

08
Features of vacuum resonant tunneling at one-and two-well barrier potentials

© M.V. Davidovich

Saratov National Research State University,
 410012 Saratov, Russia
 e-mail: davidovichmv@info.sgu.ru

Received September 16, 2021
 Revised April 27, 2022
 Accepted June 1, 2022

The features of one-dimensional vacuum tunneling and calculation of the tunneling current in barrier quantum structures with one and two wells are considered. The structures are formed by several electrodes: the cathode, the anode and two grids. The potential profiles are constructed by the method of multiple images. Equations for eigenvalues and metastable levels of a structure with an arbitrary potential profile are found. For full resonant tunneling, the metastable levels must fall into the region of the electron kinetic energy distribution at the cathode, which occurs in single-well structures in the absence of an anode voltage or at a low anode voltage compared to the voltage on the grids. A significant voltage at the anode leads to an asymmetric structure and incomplete resonant tunneling. In this case, a two-well structure with a double grid allows obtaining full resonant tunneling for a number of energies and increasing the tunneling current by orders of magnitude.

Keywords: resonant tunneling, field emission, vacuum nanotriode, potential barrier, quantum well.

DOI: 10.21883/TP.2022.09.54684.257-21

Introduction

One-well barrier potentials (Fig. 1) arise in a number of one-dimensional problems of quantum tunneling, diffraction of plane electromagnetic waves on layered structures, in models of quantum resonant tunneling diodes (RTDs), resonant tunneling transistors (RTTs), and quantum cascade lasers [1–6], as well as in problems relating vacuum nanotriodes [7–9]. Vacuum tunneling nanotriodes are very interesting objects of study for vacuum microelectronics [7,8], since they do not require heating for operation, and the emission structure can be very compact and at the same time create rather large currents necessary for the devices operation. With a short flight length up to hundreds of nm, less than the mean free path of low-energy electrons in air (on the order of 1 μm) and voltages of the order of 10–20 V, they can operate as active THz-devices and without vacuum, since ionization does not occur.

In solid-state heterostructures the implementation of multiwell and multibarrier profiles is based on the methods of doping and the creation of quantum superlattices by various methods of depositing nanosized layers with different properties, for example, of the GaAs/AlAs type. The analysis of quantum superlattices requires, generally speaking, a quantum multiple-particles approach, as well as a rigorous approach to tunneling problems [10–14]. Here, for example, one can use the methods of Green–Keldysh nonequilibrium functions together with the methods of density functional theory. Often, in the analysis of RTDs and similar structures, based on physical considerations, a rectangular one-well two-barrier or two-well three-barrier quantum potential profile is set, for which, in the presence

of varying electric potential, the non-stationary Schrodinger equation (SE) is solved and the current is determined [2–6]. To simplify, the model potential of the barriers is even given by delta functions [2]. When modeling solid-state PTTs the model rectangular potentials are often used, the vertices of which, when an anode voltage is applied, acquire bevels [1]. In more rigorous approaches, a joint self-consistent solution of a non-stationary single-particle SE with the Poisson equation (PE) [1,6] is described. In the case of vacuum gaps the density functional theory methods are suitable for distances up to fractions of nm [12–14]. In vacuum flat nanostructures with electrode sizes ranging from units to tens of nm, a rather simple theory of classical superlattices can be used. It is important that the total length of the structure be less than the free path length of electrons. Then the quantum potential can be constructed on the basis of the multiple image theory method. This approach was initially used in the analysis of vacuum field emission sources, starting with the Fowler–Nordheim theory.

Vacuum field emission sources without grids require very large fields (much more than 10¹⁰ V/m) to obtain a high current density. In such fields heating of the cathode, back bombardment, explosive emission, strong ponderomotive effect on the surface take place [15–17]. In this case, the barrier does not completely disappear, its transparency coefficient D^+ is still significantly less than unity, and the current integrated density J is significantly less than the limit values of about 10¹⁵ A/m² (for vacuum tunneling from metals). This limit value has the form [15–17] $J = em_e E_F^2 / (4\pi^2 \hbar^2)$ and is unattainable, since it is obtained from the condition $D^+(E_k) = 1$, i.e. total transparency at all kinetic energies $0 \leq E_k \leq E_F$ (here E_F is the Fermi energy).

Considering the anode to be sufficiently remote, one can obtain the barrier potential at the cathode, related to the electric field voltage E_x on its surface, in the form [15,16]:

$$V(x) = -e^2/[16\pi\epsilon_0(x + \delta)] - eEx. \quad (1)$$

Here $E = -E_x$. Further, as the barrier in a broad sense we assume any distribution of the quantum potential $V(x)$ entering SE, which may include quantum wells separated by barriers (humps). In a narrow sense, the barrier is such hump. This function will also be called the potential profile or potential diagram. On the cathode the potential depth should be equal to the work function W_c , i.e. $V(0) = -e^2/[16\pi\epsilon_0\delta] = -W_c$. From here we have a link between the introduced parameter δ and the work function of the cathode. The value $\delta = 0.1$ nm corresponds to the work function $W_c = 3.6$ eV. The maximum of V function is negative and depends on the field strength. The height of such maxima from the Fermi level of the electrode will be denoted as W . This is the work function that must be done from a given electrode in the presence of other electrodes. So, considering the anode, the above formula (1) can be modified

$$V(x) \approx -e^2/[16\pi\epsilon_0(x + \delta)] - eU_a x/d \\ - e^2/[16\pi\epsilon_0(d - x + \delta)].$$

It is approximate (the exact formula will be given below), and its accuracy is maximum either at the cathode or at the anode, and is greater, the greater d (distance cathode–anode) is. For small x and large d the last term can be neglected, and we obtain (1), where $E_x = -U_a/d$. Near the anode, the first term can be neglected. In the center, the formula loses accuracy, since it takes into account only two images, while (1) — one. Since the formula (1) for the potential was obtained by the method of images, the force of which does not act at interatomic distances due to the discrete atomic nature of the cathode, the parameter δ characterizes the size, after which the indicated image force appears, which attracts electron to the cathode. If the field is absent, the potential increases and is equal to zero at infinity. This is the free state of an electron removed from the cathode. The presence of a pulling electron field creates a barrier of a finite length. One can find the negative maximum point of the barrier, which decreases with the field increasing. This narrows the barrier. The barrier will disappear when it becomes infinitely narrow, and the maximum point is on the cathode (in this case, the turning points coincide with the cathode surface $x = 0$). It is obvious, that this will be at the field $E_c = e/(16\pi\epsilon_0\delta^2) = W/(\delta e) = 16\pi\epsilon_0 W^2/e^3$. At $\delta = 0.1$ nm this is $E_c = 3.6 \cdot 10^{10}$ V/m. At such a critical field the barrier turns into a bevel into a well (considering the finite location of the anode, the depth of the well is finite). For such structure one can calculate the transparency at the Fermi level. Due to quantum wave effects it is significantly less than unity. However, the

density of electrons riding on the barrier with the Fermi energy is equal to zero [15]. Non-zero density takes place for lower energies, but for them the transparency becomes exponentially small. Indeed, the disappearance of the barrier at V_F and a critical field allows us to approximately consider it triangular for energies below the Fermi energy with turning points $x_1 = 0$, $x_2(E) = (eE_c)^{-1}(E_F - E)$. Here, the kinetic energy is measured from the bottom of the conduction zone. The transparency of such barrier in the quasiclassical approximation is calculated exactly and is given by the formula [18]:

$$D^+(E) = D_0(E) \exp\left(-\frac{4\sqrt{\mu}(E_F - E)^{3/2}}{3\hbar e E_c}\right).$$

Here and below $\mu = 2m_e$ means double electron mass. Further increasing of the field just makes the slope steeper. Although the current density increases in this case, this is a dead end for its essential increasing. It leads to instabilities and to explosive emission destroying the cathode. Therefore, the approach is required to increase the current density at field strengths below the critical one, when the barrier is small but still exists.

The purpose of this paper is the theoretical study of the possibility of obtaining field emission densities significantly higher than 10^{10} A/m² in three- and four-electrode structures with resonant tunneling (RT). RT allows one to obtain the values $D^+(E_k) = 1$ at certain kinetic energies E_k of electrons riding on the barrier [1,6,19] and thereby approaching the maximum achievable current density at average values of the normal component of the electric field $E_x < 10^{10}$ V/m. The analysis of tunneling through one-well barrier profiles of the quantum potential $V(x)$ (Fig. 1) is important for the creation of such high-current field emission sources with the pull grid and the anode. However, two-well and multi-well barrier structures are of interest for obtaining more high-current field emission sources. Such vacuum structures require the use of four or more electrodes with dimensions on the order of nanometers. When several grid electrodes come into contact with the cathode, a heterostructure arises, which should be considered as a quantum superlattice, since a single emitter is formed. Semiconductor quantum heterostructures to determine the quantum potential distribution require quantum consideration [10–14], for example, based on the density functional theory method (DFTM) [11]. We consider the vacuum heterostructures with nanosized vacuum gaps and highly conductive electrodes, for which the quantum potential can be obtained by the classical electrostatic approach based on images method. This, taking into account the sufficient length of the structures, is justified, among other things, by the fact that sufficiently large potentials of the order of tens of volts are applied to the electrodes, which significantly exceeds the contribution from possible quantum effects. This approach, which is incomparably simpler than the DFTM, makes it possible to find a very accurate potential distribution. The method of images, in particular, makes

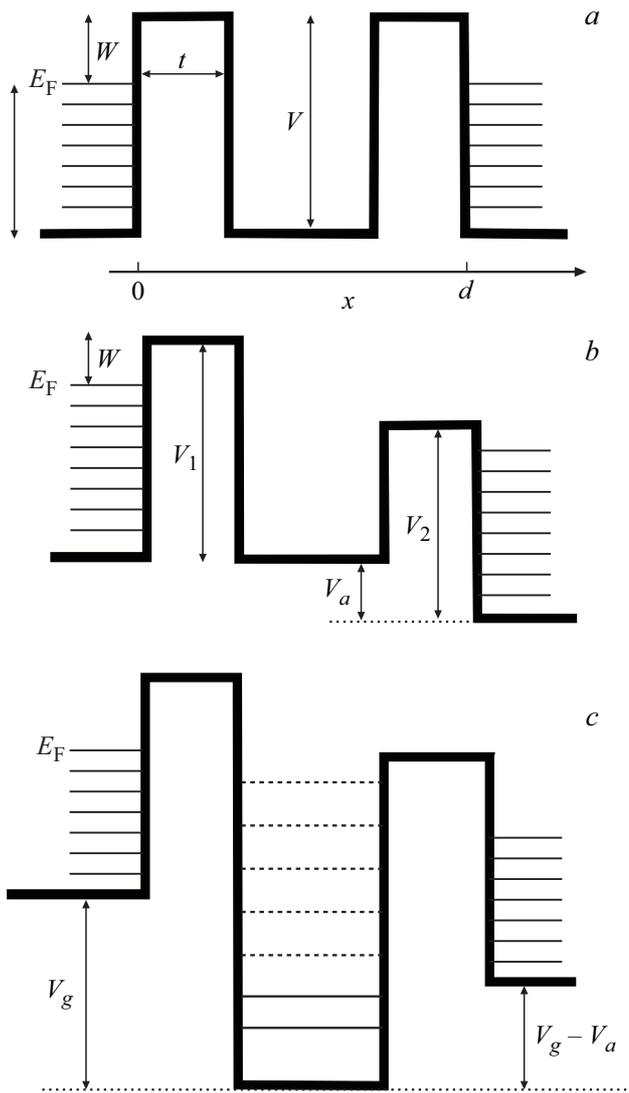


Figure 1. Model view of one-well barrier potential in vacuum nanotriode. Explanations on the conditions under which the potential has the form *a*, *b* or *c*, and also on the designations — in the text.

it possible to find simple barrier profiles at one of the electrodes [15,16]. Next, for two electrodes, we use the method of multiple images. The potential of multielectrode structures is constructed sequentially, taking into account its continuity. All dimensions below are given in nm, kinetic and potential energies in eV, and electrode potentials in V.

1. Simple models for rectangular barriers

In the one-dimensional case in field emission electronics, the stationary SE $(-\hbar^2\mu\partial_x^2 + V(x) - E_k)\psi(x) = 0$ is usually solved, in which $\mu = 2m_e$, m_e is electron mass, $\hat{p}_x = \hbar\partial_x = k\hbar = \sqrt{\mu(E_k - V(x))}$ is its momentum, and the wave function is represented by overlaying of plane waves $\psi(x) = a \exp(\pm ikx)$.

The heterogeneous potential leads to the presence of waves of both directions at each point, while in the region of barriers $E_k < V(x)$ the momentum becomes imaginary. In a potential field, one can introduce a force $F(x) = -\partial_x V(x) = -eE_x(x) = e\partial_x U(x)$, where $U(x) = -V(x)/e$ is electrostatic potential, $e = 1.6022 \cdot 10^{-19}$ C is electron charge. In vacuum tunneling the potential barrier and electrostatic potential are determined by the method of multiple images relative to the electrode surfaces [20–22]. In semiconductor tunnel structures the barrier and well profiles are created by doping and are usually modeled as rectangular regions. Applying anode voltage distorts these regions and causes slopes appearance at the tops of the barriers and at the bottom of the well [1–6]. Accounting for the potential of the electron beam as the solution of the Poisson equation also leads to $V(x)$ change. Such a solution is expedient at a high beam density. In vacuum structures the grid potential is constant, the grid screens the field, so the well bottom is flat. An example of accurate calculation of barrier profiles by the method of images is given in Fig. 2. For them, a rigorous solution of SE is obtained below. Further, the method of matrices transfer is used for this. With one-dimensional tunneling, the method of series can also be used. As can be seen, the barrier profiles are close to trapezoidal with a trapezium in the form of a triangle on a rectangular base. The voltage increasing on the electrodes brings the barrier profiles closer and closer to triangular ones. For the analytical study it is convenient

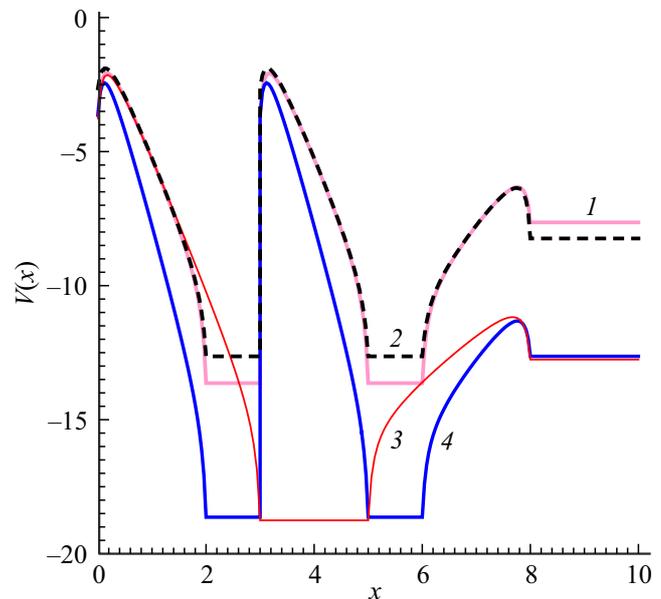


Figure 2. Profiles of complex barriers (eV) in two-well (curves 1, 2, 4) and one-well (curve 3) structures with two grids at the same potential. Structure parameters (eV): $V_g = 10$, $V_a = 5$ (curves 1, 2); $V_g = 15$, $V_a = 10$ (curves 3, 4); $W_k = 3$, $W_g = 4$ (curves 1, 3, 4); $W_k = 3.6$, $W_g = 3$ (curve 2); $t_1 = t_2 = t_3 = 2$ nm, $t_g = 1$ nm (curves 1, 2, 4); $t_1 = t_2 = 3$ nm, $t_g = 2$ nm (curve 3); $W_a = 4.5$.

to replace complex barriers with rectangular ones, for example, by adjusting their height so that the area is preserved. During tunneling of one particle it is convenient to normalize the wave function to the delta function [19,23]. In the problems of emission electronics there is usually an incident particles flow described by the wave function $\psi^+(x) = a(x) \exp(ik_0x)$, so the current density $J^+(x)$ can be introduced in terms of the density of probability flow [1–6,18]. Outside the barrier structure $V(x) = 0$, $k_0 = \sqrt{\mu E_k}/\hbar$, $a = a_0 = \text{const}$, and the normalization to the incident flow with unit density of particles is convenient $\rho^+(x) = |\psi^+(x)|^2 = a_0^2 = 1$, $x < 0$ [1–6,15–17,21–23]. Since there is a reflected wave from the structure, the total density to the left of the barrier is

$$\rho(x) = |\exp(ik_0x) + R \exp(-ik_0x)|^2 = 1 + |R|^2 + 2|R| \cos(2k_0x - \varphi_R).$$

Here phase φ_R and modulus $|R|$ of the reflection coefficient are introduced. During RT the reflection coefficient is zero. The presence of high electrons density in the region of the structure (outside the metal electrodes) requires the solution of the Poisson equation $\partial_x^2 \Phi(x) = e\rho(x)/\varepsilon_0$ there. For such solution, zero boundary conditions on the cathode and anode should be used, and the result of the solution should be applied to the electric potential: $\tilde{U}(x) = U(x) + \Phi(x)$, which changes the profile of the barrier structure.

Let us consider tunneling through the symmetrical barrier Fig. 1, *a*. Due to the symmetry the tunneling from the cathode and anode are the same, so the total current is zero. The current from the cathode arises at positive voltage $U_a = -V_a/e$ on the anode, when the barrier structure becomes asymmetric (Fig. 1, *b, c*). The wave function of the electrons flow riding on the barrier will be designated as $\psi_0^+(x) = a_0 \exp(ik_0x)$, and the flow passing to the anode as $\psi_a^+(x) = a_0 T \exp(ik_a(x-d))$, where $k_a = \sqrt{\mu(E_k + V_a)}/\hbar$. Tunneling through the barrier is a quantum effect, and the electrons that have passed through the barriers do not lose energy, but before getting on the anode they move classically and are accelerated. They transfer the excess energy eU_a to the phonons of the anode crystal over a short free path length, acquiring the Fermi energy distribution on the anode. For electrons falling on the anode surface from its depth the transparency of the barrier at a significant voltage U_a is exponentially less than $D^-(E_k) \ll D^+(E_k)$, and the reverse current can be neglected even at anode voltages of the order of several V. Here, the kinetic energy of electrons both on the cathode and on the anode is measured from the bottom of their conduction zone $V_0 = -W - E_F$ and $V_{0a} = -W - E_F - eU_a$, respectively. Let us introduce the impedances $z_0 = 1/k_0$, $z_n = 1/k_n$, $z_g = 1/k_g$, $z_a = 1/k_a$ (note that in the paper [19] the reverses of impedances are taken) and designate $\theta_n = \tan(k_n t_n)$, $\theta_g = \tan(k_g t_g)$. Here t_n are barrier widths, which can generally be different, and t_g is well width corresponding to the grid area. For the symmetrical case $t_{1,2} = t$, from the condition of the absence

of the reflection coefficient $R = 0$ or the equality of the input normalized impedance to unity $\theta_g \theta = 2Z/(1 + Z^2)$ follows, or with the notation $Z = z/z_0 = k_0/k = -i\eta$ — transcendental equation

$$\tan(t_g \hbar^{-1} \sqrt{\mu E_n}) \tanh(t \hbar^{-1} \sqrt{\mu(V - E_n)}) = \frac{2\sqrt{V/E_n - 1}}{2 - V/E_n}. \quad (2)$$

The presence of roots $E_n > 0$ of equation (2) means RT through the symmetrical barrier when the kinetic energy coincides with the level: $E_k = E_n$. Such levels are possible if $V/2 < E_n < V$. For a wide barrier with the upper level $E_n \approx V$, we replace the hyperbolic tangent by unity, and the tangent by its argument with a small right-hand side, i.e., we have

$$E_n \approx 2V \sqrt{\frac{\hbar^4}{t_g^4 V^2 \mu^2} + \frac{\hbar^2}{t_g^2 V \mu} - \frac{2\hbar^2}{t_g^2 \mu}}.$$

The levels for the structure (Fig. 1, *a*) are metastable because the wave function „leaks“ through the barriers. It is easy to see that such complex quasilevels are determined from the complex equations

$$\begin{aligned} k_0 t_g &= \pm 2 \arctan\left(\frac{Z + iZ^2\theta}{\theta - iZ}\right)^{\pm 1} + 2n\pi \\ &= \pm 2 \arctan\left(\frac{-ik_0/\kappa + (k_0/\kappa)^2 \tanh(t\kappa)}{i \tanh(t\kappa) - k_0/\kappa}\right)^{\pm 1} + 2n\pi \end{aligned} \quad (3)$$

and are metastable. This means a finite lifetime of the level: the electron is emitted from it and goes to infinity, and the level width is determined by $-\text{Im}(E_n)$. The signs in (3) correspond to different symmetry conditions for the wave function [24]. For wide barriers $\tanh(t\kappa) \approx 1$, and neglecting the small terms in (3), we get $k_0 t_g = \mp 2 \arctan(k_0/\kappa)^{\pm 1} + 2n\pi$, which matches the results from [24]. For an infinitely deep well (infinitely high and wide barriers), with the upper sign we have $E_n = (2n\pi/t_g)^2 \mu^{-1} \hbar^2$, and with the lower sign $E_n = ((2n+1)\pi/t_g)^2 \mu^{-1} \hbar^2$, $n = 1, 2, \dots$, i.e., these signs correspond to even and odd wave functions relative to the center of the well. The characteristic equation for the asymmetric well (Fig. 1, *c*) is obtained similarly by transforming the impedance on the anode $z_a = 1/k_a = \hbar/\sqrt{\mu(E + V_a)}$ to the cathode and equating the transformed impedance to the cathode impedance $z_0 = 1/k_0$. We do not present the corresponding formulas because of their cumbersome form. In this case, in addition to metastable levels, at $V_g > V_a$ the own stable energy levels can appear (in Fig. 1, *c*, shown by solid lines).

Note that SE is, in fact, the Helmholtz equation, and quantum tunneling is equivalent to electromagnetic scattering of the plane wave on a dielectric structure with permittivity $\varepsilon(x) = \sqrt{1 - V(x)/E_k}$ [23]. The case $V = 0$ corresponds to the motion of a wave with frequency $\omega = E_k/\hbar$ in vacuum. The case $E_k < V(x)$ (scattering on well or above barrier) corresponds to motion through a collisionless plasma with $\varepsilon(x) < 0$ at $\omega < \omega_p$, and the

case $E_k > V(x)$ — motion through the dielectric layer $\epsilon(x) > 0$. Also note that there is a complete analogy between the problems of scattering and the determination of quantum levels for barrier structures with plane waves diffraction and with the existence of plasmons in plane-layered optical structures. In particular, Fig. 1, *a* corresponds to a structure in the form of a dielectric plate with permittivity $\epsilon > 1$ and plates of two plasma layers with permittivity $\epsilon < 0$. In the case of imposing the condition $R = 0$, we obtain the characteristic equation for the complex resonant frequencies ω_n of such an open resonator. The thickness of the plasma layers increasing leads to the metastable levels transformation into stable intrinsic levels. In this case, the RT regions at wide plasma layers narrow down to spectral lines. Fig. 1, *c* corresponds to a vacuum-type structure, a plasma layer with negative permittivity, a dielectric layer with permittivity $\epsilon_1 > 1$, a plasma layer with permittivity $\epsilon < 1$, a dielectric substrate (half-space) with permittivity $\epsilon_2 < \epsilon_1$. In the case of photons tunneling through plasma layers, the ratio between their frequency and the plasma frequencies of the layers ω_{pm} is important. For the structure in Fig. 1, *c* it is convenient to count the energy from the bottom of the well. Then the problem for the dielectric resonator corresponds to resonances in the dielectric plate surrounded by plasma layers between two plasma half-spaces. The case $\omega_n < \omega_{pm}$ corresponds to eigenfrequencies, i.e. they are below all plasma frequencies. For higher frequencies $\omega_n > \omega_{pm}$ the resonator becomes open, and the levels — quasi-intrinsic (metastable). During diffraction on the structure the resonant passage of photons with energy coinciding with $\hbar \text{Re}(\omega_n)$ is observed. During diffraction of plane electromagnetic waves the cases of incidence at an angle at which the polarization of the waves is important (*p* or *s*) are possible. Then the waves can be classified as *E* (*TM*) or *H* (*TE*). The only difference is in the normalized wave impedances: $z_n^e = \sqrt{\epsilon_n - k_\perp^2/k_0^2}/\epsilon_n$ and $z_n^h = 1/\sqrt{\epsilon_n - k_\perp^2/k_0^2}$, which coincide at normal incidence ($k_\perp = 0$). The condition $R = 0$ for the case of incidence at an angle is the condition for the existence of surface waves (plasmons) in a layered structure [25]. In model problems of plasmonics the dissipation was not taken into account. Its accounting will lead to the fact that all resonant frequencies become complex during tunneling and scattering of photons.

Although during quantum tunneling, electrons run from the depth of the cathode on the boundary at various angles, the density is usually determined for electrons, which have the value of the normal speed component $v = v_x$ within the given limits $dn(v_x) = (4\pi^2\hbar)^{-3} m_e^3 v_x (v_F^2 - v_x^2) dv_x$ [15]. During RT the condition $R = 0$ also leads to the appearance of electron waves along the surfaces, if we assume that the structure is transversely unrestricted. However, these waves move at all angles. During conventional tunneling the transverse speed components are small, since tunneling occurs mainly for large normal speed components. Therefore,

Table 1. Parameters of some metals and value δ

Metal	Work function, eV	Potential of ionization, eV	Atomic radius, covalent radius, nm	Radius of ion, nm	δ , nm
Ba	2.52	5.21	0.222, 0.198	0.134 (+2e)	0.143
Be	3.92	9.32	0.112, 0.090	0.034 (+2e)	0.091
Cs	1.81	3.89	0.267, 0.235	0.167 (+1e)	0.199
Ka	2.25	4.34	0.235, 0.203	0.133 (+1e)	0.160
Na	2.28	5.14	0.190, 0.154	0.097 (+1e)	0.158
Li	2.49	5.39	0.145, 0.134	0.076 (+1e)	0.145
Cu	4.53	7.72	0.128, 0.117	0.077 (+1)	0.079
Mo	4.2	7.10	139, 130	0.070 (+4)	0.085
W	4.54	7.98	137, 170	0.070 (+4)	0.079

during RT one should expect the increasing of electrons scattering by transverse speeds in the output beam.

2. Modeling of the quantum potential and tunneling in complex multiwell barrier structures

The model problems considered above make it possible to qualitatively estimate tunneling. Actual energy diagrams $V(x)$ of barrier structures are obtained using the method of multiple images [20–22]. The potential profiles with one and two wells for the structure cathode–grid–anode and cathode–grid–grid–anode are shown in Fig. 2. To calculate them we used the approach described in [22] and taking into account the work function of the electrodes (cathode, anode, and grids). Further indices *c*, *g*, *a* correspond to the cathode, grid and anode. The work function of the cathode is defined as $W_c = e^2/(16\pi\epsilon_0\delta)$. Similarly, the work functions of the grids and the anode are determined through the parameters δ_g and δ_a . These small dimensions (of the order of Å) correspond to the distances at which the image forces cease to act due to the atomic structure of matter. They correlate well with the radii of the crystal lattice ions in the case when one atom donates one electron to the conduction zone. In the general case, they vary from the indicated radius to the covalent radius. Table 1 shows the corresponding data. The values of the work functions of the cathode shown in Fig. 2 are somewhat lower than for commonly used metals. They can correspond to the deposition of a thin (on the order of nm) oxide or dielectric film on a copper or tungsten surface, for example, in the form of nanodiamond clusters [22]. Magnetron sputtering of diamond-graphite films several nm thick makes it possible to significantly reduce the work function and to obtain low-threshold emission. However, this is unimportant for RT, since the result weakly depends on the cathode material as

a very high potential on the grid is used. The work function taken for the anode may correspond to Cu or W. Grids should be made of multilayer graphene (graphite) [26].

The work of moving electron to point x is defined as $W(x, \delta, \delta_a, d) - W(0, \delta, \delta_a, d)$. Assuming that it is independent of the distances δ and δ_a , which are of the order of the radius of the crystal lattice ion near the cathode and anode boundaries, we obtain the function $W(x, \delta, \delta_a, d)$. It looks like this [22]:

$$W = -\frac{e^2}{16\pi\epsilon_0} \left\{ \frac{1}{x + \delta} \left(1 + \frac{2xd}{(d-x+\delta_a)(d+x)} \right) - \frac{2}{d} + \frac{2x^2}{d^3} \sum_{n=2}^{\infty} \frac{1}{(n^2 - (x/d)^2)n} \right\}, \quad (4)$$

takes into account an infinite series of images relative to the cathode and anode and gives the potential energy of electron at point x . On the cathode

$$W(0, \delta, \delta_a, d) = -e^2(16\pi\epsilon_0)^{-1} \{\delta^{-1} - 2d^{-1}\}.$$

For a large distance d we have

$$W(0, \delta, \delta_a, d) = -e^2(16\pi\epsilon_0\delta)^{-1} = -W_0.$$

On the anode at large distance

$$W(d, \delta, \delta_a, d) = -e^2(16\pi\epsilon_0\delta_a)^{-1} = -W_a.$$

In the case of coincidence of the work functions of the cathode and the anode $W = W_a$ one should put $\delta_a = \delta$. Note that in the paper [20] (formula (31)) there is a function of type (4) (for $\delta = \delta_a = 0$) called „image potential“, but it differs from (4) by a factor of 2. That's really the potential for perfect images. But when obtaining the potential function, one should calculate the work on moving the electron. Displacement of electron by dx leads to a displacement of all its images relative to it by $\pm 2dx$, i.e. the force is defined as half of the derivative of the indicated potential with a minus sign. This proves that (4) is the potential of the electron. It can be seen from Table 1 that for a number of metals the value δ is very close to the ion radius, in some cases it is close to the covalent radius, and in the general case it lies between them, i.e., it does not exceed half of the crystal lattice constant. In fact, in (4) we count the cathode surface from the surface formed by the electron shells of the first layer of ions or close to it (and not from the centers of atoms). The monograph [15] presents the following model. There is a double electric (dipole) layer formed by a cloud of electrons escaping from the metal but not passing through the barrier. It is assumed that the field of this layer is constant, the layer has a thickness not exceeding the lattice constant, the image force begins to act from its outer boundary, and the constant force inside the layer smoothly transforms into the image force. This gives a result by two times exceed the result of (4) (at least for one image). To connect the formula with the work function,

it is necessary to double δ also. The model in [15] has a number of drawbacks. The first one is the constancy of the field in the double layer. In fact, the field is strongly heterogeneous, as is the density of the cloud. Second — the force of images does not act inside the cloud. Third — the size of the cloud is constant. Note that the electrons density in front of the barrier (in its initial part) can be found if the barrier is known. There are a number of papers [27–33], in which the course of the potential curve near the cathode and its effect on emission were calculated, taking into account quantum approaches, exchange interaction, and the double layer. Using the Hartree–Fock method [28], it is shown that the initial section of the barrier is different for different energies. Exchange interactions and the formation of a double electric layer can lead to nonmonotonic behavior of the potential course [29,31–33]. These effects, however, are small in structures with deep potential wells. In the paper [30], based on the Seitz potential, the discontinuity of the potential curve was eliminated, and it was reduced to the form (1). The introduction of δ parameter seems to be the simplest and most effective method for accounting the potential near the electrodes. Far from the electrodes it is important to take into account the space charge.

Let's consider a number of examples of defining δ . The experimentally determined work function of copper lies within 4.53–5.10, which corresponds to $0.0706 \leq \delta \leq 0.0795$. For copper, the atomic radius is $r_a = 0.128$, the covalent radius is 0.117, the ion radius ($+1e$) is $r_i = 0.077$, and the ionization energy ($+e$) 7.72 by about one and a half times exceeds the work function. For copper δ is approximately equal to the ion radius and slightly less than the covalent radius. For silver the experimentally determined work function is 4.52–4.74, which corresponds to area $0.0759 \leq \delta \leq 0.0796$. For silver, atomic radius is 0.144, covalent radius is 0.134, ion radius $+2e$ is equal to 0.089, ion radius $+e$ is equal to 0.126, ionization energy is 7.57. Here the parameter δ is less than the radius of ion $+e$ and is close to the radius of the ion $+2e$. The work function of potassium is 2.28, which corresponds to $\delta = 0.157$. For potassium the atomic radius is 0.235, the covalent radius is 0.203, the ion radius is 0.133, the ionization energy is 4.34. Here δ is slightly larger than the radius of ion. A similar result takes place for other alkali metals. For a large number of metals δ is in the region between the ion radius and the covalent radius, i.e. less than the half-period of the crystal lattice. At such distance the electron is strongly bound to one lattice atom, interacts with its electron shell, and further approach does not lead to an attractive force, while the energy of electron removal from the crystal is close to the work function. Since with such electron removal by x the positively charged electron cloud is shifted by $-x$, the force of the images acts at a distance of $2x$ [15–17], which explains the value W_0 . It is two times less than the ionization energy of atom of radius δ with charge $+e$. Assuming that the parameter δ is approximately equal to the radius of the ion $+e$, we obtain $W_0 \approx V_{ia}r_a/(2r_i)$,

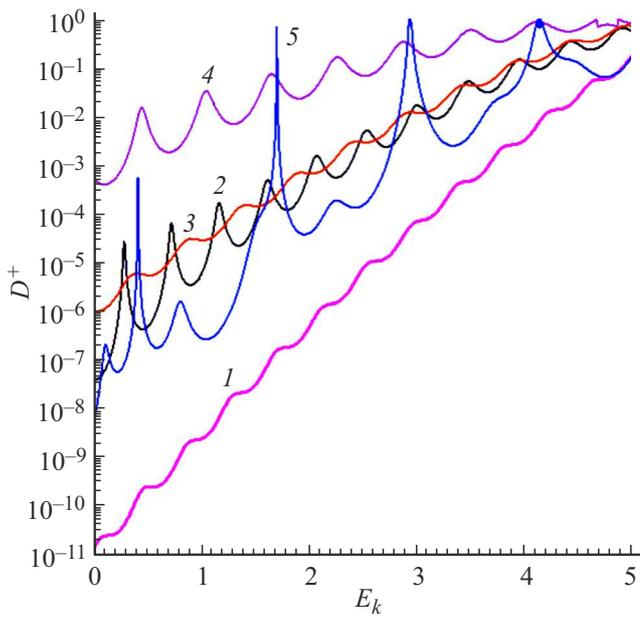


Figure 3. Coefficient of tunneling from the cathode depending on the kinetic energy of electrons (eV) in one-well structures (curves 1–4) and two-well structures (curve 5) of the same length 8 nm. Structure parameters (in eV): $W_k = 3$, $W_g = 4$, $W_a = 4.5$; $V_a = 5$ (curves 1, 4, 5); $V_a = 10$ (curves 2, 3); $V_g = 25$ (curves 1, 2, 5); $V_g = 25$ (curve 3); $V_g = 45$ (curve 4).

where V_i is the ionization potential of the atom. Such formula correlates well with the ionization potentials of metals, and a simple introduction of δ parameter takes into account the discreteness of the medium.

In the given examples (Fig. 2, 3) the coinciding work functions of the cathode W_0 and the anode W_a and differing for grids W_g are used. The influence of the work function of the grids is small, since it can be corrected by the voltage applied to them. In the regions of the grids, the cathode, and the anode the potential is constant, and barriers appear between them. Note that the barrier height depends not only on the work functions and potentials, but also on its length, namely, it decreases as the electrodes approach each other. In the case when the grid potential is greater than the anode potential, a well corresponds to it, while RT is possible (Fig. 3, 4).

Tunneling was calculated by the matrix method using transfer matrices

$$\hat{T}_j = \begin{bmatrix} \cos(k_j \Delta x_j) & -ik_j^{-1} \sin(k_j \Delta x_j) \\ -ik_j \sin(k_j \Delta x_j) & \cos(k_j \Delta x_j) \end{bmatrix},$$

where $k_j = \sqrt{\mu(E_k - V_j - V_0)}/\hbar$ for above barrier regions and $k_j = i\sqrt{\mu(V_j + V_0 - E_k)}/\hbar$ for subbarrier areas. The potential profile is described by a piecewise constant (step) function with potential values V_j :

$$V(x) = \sum_{j=1}^N V_j u_j(x).$$

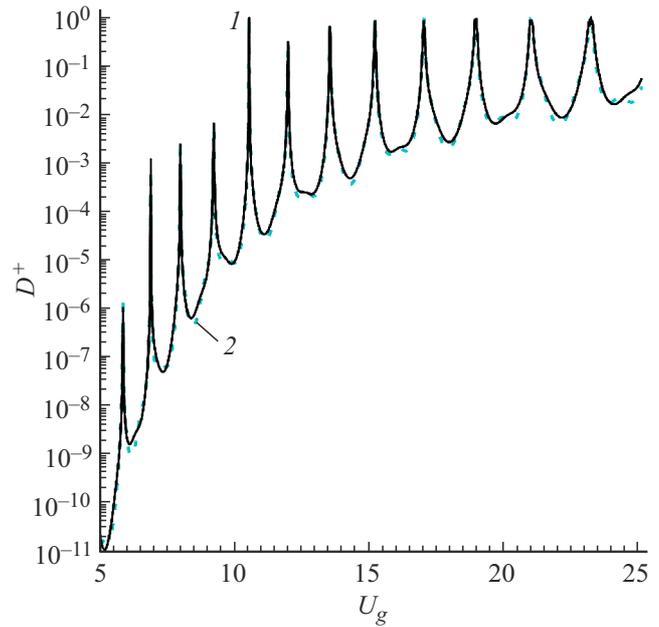


Figure 4. Tunneling coefficient at $E = E_F$ for structure $t_1 = t_2 = t_3 = 2$, $t_g = 1$, $\delta = 0.1$, $\delta_g = 0.15$ (nm) depending on the grid potential U_g (V): curve 1 — $U_a = 8$ V; curve 2 — $U_a = 5$ V.

The functions $u_j(x)$ are equal to one on the corresponding step and zero outside them. Since $V_j < 0$, for the level $V_j = -V_0$, corresponding to the bottom of the conduction zone, all kinetic energies are above the barrier, and for levels above the Fermi level $V_j > -V_F$ all motions are below the barrier.

The program for solving SE consisted in construction of the complete matrix $\hat{T} = \hat{T}_1, \hat{T}_2, \dots, \hat{T}_N$ by multiplying the domain matrices and calculating the transmission coefficient from the system of equations $1 + R = (T_{11} + ik_a T_{12})T$, $1 - R = (-iT_{21}/k_0 + T_{22}k_a/k_0)T$. Adding them up, we get the result

$$T = 2/[T_{11} + T_{22}(k_a/k_0) + i(T_{12}k_a - T_{21}/k_0)],$$

$$R = 1 + (iT_{21}k_0 - T_{22}k_a/k_0)T.$$

The system of equations was obtained from the representation of the wave function on the cathode in the form

$$\psi(x) = A(\exp(ik_0x) + R \exp(-ik_0x)),$$

and on the anode — in the form $\psi(x) = AT \exp(ik_a(xd))$. Here we use the designations $k_0 = \sqrt{\mu E}/\hbar$ on the cathode and $k_a = \sqrt{\mu(E + eU_a)}/\hbar$ on the anode, i.e. these values are proportional to electron pulses on the cathode and anode (before pulse relaxation). Usually, tunneling is considered to be a process without loss of energy and pulse, i.e. at the turning point the pulse is $k_0\hbar$. However, after tunneling the pulse of the electron on the anode increases due to acceleration as result of the above barrier motion. Indeed, in this region, an accelerating force acts on the electron with

actual pulse, i.e. the motion is quasiclassical. Therefore, at the cathode boundary the wave function should be taken in the above form. After getting the anode, over the free path length, the excess pulse relaxes to $k_0\hbar$. In this case the electron gives energy eU_a to the phonons of the anode lattice. Since also

$$Z = (1 + R)/(1 - R) \\ = (T_{11} + ik_a T_{12})/(-iT_{21}/k_0 + T_{22}k_a/k_0),$$

you can find the reflection coefficient from the barrier in another way: $R = (Z - 1)/(Z + 1)$. In this case $|R|^2 + k_a/k_0|T|^2$, and Z, R depend not only on the shape of the barrier, but also on the anode voltage. Note that if we ignore the electron acceleration and take the wave function on the anode in the form $\psi = AT \exp(ik_0(xd))$, then we get the result $|R|^2 + |T|^2 = 1$. The reflection coefficient and the transmission coefficient included in it are different and correspond to the area of motion up to the last turning point. Since pulse relaxation and anode heating occur at $x > d$, these effects do not affect the wave function $AT \exp(ik_a(xd))$, which determines the passed particles flow. The calculated coefficient of tunneling or transparency of the barrier $D^+(E) = |T|^2$ depends on the kinetic energy of the electron in the cathode region E (we will omit the index k). The quasiclassical WKB approach considers only subbarrier motions in the regions between turning points and is not suitable for complex structures with quantum wells. Besides, in order to calculate the integrals, the barrier is usually assumed to be wide, and the reflected evanescent wave is neglected [18]. This explains why the Fowler–Nordheim formula cannot be applied even in cathode–anode structure with one hump and one well at high anode voltage, when the hump practically disappears, and only one well remains.

In the calculations, respectively, 101 and 152 steps were used for one-well and two-well barriers, i.e. each barrier was approximated by 50 steps, while one was sufficient for each well. The calculated coefficient of barrier transparency (tunneling) $D^+(E_k) = |T|^2$ depends on the kinetic energy of the electron in the cathode region E_k . The transparency of the emission structure has a resonant nature: some maxima are less than unity, but for some configurations (Fig. 3, curve 5) at several energies they can reach unity. The peaks have different widths, and the variation between the maxima and minima can reach several orders of magnitude. Transparencies are the higher, the smaller the total width of the humps is. The frequency of the maxima is the greater, the wider the structure is, and, other things being equal, is proportional to the total width of the quantum wells. Fig. 4 shows the dependence of the transparency at the Fermi level on the grid voltage for a two-well potential. It has the same resonant nature for lower energies.

The calculation of the total tunneling current requires accounting of the electrons distribution by speeds (energies), i.e., integrating $D^+(E)n(E)dE$ from zero to the Fermi energy. In this case, the optimal profile should

have a sufficiently large number of metastable negative levels V_n somewhat below the Fermi level of the cathode $V_c = V_F = -W$ and above the bottom of the cathode conduction zone V_0 , wherein with not too small width of spectral lines. The kinetic energy corresponding to the level is $E_{kn} = V_0 - V_n$. The bottom of the anode conduction zone corresponds to a deeper level $V_{0a} = -W - E_F - eU_a$, the Fermi level on anode $-V_a = -W - eU_a$, and the well on the grid — level $V_g = -W_g - eU_g$. Here we take the case of identical work functions of the cathode and anode, and for the grid it is equal to W_g . We do not consider the band structure of the grid. In the conduction zone of a thin grid there are many discrete levels below V_g . Above this level at $U_g > U_a$ stable levels can exist, above which metastable levels lie. The more such closely spaced levels with overlapping spectral lines are, and the closer they are to the Fermi level, the greater integral current from the structure can be obtained. Consequently, the problem of increasing the emission current is reduced to constructing such barrier structure and determining its metastable levels. It can be formulated like this. Let $\hat{T}(E_n)$ be the transfer matrix. Taking into account the relation $k_{an}^2 = k_{0n}^2 + \mu eU_a/\hbar^2$, we have

$$k_{0n} = \frac{\sqrt{\mu(E_n - V_c)}}{\hbar} = \frac{iT_{21} - k_{an}T_{22}}{T_{11} + ik_{an}T_{12}}. \quad (5)$$

Considering $1 = T_{11} + ik_{an}T_{12}$, $k_0 = iT_{21} - k_{an}T_{22}$, $\det(\hat{T}) = 1$, we have $T_{22} = 1 - (k_{0n}/k_{an})(1 - T_{11})$. The level is metastable if $V_a < \text{Re}(E_n)$, and stable if $E_n < V_a$.

The characteristic equation (5) determines all the complex levels E_n . If the level is actual and negative, then it is determined from the actual equation $T_{22} = 1 - (k_{0n}/k_{an})(1 - T_{11})$. The shown equations are transcendental, since the matrix depends on the level energy. They also make it possible to determine metastable levels that correspond to the conduction zone of the cathode. They are of interest for RT. The following method for determining complex roots is convenient. We find the real roots of the equation (5) when operating with its real part. Further, these real roots are used as initial approximations in the iterative procedure according to formula (5). Such a procedure requires calculation at each step of the transfer matrix with complex energy.

3. Comparison of one-well and two-well barriers

Fig. 3 shows the calculations of transparency $D^+(E)$ for one-well and two-well potentials, and Fig. 4 — $D^+(U_g)$ for $E_k = E_F$ for two-well potential, while the depth of the wells changes. Transparency weakly depends on the anode voltage. The two-well potential is created by a dual grid, the electrodes of which are separated by a vacuum gap or a gap filled with a dielectric, but are at the same electric potential. It is convenient to set $U_g > U_a$. The presence of the second well leads to a doubling of the levels [19] and to the total current increasing (Fig. 5). The presence of

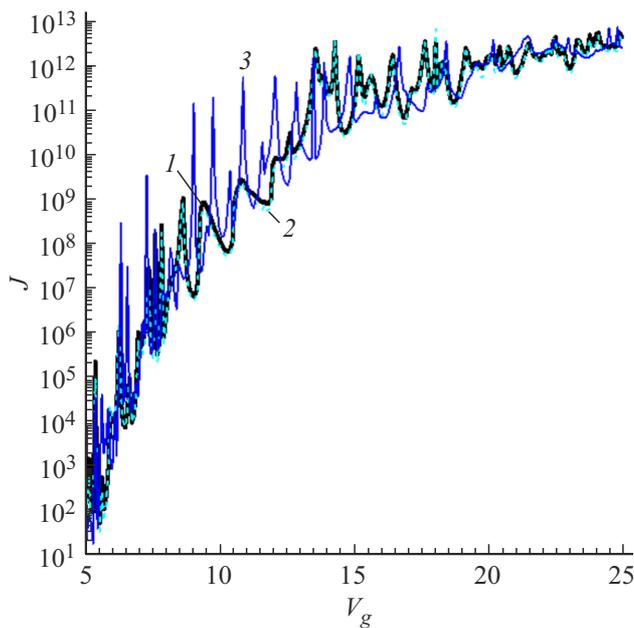


Figure 5. Volt-grid characteristics for the structure in Fig. 2, curve 2 for different U_a (V): 11 (curve 1); 9 (curve 2); 5 (curve 3).

quasi-periodic heterostructure can lead to the appearance of zones containing many closely spaced levels. However, the use of such structures is limited by the circumstance that their total length must be substantially less than the free path length in the grid material, i.e., the electron emitted from the cathode must not lose pulse due to scattering on phonons. Since at room temperature the free path length in metals is on the order of tens of nm, the length limitation is several nm. With such length of the grid structure it is already difficult to use more than two electrodes. The use of ultralow temperatures greatly complicates the design of the field emitter. In one-well asymmetric structure the quasi-levels coinciding with kinetic energies do not arise, and there is no complete resonant tunneling. The appearance of resonant peaks is due to the interaction with low-lying levels (incomplete extinguishing of the reflected wave by partial reflections from heterogeneities). The strongly irregular nature of the grid current-voltage curve can be explained by the fact that the level change of the bottom shifts the resonant levels, and maxima with minima appear at other close potential values. The anode current-voltage curves also demonstrate oscillations, but their peak-to-peak is few times maximum, and the period is about 2–3 V.

Table 2 shows the results of calculation of the complex levels $E_n = E'_n - iE''_n$ for the two-well structure. It can be seen that the E'_n values correlate well with the energies E_k at which the barrier transparency maxima occur (Fig. 3). Note that metastable levels play the main role in transient processes [1–6,19,34], which is important for pulsed sources and devices in pulsed mode. The considered two-well structures make it possible to achieve complete RT, while in one-well structures and structures with $U_g < U_a$ incomplete

Table 2. Metastable energy levels for the structure in Fig. 3, curve 5

Level number	E'_n , eV	E''_n , eV
1	4.1921	0.2012
2	2.8932	0.1213
3	1.7055	0.0347

RT is achieved [22]. Fig. 5 shows the results of calculating the total current $J = J^+ - J^-$ for one-well and two-well structures using the formula [15,16]:

$$J^\pm = \frac{em_e}{2\pi^2\hbar^2} \int_0^{E_F} D^\pm(E)(E_F - E)dE. \quad (6)$$

At the anode voltages used, the reverse current is negligible and $J \approx J^+$. The integral (6) was calculated numerically by the method of averages using 300 points. Actually low energies practically do not contribute to it. The results show a strong irregularity of the current-voltage curve with peak-to-peak up to several orders of magnitude and periods of up to fractions of volt. For Fig. 3, 4 the Fermi energy on the cathode $E_F = 5$ eV was taken, for Fig. 5 — $E_F = 10$ eV.

Based on the formula (6) it is possible to estimate the current during RT. Let there are energies E_n of RT and regions Δ_n around them, where $D^+ = 1$ can be considered (for Δ_n one can take the half-width of resonances). Due to the exponential smallness of D^+ , in the rest of the region we assume $D^+ = 0$. Then (6) is integrated, and we get

$$J^\pm = \frac{em_e}{2\pi^2\hbar^3} \sum_{n=1}^N (E_F - E_n)\Delta_n. \quad (7)$$

Formula (7) answers the question about the maximum current: it is the greater, the more RT levels there are, and the wider they are, and they should be located significantly below the Fermi level.

4. Correction of potential profile based on the solution of the Poisson equation

A high current density leads to significant density of uncompensated negative charge in the cathode–anode space, which can change the distribution of potential V . Let us introduce the matrices $\hat{t}_N = \hat{T}_N$, $\hat{t}_{N-1} = \hat{T}_{N-1}, \hat{T}_N, \dots, \hat{t}_1 = \hat{T}_1, \hat{T}_2, \dots, \hat{T}_N = \hat{T}$. They relate the amplitudes of forward and backward waves $\psi_j(x) = a_j^+ \exp(ik_jx) + a_j^- \exp(-ik_jx)$:

$$a_j^+ + a_j^- = (t_j^{11} + ik_a t_j^{12})T,$$

$$a_j^+ - a_j^- = (-it_j^{21}/k_j + t_j^{22}k_a/k_j)T,$$

whence we have

$$a_j^+ = [t_j^{11} + t_j^{22}k_a/k_j + ik_a t_j^{12} - it_j^{21}/k_j]T/2,$$

$$a_j^- = [t_j^{11} - t_j^{22}k_a/k_j + ik_a t_j^{12} + it_j^{21}/k_j]T/2,$$

$j = 1, 2, \dots, N$. Determining these amplitudes, we calculate the densities $\rho_j = \rho(x_j) = |\psi_j(x_j)|^2$. Let us go from discrete density values to continuous ones. In the relation to structure, we have the decomposition

$$\rho(x) = \sum_{n=0}^{\infty} [\alpha_n \cos(2n\pi x/d) + \beta_n \sin(2n\pi x/d)] \quad (8)$$

with coefficients

$$\alpha_0 = \frac{2}{d} \sum_{j=1}^N \rho_j \Delta_j, \quad \beta_0 = 0,$$

and for $n = 1, 2, \dots$ we have

$$\begin{aligned} \alpha_n &= \frac{2}{d} \sum_{j=1}^N \rho_j \int_{x_j - \Delta_n/2}^{x_j + \Delta_n/2} \cos\left(\frac{2n\pi x}{d}\right) dx = \frac{2}{n\pi} \sum_{j=1}^N \rho_j \\ &\times \sin\left(\frac{2n\pi \Delta_j}{d}\right) \cos\left(\frac{2n\pi x_j}{d}\right), \\ \beta_n &= \frac{2}{d} \sum_{j=1}^N \rho_j \int_{x_j - \Delta_n/2}^{x_j + \Delta_n/2} \sin\left(\frac{2n\pi x}{d}\right) dx = \frac{2}{n\pi} \sum_{j=1}^N \rho_j \\ &\times \sin\left(\frac{2n\pi x_j}{d}\right) \sin\left(\frac{2n\pi \Delta_j}{d}\right). \end{aligned}$$

The value (8) determines the density of particles in the flow, and the value $-\epsilon\rho(x)$ is the charge density included in the PE $\partial_x^2 \Phi(x) = e\rho(x)/\epsilon_0$. To solve PE, one can use discrete difference methods, but we will apply the series method, presenting the potential in the form

$$\Phi(x) = \sum_{n=0}^{\infty} [f_n \cos(2n\pi x/d) + g_n \sin(2n\pi x/d)]. \quad (9)$$

Accordingly, we have

$$f_n = \frac{e\alpha_n}{\epsilon_0(2n\pi/d)^2}, \quad g_n = \frac{e\beta_n}{\epsilon_0(2n\pi/d)^2}, \quad n = 1, 2, \dots$$

Potential (9) is periodic. To determine the last unknown coefficient f_0 , one should impose the boundary condition $\Phi(0) = \Phi(d) = 0$, whence we have

$$\begin{aligned} f_0 &= -\sum_{n=1}^{\infty} f_n = -\frac{ed^2}{2\pi^2\epsilon_0} \sum_{n=1}^{\infty} \frac{1}{n^3} \sum_{j=1}^N \rho_j \\ &\times \sin\left(\frac{2n\pi \Delta_j}{d}\right) \cos\left(\frac{2n\pi x_j}{d}\right). \end{aligned} \quad (10)$$

The incoming series converge quite well, so that in (10) it is sufficient to use several tens of terms. In (9) it is sufficient to use about hundreds of terms.

Let us consider a single-barrier structure with barrier height V_0 from the zero level of kinetic energy and a well depth $V_1 > 0$. To the left of the barrier the wave function has the form

$$\psi(x) = A[\exp(ik_0x) + R \exp(-ik_0x)],$$

and to the right of it —

$$\psi(x) = AT \exp(ik_1x).$$

Here $k_0 = \sqrt{\mu E_k}/\hbar$, $k_1 = \sqrt{\mu(E_k + V_1)}/\hbar$, A defines the particle flow. In the barrier area

$$\psi(x) = A[a^+ \exp(ikx) + a^- \exp(ikx)].$$

In this wave function $k = \sqrt{\mu(E_k - V_0)}/\hbar = i\kappa$, i.e.

$$\psi(x) = A[a^+ \exp(-\kappa x) + a^- \exp(\kappa x)].$$

We assume that $E_k < E_F < V_0$. Sewing the wave functions and their derivatives, we obtain the equations $1 + R = a^+ + a^-$, $1 - R = i\kappa(a^+ - a^-)/k_0$, $a^+ \exp(-\kappa d) + a^- \exp(\kappa d) = T$, $-a^+ \exp(-\kappa d) + a^- \exp(\kappa d) = ik_1 T/\kappa$. These equations have simple solutions:

$$a^+ = T(1 - ik_1/\kappa) \exp(\kappa d)/2,$$

$$a^- = T(1 + ik_1/\kappa) \exp(-\kappa d)/2,$$

$$T(\kappa) = \frac{2}{\cosh(\kappa d)(1 + k_1/k_0) - i(k_1/\kappa - \kappa/k_0) \sinh(\kappa d)}. \quad (11)$$

Denote $T = |T|^2 \exp(i\varphi)$, $\varphi_0 = \arctan(k_1/\kappa)$. Then the particles density

$$\rho(x) = |\psi|^2 = A^2 \tilde{\rho}(x)$$

is

$$\tilde{\rho}(x) = |a^+|^2 \exp(-2\kappa x) + |a^-|^2 \exp(2\kappa x) + 2|a^+|^2 \cos(2\varphi_0)$$

or

$$\begin{aligned} \tilde{\rho}(x, E) &= |T(E)|^2 [1 + (k_1/\kappa)^2] \\ &\times [\cosh(2\kappa(x - d)) + \exp(2\kappa d) \cos(2\varphi_0)]/2. \end{aligned} \quad (12)$$

Here $\kappa = \sqrt{\mu(V_0 - E)}/\hbar$, so (12) can be considered as a function of energy. Formula (12) was obtained for a one-speed flow with energy E and unit density of particles in the flow. Indeed, the wave function $A \exp(ik_0x)$ is used. For $A = 1$, such a normalization for plane wave gives a probability unit density and incident flow probability density $j^+ = k_0 \hbar/m_e = v$ equal to its speed. Accordingly, the flow at a length of v contains one particle per second, i.e. the particle passes through the x plane every second with

speed v . It is convenient to interpret such normalization as unit density of particles in one-speed flow [23]. If A^2 has the dimension „number of particles per unit length“, then $J = -eAv$ has the dimension of current, and if A^2 has the dimension „number of particles per unit volume grqq“, then $J = -eAv$ has the dimension of current density.

When tunneling from metal the distribution of electrons by energies should be taken into account. The wave function must be taken in the form of wave package

$$\Psi(x) = \Psi^+(x) + \Psi^-(x) = \int_0^{k_F} A(k) \times [\exp(ikx) + R(k) \exp(-ikx)] dk.$$

It's multi-speed. We consider the amplitude to be actual. The incident electron flow for the speed range $dv = (\hbar/m_e)dk$ has a particle bulk density (m^{-3}) [15]:

$$dn = \frac{m_e^3}{4\pi^2\hbar^3} (v_F^2 - v^2)dv = \frac{(k_F^2 - k^2)}{4\pi^2} dk.$$

Obviously it must be

$$|\Psi^+(x)|^2 = \int_0^{k_F} A^2(k)dk = \int_0^{k_F} \frac{(k_F^2 - k^2)}{4\pi^2} dk = \frac{k_F^3}{6\pi^2} = \frac{(\mu E_F)^{3/2}}{6\pi^2\hbar^3}. \quad (13)$$

We have obtained that (13) is equal to $N/2$, where N is the concentration of free electrons, i.e. the formula considers that half of the electrons moving towards the cathode boundary. In the absence of voltages they are completely reflected, i.e. exactly N electrons move in both directions. Now the total particles density can be written as

$$P(x) = \frac{m_e^3}{2\pi^2\hbar^3} \int_0^{v_F} (v_F^2 - v^2) \tilde{\rho}(x, v) dv = \frac{m_e^{3/2}}{2^{3/2}\pi^2\hbar^3} \int_0^{v_F} (E_F^2 - E^2) \tilde{\rho}(x, E) E^{-1/2} dE. \quad (14)$$

Expression (12) should be considered in the first integral (14) as a function of speed with the replacement $E = m_e v^2/2$. Taking into account (11) and (12), the integral (14) obtained for a rectangular barrier must already be calculated numerically.

From the form of the function (12) we can conclude that the density of the initial part of the barrier is low and increases towards its end. The density increasing means the electrostatic potential decreasing and the quantum potential increasing. This causes the flat top of the barrier to be beveled upward. The actual potential in the diode structure at a high anode voltage is close to triangular, so the influence of the space charge will lead to the transformation of an almost triangular barrier into a more rectangular one. Its peak as a whole will also rise. Assuming that V_0 and V_1

are determined by the anode potential, these effects can be compensated by its potential increasing. In the case of a multiwell potential formed by several electrodes, these effects are easily compensated by changing the potentials of the grid electrodes.

The paper [27] considers the space charge effect in solving PE with the following boundary conditions: $\Phi(0) = 0$, $\Phi(d) = U_a$, $\Phi'(0) = U_a/d$. Assuming that in the diode structure, electrons leave the cathode mainly with the Fermi speed, and over the length x acquire the speed $v = \sqrt{v_F^2 + 2e\Phi(x)/m_e}$, we obtain at constant current the density

$$\rho(x) = \frac{J/e}{\sqrt{v_F^2 + 2e\Phi(x)/m_e}}.$$

In resonant tunneling, the escape speed is lower and the density is higher. The relation of U_a and J obtained in the paper [27] has the form (in SI system)

$$\frac{U_a}{d} = \frac{U_a}{d} \left[1 - J \frac{4}{3} \left(\frac{m_e}{2e} \right)^{1/2} \frac{d^2}{\epsilon_0 U_a^{3/2}} \right].$$

For $J = 10^{10}$ A/m², $d = 10$ nm, $U_a = 20$ V the second term in square bracket is 0.085, and the bracket itself is equal to 0.915, i.e. this case corresponds to the field decreasing on the cathode by less than 10%. This limits the current, but, as can be seen from the formula, the field decreasing can be compensated by the anode voltage increasing.

When solving the Poisson equation we used the Fourier series decomposition formula and the Fourier series method to determine its coefficients. It implies a periodic continuation of the potential, which is quite natural under the boundary condition $\Phi(0) = \Phi(d)$, but it must be imposed. You can use the odd continuation formula

$$\Phi(x) = \sum_{n=1}^{\infty} u_n \sin(n\pi x/d). \quad (15)$$

It already satisfies zero boundary conditions. The representation (8) can be used to determine the coefficients. But it is better to take it in the form

$$\rho(x) = \sum_{n=0}^{\infty} \gamma_n \cos(n\pi x/d).$$

Then the coefficients are given by the expressions

$$u_n = -\frac{ed^2}{\pi^3 \epsilon_0 n^2} \sum_{m=0}^{\infty} \gamma_m v_{nm},$$

$$\gamma_n = \frac{2}{d(1 + \delta_{n0})} \sum_{j=1}^N \rho_j \int_{x_j - \Delta_j/2}^{x_j + \Delta_j/2} \cos\left(\frac{n\pi x}{d}\right) dx$$

$$= \frac{4}{n\pi(1 + \delta_{n0})} \sum_{j=1}^N \rho_j \sin\left(\frac{n\pi \Delta_j}{2d}\right) \cos\left(\frac{n\pi x_j}{d}\right).$$

Here $v_{nm} = 0$ if n and m are of the same parity, and $v_{nm} = 4n/(m^2 - n^2)$ if n and m of different parity.

Consider the question: is it possible to use the static formulas obtained above to correct the potential. Let now tunneling current flow through the structure with density J . Tunneling is the process without energy loss. If the electron riding on the barrier has speed v_0 , then it has the same speed at the exit of the barrier. Further, when moving above the barrier, it, generally speaking, accelerates and acquires the speed $v(x) = \sqrt{v_0^2 + 2eU(x)/m_e}$. At the considered voltages it is equal to or slightly greater than the Fermi speed. The current is determined as $J = -e\rho v$. Since $\text{div} J = 0$, there is no change in the space charge with time, i.e., it does not accumulate. But it presents: $-e\rho = J/v$. Taking the average speed in the form $\bar{v}(x) = \sqrt{2eU_a/m_e}$, approximating the potential in (15) with one first term and averaging, we obtain the rise of the barrier at the center:

$$\Phi(d/2) = \frac{d^2 J}{2\pi\epsilon_0\sqrt{2eU_a/m_e}}.$$

Taking quite technically achievable currents with density $J = 10^{10} \text{ A/m}^2$, $U_a = 9.1 \text{ V}$, $d = 10 \text{ nm}$, we get $\Phi(d/2) = 0.010 \text{ V}$. Thus, the density $J = 10^{12} \text{ A/m}^2$ is quite achievable. In a stationary single-speed beam forces do not act on electrons. More precisely, they do not affect the longitudinal motions, and the beam moves as a whole. The space charge is important when multi-speed longitudinal motions occur in different parts. If the beam is transversely limited, then separating transverse forces act. In a relativistic beam due to the pinch effect they almost do not work. When are space charge forces very important? Obviously, this will be in the non-stationary beam $J(x, t)$. The given stationary solutions should be modified. Let nonstationary beam appear at the moment $t = 0$ (at this moment voltage is applied to the electrodes). We write the representation of the current density in the form

$$J(x, t) = \sum_{n=0}^N b_n(t) \cos(n\pi x/d). \quad (16)$$

At each moment it is determined from the solution of the SE with the potential given at that moment. The field between the cathode and the anode is determined by the vector-potential with one component that satisfies the wave equation

$$\partial_x^2 A_x - c^{-2} \partial_t^2 A_x = -J(x, t). \quad (17)$$

For the vector-potential we take the decomposition

$$A_x(x, t) = \sum_{n=0}^N a_n(t) \cos(n\pi x/d). \quad (18)$$

Then for the decomposition coefficients we obtain the differential equations

$$a_n''(t) + (n\pi c/d)^2 a_n(t) = c^2 b_n(t). \quad (19)$$

Of these, for given right-hand sides one can determine $a_n(t)$, find the vector-potential, and find the scalar potential from the equation

$$\partial_x A_x(x, t) + \epsilon_0 \partial_t \Phi(x, t) = 0. \quad (20)$$

Integrating, for the scalar potential we have

$$\Phi(x, t) = \epsilon_0^{-1} \sum_{n=0}^N \frac{n\pi}{d} \sin\left(\frac{n\pi x}{d}\right) \int_0^t a_n(t') dt'. \quad (21)$$

It is equal to zero at the initial moment, and on boundaries satisfies the zero boundary conditions $\Phi(0, t) = \Phi(d, t) = 0$. The scalar potential can also be sought from the wave equation

$$\partial_x^2 \Phi(x, t) - c^{-2} \partial_t^2 \Phi(x, t) = -\frac{e\rho(x, t)}{\epsilon_0} = \frac{1}{\epsilon_0} \int_0^t \partial_x J(x, t') dt'. \quad (22)$$

Differential equations (19) are solved by the Fourier method. With frequency notation $\omega_n = n\pi c/d$ this gives

$$a_n(t) = \frac{c^2}{\omega_n} \int_0^t b_n(t') \sin(\omega_n(t - t')) dt'. \quad (23)$$

Really,

$$a_n'(t) = c^2 \int_0^t b_n(t') \cos(\omega_n(t - t')) dt',$$

$$a_n''(t) = -c^2 \omega_n \int_0^t b_n(t') \sin(\omega_n(t - t')) dt' + c^2 b_n(t),$$

which satisfies equation (19). Thus, we get the potential

$$\begin{aligned} \Phi(x, t) &= \frac{c}{\epsilon_0} \sum_{n=0}^N \sin\left(\frac{n\pi x}{d}\right) \int_0^t \int_0^{t'} b_n(t'') dt'' dt' \\ &= \sum_{n=0}^N u_n(t) \sin\left(\frac{n\pi x}{d}\right). \end{aligned} \quad (24)$$

This potential should be added with the alternating potential on the electrodes. In particular, we can assume that $U_a(t) = U_{0a} + u_a(t)$, $U_g(t) = U_{0g} + u_g(t)$, $u_a(0) = u_g(0) = 0$. For the modified potential, we solve the non-stationary SE, from which we find a new value of the current density. We continue the process until convergence, after which we change the time, i.e. the potentials on the electrodes, and repeat everything for a new moment in time. The process of solving non-stationary SE is considered in the paper [6]. Since it is relativistically non-covariant, a change in the wave function in some region leads to this perturbation propagation at the next infinitely small

moment of time over the entire infinite region, which corresponds to unlimited propagation of perturbations spreading speed. However, at a large distance these perturbations are infinitely small, since the propagatory Green's Function (GF) is similar to the Gaussian function. The equations in [6] are of the Lippmann–Schwinger type and require the construction of a complete GF. It is easy to construct such GF in the region $0 < x < d$ by specifying particle flows from the left and from the right. Then the main task is to determine the waves diverging from the structure. The above equations (16)–(24), in fact, are the equations of the Fabry–Perrot resonator excitation. They are relativistically covariant. It makes no sense to use them in conjunction with the Dirac equation, since at the considered voltages the flow is not relativistic, and the speeds are small. The spin accounting would lead to the need for multiparticle consideration, and the need to introduce only the vector-potential complicates the description of the barriers.

5. Effects associated with high current density

Experiments on simple cathode–anode structures showed the possibility of obtaining densities up to 10^{15} A/m^2 [17], i.e. almost up to limit. When limit currents are reached, there are a number of negative effects. The tunnel current flows through the structure and closes from the anode through the power supply to the cathode. We assume that there is no grid current. In the extraordinary majority of works relating field emission from metals the electrons in the metals are considered as free electrons with dispersion $E = p^2/(2m_e)$. This is associated with very wide conduction zone of about $E_F + W_0$, which is usually larger than 10 eV, while the electron accelerated from the Fermi level can acquire energy up to several eV without significant distortion of the dispersion. The internal field accelerating the electron is $E_J = J/\sigma$. Substituting the conductivity of copper, we obtain at $J = 10^{13} \text{ A/m}^2$ the value $E_J = 1.75 \cdot 10^6 \text{ V/m}$, which, at the free path length 40 nm will lead to acquired energy of 0.07 eV. Therefore, the effects associated with the nonparabolic nature of the dispersion, i.e. with a complicated dispersion law and the structure of the Fermi surface, can be ignored. Thus, the main effect of a strong current is to change the configuration of potentials, which can be corrected by the potentials increasing on the electrodes. The effects of potentials change were considered, for example, in papers [35,36].

The most negative effect during resonant tunneling is the cathode heating due to the Nottingham effect and due to the Joule heat. The first one is very strong, since the transparency regions can be significantly below the Fermi level. Besides, it is almost surface, since heat is released approximately at the cathode–vacuum boundary along the free path length. The free path length is related to the pulse relaxation, and the heat release due to the electron transition from the Fermi level to the hole formed due

to tunneling occurs at the free path lengths by energies, which can reach 100 nm. A possible solution of this problem are low and ultralow temperatures, as well as ultrapure cathode materials. In this case, the free path length increases by orders of magnitude, i.e., the effect becomes to volumetric, but it exists even at ultralow temperatures. Besides, the free path length increasing makes it possible to make more extended multielectrode structures. A nanosized structure can be considered flat if its transverse dimensions are 2–3 orders of magnitude larger. It is advisable to place such a structure on a massive metal base with even larger transverse dimensions — thermostat. Stable sources are important for emission electronics. The thermostable mode for a given current can be calculated quite well by selecting the length and difference of surface temperatures of the cathode and thermostat.

Conclusion

Thus, the paper shows the possibility to achieve at average electric fields of about several V/nm in one-well and two-well resonant-tunneling structures of current densities by 1–2 orders of magnitude higher than those achievable in much stronger fields. Specific triode nanostructures of field emission electronics are proposed, in which the said profiles of the quantum potential and emission currents are realized. The work of structures is based on the appearance of metastable levels in the energy range below the Fermi energy on the cathode, and their creation requires the use of nanotechnologies to make electrodes of cathode, grids and anode at distances of about 10 nm. Two-well structures with a dual grid are more promising. The use of grids in microstructural emitters is a solved problem even for nonplanar structures [37]. In flat nanostructured emitters on the grids a constant potential must be provided, and the grid itself must be very strong. Promising materials for grids can include multilayer graphene and woodpile-like structures made of carbon nanotubes with a metallic type of conductivity. Such structures can be fixed by laser welding. They are very durable and have high thermal conductivity. At grid size of several nm, electron scattering can be neglected.

Acknowledgments

The author wishes to thank O.E. Glukhova and R.K. Yafarov for discussing a number of issues.

Funding

The study was supported by grant from the Russian Science Foundation №21-19-00226.

Conflict of interest

The author declares that he has no conflict of interest.

References

- [1] K.S. Grishakov, V.F. Elesin. *Semiconductors*, **50** (8), 1092 (2016). DOI: 10.1134/S1063782616080121
- [2] V.F. Elesin. *ZhETF*, **145** (6), 1078 (2014) (in Russian) [V.F. Elesin. *JETP*, **118** (6), 951 (2014)]. DOI: 10.1134/S1063776114060041
- [3] V.F. Elesin, B.V. Kopaev. *JETP*, **96** (6), 1149 (2003). DOI: 10.1134/1.1591227
- [4] V.F. Elesin, I.Yu. Kateev, M.A. Remnev. *Semiconductors*, **43**, 257 (2009). DOI: 10.1134/S1063782609020250
- [5] V.F. Elesin, I.Yu. Kateev. *Semiconductors*, **39** (9), 1071 (2005). DOI: 10.1134/1.2042601
- [6] M.V. Davidovich. *JETP Lett.*, **110** (7), 472 (2019). DOI: 10.1134/S0021364019190068
- [7] D.I. Trubetskoy, A.G. Rozhnev, D.V. Sokolov. *Lektsii po sverkhvysokochastotnoy vakuumnoy mikroelektronike* (GosUNTs „Koledzh“, Saratov, 1996) (in Russian)
- [8] C. Bower, D. Shalóm, W. Zhu, D.I. López, G.P. Kochanski, P.L. Gammel, S. Jin. *IEEE Trans. ED*, **49** (8), 1478 (2002). DOI: 10.1109/TED.2002.801247
- [9] R. Riccitelli, F. Brunetti, C. Paoloni, G. Ulisse, A. Di Carlo, V. Krozer. *IEEE International Vacuum Electronics Conference* (Monterey, IEEE Publisher, USA, 2008), p. 382–383. DOI: 10.1109/IVELEC.2008.4556544
- [10] J.W. Gadzuk. *Surf. Sci.*, **15**, 466 (1969). DOI: 10.1016/0039-6028(69)90135-6
- [11] M.B. Partenskii. *Sov. Phys. Usp.*, **22**, 330 (1979). DOI: 10.1070/PU1979v022n05ABEH005498
- [12] C. Berthod, T. Giamarchi. *Phys. Rev. B*, **84**, 155414 (2011). DOI: 10.1103/PhysRevB.84.155414
- [13] T.N. Todorov, G.A.D. Briggs, A.P. Sutton. *J. Phys. Condens. Matter*, **5**, 2389 (1993).
- [14] W. Wang, Z. Lia. *J. Appl. Phys.*, **109** (11), 114308 (2011). DOI: 10.1063/1.3587186
- [15] D.I. Proskurovsky. *Emissionnaya elektronika* (Izd-vo TGU, Tomsk, 2010) (in Russian)
- [16] G.N. Fursey. *Field Emission in Vacuum Micro-Electronics* (Kluwer Academic Plenum Publishers, Springer, NY, 2005)
- [17] G.N. Furosey. *Avtoelektronnaya emissiya* (Lan², SPb, 2012) (in Russian)
- [18] A.S. Davydov. *Kvantovaya mekhanika* (Nauka, M., 1973) (in Russian)
- [19] E.A. Nelín. *Phys. Usp.*, **50**, 293 (2007). DOI: 10.1070/PU2007v050n03ABEH006091
- [20] J.G. Simmons. *J. Appl. Phys.*, **34** (6), 1793 (1963). DOI: 10.1063/1.1702682
- [21] M.V. Davidovich, R.K. Yafarov. *Tech. Phys.*, **63** (2), 274. DOI: 10.1134/S106378421802010X
- [22] M.V. Davidovich, R.K. Yafarov. *Tech. Phys.*, **64** (8), 1210. DOI: 10.1134/S106378421908005X
- [23] M.V. Davidovich. *J. Exp. Theor. Phys.*, **130**, 35 (2020). DOI: 10.1134/S1063776119120161
- [24] L.D. Landau, I.M. Lifshits. *Kvantovaya mekhanika. Nerelyativistskaya teoriya* (GIFML, M., 1963) (in Russian)
- [25] M.V. Davidovich. *Quant. Electron.*, **47** (6), 567 (2017). DOI: 10.1070/QEL16272
- [26] M.V. Davidovich, I.S. Nefedov, O.E. Glukhova, M.M. Slepchenkov. *J. Appl. Phys.*, **130**, 204301 (2021). DOI: 10.1063/5.0067763
- [27] T.E. Stern, B.S. Gossling, R.H. Fowler. *Proc. R. Soc. Lond., A* **124**, 699 (1929). DOI: 10.1098/rspa.1929.0147
- [28] J. Bardeen. *Phys. Rev.*, **49** (9), 656 (1936). DOI: 10.1103/PhysRev.49.653
- [29] T.L. Juretschke, *Phys. Rev.*, **93** (5), 1140 (1953). DOI: 10.1103/PhysRev.92.1140
- [30] G.F. Vasiliev. *Radiotekhnika i elektronika*, **5**, (11) 1857, (1960) (in Russian)
- [31] T.L. Loucks, P.H. Catler. *J. Phys. Chem. Sol.*, **25** (1), 105 (1964). DOI: 10.1016/0022-3697(64)90166-0
- [32] P.H. Cutler, J.J. Gibbons. *Phys. Rev.*, **111** (2), 394 (1958). DOI: 10.1103/PhysRev.111.394
- [33] G.G. Belford, A. Kuppermann, T.E. Phipps. *Phys. Rev.*, **128** (2), 524 (1962). DOI: 10.1103/PhysRev.128.524
- [34] A.I. Baz', Ya.B. Zel'dovich, A.M. Perelomov. *Rasseyanie, reaktivnaya i raspady v nerelyativistskoy kvantovoy mekhanike* (Nauka, M., 1971) (in Russian)
- [35] F.I. Itskovich. *ZhETF*, **50** (5), 1425 (1966) (in Russian) [F.I. Itskovich. *ZETF*, **50** (5), 1425 (1966).]
- [36] F.I. Itskovich. *ZhETF*, **52** (6), 1720 (1967) (in Russian) [F.I. Itskovich. *ZETF* **52** (6), 1720 (1967).]
- [37] N.S. Xua, S. Ejaz Huqb. *Mater. Sci. Engineer. R Reports*, **48** (2), 47 (2005). DOI: 10.1016/j.mserr.2004.12.001