transition $M_{11}$ of metallic SWCNTs. The zero of the energy axis corresponds to the Fermi level.**

In [33], the intense photoluminescence of semiconducting SWCNTs in the near infrared was first discovered, which arose after the excitation in the visible range by pulse laser radiation.* Observation of such luminescence allowed to suggest a method for determining the chiral indices *n* and *m* in the semiconducting SWCNTs. The sketch of spectroscopic transitions in semiconducting SWCNTs is shown in Fig. 9. The laser radiation excites the transition $S_{22}$. After that, there is a rapid non-radiative relaxation of the electron into the lowest state of the conduction band and of the hole in the topmost state of the valence band. Afterwards the radiative transition $S_{11}$ (luminescence) occurs, resulting in recombination of electrons and holes. The wavelengths $\lambda_{22}$ and $\lambda_{11}$ of transitions $S_{11}$ and $S_{22}$ for SWCNTs are determined by chiral indices *n* and *m*. (Equations (8) are approximate. In fact, $E_{11}^{S}$ as well as $E_{22}^{S}$ depend on indices (*n*, *m*) in a more complicated fashion than (8) show.) Therefore, varying the excitation wavelength $\lambda_{22}$ and recording the luminescence wavelength $\lambda_{11}$, it is possible, using the data for $\lambda_{11}$ and $\lambda_{22}$, to determine the relative quantity of various semiconducting SWCNTs (with various chiral indices *n* and *m*) for an unknown mixture [34, 43, 36, 31, 37].

Another important method of characterization of the carbon nanotubes is Raman spectroscopy [38]. In [39], the first observation of Raman spectra from individual spatially isolated free-standing SWCNTs with identified (*n*, *m*) indices was presented. The intensity of the Raman scattering signal undergoes a sharp increase when the frequency of the incident or the scattered photon is in resonance with the frequency of the electronic molecular transition. It is the existence of such resonance that allowed for observation of Raman scattering from individual free-standing

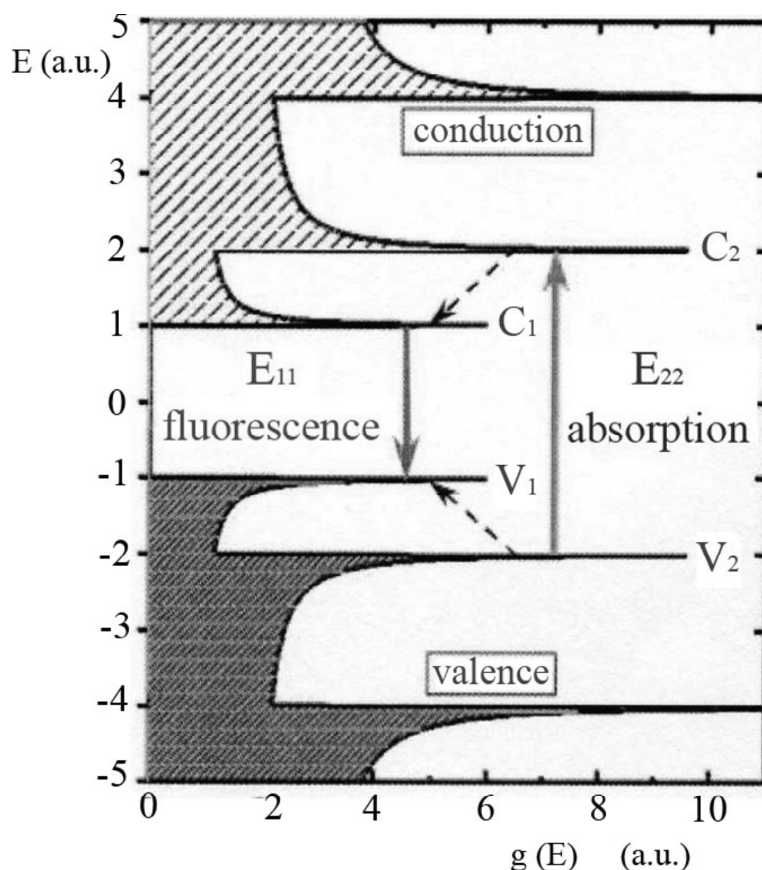


**Figure 9: Schematic density of electronic states for a single-wall carbon nanotube (SWCNT). Solid arrows indicate the optical excitation and emission transitions of interest; dashed arrows indicate nonradiative relaxation of the electron (in the conduction band) and hole (in the valence band) before emission. The zero of the energy axis corresponds to the Fermi level. After [34].**

*We would like to note that one of the co-authors of papers [26, 33, 34] Richard E. Smalley was awarded the Nobel Prize in Chemistry in 1996 with Robert F. Curl Jr. and Sir Harold W. Kroto "for their discovery of fullerenes". Fullerenes are another form of carbon. $C_{60}$ fullerene molecule is a closed sphere, made up of regular pentagons and hexagons with carbon atoms at the vertices.

SWCNTs in [39 – 41]. On the other hand, the resonance (the consequence of which is a dramatic increase in the Raman signal) allows estimating with a good accuracy the transition frequencies in index-identified individual SWCNTs [40, 42]. Independent identification of $(n, m)$ indices in SWCNTs was carried out in [39] using the electron diffraction (see also [43]). The Raman spectra of SWCNTs consist of two main peaks at substantially different frequencies. First, the so-called radial breathing mode (RBM). It corresponds to radial expansion-contraction of the nanotube; during it, all carbon atoms move coherently in the radial direction. The RBM appears only in carbon nanotubes, so its observation in the Raman spectrum is direct evidence that the sample contains SWCNTs. The typical RBM range is 100 – 500 $cm^{-1}$. The RBM frequency is inversely proportional to the SWCNT diameter and is expressed as (see [39])

$$\omega_{RBM}(cm^{-1}) = \frac{204}{d(nm)} + 27(cm^{-1}) .$$

Another peak in the Raman spectrum of SWCNTs is the G band spectra (see, e.g., [40]). This band is split into many features around 1580 $cm^{-1}$. Both peaks - the RBM and the G band spectra - are the first-order Raman peaks. Note that in graphene the only first-order Raman peak is the G peak.

Carbon nanotubes have interesting electrical properties. For instance, in [44] ballistic electrons transport of in nanotubes was experimentally established, i.e., the transport without electron scattering and energy loss. It was discovered that the conductance of the carbon nanotube is equal to the conductance quantum $2e^2/h = (12.9\ k\Omega)^{-1}$. We would like to note that a more detailed analysis of quantization of conductance in low-dimensional systems follows in the next section. Carbon nanotubes can find wide application in vacuum electronics. Electron field emission is used in electron guns in electron microscopes, flat panel displays, microwave amplifiers. Electron field emission from carbon nanotubes was first demonstrated in 1995 [45] and has been studied intensively since then. For practical applications, materials that emit electrons must have low threshold emission fields and should be stable at high current density. Carbon nanotubes possess these properties. Carbon nanotubes are a promising material for nanoelectronics. In particular, active work is being done to create a variety of modifications of field-effect transistors containing carbon nanotubes (see, for example, [46 – 48]). Many experiments show (see, for example, [49, 50]) that carbon nanotubes can be an effective medium for the storage of hydrogen, which is an environmentally friendly source of energy.

## 3. BALLISTIC TRANSPORT OF ELECTRONS IN LOW-DIMENSIONAL STRUCTURES. CONDUCTANCE QUANTIZATION

The electrical conductance in semiconductor nanostructures has interesting features. The conductance of a normal macroscopic conductor is given by

$$G = \sigma \frac{S}{L} \tag{9}$$

where $\sigma$ is the conductivity, $L$ is the length, $S$ is the cross-sectional area ($S = \pi r^2$, $r$ is radius of the conductor cross-section). Formula (9) is valid when the mean free path of the electron $l$ is much smaller than $L$ and $r$. In this case, the motion of electrons through a conductor is diffusive. In another case $l > r$ and $l > L$ electrons pass from one contact to another in ballistic regime, without experiencing collisions with impurities and lattice atoms.

For simplicity, let us consider the situation when a conductor, through which electrons are moving in ballistic regime, is a one-dimensional structure such as the quantum wire. Let us assume that such quantum wire is connected to the two metal contacts, $a$ and $b$, characterized by the Fermi level $E_{F\alpha}$ and $E_{F\beta}$. Let us also assume that the system temperature does not exceed a few Kelvins. In this case, the energy of almost all electrons in the metal contacts is less than the energy of the corresponding Fermi levels $E_{F\alpha}$ and $E_{F\beta}$. If we apply a weak voltage $V$ between the metal contacts, there will be a displacement of the energy levels of electrons of one metal ($\alpha$) relative to the energy levels of electrons of the other metal ($\beta$), while $E_{F\alpha} - E_{F\beta} = eV$. The current will be created only by the electrons with

energies between $E_{F\alpha}$ and $E_{F\beta}$. The amount of current can be determined as $I = eNv$, where $N$ is the electron density, $v$ is their velocity. Expressing the concentration of electrons $N$ for the quantum wire through the density of states $n_{1D}(E)$

$$n_{1D}(E) = \frac{2}{\pi\hbar v} \tag{10}$$

and considering that the energy of the electrons are in a small range $eV$, for the value of the current we have

$$I = e\frac{n_{1D}}{2}eVv = \frac{e^2}{\pi\hbar}V. \tag{11}$$

In (11) the factor ½ takes into account the fact that only half of the electrons are moving in the same direction.

From (11) we obtain the conductance $G = I/V$:

$$G = \frac{e^2}{\pi\hbar} = \frac{2e^2}{h}$$

i.e. the conductance of the quantum wire is independent of its length. The value of $G$ is related to the quantum unit of resistance $R_0$ ($R_0 = h/e^2 = 25.813$ kΩ) by the ratio: $G = 2/R_0$. The generalization of the results is related to the possible presence of several energy subbands for the quantum wire in the range from $E_{F\alpha}$ to $E_{F\beta}$.

Taking into account the presence of multiple channels in electron transport in ballistic regime gives the following formula for the electrical conductance:

$$G = \frac{2e^2}{h}\sum_{n,m=1}^{N_C} T_{mn} \tag{12}$$

where $T_{mn}$ are the transmission coefficients, and $N_C$ is the total number of channels contributing to conductance. It is important to note that for the quantum wire the channels are sub-bands defined by (4), (5). Equation (12) takes into account the fact that the electrons may be injected into any channel $n$, and then, after interaction with the scattering center, occur in the other channel $m$. If during the ballistic transport of electrons there is no channel mixing, then $T_{mn} = \delta_{mn}$, and formula (12) transforms into the formula

$$G = \frac{2e^2}{h}N_C. \tag{13}$$

One should note that significant theoretical results in the field of quantized conductance for ballistic electron transport were obtained in [51, 52]. The first experimental observation of quantized conductance was performed in [53, 54].

## 4. LASER PHYSICS

### Semiconductor diode lasers. Improving characteristics of semiconductor diode lasers with the help of semiconductor nanostructures.

In order to illustrate the principle of the semiconductor diode laser, let us consider Fig. 10. This figure shows the valence band V and the conduction band C of a semiconductor. Between these bands is a band gap of width $E_g$. We first assume that the semiconductor is at temperature $T = 0$ K. In this case, the semiconductor conduction band is empty and the valence band is completely filled with electrons. Let us assume that there is a pump mechanism, which throws some electrons from the valence band to the conduction band. In this case, after a short time (of the order of 1 ps), due to relaxation, electrons in the conduction band move to the lowermost unoccupied energy levels of this band. Similarly, electrons which are on top of the valence band make a transition to the lowermost available

levels of the valence band. Thus, such electrons leave "holes" in the upper part of the valence band. We would like to introduce the concept of the Fermi level of a semiconductor $E'_{Fc}$ for the conduction band and the Fermi level of a semiconductor $E'_{Fv}$ for the valence band. The Fermi level ($E'_{Fc}$ or $E'_{Fv}$) is the energy level, for which at temperature $T = 0$ K, the following takes place: above the Fermi level in the conduction band ($E'_{Fc}$ case) or in the valence band ($E'_{Fv}$ case) all the electron energy levels are unoccupied, and below which all energy levels are fully occupied by electrons. Radiation can occur when an electron from the conduction band goes down to the valence band, recombining itself with a hole. Firstly, there occurs spontaneous recombination radiation. Such radiation is used in light-emitting diodes. Secondly, under certain conditions, there may appear forced recombination radiation, leading to lasing (the active lasing medium must be in the cavity). It can be shown that for lasing to take place, it is necessary that the following condition be satisfied:

$$E_g \leq \hbar\omega \leq E'_{Fc} - E'_{Fv} \tag{14}$$

where ω is the frequency of the emitted photons. When the condition (14) is satisfied, the probability of stimulated emission of a photon in a semiconductor is higher than the probability of absorption of a photon. We would like to emphasize that the condition of the possible onset of lasing (14) is valid for any temperature, not only for $T = 0$ K (in the latter case this condition can be understood using Fig. 10b). A necessary condition for (14) is the inequality: $E_g \leq E'_{Fc} - E'_{Fv}$. The Fermi energies $E'_{Fc}$ and $E'_{Fv}$ depend on pump intensity. The more intense the pumping is, the greater is the value $E'_{Fc}$ and the smaller is $E'_{Fv}$. Pumping of the semiconductor laser can be accomplished in several ways, including excitation by the radiation of another laser and electron beam pumping. However, the main type of semiconductor lasers at the moment are double heterostructure injection lasers.

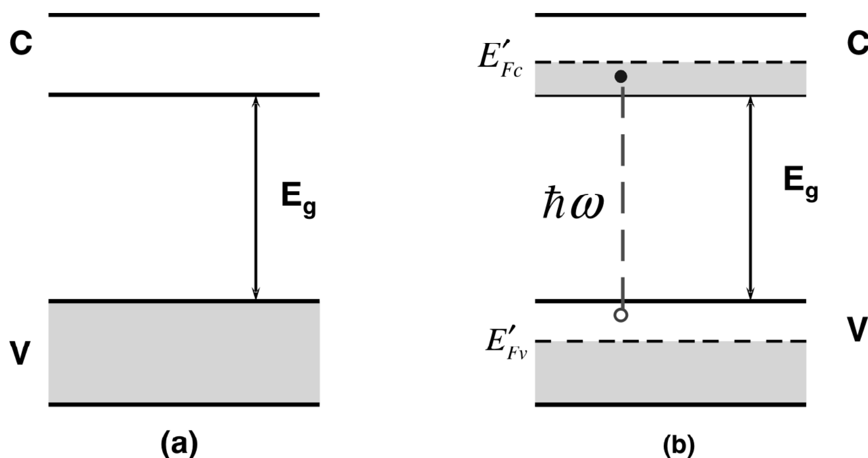


**Figure 10: The principle of operation of a semiconductor diode laser.**

The concept of double heterostructure laser was formulated in 1963 independently by Zh. I. Alferov and R. F. Kazarinov [55], and H. Kroemer [56]*. About 5 years after papers [55, 56] were published, very fast progress started in the experimental implementation of double heterojunction lasers (see, e.g., [56 – 59]).

Figure 11 illustrates a structure of a double heterojunction laser and the dependence of the refractive index of the material on the coordinate. As the active layer, intrinsic semiconductor GaAs is most commonly used, and as the outer layers – impurity semiconductors of p- and n-type $Al_{0.3}Ga_{0.7}As$. A typical thickness $d$ of the active layer is in the range of 100 to 200 nm. Figure 12 shows the energy diagram of a double heterojunction structure. During forward bias into the active layer (GaAs), electrons are injected from the n-$Al_{0.3}Ga_{0.7}As$ layer and the holes from the p-$Al_{0.3}Ga_{0.7}As$ layer, which is shown in Figure 12 with the curved arrows. The characteristic features of this structure

*Zhores I. Alferov and Herbert Kroemer were awarded in 2000 the Nobel Prize in Physics "for developing semiconductor heterostructures used in high-speed-and opto-electronics".

are as follows. First, the refractive index of the active layer ($n_1$ = 3.6 for GaAs) is much greater than the refractive index of the outer layer ($n_2$ = 3.4 for $Al_{0.3}Ga_{0.7}As$). Therefore, the light travels along the active layer as in the waveguide, due to the effect of total internal reflection of a significant part of the induced photons (photon localization effect). Second, the band gap width $E_{g1}$ of the active region is substantially smaller than the width $E_{g2}$ of the band gap of the outer layers . ($E_{g1} \approx 1.5$ eV for GaAs and $E_{g2} \approx 1.8$ eV for $Al_{0.3}Ga_{0.7}As$). Therefore, the electrons and holes cannot leave the active layer, since it is limited by potential barriers, and all the recombination processes occur in the active layer.

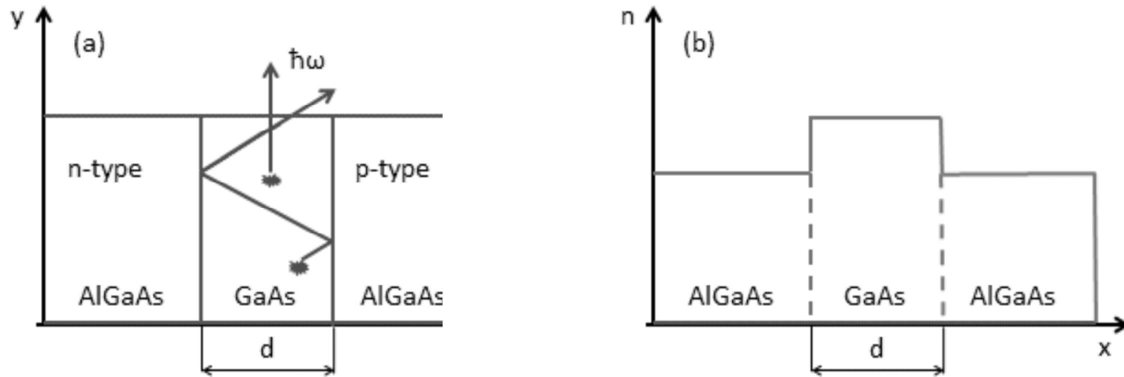


**Figure 11: Geometric characteristics (a) and the refractive index profile (b) of a double heterostructure semiconductor diode laser.**

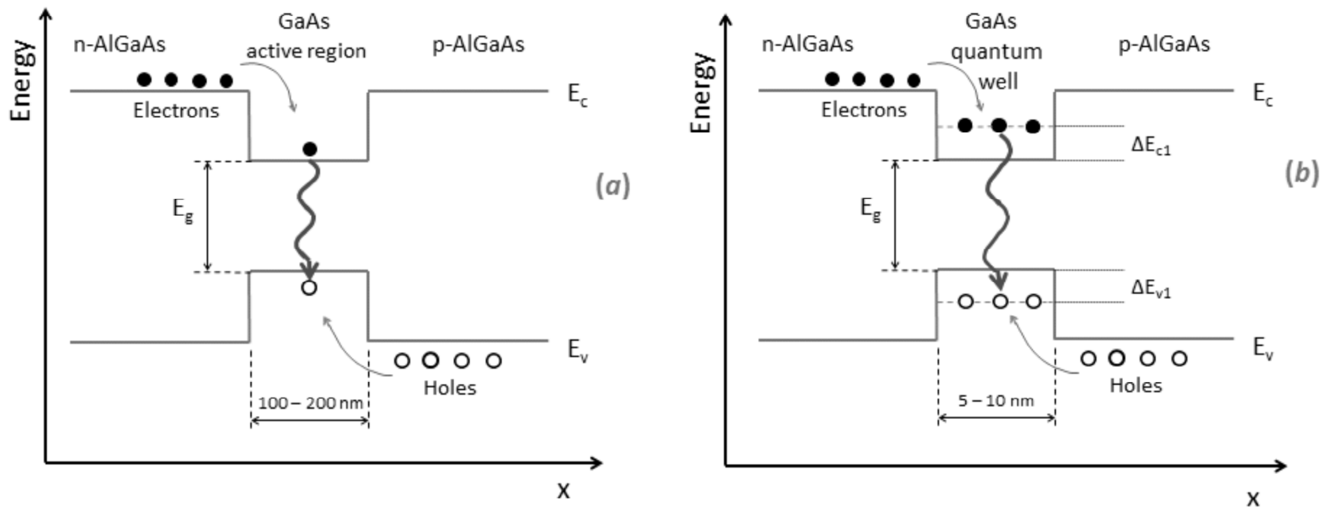


**Figure 12: Energy characteristics of semiconductor diode laser on a double heterojunction. x-coordinate dependence of the valence band edge ($E_V$) and conduction band edge($E_C$) for (a) the case of a wide layer of GaAs (*d* = 100 - 200 nm) and (b) a nanoscale thickness of the GaAs (*d* = 5 - 10 nm). The wavy arrows indicate the laser transitions.**

If the active layer thickness $d$ is reduced to a size comparable to the de Broglie wavelength, the active layer will be a quantum well (Fig. 12b). In this case, the electrons and holes injected under forward bias will be located on the energy levels of the quantum well (the electrons in the conduction band and the holes in the valence band) . Recombination transitions give radiation of frequency

$$\omega = \hbar^{-1}(E_g + \Delta E_{c1} + \Delta E_{v1})$$

where $\Delta E_{c1}$ is the energy level (measured from the bottom of the conduction band) at which electrons are collected, and $\Delta E_{v1}$ is the absolute value of the energy level in the valence band (measured from the top of the valence band), the recombination transition of the electron onto which will lead to the emission of a photon. The first experimental works where quantum well lasers were implemented, were papers [61 – 63]. One of the advantages of quantum-well

lasers is the possibility of radiation frequency tuning by varying the thickness of the active layer *d*. With a decrease of *d*, the distance from the edges of the bands (the conduction band and valence band) to the first dimensional energy levels increases, and this increases the frequency of the radiation. Another advantage of quantum-well lasers is that their threshold current density value $I_{th}$ (i.e., the current density at which lasing begins) is about 4 – 5 times less than for conventional double heterostructure lasers. For quantum-well lasers, the typical $I_{th}$ value is about 200 A/cm$^2$.

### Quantum Cascade Lasers

The semiconductor diode lasers described above are based on radiative recombination of electrons and holes in the active layer through the band gap lying between the conduction and valence bands. The electrons and holes are injected into the active layer through forward bias in the p-n transition. For these lasers, the laser photon energy is mainly determined by the band gap width of the semiconductor constituting the active layer.

Quantum cascade lasers (QC lasers) are fundamentally different from the above semiconductor lasers. They are not based on radiative recombination of electrons and holes, but use optical transitions between the electronic states lying in the tunnel-coupled quantum wells. In other words, in contrast to the semiconductor diode lasers, QC lasers use only one type of carriers – electrons, and the working energy levels are those related to the conduction band of the semiconductor. In QC lasers injectors, through which the pumping of the upper laser level and the draining of the population of the lower laser level are carried out, are superlattices.

In the pioneering work [64] it was suggested to use internal transitions between minibands and the effect of tunneling in semiconductors with a superlattice for amplification of electromagnetic radiation. Work [64] was published shortly after the publication of [3] devoted to the study of the properties of superlattices.

The first experimental demonstration of the QC laser was introduced in article [65]. Figure 13 illustrates the principle of the QC laser. To operate a QC laser, a sequence of steps is used, each of which includes the injection region and the active region separated by a tunneling barrier. Two such steps are shown in the figure. A characteristic feature of the active region (see Fig. 13) is three energy levels (three subbands) 1, 2 and 3, appearing in a system of three tunnel-coupled quantum wells. The electrons, because of tunneling, are injected in subband 3, which is the upper laser level. Lasing occurs at the transition between subbands 2 and 3 (in transition $3 \rightarrow 2$). The structure shown in Fig. 13 was grown in work [65] using molecular beam epitaxy. The quantum wells and barriers in the active region are formed by the materials GaInAs and AlInAs, respectively. Field injector consists of AlInAs-GaInAs superlattice with a constant period, which is shorter than the de Broglie wavelength for thermal electrons. Each region of the injection has n- type doping, which ensures the supply of electrons from the injection region into the active region. For the chart of the conduction band to have the shape of a staircase (as depicted in the figure), the electric field strength of about $10^5$ V/cm was applied to the system.

For lasing to occur at the $3 \rightarrow 2$ transition, it is necessary that population inversion occurs in this transition. Therefore, the thicknesses of the layers in the active region are chosen so that, firstly, $2 \rightarrow 1$ transition would be fast and non-radiative. In other words, it is necessary for the lifetime of electrons at level 2 to be small and for this level to empty itself rapidly. Secondly, we need the most probable radiative transition from level 3 to be the transition to level 2, and the lifetime of electrons at level 3 to be significantly longer than the lifetime at level 2. For continuous operation of a QC laser, it is necessary to continuously inject electrons to levels 3 and remove them from levels 1. Both processes (injection and removal) are done through the injection region by means of resonant tunneling through the entrance and exit barriers. The injection regions are superlattices. The thickness of the superlattice layers is selected in such way that at the output from the active region, electrons from levels 1 and 2 could tunnel into the miniband of the superlattice, and next to level 3 there would be a miniband gap of the superlattice that would make tunneling of electrons from level 3 into the superlattice impossible. The electric field is chosen in such way that the bottom of the miniband of the superlattice (on which electrons accumulate from tunneling from levels 1 and 2) would coincide with the level 3 of the next active region. Electrons which emitted radiation in the active region, tunnel into the superlattice through the output potential barrier and fly through the superlattice by resonance tunnel transitions. Thus they transfer excess energy to the atoms of the superlattice, transitioning onto the energy level near the bottom of the miniband of the superlattice. Then they tunnel through the potential barrier and get to level 3 of the next active region.

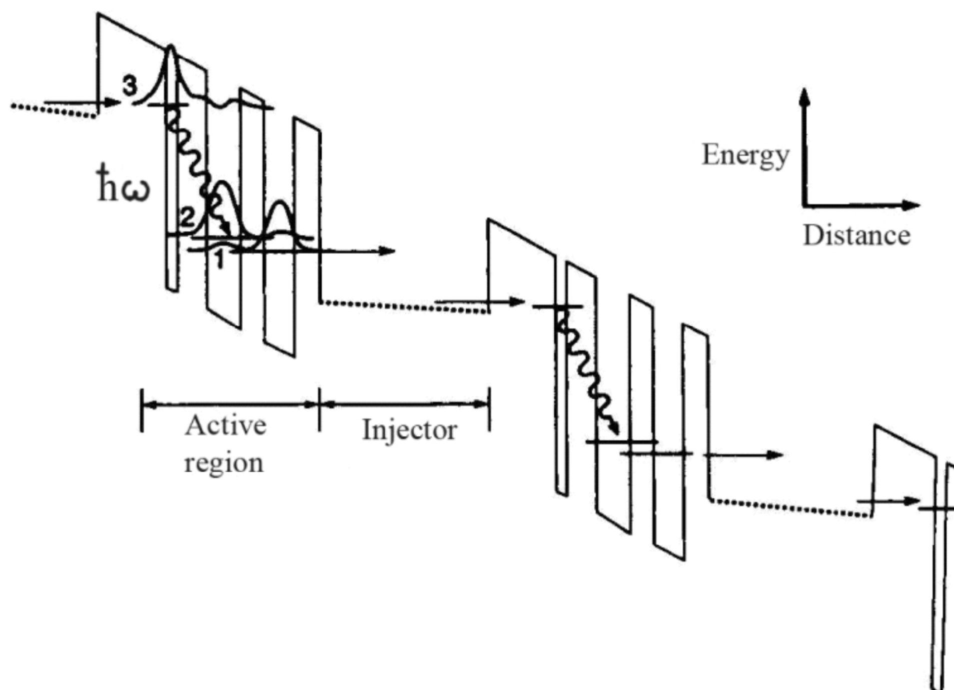


**Figure 13: Diagram illustrating the principle of the QC laser. Electrons are injected through the AlInAs barrier with a width of 4.5 nm to the level *n* = 3 of the active region. The active region contains GaInAs quantum wells of widths 0.8 nm, 3.5 nm and 2.8 nm. These wells are separated by AlInAs barriers of widths 3.5 nm and 3.0 nm. This relation between the widths of the wells and barriers provides a reduction in the spatial overlap between the wave functions of the states belonging to the levels *n* = 3 and *n* = 2. The electrons leave the third quantum well (width 2.8 nm) of the active region by tunneling through the barrier AlInAs of width 3.0 nm. The figure shows the squares of the absolute values of the wave functions for levels 1, 2 and 3. The calculated energy differences are as follows: $E_3 - E_2$ = 295 meV and $E_2 - E_1$ = 30 meV. The wavy arrow indicates the laser transition. After [65].**

Choosing the widths of wells and barriers in the active region (see caption to Fig. 13) provides a high probability of finding electrons at level 3 in the first and second quantum wells, and at levels 1 and 2 - in the second and third wells. This improves efficiency of electron injections into the active region and the drain from active region. In paper [65] the QC laser included 25 periods (steps), each of which contained the active region plus injector. Note that if the QC laser contains *N* steps, then each electron, having gone through all the steps (*N* steps), will create *N* laser photons.

Importantly, using the same materials forming the heterostructure of the QC laser, it is possible to create QC lasers over a wide wavelength range - from mid-infrared to submillimeter wave region. This is achieved by varying the structure of regions that make up the QC laser, including the width of the quantum wells and barriers.

The main applications of QC lasers are associated with spectroscopic detection of gas traces (environmental monitoring, medical respiratory monitoring, regulating industrial processes and the security sector).

## 5. NANOOPTICS

There are two ways of studying nanoscale objects by optical methods. One way is to use a probe radiation with a short wavelength (radiation in the far ultraviolet or X-ray range).

In nanooptics another way is realized - optical radiation range is used, with the possibility of overcoming the classical diffraction limit, formulated by E. Abbe and Lord Rayleigh in the 19th century. This limit defines the limits of the resolution of conventional optical devices. As it is well known (see, e.g., [66]), the corresponding criterion in the field of optical microscopy is expressed by the following inequality

$$\Delta x \geq 0.61\frac{\lambda}{NA} \tag{15}$$

where $\Delta x$ is the minimum distance between resolvable points, $\lambda$ is the wavelength of the radiation, $NA$ is the numerical aperture, equal to the product of the refractive index and the sine of the angle of aperture.

It is possible to overcome the specified classical diffraction limit with a substantial increase in the wave vectors of electromagnetic waves that can be achieved through the use of the dispersion properties of the medium in which radiation occurs. For this purpose, one can use evanescent (decaying) waves on the interfaces. In this case, the wave, which is attenuated in one direction, can have a significant increase in the wave vector in another direction.

A condition similar to (15) may be obtained using the Heisenberg uncertainty principle

$$\Delta x \cdot \Delta p_x \geq 0.5\hbar, \quad \Delta p_x = \hbar \Delta k_x \tag{16}$$

where $\Delta k_x$ is the dispersion of the projection of the wave vector on the axis $x$. From (16) it follows that for minimizing $\Delta x$ it is necessary to have the greatest possible dispersion of the x-projection of the wave vector, which, in turn, can be estimated as the length of the wave vector

$$k = \left(k_x^2 + k_y^2 + k_z^2\right)^{1/2} = \frac{2\pi}{\lambda}$$

where $\lambda$ is the wavelength of the radiation. Substituting this expression into (16) and taking into account $(\Delta k_x)_{\max} \approx (k_x)_{\max} \approx k$, we obtain the condition similar to criterion (15). Let us suppose now that we have an inhomogeneous medium. If in such medium the components $k_y$ and $k_z$ are imaginary (wave attenuation occurs along the axes $y$ and $z$), then for a fixed wavelength $\lambda$ it is possible to increase the dispersion of the projection of the wave vector $\Delta k_x$, because now $(\Delta k_x)_{\max} > k = 2\pi/\lambda$. The last inequality shows the possibility of reducing the minimum possible resolution of the distance along the x-axis.

It should be noted that the evanescent wave can be created when, for example, radiation is directed through an aperture with a radius smaller than the wavelength. Formation of images using evanescent waves is the basis of near-field optical scanning microscopy, the idea of which was proposed by the Irish scholar E.H. Synge in 1928 [67]. In this paper it is proposed to use a pinhole of a small diameter (less than $\lambda$) in the opaque plate, which is illuminated with light on one side, and on the other side abuts on the test sample, and is therefore not influenced by diffraction. Light transmitted through the hole and sample is collected by the microscope and registered by the photodetector. In order to obtain the image of the surface, the pinhole is moved along the sample surface. The resolution of such images is not limited by the wavelength of the radiation $\lambda$, but by the size of the pinhole. The first experimental realization of such a device was made in 1972 in experiments with radiation in the microwave range [68], where the resolution capacity $\lambda/60$ was detected. And in the mid-1980s, the super-resolution was demonstrated in the optical frequency range. The achieved resolution capacity was $\lambda/20$ [69, 70].

A key theme of nanooptics is the localization of light. The scope of research of nanooptics comprises the dispersion properties of electromagnetic waves, propagating in highly inhomogeneous media. An important role here is played by the dependence of the optical effects on the geometry of the media in which the waves propagate.

Surface plasmons can appear due to the interaction of the electric field of the electromagnetic wave incident on the surface of the conductor (e.g., metal) with surface charges. Surface plasmons are plasma waves propagating along the conductor surface. They are possible due to a gas of free electrons in a metal. A pioneering work studied these waves was paper [71]. The wave vector of the surface plasmon has the form

$$k_{SP} = \frac{\omega}{c}\sqrt{\frac{\varepsilon_1 \varepsilon_2}{\varepsilon_1 + \varepsilon_2}}. \tag{17}$$

Here $\varepsilon_1$ is the dielectric permittivity of the dielectric, and $\varepsilon_2$ is the dielectric permittivity of the metal. Equation (17) follows from the solution of the Maxwell equations. It represents a dispersion relation for the surface plasmon allowing for the frequency dependence of the dielectric permittivity of the metal $\varepsilon_2$. The important conditions for the existence of surface plasmons are related to the following inequalities:

$$\varepsilon_2' < 0, \qquad |\varepsilon_2'| > \varepsilon_1.$$

Here $\varepsilon_2'$ is the real part of the dielectric permittivity of the metal. It is assumed that there is no dissipative loss in the dielectric, and consequently, the value $\varepsilon_1$ is real and positive, $\varepsilon_1 \geq 1$; also the smallness of dissipative losses in the metal is assumed: $|\varepsilon_2''| << |\varepsilon_2'|$, where $\varepsilon_2''$ is the imaginary part of the metal permittivity. For the dielectric permittivity of the metal the following formula can be used:

$$\varepsilon_2(\omega) = 1 - \frac{\omega_{pe}^2}{\omega(\omega + i\nu)}$$

where $\omega_{pe}$ is the electron plasma frequency for the conduction electrons in the metal, $\omega_{pe} = (4\pi N_e e^2/m_e)^{1/2}$, and $\nu$ is the frequency of collisions of conduction electrons in the metal with ions, resulting in loss of directional velocity. The intensity of the electromagnetic field as a function of the z-coordinate (z-axis is orthogonal to the dielectric-metal interface) decays exponentially in the dielectric as well as in the metal. This decay in the metal occurs at a significantly smaller scale than in the dielectric. The electric field in the z-direction is evanescent. This field is a consequence of the non-radiative nature of surface plasmons that prevents the power of the wave from leaving the surface. Figure 14 schematically shows the propagation of surface waves along a dielectric-metal interface (along the x-axis). The value $k_{SP}$ is complex, because the permittivity of the metal $\varepsilon_2$ is complex. The imaginary part of $k_{SP}$ determines the propagation length of the surface plasmon.

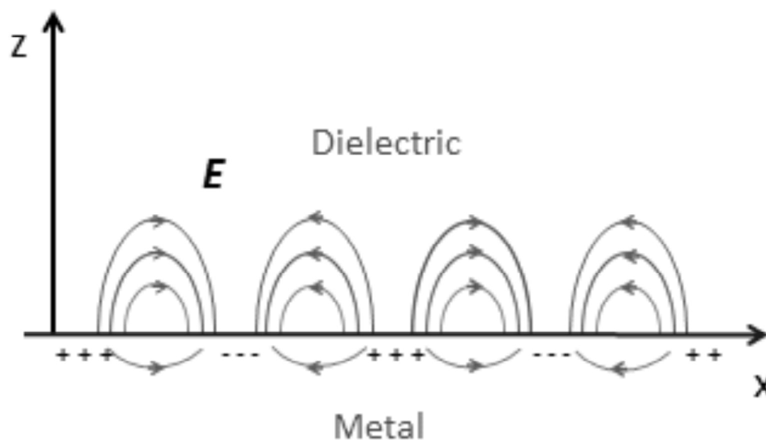


**Figure 14: Electric field lines and the distribution of charges due to surface plasmon propagating along a metal-dielectric interface.**

Paper [72] was the first to describe the phenomenon of surface plasmon amplification by stimulated emission of radiation (SPASER). An application of this phenomenon is spaser, the nanoplasmonic analogue of a laser. In spaser, a nanoparticle plays the role of resonant cavity, and surface plasmons play the role of photons. The first experimental work in which spaser was implemented, was [73]. In this paper the implementation of spaser was based on the use of 44-nm-diameter nanoparticles with a gold core and dye-doped silica shell. For stimulated emission, pumping by an external source of electromagnetic radiation was used. However, a single spaser emits very little light, and it does not have collimation into a narrow beam. Paper [74] introduced a further development of the concept of spaser based on combining the metamaterial and spaser radiation. Recently in [75], a plasmonic laser was demonstrated, which consists of a periodic array of metal nanoparticles (Au or Ag nanoparticles) surrounded with a nanolayer of organic dye. This organic dye is used as the gain medium. This plasmon laser has some features

in common with the lasing spaser, which was proposed in [74], since it is based on a periodic array (two-dimensional metamaterial). Simultaneously and independently of [75], paper [76] also demonstrated another lasing plasmonic array. It represented a metal hole array with a semiconductor (InGaAs) gain layer placed in close (20 nm) proximity of the metal hole array. Lasing spasers can be widely used in various fields, including ultrafast information processing and displays.

Generation of strong electric fields at the surface of metallic nanoparticle or at nanofeatures of metallic microparticles due of surface plasmon resonance can cause a huge increase in the intensity of Raman scattering of molecules in the strong electric field (surface-enhanced Raman scattering (SERS)). This allows one, in particular, to observe Raman scattering from a single molecule attached to a metal colloidal particle [77, 78].

Instead of metallic nanoparticle one can also use the SERS-active tip of a scanning probe microscope. This tip can be made by its coating with a SERS-active metal. In this case, the appropriate research method based on Raman spectroscopy is called tip-enhanced Raman scattering (TERS). The combination of the method of Raman scattering with scanning probe microscopy (e.g., atomic force microscopy or scanning tunneling microscopy) allows for mapping of molecules with very high spatial resolution. Recently in paper [79] the authors reported the visualization of individual molecules using the TERS (using scanning tunneling microscopy) and achieving a record sub-nm spatial resolution.

In nanooptics a number of important areas can be identified: studies of overcoming the diffraction limit for light (localization of light; apertures, fibers; optical near-field microscopy); studies of applications of highly-focused light (confocal microscopy; multiphoton microscopy); studies of the use of resonant phenomena in highly inhomogeneous media (generation of plasmons, surface plasmon-polaritons, the use of optical properties of microresonators formed by dielectric spheres), studies of interaction of light with systems of nanometer dimensions, or structural elements which have nanometer dimensions (quantum wells, quantum wires, quantum dots, carbon nanotubes, graphene, DNA, proteins, metamaterials, photonic crystals).

## 6. CONCLUSIONS

Thus, nanophysics has come a long way since Feynman's lecture (1959). At present, it covers a wide scope of activvities, including a large number of new effects, principles, methods and materials that are used to create new devices. Some of the devices already have some commercial success, such as quantum cascade lasers. Overall, however, we are unable to say at the moment which new devices will fall into mass production. There are many technical problems to be solved. Many methods of manufacturing nanodevices are very costly. However, the general technological trend is miniaturization. Therefore, it is inevitable that a number of nanodevices will have a wide commercialization. Perhaps modern semiconductor microelectronics, the basic material of which is silicon, will be replaced by nanoelectronics, the elements of which will be nano-objects, operating on new principles. Also, possibly in the future there will be nanodevices in mass production that will function like biological systems.

### *References*


[1] A. Nakamura, T. Tokizaki, H. Akiyama, T. Kataoka, *Journal of Luminescence*, v. 53, p. 105, (1992).

[2] S. Baskoutas, A.F. Terzis, *J. Appl. Phys.*, v. 99, p. 013708, (2006).

[3] L. Esaki, R. Tsu, *IBM J. Res. Develop.*, v. 14, p. 61, (1970).

[4] L. Esaki, Journal de Physique, Colloque C5, supplement au n°4, Tome 45, p. C5-3, (1984).

[5] K. Leo, *Semicond. Sci. Technol.*, v. 13, p. 249, (1998).

[6] M. Helm, *Semicond. Sci. Technol.*, v. 10, p. 557, (1995).

[7] A. Wacker, *Phys. Rep.*, v. 357, p. 1, (2002).

[8] L.L. Bonilla, H.T. Grahn, *Rep. Prog. Phys.*, v. 68, p. 577, (2005).

[9] K.S. Novoselov, A.K. Geim, S.V. Morozov, D. Jiang, Y. Zhang, S.V. Dubonos, I.V. Grigorieva, A.A. Firsov, Science, v. 306, p. 666, (2004).

[10] R.R. Nair, P. Blake, A.N. Grigorenko, K.S. Novoselov, T.J. Booth, T. Stauber, N.M.R. Peres, A.K. Geim, Science, v. 320, p. 1308, (2008).

[11] K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, M. I. Katsnelson, I. V. Grigorieva, S. V. Dubonos, and A. A. Firsov, *Nature*, v. 438, p. 197, (2005).

[12] Y. Zhang, Y.-W. Tan, H.L. Stormer, and P. Kim, *Nature*, v. 438, p. 201, (2005).

[13] K. V. Klitzing, G. Dorda, M. Pepper, *Phys. Rev. Lett.*, v. 45, p. 494, (1980).

[14] K. Nakata, M. Fujita, G. Dresselhaus, M.S. Dresselhaus, *Phys. Rev. B*, v. 54, p. 17954, (1996).

[15] X.L. Li, X.R. Wang, L. Zhang, S.W. Lee, H.J. Dai, *Science*, v. 319, p. 1229, (2008).

[16] 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*, v. 323, p. 610, (2009).

[17] E. McCann, V.I. Fal'ko, *Phys. Rev. Lett.*, v. 96, p. 086805, (2006).

[18] E.V. Castro, K.S. Novoselov, S.V. Morozov, N.M.R. Peres, J. Dos Santos, J. Nilsson, F. Guinea, A.K. Geim, A.H. Castro Neto, *Phys. Rev. Lett.*, v.99, p. 216802, (2007).

[19] F. Schedin, A.K. Geim, S.V. Morozov, E.W. Hill, P. Blake, M.I. Katsnelson, K.S. Novoselov, *Nature Materials*, v. 6, p. 652, (2007).

[20] A.C. Ferrari, J.C. Meyer, V. Scardaci, C. Casiraghi, M. Lazzeri, F. Mauri, S. Piscanec, D. Jiang, K.S. Novoselov, S. Roth, A.K. Geim, *Phys. Rev. Lett.*, v. 97, p. 187401, (2006).

[21] A.C. Ferrari, *Solid State Communications*, v. 143, p. 47, (2007).

[22] A.C. Ferrari, J. Robertson, *Phys. Rev. B*, v. 61, p. 14095, (2000).

[23] A. Reina, X. Jia, J. Ho, D. Nezich, H. Son, V. Bulovic, M.S. Dresselhaus, J. Kong, *Nano Lett.*, v. 9, p. 30, (2009).

[24] S. Iijima, *Nature*, v. 354, p. 56, (1991).

[25] R. Saito, M. Fujita, G. Dresselhaus, M.S. Dresselhaus, *Appl. Phys. Lett.*, v. 60, p. 2204, (1992).

[26] J. W. G. Wildöer, L.C. Venema, A.G. Rinzler, R.E. Smalley, C. Dekker, *Nature*, v. 391, p. 59, (1998).

[27] T.W. Odom, J.-L. Huang, P. Kim, C.M. Lieber, *Nature*, v. 391, p. 62, (1998).

[28] K. Ichimura, M. Osawa, K. Nomura, H. Kataura, Y. Maniwa, S. Suziki, Y. Achiba, *Physica B*, v. 323, p. 230, (2002).

[29] C.T. White, J.W. Mintmire, Nature, v. 394, p. 29 (1998); J.W. Mintmire, C.T. White, *Phys. Rev. Lett.*, v. 81, p. 2506, (1998).

[30] M.E. Itkis, D.E. Perea, R. Jung, S. Niyogi, R.C. Haddon, *J. Am. Chem. Soc.*, v. 127, p. 3439, (2005).

[31] K. Iakoubovskii, N. Minami, S. Kazaoui, T. Ueno, Y. Miyata, K. Yanagi, H. Kataura, S. Ohshima, T. Saito, *J. Phys. Chem. B*, v. 110, p. 17420, (2006).

[32] Y. Miyata, K. Yanagi, Y. Maniwa, H. Kataura, *J. Phys. Chem. C*, v. 112, p. 13187, (2008).

[33] M.J. O'Connell, S.M. Bachilo, C.B. Huffman, V.C. Moore, M.S. Strano, E.H. Haroz, K.L. Rialon, P.J. Boul, W.H. Noon, C. Kittrell, J. Ma, R.H. Hauge, R.B. Weisman, R.E. Smalley, *Science*, v. 297, p. 593, (2002).

[34] S.M. Bachilo, M.S. Strano, C. Kittrel, R.H. Hauge, R.E. Smalley, R.B. Weisman, *Science*, v. 298, p. 2361, (2002).

[35] R.B. Weisman, S.M. Bachilo, *Nano Lett.*, v. 3, p. 1235, (2003).

[36] T. Okazaki, T. Saito, K. Matsuura, S. Ohshima, M. Yumura, S. Iijima, *Nano Lett.*, v. 5, p. 2618, (2005).

[37] D.A. Tsyboulski, J.-D.R. Rocha, S.M. Bachilo, L. Cognet, R.B. Weisman, *Nano Lett.*, v. 7, p. 3080, (2007).

[38] M.S. Dresselhaus, G. Dresselhaus, R. Saito, A. Jorio, *Phys. Rep.*, v. 409, p. 47, (2005).

[39] J.C. Meyer, M. Paillet, T. Michel, A. Moréac, A. Neumann, G.S. Duesberg, S. Roth, J.-L. Sauvajol, *Phys. Rev. Lett.*, v. 95, p. 217401, (2005).

[40] M. Paillet, T. Michel, J.C. Meyer, V.N. Popov, L. Henrard, S. Roth, J.-L. Sauvajol, *Phys. Rev. Lett.*, v. 96, p. 257401, (2006).

[41] A. Débarre, M. Kobylko, A.M. Bonnot, A. Richard, V.N. Popov, L. Henrard, M. Kociak, *Phys. Rev. Lett.*, v. 101, p. 197403, (2008).

[42] T. Michel, M. Paillet, J.C. Meyer, V.N. Popov, L. Henrard, J.-L. Sauvajol, *Phys. Rev. B*, v. 75, p. 155432, (2007).

[43] J.C. Meyer, M. Paillet, G.S. Duesberg, S. Roth, *Ultramicroscopy*, v. 106, p. 176, (2006).

[44] S. Frank, P. Poncharal, Z.L. Wang, W.A. de Heer, *Science*, v. 280, p. 1744, (1998).

[45] A.G. Rinzler, J.H. Hafner, P. Nikolaev, P. Nordlander, D.T. Colbert, R.E. Smalley, L. Lou, S.G. Kim, D. Tománek, *Science*, v. 269, p. 1550, (1995).

[46] S. J. Tans, A.R.M. Verschueren, C. Dekker, *Nature*, v. 393, p. 49, (1998).

[47] S.J. Wind, J. Appenzeller, R. Martel, V. Derycke, Ph. Avouris, *Appl. Phys. Lett.*, v 80, p. 3817, (2002).

[48] A. Javey, J. Guo, O. Wang, M. Lundstrom, H. Dai, *Nature*, v. 424, p. 654, (2003).

[49] A.C. Dillon, K.M. Jones, T.A. Bekkedahl, C.H. Kiang, D.S. Bethune, M.J. Heber, *Nature*, v. 386, p. 377, (1997).

[50] C. Liu, Y.Y. Fan, M. Liu, H.T. Cong, H.M. Cheng, M.S. Dresselhaus, *Science*, v. 286, p. 1127, (1999).

[51] R. Landauer, IBM *J. Res. Dev.*, v. 1, p. 223 (1957); R. Landauer, *Phys. Lett.*, v. 85A, p. 91 (1981); R. Landauer, *Phil. Mag*., v. 21, p. 863, (1970).

[52] M. Büttiker, Y. Imry, R. Landauer, and S. Pinhas, *Phys. Rev. B*, v. 31, p. 6207, (1985).

[53] B.J. van Wees, H. van Houten, C.W.J. Beenakker, J.G. Williamson, L.P. Kouwenhoven, D. van der Marel, and C.T. Foxon, *Phys. Rev. Lett.*, v. 60, p. 848, (1988).

[54] D.A. Wharam, M. Pepper, H. Ahmed, J.E.F. Frost, D.G. Hasko, D.C. Peacock, D.A. Ritchie, G.A.C. Jones, *J. Phys. C*, v. 21, p. L209, (1988).

[55] Zh.I. Alferov, R.F. Kazarinov, “Semiconductor laser with electric pumping”, Inventor’s Certificate No. 181737 (in Russian), Application No. 950840, priority as of March 30, 1963.

[56] H. Kroemer, *Proc. IEEE*, v. 51, p. 1782, (1963).

[57] Zh.I. Alferov, V.M. Andreev, V.I. Korol’kov, E.L. Portnoi, D.N. Tret’yakov, *Sov. Phys. Semicond*., v. 2, p. 843, (1969).

[58] Zh.I. Alferov, V.M. Andreev, V.I. Korol’kov, E.L. Portnoi, D.N. Tret’yakov, *Sov. Phys. Semicond*., v. 2, p. 1289, (1969).

[59] Zh.I. Alferov, V.M. Andreev, E.L. Portnoy, M.K. Trukan, *Sov. Phys. Semicond*., v. 3, p. 1107, (1970).

[60] I. Hayashi, M.B. Panish, P.W. Foy, S. Sumski, *Appl. Phys. Lett.*, v. 17, p. 109, (1970).

[61] J.P. van der Ziel, R. Dingle, R.C. Miller, W. Wiegmann, W.A. Nordland, Jr., *Appl. Phys. Lett.*, v. 26, p. 463, (1975).

[62] R.D. Dupuis, P.D. Dapkus, N. Holonyak, Jr., E.A. Rezek, R. Chin, *Appl. Phys. Lett.*, v. 32, p. 295, (1978).

[63] W.T. Tsang, *Appl. Phys. Lett.*, v. 40, p. 217, (1982).

[64] R.F. Kazarinov and R.A. Suris, *Sov. Phys. Semicond*. v. 5, p. 707, (1971).

[65] J. Faist, F. Capasso, D.L. Sivco, C. Sirtori, A.L. Hutchinson, and A.Y. Cho, *Science*, v. 264, p. 553, (1994).

[66] M. Born, E. Wolf, Principles of Optics. Electromagnetic Theory of Propagation, Interference and Diffraction of Light. Fourth Edition, 1970, Pergamon press, Oxford.

[67] E.H. Synge, *Philosophical Mag*., v.6, p. 356, (1928).

[68] E.A. Ash, G. Nicholls, *Nature*, v. 237, p. 510, (1972).

[69] D.W. Pohl, W. Denk, M. Lanz, *Appl. Phys. Lett.*, v. 44, p. 651, (1984).

[70] A. Lewis, M. Isaacson, A., Harootunian, A. Muray, *Ultramicroscopy*, v. 13, p. 227, (1984).

[71] R.H. Ritchie, *Phys. Rev.*, v.106, p. 874, (1957).

[72] D.J. Bergman, M.I. Stockman, *Phys. Rev. Lett.*, v. 90, p. 027402, (2003).

[73] M.A. Noginov, G. Zhu, A.M. Belgrave, R. Bakker, V.M. Shalaev, E.E. Narimanov, S. Stout, E. Herz, T. Suteewong, U. Wiesner, *Nature*, v. 460, p. 1112, (2009).

[74] N. I. Zheludev, S. L. Prosvirnin, N. Papasimakis, V.A. Fedotov, *Nature Photonics*, v. 2, p. 351, (2008).

[75] W. Zhou, M. Dridi, J.Y. Suh, C.H. Kim, D.T. Co, M.R. Wasielewski, G.C. Schatz, T.W. Odom, *Nature Nanotechnology*, v. 8, p. 506, (2013).

[76] F. v. Beijnum, P.J.v. Veldhoven, E.J. Geluk, M.J.A.d. Dood, G.W.’t Hooft, M.P.v. Exter, *Phys. Rev. Lett.*, v. 110, p. 206802, (2013).

[77] K. Kneipp, Y. Wang, H. Kneipp, L.T. Perelman, I. Itzkan, R.R. Dasari, and M.S. Feld, *Phys. Rev. Lett.*, v. 78, p. 1667, (1997).

[78] S. Nie, S.R. Emory, *Science*, v. 275, p. 1102, (1997).

[79] R. Zhang, Y. Zhang, Z.C. Dong, S. Jiang, C. Zhang, L.G. Chen, L. Zhang, Y. Liao, J. Aizpurua, Y. Luo, J.L. Yang, J.G. Hou, *Nature*, v. 498, p. 82, (2013).