Quantum dot heterostructures: fabrication, properties, lasers (Review)

N. N. Ledentsov, V. M. Ustinov, V. A. Shchukin. P. S. Kop'ev, and Zh. I. Alferov

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia

D. Bimberg

Institut für Festkörperphysik, Technische Universität Berlin, 10623 Berlin, Germany

(Submitted November 5, 1997; accepted for publication November 13, 1997)
Fiz. Tekh. Poluprovodn. **32**, 385–410 (April 1998)

In the present review we summarize original results where 1) we have experimentally discovered a novel class of spontaneously ordered nanostructures, namely equilibrium arrays of three-dimensional, coherently strained islands on crystal surfaces; 2) we have developed a theory of spontaneous formation of semiconductor nanostructures in heteroepitaxial systems; 3) we have experimentally demonstrated the existence of a novel class of semiconductor heterostructures, namely perfect quantum dots having an atom-like energy spectrum; we have performed a detailed investigation of the optical properties of quantum dots; 4) we have fabricated quantum dot-based injection lasers demonstrating unique charactristics, namely high-temperature stability of the threshold current and ultra-high material gain. © *1998 American Institute of Physics.* [S1063-7826(98)00104-5]

1. INTRODUCTION

Heterostructures[1] with spatially limited charge carriers in all three dimensions (quantum dots) realize the limiting case of size quantization in semiconductors, when modifications of the electronic properties of the material are most strongly expressed. The electron spectrum of an ideal quantum dot (QD) is a set of discrete levels separated by regions of forbidden states, and corresponds to the electron spectrum of an isolated atom, although a real quantum dot can consist of hundreds of thousands of atoms. Thus, the unique possibility of modeling the experiments in atomic physics with macroscopic objects presents itself. From the device point of view, the atom-like electron spectrum of charge carriers in quantum dots in the case where the distance between the levels is noticeably greater than the thermal energy makes it possible to eliminate a major problem of contemporary micro- and optoelectronics, namely ''smearing'' of charge carriers over an energy window of the order of kT, leading to a degradation of device properties with increase of the working temperature. In addition, all of the material characteristics that are most important for applications, e.g., the radiative recombination time, the energy relaxation time between electron sublevels, the Auger recombination coefficients, etc. turn out to depend radically on the geometric size and shape of the quantum dot, which makes it possible to use the same semiconductor system to realize devices that impose substantially different requirements on the active medium.

Over many years attempts have been made worldwide to fabricate quantum dots and devices based on them by ''traditional means,'' e.g., by selective etching of quantum-well structures,[2] growth on profiled substrates and on cleavage faces,[3] or condensation in glassy matrices.[4] Within the scope of these efforts device-oriented structures have not been created, but the fundamental possibility of realizing an atom-like spectrum of the density of states in a macroscopic semiconductor structure has not been explicitly demonstrated.

A qualitative breakthrough in this case involves the use of self-organizing effects of semiconductor nanostructures in heteroepitaxial semiconductor systems. This effort led to the realization of ideal heterostructures of quantum dots with high crystalline perfection, high quantum yield of radiative recombination and high size homogeneity ($\sim$10%). In such fabricated structures unique physical properties were demonstrated for the first time which had been expected for ideal quantum dots for many years, effects associated with energy relaxation and radiative recombination of nonequilibrium carriers, etc. and optoelectronic devices such as quantum-dot injection heterolasers were constructed for the first time.

2. SPONTANEOUS ORDERING OF SEMICONDUCTOR NANOSTRUCTURES. FORMATION OF QUANTUM DOTS

2.1. General information

The spontaneous appearance of periodically ordered structures on semiconductor surfaces and in epitaxial semiconductor films encompasses a wide range of phenomena in solid state physics and in semiconductor technology. The sponataneous appearance of nanostructures belongs to a wider class of fundamental phenomena of self-organization in condensed media. The explosion of interest in this field has been due to the need to fabricate semiconductor nanostructures with characteristic dimensions of 1–100 nm, and the spontaneous ordering of nanostructures makes it possible

 © 1998 American Institute of Physics

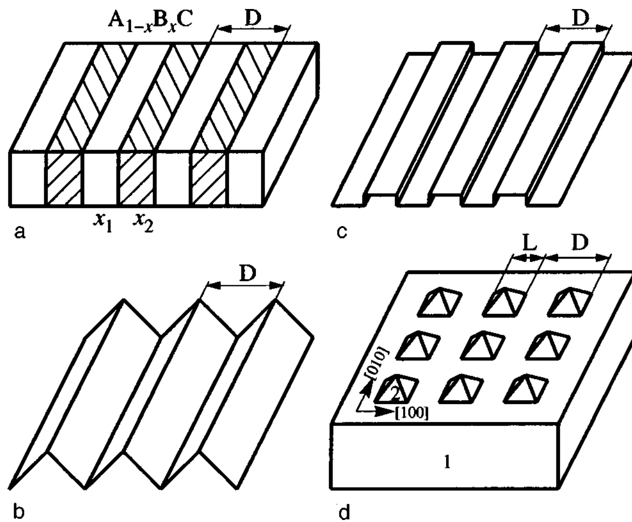


FIG. 1. Different classes of spontaneously arising nanostructures: a—structures with modulation of composition of the solid solution; b—periodically faceted surfaces; c—periodic structures of planar elastic domains; d—ordered arrays of three-dimensional, coherently strained islands (*2*) on substrate (*1*).

to obtain inclusions of narrow-band semiconductors in a wide-band matrix and thereby create a localizing potential for current carriers. Periodic structures of such inclusions can form superlattices consisting of quantum wells, quantum wires, or quantum dots. The phenomenon of spontaneous formation of nanostructures creates a basis for a new technology of fabricating ordered arrays of quantum wires and quantum dots—a basis for a new generation of opto- and micro-electronics.

In the context of the physical mechanisms of the spontaneous appearance of ordered nanostructures it is customary to distinguish two fundamental possibilities. First, ordered nanostructures can arise in closed systems, e.g., when annealing samples or upon extended interruption of growth. Such structures are equilibrium structures, and a thermodynamic approach is used to describe them. Second, ordered structures can arise in open systems during crystal growth. These structures are not equilibrium structures, and a kinetic treatment must be used for their description. In our studies we understand, ''self-organization'' of nanostructures in a broad sense, as any sort of appearance of macroscopic order in an originally homogeneous system.[5] Such a use of the term encompasses both equilibrium phenomena and nonequilibrium processes, and also combinations of the two. Such an approach makes it possible to analyze from a unified position the various mechanisms of spontaneous appearance of nanostructures, where, as a rule, equilibrium is only partly established (e.g., equilibrium may be established on the surface but not in the volume).

In this section we consider spontaneously ordered nanostructures, among which we can distinguish four large classes, as shown in Fig. 1. These nanostructures are

—structures with periodic modulation of content in epitaxial films of semiconductor solid-solutions;

—periodically faceted surfaces;

—periodic structures of planar domains (e.g., monolayer-high islands);

—ordered arrays of three-dimensional, coherently strained islands in heteroepitaxial mismatched systems.

Although the reason for the instability of the homogeneous state is different for each class of nanostructures, the reason for ordering in the inhomogeneous state is the same for all classes of nanostructures. In all these systems neighboring domains differ in their lattice constant and/or surface structure and, consequently, the domain boundaries are sources of long-range strain fields. This makes it possible to use a unified approach to all four classes of ordered nanostructures and treat them as equilibrium structures of elastic domains corresponding to a free-energy minimum. Until recently, the domain structures depicted in Figs. 1a–c have been traditionally considered outside the scope of semiconductor nanostructures. The unified approach developed in Refs. 6–9 allows one to delineate the main regularities of formation of ordered structures in these simpler cases (Figs. 1a–c) and then, on the one hand, apply them to a description of arrays of three-dimensional coherently strained islands (Fig. 1d) and, on the other hand, use them in the development of a new technology of fabricating semiconductor nanostructures.

The applicability of the thermodynamic approach to processes taking place during molecular-beam epitaxy (MBE) of III–V semiconductors was grounded earlier in a number of works (see, e.g., Refs. 10–14). It is on this basis that the thermodynamic approach to the description of processes of spontaneous formation of nanostructures during MBE of III–V semiconductors is presently founded.

2.2. Concentration elastic domains in semiconductor solid solutions

The possibility of the spontaneous formation of structures with modulated composition in solid solutions is connected with the instability of a homogeneous solid solution to spinodal decomposition.[15,16] For the solid solution $A_{1-c}B_cC$ this instability means that a solid solution with some inhomogeneous composition profile $c(\mathbf{r})=\bar{c}+\delta c(\mathbf{r})$ has a lower free energy than the corresponding homogeneous solid solution with the composition $c(\mathbf{r})=\bar{c}$. The change in the free energy of the system due to fluctuations of the composition $\delta c(\mathbf{r})$, is

$$\delta F=\int\{[H(\bar{c}+\delta c(\mathbf{r}))-TS_i(\bar{c}+\delta c(\mathbf{r}))]-[H(\bar{c})-TS_i(\bar{c})]\}dV+E_{\text{elastic}}, \quad (1)$$

where H is the enthalpy, S_i is the entropy of mixing, T is the temperature, and E_{elastic} is the elastic energy. The instability of the homogeneous solid solution to fluctuations of the composition arises when the enthalpy of formation of the solid solution $A_{1-c}B_cC$ from the binary components AC and BC is positive, $\Delta H_{\text{formation}}=H(A_{1-c}B_cC)-(1-c)H(AC)-cH(BC)>0$, which is valid for all ternary solid solutions of III–V semiconductors. Thus, at $T=0$ a two-phase mixture of the pure ingredients AC and BC has a lower free energy

than the homogeneous solid solution $A_{1-c}B_cC$, and the latter turns out to be unstable. At finite T the contribution S_i to the free energy corresponds to complete mixing of the components and stabilizes the homogeneous solid solution.

The elastic energy is due to the dependence of the equilibrium lattice parameter of the solid solution, a, on the composition c according to Végard's law. Regions of the solid solution with different compositions have different values of the equilibrium lattice constant. Mating of the two regions takes place by way of elastic deformation, and the elastic energy is associated with this elastic deformation. It is specifically the elastic energy that depends on the spatial profile of the composition that defines the ''soft mode'' corresponding to the most unstable fluctuations of the composition. In a bulk sample the soft mode is due only to elastic anisotropy of the crystal and can be represented as a composition wave with wave vector parallel to a direction of lowest stiffness of the crystal.[15] For most cubic materials these are the [100], [010], and [001] directions. The final state of the decaying solid solution is a 1D layered structure of elastic concentration domains alternating along one of the lowest-stiffness directions.[16]

Since $E_{\text{elastic}}=0$ in a homogeneous solid solution and $E_{\text{elastic}}>0$ in the inhomogeneous solid solution, the elastic energy stabilizes the homogeneous solid solution. The calculation of the critical temperatures of instability of the solid solution to spinodal decomposition carried out by Stringfellow[17] showed that bulk crystals of ternary solid solutions of III–V semiconductors are stable to spinodal decomposition at all temperatures.

To investigate the possibilities of spontaneous formation of nanostructures, Ipatova *et al.*[18,19] constructed a theory of spinodal decomposition in epitaxial films of solid solutions, where relaxation of strains near a free surface should decrease the effect of elastic stabilization of the homogeneous solid solution. For a solid solution in an epitaxial film on a (001) substrate of cubic crystal lattice-matched with the homogeneous solid solution with composition $\bar{c}$, they found a ''soft mode'' of the composition fluctuations. They showed that the soft mode is localized near the free surface $z=0$ and that it falls off exponentially into the film: $\delta c(\mathbf{r})\sim\exp(-|kz|)\exp(i\mathbf{k}_\parallel\mathbf{r}_\parallel)$, and the wave vector in the surface plane $k_\parallel=(k_x,k_y)$ is aligned with the lowest-stiffness axis [100] (or [010]). Because of relaxation of the strains near the free surface, the elastic energy associated with the ''free mode'' is decreased by a factor of $\approx 1/3$ in comparison with the elastic energy in a bulk sample, and many solid solutions turn out to be unstable to spinodal decomposition already at 450–500 °C. They found the equilibrium structure with composition modulation for which the composition was modulated in the surface plane in either the [100] or the [010] direction, and the amplitude of modulation is maximum on the free surface and falls off into the film. These studies showed that relaxation of strains near a free surface facilitates the spontaneous formation of nanostructures, and the elastic anisotropy of the material determines the orientation of the nanostructures.

2.3. Periodically faceted surfaces

The reason for spontaneous faceting of a flat crystal surface is the orientational dependence of the surface free energy. If the flat surface has a large specific surface energy, then it spontaneously transforms into a structure of ''hills and troughs.''[20] This reduces the total free energy of the surface in spite of increasing its area.

The explanation for the appearance of a periodically faceted structure is linked to the concept of capillary phenomena on solid surfaces.[21,22] Since atoms on a solid surface are located in a different environment than atoms in the interior, the equilibrium distance between surface atoms differs from the equilibrium distance between atoms in the interior. Therefore, microscopically the surface of a crystal can be considered as an elastically strained layer. In the phenomenological description the surface possesses a surface tension, and the dependence of the surface energy γ on the deformation tensor $\varepsilon_{\alpha\beta}$ contains a nonvanishing linear component in the expansion

$$\gamma(\mathbf{m};\varepsilon_{\alpha\beta})=\gamma_0(\mathbf{m})+\tau_{\alpha\beta}(\mathbf{m})\varepsilon_{\alpha\beta}+\frac{1}{2}S_{\alpha\beta\varphi\phi}(\mathbf{m})\varepsilon_{\alpha\beta}\varepsilon_{\phi\varphi}+\dots \tag{2}$$

Here $\mathbf{m}$ is the surface normal, and $\tau_{\alpha\beta}(\mathbf{m})$ is the surface stress tensor. The quadratic expansion coefficients $S_{\alpha\beta\varphi\phi}(\mathbf{m})$ have the meaning of excess surface elastic moduli, which can be both positive and negative.

Surface tension on the surface of a crystal leads to capillary effects analogous to the Laplace pressure under a curved liquid surface.[21] The discontinuity in the surface stress tensor at a crystal edge leads to the appearance of elastic deformation fields in the crystal. The presence of a linear term in the deformation in the expansion of the surface energy (2) leads to the result that the contribution of the surface tension to the crystal energy is negative. In other words, surface tension-induced relaxation of the elastic energy occurs.

A theory of spontaneous formation of periodically faceted surfaces was constructed by Andreev[23] and Marchenko.[24] This approach was later extended to faceted surfaces in heteroepitaxial strained systems and is one of the starting points for our consideration of arrays of three-dimensional strained islands.

According to Marchenko,[24] the total energy of a periodically faceted surface with period D, defined per unit area of the initially flat surface, is

$$E=E_{\text{facets}}+E_{\text{edges}}+\Delta E_{\text{elastic}}, \tag{3}$$

where $E_{\text{facets}}=\text{const}(D)$ is the surface free energy of the sloping faces, $E_{\text{edges}}=C_1D^{-1}$ is the short-range component of the energy of the edges, and $\Delta E_{\text{elastic}}=-C_2D^{-1}\ln(D/a)$ is the elastic relaxation energy. Since deformations are created by line sources (edges), $\Delta E_{\text{elastic}}$ depends logarithmically on D, and the total energy (3) always has a minimum at some optimal period D_{opt}.

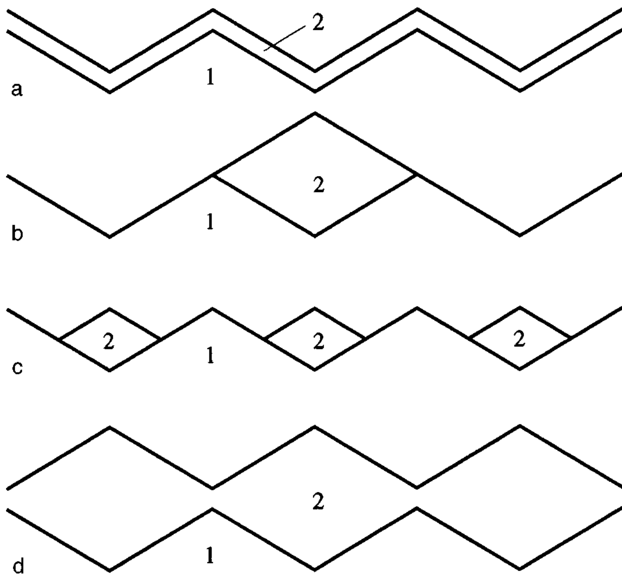


FIG. 2. Possible heteroepitaxial structures on a faceted surface. *1*—substrate material, *2*—deposited material, a—uniform coating; b—system of isolated ''thick'' clusters; c—system of ''thin clusters,'' d—heterophase system with a large amount of coating. The periodically faceted surface structure has been reconstructed. ''Hills'' on the surface of the heterophase system are located above the ''troughs'' of the substrate and vice versa. The heterophase system contains a continuous layer of material *2* with periodically modulated thickness.

2.3.1. Heteroepitaxial structures on periodically faceted substrates

Periodically faceted surfaces afford the possibility of direct fabrication of ordered arrays of quantum wires in the case where growth of material 2, which is deposited on the surface of material 1, takes place in ''troughs.'' To examine this possibility, an original theoretical study of possible structures of a heteroepitaxial system formed upon deposition of material 2 on a faceted surface of material 1 was performed.[25] This study considered a situation in which, first of all, the two materials are almost lattice-matched (an example is the system GaAs/AlAs) and, second, both the surface of material 1 and the surface of material 2 are unstable to faceting. At temperatures significantly below the melting point such an instability should hold for vicinal and high-index surfaces. Faceting has been experimentally observed for surfaces vicinal to (001) in GaAs,[26,27] in AlAs,[26], on (311) surfaces in GaAs and AlAs,[28] and on the (775) surface in GaAs.[29] For this situation several possibilities for the morphology of a heterophase system were considered (shown in Fig. 2) and the energies were compared.

The total energy of the heterophase system is[25]

$$E=E_{\text{facets}}+E_{\text{edges}}+\Delta E_{\text{elastic}}+E_{\text{interface}}. \tag{4}$$

This expression contains the interface energy, in addition to the three contributions present in the energy of a faceted surface of one material. The comparison of energies for several different types of heteroepitaxial structures carried out in Ref. 25 yields the following conclusion.

The choice between the two possible growth regimes is determined by whether the deposited material 2 wets the faceted substrate 1. If it wets it, a uniform coating of the periodically faceted substrate is formed (Fig. 2a). AlAs deposited on a periodically faceted GaAs(001) vicinal surface, misoriented by 3° in the $[1\bar{1}0]$ direction may serve as an example.[30]

If the deposited material does not wet the substrate, then isolated clusters of deposited material are formed in the ''troughs'' of the periodically faceted surface (Fig. 2b). Such a situation is realized when GaAs is deposited on an AlAs(001) vicinal surface, misoriented by 3° in the $[1\bar{1}0]$ direction, and also in the deposition of GaAs/AlAs(311) and AlAs/GaAs(311). Experimental studies of the system GaAs/AlAs(311) (Refs. 31 and 32) have demonstrated the possibility of direct fabrication of isolated clusters of GaAs on AlAs and subsequent formation of quantum wires during growth of the structure.

In the case of a nonuniform cluster coating, the periodic faceting of the surface is reconstructed after deposition of several monolayers. Thus, ''hills'' on the surface of the deposited material are formed above ''troughs'' of the substrate and vice versa, and a continuous layer with modulated thickness is formed (Fig. 2d). Thus, formation of clusters during heteroepitaxial growth makes it possible to directly fabricate isolated quantum wires, superlattices of quantum wires, and quantum wells with modulated thickness. Thus, periodic modulation of the thickness of a quantum well was observed by transmission electron microscopy (TEM) for a GaAs well in an AlAs matrix for the middle orientation of the (775) face.[29] The modulation period was 12 nm and the modulation amplitude was 1.2 nm.

2.3.2. Faceted surfaces in a heteroepitaxial mismatched system

A theoretical study of the effect of heteroepitaxial strain on the faceting of a surface was carried out in Refs. 33 and 34. A heteroepitaxial mismatched system was considered for the case where the surface of the film material is unstable to faceting. In this case, two sources of deformations are present in the system: first, lattice mismatch between the film and substrate and, second, the discontinuity in the surface stress tensor at the edges. The dependence of the faceting period on the mismatch was found. It was shown that as the mismatch is increased, the faceting period at first grows, but after that the possibility of equilibrium faceting disappears. The method for calculating the deformation fields and the elastic energy developed for such a system was later applied to a system of three-dimensional strained islands.

2.4. Surface structures of flat elastic domains

Structures of flat domains on a surface arise when different phases can coexist on the surface.[35,36] Examples are the surface reconstruction phases (2×1) and (1×2) on Si(001), islands of monolayer height in heterophase systems, etc. Thus, neighboring domains have different values of the surface stress tensor $\tau_{\alpha\beta}$, and effective forces on the crystal arise. These forces create a field of elastic deformations and contribute to the relaxation of the elastic energy. The total energy of the system of flat domains is[36]

$$E = E_{\mathrm{surface}} + E_{\mathrm{boundaries}} + \Delta E_{\mathrm{elastic}}. \qquad (5)$$

The surface energy E_{surface} does not depend on the period of the structure D, the energy of the domain boundaries is $E_{\mathrm{boundaries}} = C_1 D^{-1}$, the elastic relaxation energy has the form $\Delta E_{\mathrm{elastic}} = -C_2 D^{-1} \ln(D/a)$, and the total energy (2.5) always has a minimum for some optimal period D_{opt}.

2.4.1. The heteroepitaxial system InAs/GaAs(001) with submonolayer and monolayer coatings

In Refs. 37–39 the photoluminescence method was used to study grown-on layers of InAs(001) in a GaAs matrix, grown out to submonolayer and monolayer heights. For the system with submonolayer coating we observed a narrow line, which shifted toward shorter wavelengths relative to the luminescence line of the structure with one InAs monolayer. The anisotropy of the spectrum gives evidence of the non-equivalence of the [110] and $[1\bar{1}0]$ directions. The authors concluded that InAs forms islands which are stretched in the $[\bar{1}10]$ direction and which have identical width. Later the conclusion of the existence of nanosize islands for a submonolayer coating in InAs/GaAs(001) was confirmed by scanning tunneling microscopy,[40] where it was shown that the width of the islands is 4 nm.

The effect of prolonged interruption of growth on the structure of an InAs coating deposited on GaAs(001) was investigated in Ref. 41. For a system with submonolayer coating, for prolonged interruption of growth following growing-on of structures and measurement of the photoluminescence spectra it was established that an InAs monolayer coating is unstable and decays into surface segments with submonolayer coating and into islands a few monolayers in height. For systems with submonolayer coating it was shown that the photoluminescence spectra are not fundamentally changed by using the method of interruption of growth. This indicates that a heterophase system of islands of monolayer height for a submonolayer coating is stable.

The results of Refs. 38, 39, and 41 demonstrate that the system of islands of monolayer height in the system InAs/GaAs is an equilibrium system. This interpretation was advanced as a counterweight to kinetic models of the formation of islands for a submonolayer coating.[42] Thus, experiments on annealing for submonolayer and monlayer coatings testify to the stability of arrays of islands of monolayer height and confirm the applicability of the thermodynamic approach to describe heterophase systems of III–V semiconductors for a submonolayer coating.

Surface morphology studies aided by scanning tunneling microscopy[43] have shown that growth of one monolayer on the singular surface (001)GaAs and 30-s interruption of growth lead to the appearance of a one-dimensional corrugated structure which can be described as an array of ''quantum wires'' with 30-nm width, oriented in the [100] direction. Increase in the amount of deposited material leads to the appearance of a more complicated ''parquet structure,'' in which relief is present in the two dimensions [100] and [010]. Increasing the interruption of growth from 2 to 30 s makes the relief more visible.

Deposition of 1 and 1.5 monolayers of InAs on a GaAs vicinal surface, misoriented by 3° in the [110] direction relative to the (001) surface, leads to a more complicated corrugated surface on which relief is present in the [100] direction as well as in the direction perpendicular to the steps.

The stability of these surface structures upon interruption of growth indicates that the equilibrium porphology of the heteroepitaxial system InAs/GaAs(001) was observed for intermediate (subcritical) thickness of the deposited InAs layer. The symmetry of such equilibrium structures is determined by the symmetry of the surface stress tensor $\tau_{\alpha\beta}$ and by the symmetry of the bulk elastic moduli of the substrate. The principal axes of the surface stress tensors of GaAs and InAs are $[1\bar{1}0]$ and [110], while the bulk elastic properties of the materials are determined by the directions of the lowest-stiffness axes: [100] and [010]. Thus, there is a transition here from orientation of the elastic domains for submonolayer coating in the $[1\bar{1}0]$ direction, observed in Refs. 38 and 40, to orientation in the [100] and [010] directions for 1.0–1.5 monolayer coatings. This transition may be interpreted as the result of growth of the lattice mismatch contribution to the strain field in comparison with the contribution due to the discontinuity in $\tau_{\alpha\beta}$ at the boundary of the two phases. This growth is due to an increase in the volume of deposited InAs. On vicinal surfaces an additional contribution to the strains appears due to the steps.

2.5. Ordered arrays of three-dimensional, coherently strained islands

2.5.1. General morphology of mismatched heteroepitaxial systems

In heteroepitaxial growth it is customary to distinguish three regimes:

The Frank-van der Merwe regime, in which layer-by-layer (two-dimensional) growth of material B on substrate A is realized;

The Volmer–Weber regime, in which island (three-dimensional) growth of B on an open surface of the substrate A occurs;

The Stranski–Krastanow regime, in which layer-by-layer growth of B and A is first realized, followed by formation of three-dimensional islands of B on the open substrate.

In lattice-matched heteroepitaxial systems the growth regime is determined only by the relative values of the energies of the two surfaces and the interface energy. If the sum of the surface energy of the epitaxial layer γ_2 and the interface energy γ_{12} is less than the surface energy of the substrate: $\gamma_2 + \gamma_{12} < \gamma_1$, i.e., if the deposited material 2 wets the substrate, then the Frank-van der Merwe regime arises.

Variation of the sum $\gamma_2 + \gamma_{12}$ can lead to a transition from the Frank-van der Merwe regime to the Volmer–Weber regime.

In a heteroepitaxial system in the presence of lattice-mismatch between the deposited material and the substrate, initial growth can take place in a layer-by-layer mode. However, a thicker layer has a larger elastic energy, and there is a tendency for the elastic energy to decrease via formation of isolated islands. In these islands relaxation of the strains

takes place with a corresponding decrease of the elastic energy. The Stranski–Krastanow growth regime arises in this manner.

Experiments on InAs/GaAs(001) (Ref. 44) and on Ge/Si(001) (Refs. 45 and 46) have in fact demonstrated the possibility for the formation of three-dimensional, coherently strained, i.e., dislocation-free, islands. Theoretical studies[47,48] have shown that formation of three-dimensional, coherently strained islands leads to a decrease of the elastic energy and for not too large a volume of the island (up to $\approx 10^6$ atoms) is more favored than formation of islands with dislocations. However, it has been traditionally assumed that in a system of three-dimensional islands ripening (coalescence) must invariably take place, when large islands grow due to diffusive redistribution of material, which leads to diminution and disappearance of small islands,[49] and finally islands of such size are formed in which formation of mismatch dislocations is energetically favored. Such a coexistence of coherent islands and islands with dislocations was observed in Ref. 50.

However, subsequent experimental studies of arrays of coherently strained islands in InGaAs/GaAs(001) and InAs/GaAs(001) unexpectedly showed that a narrow size distribution of the islands is possible.[51,52] In addition to a narrow size distribution of the islands, the detection of a correlation in the positions of the islands, which is characteristic of a square lattice, was reported in Refs. 53–59. They showed that upon interruption of growth the dimensions of the islands and their mutual arrangement reach a limiting value and thereafter do not change further with time. The experimental results of those studies indicate the existence of a new class of spontaneously ordered nanostructures—ordered arrays of three-dimensional, coherently strained islands on a surface.

2.5.2. Equilibrium state in a system of coherently strained, three-dimensional islands

In connection with the experimental data of Refs. 53–59, a theoretical study of equilibrium in a system of three-dimensional, coherently strained islands was carried out in Refs. 60–65. In these studies a heterophase system was considered. It consisted of Q monolayers of material 2 deposited on a substrate of material 1 under conditions of growth interruption when the total amount of material 2 is fixed. If the thickness Q exceeds the critical thickness of the wetting layer Q_c, then the excess amount of material ($Q-Q_c$ monolayers) forms islands. To examine the fundamental possibility of formation of an ordered array of islands, it can be assumed that all the islands have the same shape and size and that they form a periodic superlattice on the surface. Thus, the change in the total energy of the system per unit surface area when part of the material from the flat strained layer goes into the islands is written in the form

$$\Delta E=\frac{\tilde{E}_{\text{island}}}{A_0}+\frac{\tilde{E}_{\text{interaction}}}{2A_0}. \tag{6}$$

Here $\tilde{E}_{\text{island}}$ is the change in the energy of the system associated with formation of one island, $\tilde{E}_{\text{interaction}}$ is the energy of interaction of one island with the others, and A_0 is the surface area underlying one island. $Q-Q_c$ monolayers with area A_0 form one island with volume $\tilde{V}_{\text{island}}$, and the condition for conservation of the amount of material is

$$(Q-Q_c)aA_0=\tilde{V}_{\text{island}}, \tag{7}$$

where a is the lattice constant.

The change in the energy of the system due to the formation of one island can be represented as the sum of three contributions

$$\tilde{E}_{\text{island}}=\Delta\tilde{E}_{\text{elastic}}+\Delta\tilde{E}_{\text{surface}}+\tilde{E}_{\text{edges}}. \tag{8}$$

A system of strained islands has two peculiarities distinguishing it from other classes of nanostructures. First, in a system of islands there are two sources of strain fields: on the one hand, lattice mismatch between the deposited material and the substrate, and on the other hand, the discontinuity in the surface stress tensor at the edges of the islands. Correspondingly, the elastic energy is equal to the sum of the bulk elastic relaxation energy, the elastic relaxation energy at the edges, and the interaction energy of the two elastic fields. Second, the deformational dependence of the surface energy due to capillary effects is important. Thus the change in the energy of the system upon the apperance of one island is[60]

$$\tilde{E}_{\text{island}}(L)=-f_1\lambda\varepsilon_0^2L^3+(\Delta\Gamma)L^2 -\frac{f_2\tau^2}{\lambda}L\ln\left(\frac{L}{a}\right)+f_3\eta L. \tag{9}$$

Here the first term is the bulk elastic relaxation energy, $\Delta E^V_{\text{elastic}}$, which is always negative. The second term is the variation of the surface energy of the system allowing for renormalization of the surface energy in the deformation field. The variation in the surface energy contains contributions due to the appearance of tilted faces of the islands, the appearance of an interface between the island and the substrate, and the disappearance of the flat segment of the wetting layer. It also contains linear and quadratic terms in the deformation due to renormalization of the surface energy in the deformation field. Key here for further analysis is that the quantity $(\Delta\Gamma)$ can be both positive and negative. The third term in expression (9) is the contribution of the edges of the island to the elastic relaxation energy $\sim -L\ln L$; this contribution is always negative. And finally, the fourth term in expression (9) is the short-range contribution to the energy of the edges, where η is the characteristic energy per unit length of an edge. The coefficients f_1, f_2, and f_3 are determined only by the shape of an island and do not depend on its size.

For a dilute array of islands, when the mean distance between the islands is large in comparison with the size of an island L, the migration time of the atoms over the surface of one island is significantly less than the migration time between islands. Therefore the equilibrium shape of any one island is established faster than the equilibrium structure of the entire array of islands, and for islands of any volume, there exists an equilibrium shape. This shape is determined by the minimum of the energy (9) under the condition of fixed volume of the island. At temperatures far from the melting point, the equilibrium shape contains only faces with small surface energy; as a rule, these are faces withe low

Miller indices. To analyze the possibility of formation of an equilibrium array of three-dimensional strained islands, it is sufficient to assume that the equilibrium shape of an island does not depend on its volume and that it coincides with the experimentally observed shape of an island—a pyramid with a square base. Thus, the quantity $(\Delta\Gamma)$ entering into expression (9) is expressed in terms of the surface energies of the different surfaces as follows:

$$(\Delta\Gamma)=\frac{\gamma_2}{\cos\vartheta}+\gamma_{12}-W-g_1\tau\varepsilon_0-g_2S\varepsilon_0^2. \quad (10)$$

Here γ_2 is the specific surface energy of the lateral faces of the pyramid, ϑ is the tilt angle of the lateral faces of the pyramid, γ_{12} is the specific surface energy of the interface, W is the specific surface energy of the wetting layer, $(-g_1\tau\varepsilon_0)$ is the interaction energy of the two elastic fields, which defines the renormalization of the surface energy that is linear in the mismatch ε_0, and $(-g_2S\varepsilon_0^2)$ is the renormalization of the surface energy that is quadratic in the mismatch.

The equilibrium state in the system of islands is reached as a result of exchange of material between the islands by migration over the surface of the wetting layer. For a dilute system of islands, the elastic interaction between islands can be ignored. Thus, substituting expression (9) into Eq. (6), we obtain

$$\Delta E=6(Q-Q_c)a\cot\vartheta\left[-f_1\lambda\varepsilon_0^2+\frac{\Delta\Gamma}{L}-\frac{f_2\tau^2}{\lambda L^2}\ln\left(\frac{L}{2\pi a}\right)+\frac{f_3\eta}{L^2}\right]. \quad (11)$$

Note that the bulk elastic relaxation energy $\Delta E_{\text{elastic}}^V$ [the first term in (11)] does not depend on the size of an island L. To find the minimum of ΔE from Eq. (11), it is convenient to introduce a characteristic length

$$L_0=2\pi a\exp\left[\frac{f_3\eta\lambda}{f_2\tau^2}+\frac{1}{2}\right] \quad (12)$$

and characteristic energy per unit area

$$E_0=\frac{3\cot\vartheta(Q-Q')af_2\tau^2}{\lambda L_0^2}. \quad (13)$$

Thus the energy $E'(L)$ of the sum of L-dependent terms takes the form

$$E'(L)=E_0\left[-2\left(\frac{L_0}{L}\right)^2\ln\left(\frac{e^{1/2}L}{L_0}\right)+\frac{2\alpha}{e^{1/2}}\left(\frac{L_0}{L}\right)\right]. \quad (14)$$

The behavior of the energy per unit area as a function of the diameter of a single island $E'(L)$ is governed by the controlling parameter

$$\alpha=\frac{e^{1/2}\lambda L_0}{f_2\tau^2}(\Delta\Gamma). \quad (15)$$

The physical meaning of the parameter α is that it is the ratio of the change in the surface energy of the system upon formation of one island, $(\Delta\Gamma)L^2$, to the contribution of the edges to the elastic relaxation energy, $|\Delta\tilde{E}_{\text{elastic}}^{\text{edges}}|$. The dependence of the energy of a dilute array of islands per unit surface area on the size of an island is plotted in Fig. 3 for different values of α. Analysis of this dependence allows us to determine whether there is a thermodynamic tendency in the system of islands toward ripening. If $\alpha>1$, then the minimum of the energy $E'(L)$ corresponds to islands whose size tends to infinity. Physically, this means that the change in the surface energy of the system due to formation of one island is positive and large. Therefore, it turns out to be energetically favorable to unite all the islands into one; i.e., there is a tendency in the system toward ripening. If $\alpha\leqslant 1$, then the minimum of the energy $E'(L)$ is reached at some optimum diameter of the islands. In this case, the change in the surface energy of the system due to formation of one island is either positive and small or it is negative. Thus, unification of the islands into one is not energetically favorable, and there is no tendency in the system toward ripening.

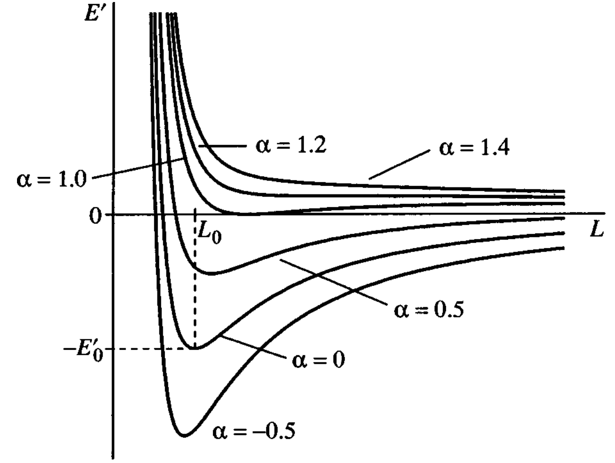


FIG. 3. Energy of a dilute array of three-dimensional, coherently strained islands per unit surface area as a function of island size. The parameter α is the ratio of the variation of surface energy upon formation of an island to the contribution of the edges of the island to the elastic relaxation energy. If $\alpha>1$, there is present in the system a thermodynamic tendency for the islands to coalesce. If $\alpha\leqslant 1$, there exists an optimal size of the islands, and the system of islands is stable to ripening.

Note that the change in the surface energy due to formation of an island $(\Delta\Gamma)$ can be negative even when the flat surface of deposited material itself is stable to spontaneous faceting, as is, for example, the case for the (001) surface of InAs. Upon the formation of an island, the appearance of tilted faces with area greater than the area of the base of the island is accompanied by the disappearance not of segments of flat surface of deposited material, but of segments of the wetting layer, whose surface energy can be strongly renormalized, first of all, because of chemical bonds with the substrate and, second, because of the deformational dependence of the surface energy.

For a dense array of islands, where the distance between the islands is comparable with the diameter of a single island, their elastic interaction due to penetration into the substrate of the inhomogeneous strain field created by the islands becomes important. The main interaction at large distances is the dipole–dipole interaction, $U\approx F(\varphi)/r^3$, where $F(\varphi)$ is a factor that depends on the azimuthal angle

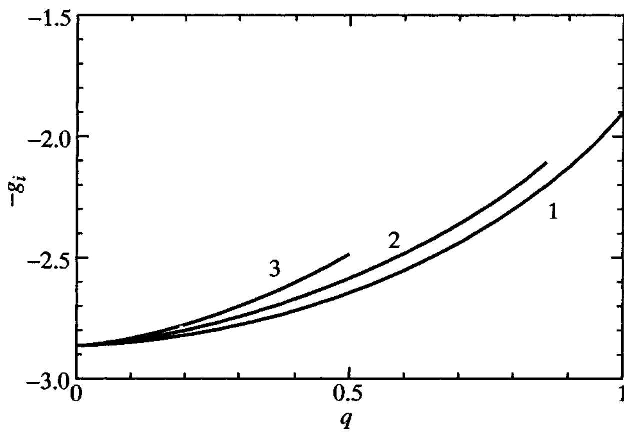


FIG. 4. Energy per unit area $\Delta E^V_{\text{elastic}}+E_{\text{interaction}}=(Q-Q')a\check{\lambda}\varepsilon_0^2\vartheta_0$ $(-g_i(q))$ for different arrays of interacting, coherently strained islands as a function of the fraction of surface q covered with islands. *1*—two-dimensional square lattice of pyramids with primitive translation vectors (1,0,0) and (0,1,0); *2*—two-dimensional hexagonal lattice of pyramids with primitive translation vectors $(-\frac{1}{2},-\sqrt{3}/2,0)$ and (1,0,0) and *3*—two-dimensional square lattice of pyramids in the form of a "checkerboard" with primitive translation vectors, and *3* terminate at the maximum possible covering for the array of islands of the given symmetry.

in the plane and is determined by the elastic anisotropy of the substrate. In contrast to a bulk crystal, in which the interaction of the elastic inclusions is sign-changing as a function of direction between the inclusions,[66] the interaction between islands on the surface is repulsive for any direction between the islands, which ensures stability of the array of islands.

To find the mutual arrangement of the islands in a plane posssessing minimum energy, we calculated the elastic energy for several different island superlattices. Figure 4 plots the results of this calculation for islands having the shape of pyramids with square base and small tilt angle ϑ_0 of the island faces to the (001) surface. The graph plots the sum of the bulk elastic relaxation energy and the interaction energy of the islands as functions of the fraction of surface q covered with islands, $\Delta E^V_{\text{elastic}}+E_{\text{interaction}}=(Q-Q')a\check{\lambda}\varepsilon_0^2\vartheta_0$ $(-g(q))$. Here $\check{\lambda}=(c_{11}+2c_{12})^2(c_{11}-c_{12})^3c_{11}^{-3}\times(c_{11}+c_{12})^{-1}$, and the functions $(-g(q))$ are plotted in Fig. 4 for different arrays of islands.

These results are for the energy of the islands that interact on a (001) surface of an elastically anisotropic cubic crystal. Among the various arrays of islands the minimum corresponds to the two-dimensional square lattice of islands, whose primitive translation vectors are aligned with the lowest-stiffness axes [100] and [010]. A comparison calculation, carried out for islands of the same shape but arrayed on the surface of an elastically isotropic medium, shows that in this case a hexagonal superlattice has the least energy. Thus, the energetic advantage of a square lattice over a hexagonal lattice is explained by the elastic anisotropy of the substrate, not by the shape of an individual island. In Ref. 60, it was shown that the results obtained in the approximation of a small tilt angle of the faces of the islands are preserved for a large tilt angle, in particular, for $\vartheta_0=45°$, i.e., for the experimentally observed shape of InAs/GaAs(001) islands.

The effect of the interaction between the islands on the possibility for the existence of an optimum size of the islands was examined in Ref. 60. It was shown that for $\alpha<0$ the minimum energy corresponding to finite size of the islands for a dilute array is kept for a dense array of interacting islands, and the tendency toward ripening does not arise. Thus, a theoretical treatment demonstrates the possibility for the existence of an equilibrium array of islands which have optimal size and which are ordered in a square lattice.

2.6. Technology of fabrication and structural studies of InAs/GaAs(001) three-dimensional islands

In Refs. 53–59 arrays of coherently strained InAs/GaAs(001) islands were fabricated by molecular-beam epitaxy. The surface morphology was monitored *in situ* by high-energy electron diffraction. After deposition of a fixed nominal thickness an InAs structure was grown on the GaAs substrate, yielding a system of InAs points in a GaAs matrix, which was then examined by transmission electron microscopy and photoluminescence spectroscopy.

A morphological study of the heterophase system InAs/GaAs(001) gave the following results. After deposition of a critical mean thickness of InAs(001) (1.6–1.7 monolayers) formation of three-dimensional InAs islands takes place. Deposition of 2 monolayers of InAs leads to formation of small-size islands, which for the most part do not manifest distinct crystallographic facets and which possess a wide size distribution. Increasing the thickness of the deposited InAs to 4 monolayers leads to the formation of a dense array of islands which have a square base with sides oriented along the [100] and [010] axes, and a length 140 Å. High-resolution transmission electron spectroscopy studies, observing in the plane of the surface and in cross section, revealed that the InAs islands have the shape of pyramids with (101) tilted faces and sharp edges between the faces. A statistical study of the mutual arrangement of the islands showed that the preferred directions between nearest neighbors are along the [100] and [010] axes.

2.6.1. Effect of growth interruption on the formation of islands

The application of growth interruption after InAs deposition before growing-on of the system of islands to the GaAs layer leads to dramatic changes in the morphology of the heterophase system in comparison with rapidly grown-on samples.[6] Interruption of growth at 40 s (10 s) is sufficient to allow the islands to reach a size of ~ 140 Å for 2.5 monolayers (3 monolayers) of InAs. For very long interruption of growth (100 s) the islands attain the same size even after deposition of 2 monolayers of InAs. These results indicate that the size of the InAs islands is ≈ 140 Å, the equilibrium size, which can be achieved as a result of growth interruption.

2.6.2. Effect of arsenic pressure on the morphology of a heterophase system

To investigate the stability of the fabricated array of islands to changes in their growth conditions, the authors of Ref. 6 examined the dependence of the morphology of a heterophase system on the arsenic pressure. Deposition of

four monolayers at $T = 480\ ^\circ\mathrm{C}$ and standard arsenic pressure for molecular-beam epitaxy ($P^0_{\mathrm{As}} \approx 2 \times 10^{-6}$ Torr) leads to equilibrium dots with a high concentration per unit area ($5 \times 10^{10}\ \mathrm{cm}^{-2}$). The parameters of the array of dots do not change when the arsenic pressure is varied by ≈50% near this value. Increasing the arsenic pressure by a factor of 3 (to $3 \cdot P^0_{\mathrm{As}}$) leads to dramatic changes in the morphology of the system. The concentration of dots is decreased, and a high concentration of large (≈500–1000 Å) InAs clusters, which contain dislocations, appears.

Decreasing the arsenic pressure leads to a different change in the morphology of the heterophase system. At an arsenic pressure of $(1/6) P^0_{\mathrm{As}}$ three-dimensional islands completely disappear, and flat InAs islands with sizes ≈1000 Å are formed.

To elucidate the character of the phase transition from three-dimensional islands to flat islands, four monolayers of InAs were deposited; the indium feed was then shut off. The method of *in situ* high-energy electron diffraction was used to examine the morphology of the surface. It was found that the given phase transition at low arsenic pressures is reversible, while the transition from coherent three-dimensional islands to large-scale clusters containing dislocations, which takes place when the arsenic pressure is raised, is irreversible.

All changes in the morphology of the heterophase system attendant to varying the arsenic pressure can be explained on the basis of surface energies. It is well known that the stoichiometry of (001) surfaces of III–V semiconductors in equilibrium with the gas phase depends on the vapor pressure of the group-V elements in the gas phase. Thus, varying the As_2 pressure leads to a change in the GaAs(001) surface energy and even to a change in the surface reconstruction.[67] It is natural to expect that the surface energy of the InAs/GaAs(001) wetting layer also depends on the arsenic pressure. An increase in the surface energy of the wetting layer leads to an increase of the parameter α, which controls the dependence of the energy of the heterophase system on island size, and can induce a transition from the regime in which the island size is optimal to the regime in which there is a tendency toward ripening. The experimental data indicate that the tendency toward ripening arises when the arsenic pressure is increased from P^0_{As} to $3 \cdot P^0_{\mathrm{As}}$.

It is well known that at low arsenic pressures indium forms segregation layers on a GaAs surface. Such layers may be considered as a quasiliquid phase,[11–14] whose surface energy can differ strongly from that of crystalline InAs. Such a change in the surface energy alters the critical thickness of the strained layer at which three-dimensional islands form. In Ref. 68 it was experimentally established that growth of InAs on GaAs(001) under conditions of an indium excess takes place pseudomorphically up to a thickness of seven monolayers.

2.7. Arrays of vertically coupled quantum dots

Grown-on islands (quantum dots) of material 2 in a matrix of material 1 are coherent elastic inclusions, which create long-range strain fields over the entire grown-on heterphase system. Upon repeat deposition of material 2 onto the grown-on heterostructure a fundamentally new growth regime arises: growth in a strain field created by grown-on dots of the first layer. Since the bulk diffusion coefficients of the atoms of the main components in III–V semiconductors at MBE temperatures are several orders of magnitude less than the surface diffusion coefficients, bulk diffusion can be ignored. Therefore the distribution of the components of a heterophase system in the volume of the grown-on heterostructure are ''frozen in place.'' This distribution creates a static strain field which defines the kinetics of surface migration. The surface migration consists of diffusion and drift in the strain field.

Such growth kinetics in an open system were investigated theoretically in Refs. 69 and 70 for epitaxial growth of a solid solution. Criteria were established for amplification of composition fluctuations of a solid solution with epitaxial film thickness, which is the initial stage of the process of self-organization of structures with periodic modulation of the composition of the solid solution. It was shown that amplification of the composition fluctuations during growth in an open system arises over a wider temperature interval than spinodal decomposition under equilibrium conditions in a closed system.

If growth is interrupted during deposition of material 2 onto a structure with grown-on dots, then the deposited material will tend to come to partial equilibrium, i.e., to an equilibrium surface morphology with a ''frozen'' distribution of material in the volume of the heterophase structure.

It has been known for some time[44] that in a multilayer system of InAs dots in a GaAs matrix the dots are spatially correlated. However, in those studies the distance between the dots noticeably exceeded the size of the dots themselves. In such a case, the spatial correlation of the dots does not alter their electronic structure.

In order to acquire the possibility of control and optimization of the electronic structure of quantum dots with the aim of applying them in optoelectronics, the technology of fabricating arrays of electronically coupled quantum dots was developed.[6,71–74] With this aim, the alternating deposition of InAs and GaAs was used. The amount of deposited GaAs was chosen such that the InAs pyramid was grown-on only partly. Then, in the following deposition cycle the InAs dots, grown above the first layer of dots, should be electronically coupled with the first layer of dots.

Figure 5 shows images of vertically coupled InAs quantum dots in a GaAs matrix.[6] The images were obtained by transmission electron microscopy in the surface plane (plan view) and in cross section. Growth was realized by MBE with alternating deposition of 5.5 Å of InAs and 15 Å of GaAs. The cross section of the structure shows that a vertically coupled quantum dot consists of three regions of InAs, separated by thin (3–4 monolayers) regions of GaAs.

The dimensions of the upper part of a vertically coupled dot are larger than the dimensions of the lower part and are equal to 170 Å. The plan-view image shows that the dots have a square base, oriented in the [100] and [010] directions. A histogram of the directions between the given dot and its nearest neighbor was constructed in Ref. 73 and dem-

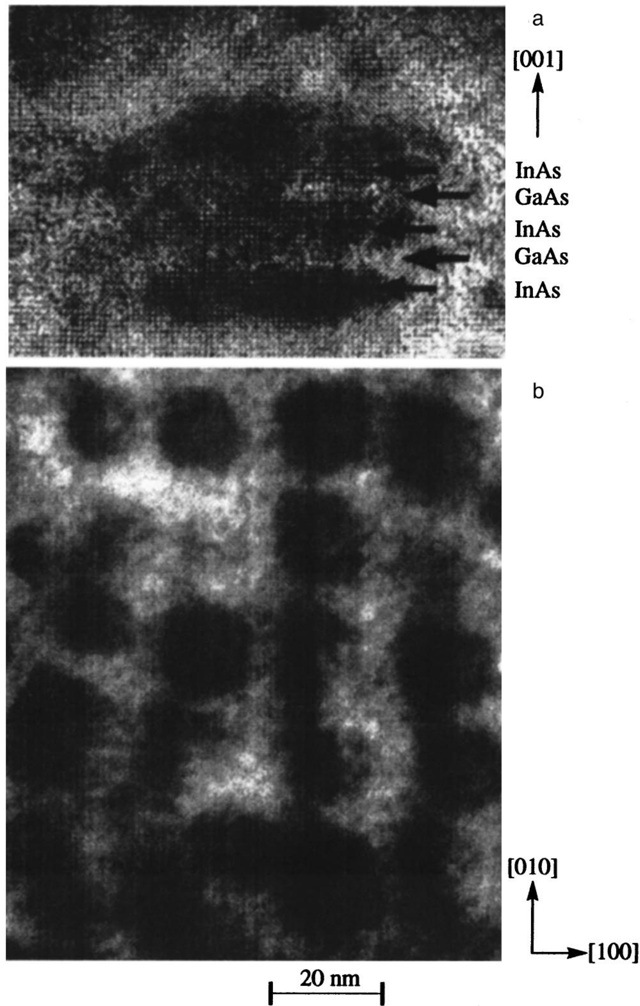


FIG. 5. TEM images of vertically coupled InAs quantum dots in a GaAs matrix: cross-sectional image above, plan-view image below.

onstrates the preferred orientation of a pair of nearest neighbors along the [100] and [010] axes.

Similar results in the growth of vertically coupled dots by gas-phase epitaxy from metallo-organic compounds were obtained in Ref. 75.

3. ELECTRONIC STRUCTURE AND OPTICAL PROPERTIES

3.1. General information

If the size of the semiconducting crystal is decreased to a few tens or few hundreds of atomic spacings, then all the fundamental characteristics of the material are radically altered as a result of quantum-well effects.[76] The limiting case of quantum-well effect is realized in structures with spatial confinement of the charge carriers in all three dimensions. These so-called ''superatoms'' or ''quantum dots'' afford the possibility of the most radical modification of the electron spectrum in comparison with the case of a bulk superconductor. According to the theoretical estimates, devices such as diode lasers using quantum dots as the active medium should possess substantially superior properties in comparison with the currently widely used quantum-well lasers.[77] Such superior properties include a substantially larger material gain, a lower threshold current density, its complete insensitivity to the lattice temperature, better dynamic characteristics and greater possibilities for controlling the radiative recombination photon energy (''color''). Experimental confirmation of the indicated advantages has become possible because of the appearance of quantum dots which satisfy very rigid requirements on their size, shape, uniformity, and density.

3.2. Requirements placed on quantum dots

3.2.1. Minimum size

The lower limit for the size of a quantum dot is defined as the size at which only one electronic level exists in the quantum dot. This critical dimension ($D_{\min}$) strongly depends on the conduction band discontinuity (ΔE_c) in the corresponding heterojunction used to fabricate the quantum dot. In a spherical quantum dot only one electronic level exists in the case in which ΔE_c exceeds the value[78,74]

$$\Delta E_c^* = \frac{\hbar^2}{2m_e^*}\left(\frac{\pi}{D_{\min}}\right)^2 \equiv \Delta E_1^{QW}, \tag{16}$$

where m_e^* is the effective electron mass, and ΔE_1^{QW} is the first level in the rectangular quantum well (QW) with infinite walls and infinite width $D_{\min}$. Assuming the conduction band discontinuity to be of the order of 0.3 eV, which is typical of direct-band quantum wells in the system GaAs–$Al_{0.4}Ga_{0.6}As$, we find that the diameter of a quantum dot should not be less than 40 Å. This is, generally speaking, the absolute lower limit for the size of a quantum dot since for quantum dots even of somewhat greater size the energy distance between the electronic level in the quantum dot and the electronic level in the matrix material will be extremely small, and at finite temperatures the thermal ejection of carriers from the quantum dot can lead to their destruction. For the system InAs–AlGaAs the conduction band discontinuity is substantially greater; however, the electron mass is smaller and thus the quantities $\Delta E_c m_e^*$ are comparable, and the critical sizes of the quantum dots are similar.

3.2.2. Maximum size

If the distance between the energy levels becomes comparable with the thermal energy (kT), then the population of the upper levels grows. For quantum dots the condition for which the population of the higher-lying levels can be disregarded is written as[78,74]

$$kT \leqslant \frac{1}{3}(E_2^{QD} - E_1^{QD}), \tag{17}$$

where E_1^{QD} and E_2^{QD} are the energies of the first and second quantum-size levels, respectively. This means that in the case of a spherical (or cubic) quantum dot, the advantages of size quantization can be completely realized if[74]

$$kT \leqslant E_1^{QW}. \tag{18}$$

This condition places an upper limit on the size of a quantum dot of the order of 120 Å in the system GaAs–AlGaAs and of the order of 200 Å for the system InAs–GaAs, because of

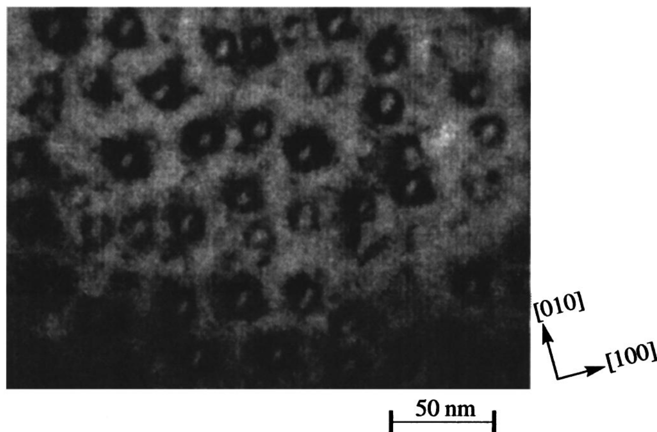


FIG. 6. Plan-view TEM image of InAs quantum dots in a GaAs matrix.

the substantially smaller effective electron mass in the latter case. Effective hole quantization requires even smaller sizes.

3.2.3. Structural perfection, density, and homogeneity

For applications to optoelectronic devices, the quantum dots should not contain dislocations or point defects, and all heterojunctions forming the quantum dot should possess a low surface recombination rate. These conditions make it preferable to use methods of direct fabrication of quantum dots. Dense arrays of quantum dots ($\sim 10^{11}$ cm^{-2}) are required to realize high modal gain in lasers. The unique advantages of quantum-dot structures can be realized only if the quantum dots are as uniform as possible in shape and size. Ordering of quantum dots in the substrate plane and the possibility of creating periodic lattices of quantum dots in all three dimensions are also desirable in a number of cases.

3.3. Formation of ordered quantum-dot arrays and demonstration of the quantum-dot electron spectrum, similar to an atomic electron spectrum

Efforts were undertaken to fabricate quantum dots by local etching or mixing of quantum wells, growth on profiled substrates, condensation in glassy matrices, etc. At the same time, these methods have not made it possible to simultaneously satisfy the requirements of Secs. 3.2.1–3.2.3 and the basic requirements of present-day semiconductor technology (planar surface, the possibility of current injection, etc.) In contrast, the use of the effect of transition to three-dimensional growth, which has been traditionally considered by technologists as extremely undesirable, has led to a breakthrough in the direct fabrication of quantum dots. The possibility of forming three-dimensional islands in a wideband matrix was demonstrated as early as 1985.[44] It did not attract great attention at that time, however, because the prospect of fabricating quantum dots of uniform size by that method raised grave doubts, and the formation of macroscopic, strongly defective clusters was assumed to be unavoidable.[50]

Significant experimental and theoretical studies (described above in Sec. 2) were needed before dense arrays of quantum dots, of high structural perfection and uniform in size and shape coould be realized in practice.[53] Figure 6 presents a plan view (surface view) of a structure with InAs quantum dots in a GaAs matrix. The formation of dense ($\sim 10^{11}$ cm^{-2}) arrays of quantum dots was first demonstrated in this work. The effect of the growth parameters and composition of the quantum dots on their optical properties was investigated and their size and shape were determined. In Refs. 56 and 58 a histogram of directions between a given dot and its nearest neighbor was constructed and spatial correlation of the dots characteristic of lateral ordering of quantum dots in a square lattice with principal axes in the [100] and [010] directions was established.

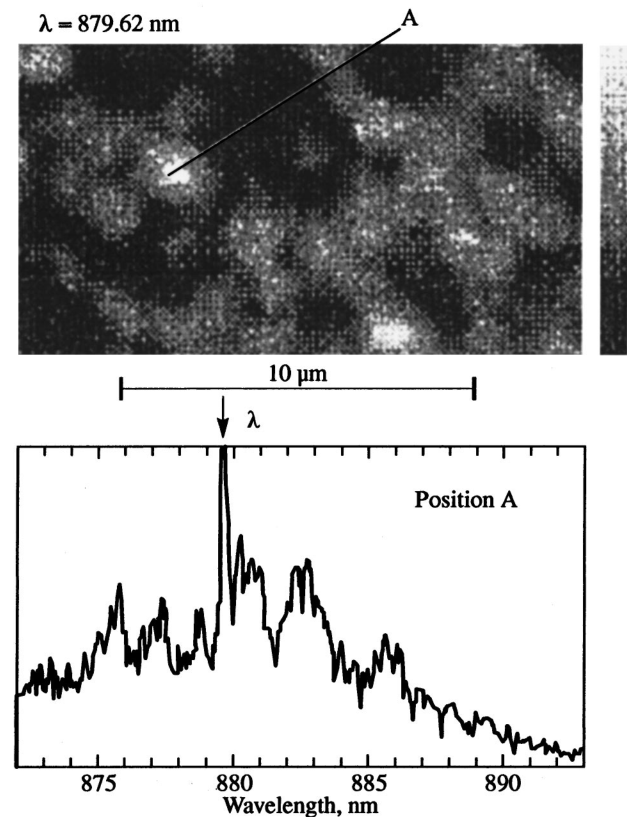


FIG. 7. Low-temperature (5 K) cathodoluminescence (CL) spectra with high spatial resolution from a structure of InAs quantum dots in a GaAs matrix. Above—monochromatic CL intensity distribution, below—CL spectrum for excitation focused at the point *A*.

The cathodoluminescence spectra of quantum dots was investigated with high spectral and lateral resolution in Ref. 53 (see Fig. 7). It was shown there that the quantum dot luminescence peak, broadened due to fluctuations of the shape and size of the islands, decays when excited by a focused electron beam into a set of super-narrow lines from isolated quantum dots.

At the same time, an article was published by the French group on observation of super-narrow lines in nanosize mesas fabricated from quantum-dot structures.[79] Note that the independence of the luminescence linewidth of a quantum dot on the observation temperature is unambiguous proof of the similarity of the electron spectrum of a quantum dot with the electron spectrum of an atom. Such independence was first demonstrated in Ref. 55.

The absorption, photoluminescence, and luminescence excitation spectra of quantum-dot array samples (Fig. 8)

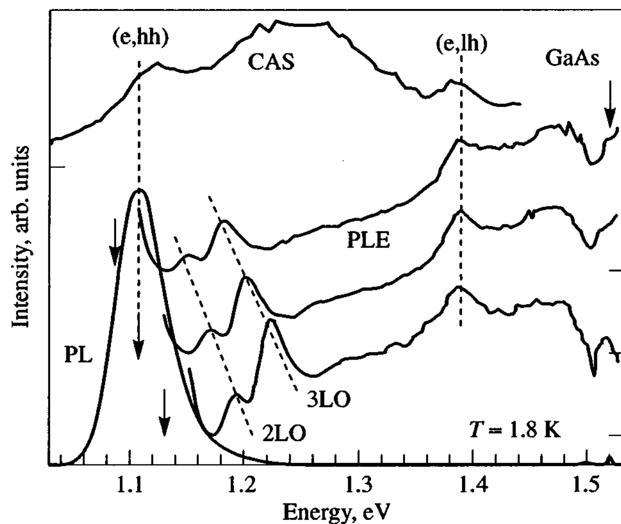


FIG. 8. Calorimetric absorption spectrum (CAS), photoluminescence (PL) spectrum, and photoluminescence excitation (PLE) spectra of InAs quantum-dot structures in a GaAs matrix.

were compared for the first time in Ref. 53. The unique features of the optical properties of quantum-dot structures were also determined for the first time in that study. Thus, it was demonstrated there for the first time that the absorption peak and the luminescence peak of the ground state of a quantum dot coincide. The ground-state peak is absent in the photoluminescence excitation spectra of quantum-dot ensembles. This has to do with the fact that the spectrum of the density of states of a quantum dot is similar to the atomic spectrum and with the absence of transport between neighboring quantum dots. We have noted that resonant excitation of the ground state of a quantum dot is possible only if the photon energy corresponds exactly to the absorption spectrum of the quantum dot. During radiative recombination a quantum dot emits a photon with exactly the same energy, and thus the quantum-dot excitation and luminescence photons can be separated only in experiments with temporal resolution. Dots having the same ground-state energy can have different excited-state energies due to fluctuations in the shape of the quantum dot and the inhomogeneity of the strain fields. Thus, in the photoluminescence excitation spectrum the closest energy to the recording energy is a feature associated with the first exciton excited state of the quantum dot. This feature of the quantum-dot photoluminescence excitation spectrum, unique in comparison with the photoluminescence excitation spectrum of quantum-well structures, makes it easy to determine the presence of quantum dots in a sample by purely optical methods by comparing the photoluminescence excitation and absorption (or optical reflection). The given work also showed for the first time that the quantum-dot excitation spectrum is modulated with an energy corresponding to the LO phonon energy in InAs. This has to do with the fact that relaxation to the ground state of a quantum dot takes place faster if the energy distance between the ground state and the excited state is equal to a multiple of the number of LO phonons, and the probability of nonradiative recombination in the excited state is lower. At the same time, the large spectral width of the modulation and the presence of efficient relaxation for nonresonant excitons gave grounds for the first time to conclude that these relaxation processes take place very rapidly, and that the so-called ''bottleneck effect'' in quantum dots[80] does not play a significant role. Later we showed that the relaxation time to the ground state does not exceed 25–40 ps, depending on the geometry of the quantum dot, and that it can be substantially less than in quantum-well structures.[74] Moreover, the relaxation time can be decreased to 13 ps in structures with coupled quantum dots. On the other hand, it should be noted that in contrast to quantum-well structures, the population time of the ground state in a quantum dot corresponds to the destruction time of an excited state in the quantum dot, and the latter time can thus be substantially longer than the destruction time of the upper subband in quantum-dot structures (0.5–1 ps). This fact underlies the great promise of using quantum-dot structures in radiation sources in the middle and far infrared operating on interlevel transitions.

3.4. Interrelationship between growth regimes and optical properties of quantum dots

3.4.1. Structures fabricated by submonolayer deposition

Altering the conditions of growth, composition, and amount of deposited material has a dramatic effect on the optical properties of quantum dots. The width of the photoluminescence spectrum of an ensemble of quantum dots varies typically in the range 100–10 meV.[6,53,56,58] The effect of the amount of deposited InAs and the growth interruption time after deposition on the concentration and size of the quantum dots and on the luminescence properties of quantum-dot structures was investigated in Ref. 6. It was shown there that submonolayer growth of InAs on GaAs(100) leads to the formation of stretched two-dimensional islands, of one monolayer height, as was first suggested in Ref. 38 and later confirmed in independent scanning electron microscopy studies.[40] The high size uniformity of the islands leads to the appearance of a narrow line in the photoluminescence spectrum on the long-wavelength side of the exciton luminescence line of GaAs. A feature of the luminescence of submonolayer inclusions distinguishing it from the photoluminescence of structures of superthin quantum wells is the absence of line broadening attendant to increases in the pump density, which is a feature of the spectrum of the island density of states.[6]

The optical properties of structures with submonolayer inclusions were investigated in Refs. 81 and 82. Such inclusions can also be interpreted as quantum-dot structures by virtue of the characteristic dimensions of the islands that localize the exciton. It was shown that inclusion of a single InAs submonolayer in a GaAs matrix leads to the appearance in the optical reflection spectrum of a distinctly expressed feature which is resonant with the photoluminescence line from the submonolayer islands. Analysis of the optical reflection spectra shows that the exciton oscillator strength of the submonolayer inclusion ($\tau_{rr} \approx 80$ ps) depends only slightly on the amount of deposited material (0.08–0.3

monolayer),[81] which cannot be explained within the framework of the narrow quantum-well model, but accords with the concept of a quantum-dot array.[83]

Magneto-optical studies of submonolayer structures were carried out by us.[82,84] We showed that the exciton binding energy in a submonolayer structure substantially exceeds the exciton binding energy in a super-narrow quantum well of analogous mean composition.[82] Efficient scattering with spin flip in submonolayer structures was first observed in Ref. 84, and the violation of the momentum selection rules necessary for such a process was interpreted as evidence of efficient localization of the exciton on the islands. The g-factors of heavy and light holes and the corresponding excitons were also determined.

The small exciton localization energy on submonolayer islands associated with their small thickness leads, however, to efficient destruction of localized states with growth of the temperature. To overcome this problem, the concept of submonolayer inclusions was extended to the system CdSe–ZnSe, where the large effective masses of the charge carriers and the large binding energy and small Bohr radius of the exciton make localization of excitons on submonolayer islands substantially more efficient.[85,86]

The large oscillator strength associated with submonolayer inclusions in systems of wideband compounds was first demonstrated in this system. Structures with a set of planes with submonolayer inclusions were also fabricated and their properties investigated. Thus, it was shown that in submonolayer superlattices lasing occurs at energies resonant with the exciton ground state.[85,86] In contrast, in quantum-well structures or in bulk material the laser photon energy is shifted from the exciton transition by the energy of one or even several LO phonons toward longer wavelengths. This is a result of the necessity of satisfying the momentum selection rules (excitons with noticeable momentum in flat quantum dots, dominant at finite substrate temperatures and high excitation densities, cannot recombine radiatively since the resulting photon must possess zero momentum, and an additional particle, e.g., a phonon, is required for momentum transfer[87]). Thus, localization of excitons in a submonolayer leads to a lifting of the momentum selection rule and the laser generation has a resonant character. This fact in combination with the large oscillator strength of the exciton transition motivated us to advance the concept of an ''exciton waveguide effect'' and the idea of a laser based on this effect, in which external optical confinement of the active medium by layers with a lower reflection coefficient is absent, and the waveguide effect is achieved as the result of a resonant modulation of the refractive index, which is associated with exciton absorption in a submonolayer superlattice and described by the Kramers–Kronig relations.[88] At present, using the indicated concept, lasers having a super-narrow buffer layer and super-low CdSe concentration in the active medium, which work without external optical confinement by material with a lower refractive index, even at room temperature, have been realized.[89]

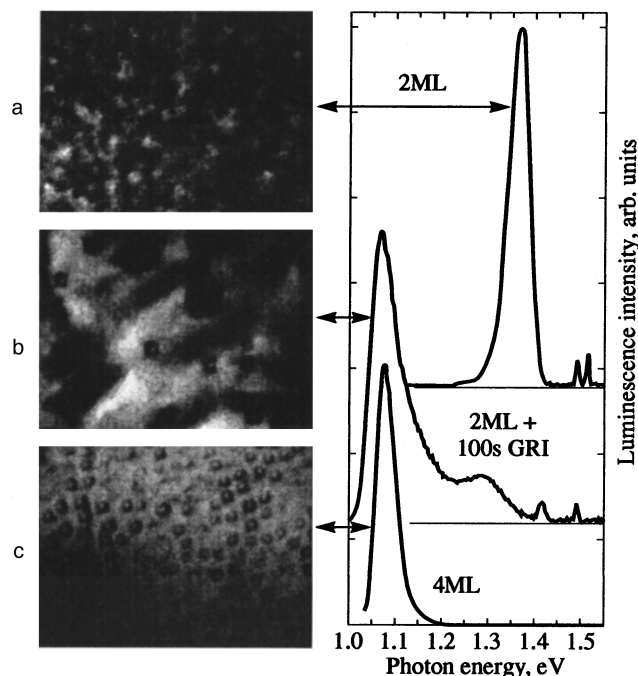


FIG. 9. Effect of interruption of growth on the structure and optical spectra of quantum dots. Plan-view TEM images and photoluminescence spectra of InAs quantum-dot structures in a GaAs matrix. All structures were fabricated by molecular-beam epitaxy at a temperature of 480 °C and an arsenic pressure of 2×10^{-6} Torr. a—deposition of two monolayers of InAs, b—deposition of four monolayers of InAs without interruption of growth; c—deposition of two monolayers of InAs in the submonolayer epitaxy regime (0.3-monolayer InAs cycles interrupted by 100-s intervals).

3.4.2. Three-dimensional islands: Effect of the amount of deposited material, arsenic pressure, and growth interruptions on the luminescence of quantum dots

The effect of growth interruption on the photoluminescence of structures with three-dimensional InAs–GaAs quantum dots was investigated in Ref. 6. It was shown there that increasing the amount of deposited InAs all the way to four monolayers leads to a shift of the quantum-dot photoluminescence line toward longer wavelengths, which is in line with the increase of the characteristic size of the quantum dots observed in TEM studies. Increasing the mean thickness of the InAs layer to more than four monolayers did not lead to further shift of the line or to an increase in the size of the quantum dots; however, mesoscopic clusters containing dislocations appeared in the structure, and the photoluminescence intensity of these structures was lower. These results confirm the presence of a ''limiting'' or ''equilibrium'' size of the quantum dots, which follows from TEM data.[6] Increasing the growth interruption time after deposition of the quantum dots leads to the result that for the same mean thickness of the InAs layer (two monolayers) the photoluminescence peak is shifted toward longer wavelengths and the size of the quantum dots reaches the indicated ''equilibrium'' value (see Fig. 9).

The effect of arsenic pressure on the optical properties and size of the quantum dots was also investigated in Ref. 6. It was shown that a substantial increase in the arsenic flux during deposition leads to a degradation of the luminescence

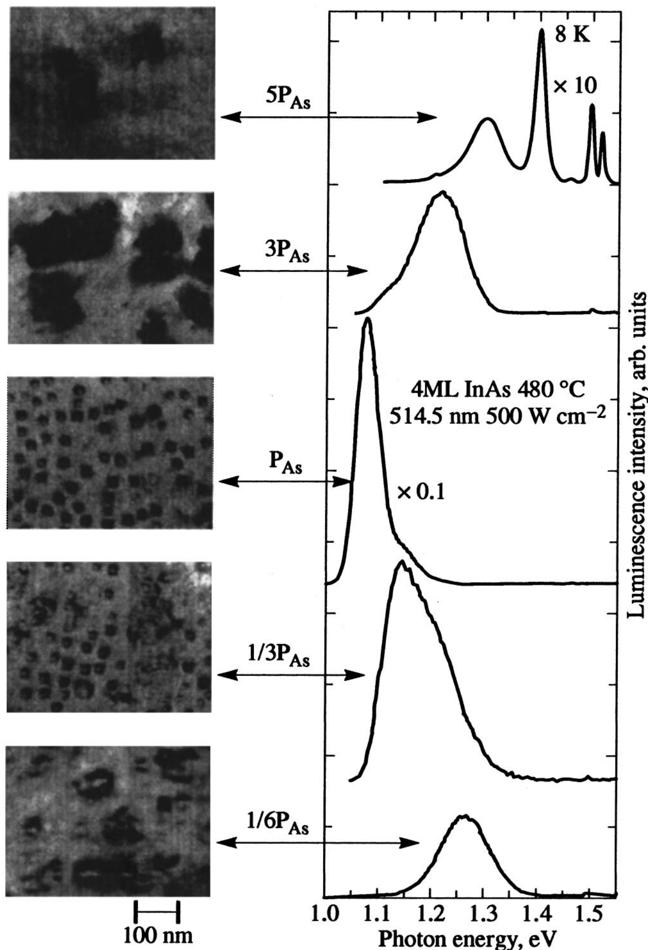


FIG. 10. Effect of arsenic pressure on the structure and optical spectra of quantum dots. Plan-view TEM images and photoluminescence spectra of InAs quantum-dot structures in a GaAs matrix. All structures were fabricated by molecular-beam epitaxy, the amount of deposited InAs was four monolayers; P_{As}—standard arsenic pressure 2×10^{-6} Torr, at which coherently strained three-dimensional islands of InAs (quantum dots) are formed. Lowering the arsenic pressure suppresses the formation of quantum dots and leads to a two-dimensional structure with a corrugated surface. Raising the arsenic pressure leads to ripening of the three-dimensional islands and formation of large InAs clusters containing dislocations.

properties and a decrease in the characteristic size of the quantum dots. TEM data indicate the appearance in the structure of mesoscopic InAs clusters that contain dislocations. At large arsenic pressures only the InAs wetting layer line ($\sim$1.4 eV, 4 K) is present in the photoluminescence spectra. In contrast, a substantial decrease in the arsenic flux leads to suppression of the formation of three-dimensional islands and the InAs distribution has a quasi-two-dimensional character with locally formed mesoscopic bulges (Fig. 10). Formation of dislocations in this case does not take place; however, the photoluminescence intensity in this case is also substantially lower while the photoluminescence line width is substantially greater than in the case of quantum dots formed at optimum arsenic pressure. Thus, the growth regimes leading to formation of ordered quantum-dot arrays are also optimal to obtain the best luminescence properties of the structure.

3.5. Geometry and spectrum of electron states of quantum dots

Under optimal deposition conditions at 460 °C pyramidal InAs quantum dots with principal axes oriented in the directions {001} are formed on the surface. These dots have a square-base ($\sim$120 Å) and height $\sim$60 Å (Ref. 90). Increasing the substrate temperature to 480 and 520 °C leads to an increase in the lateral size of the islands to 140 and 180 Å, respectively.[6] In this case the height of the quantum dots is decreased. A detailed analysis of the structural properties of the quantum dots was carried out in the original studies.[57,91,92] Thus, on the basis of molecular dynamics calculations it was shown that an adequate high-resolution cross-section TEM image of a quantum dot can be obtained only under optimal defocusing conditions, for film thickness not greater than two side lengths of the quantum dot, and under the condition that the quantum dot is completely encompassed by the film.[91,92] In the opposite case, a pyramidal quantum dot appears truncated and its real size can be either reduced or increased by the strain fields induced in the matrix by the quantum dots. TEM images (not high-resolution) of pyramidal quantum dots always have the shape of a ''lens'' or a ''truncated pyramid,'' regardless of the imaging conditions.[92]

Detailed theoretical calculations of the strain fields and the electron structure of pyramidal quantum dots allowing for piezoelectric and exciton effects were performed by us.[93–95] A comparison of these calculations with the experimental data showed that one electron level and several hole levels exist in a pyramidal quantum dot with side length 12 nm. Only $|3/2,3/2\rangle$ hole states contribute to the spectrum of localized hole states.

A cross-sectional TEM image of an InAs–GaAs quantum dot (left) and a plan-view TEM image of the sample (right) are shown in Fig. 11 (upper part), taken from Ref. 90. The lower part of Fig. 11 shows the ground-state wave function and excited-state wave functions of a hole in a pyramidal quantum dot with side length 12 nm calculated in Refs. 93–95.

Figure 12 (taken from Ref. 6) is a plot of the electron ground-state energy and the hole ground-state energy and excited-state energies as a function of the base length of the pyramid. The hole ground state has an $\sim$88% overlap with the electron ground state. In contrast to a cubic or spherical quantum dot, the wave functions of the excited hole states $|001\rangle$ ($|002\rangle$) also have a significant overlap with the electron ground state of the order of 34% (11%), thereby forming an allowed transition. The small finite overlap of the wave functions of the electron ground state $|001\rangle$ with the hole states $|010\rangle$ (0.5%) and $|110\rangle$ (2.5%) is due to piezoelectric effects on the lateral edges of pyramids of type {112}. The exciton binding energy in a quantum dot is close to 20 meV, i.e., it is greater by an order of magnitude than the exciton binding energy in bulk InAs.

Photoluminescence spectra of InAs quantum dots with base length 120 Å in a GaAs matrix are shown in Fig. 13. The band of the exciton ground state in the quantum dot, which has a full-width at half-maximum (FWHM)

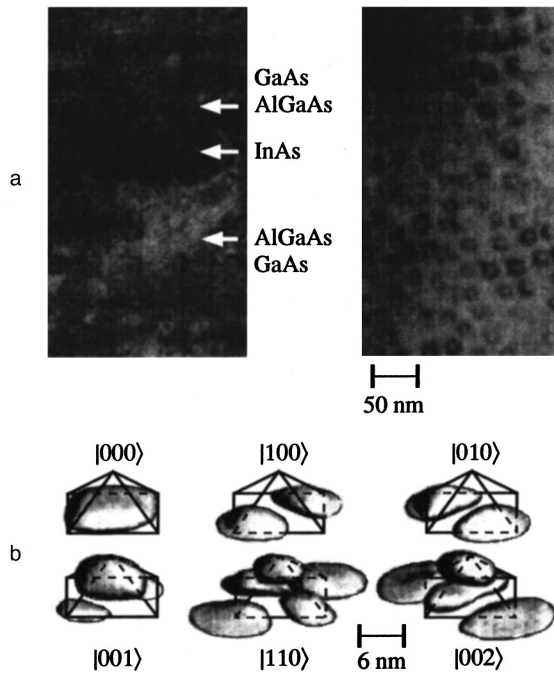


FIG. 11. TEM images of InAs quantum dots in a GaAs matrix: cross-sectional image (upper left), plan-view image (upper right). Wave functions of the holes in these quantum dots are shown below.

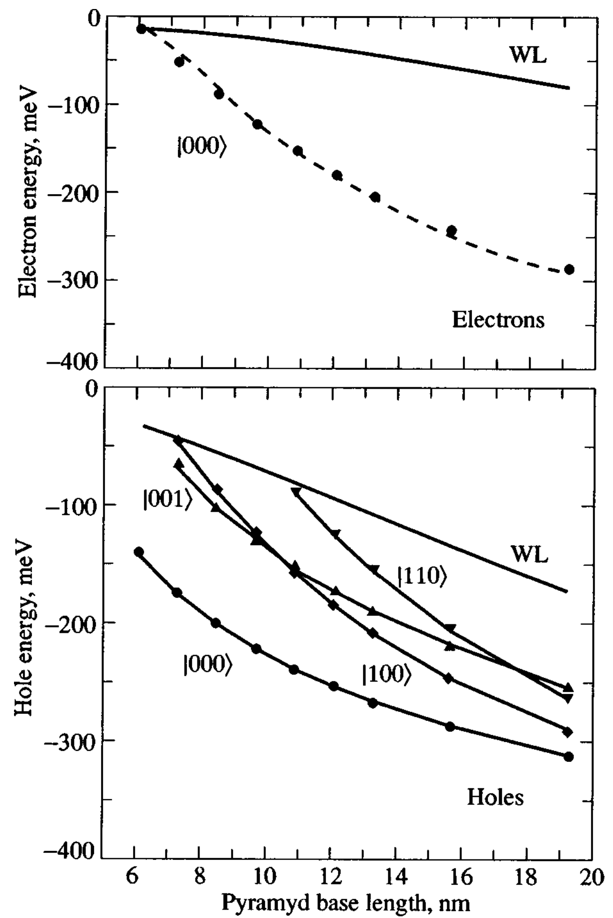


FIG. 12. Electron ground-state energy, hole ground-state energy and hole excited-state energies in an InAs quantum dot having the shape of a pyramid with square base in the (001) plane and ⟨101⟩ lateral faces and located in a GaAs matrix. Energies are plotted as a function of the base length of the pyramid.

$\sim$40–50 meV, dominates the spectra. The photoluminescence intensity of this band saturates at large excitation densities, and lines associated with the excited hole states $|001\rangle$ and $|002\rangle$ appear in the spectrum. All three peaks and the feature associated with the transition due to the InAs wetting layer are also observed in the calorimetric absorption spectrum. The calculated positions of the corresponding transitions are indicated by arrows. The width of the ground-state peak (40 meV) corresponds roughly to variation of the quantum-dot side length by 1 nm, as follows from Fig. 12. This fact indicates that the mean fluctuation of the characteristic quantum-dot side length does not exceed two monolayers.

Excited states of quantum dots appear only at large excitation energies, when the ground state of the quantum dot is saturated. This indicates that relaxation of an exciton to its ground state takes place very rapidly. The capture time in a quantum dot was estimated in Ref. 90 to be less than or approximately equal to 1 ps, and the interlevel relaxation time was estimated to be in the range 25–40 ps (Ref. 74) (depending on the presence or absence of resonance with a multiple number of *LO* phonons, respectively). The radiative recombination time of an exciton in the ground state is roughly 1.5 ns and depends weakly on the size of the quantum dot.[6,74]

3.6. Phonon spectrum and resonant luminescence of quantum dots

The photoluminescence spectra of a quantum-dot structure with resonant excitation with a photon energy of 1.165 eV, to an exciton excited state in the quantum dot are shown in Fig. 14.[6,96] For comparison, the figure also shows the pho-

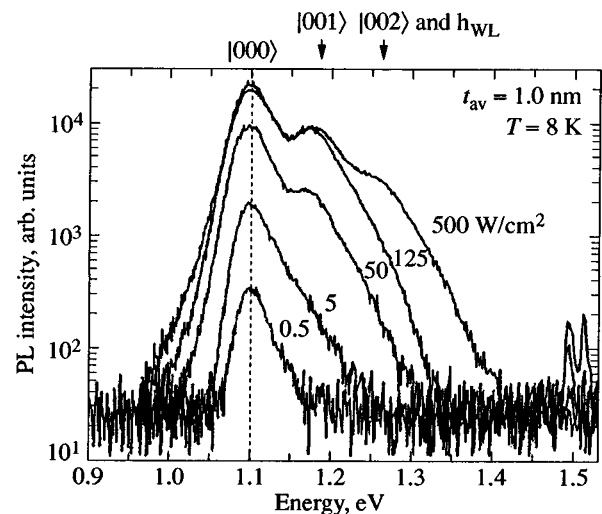


FIG. 13. Photoluminescence spectra of InAs quantum dots in a GaAs matrix, plotted as functions of the excitation power. At low power levels the exciton ground-state band dominates. At high power levels the intensity of this band saturates and lines appear in the spectrum. These lines are associated with excited states of the exciton with participation of a heavy hole.

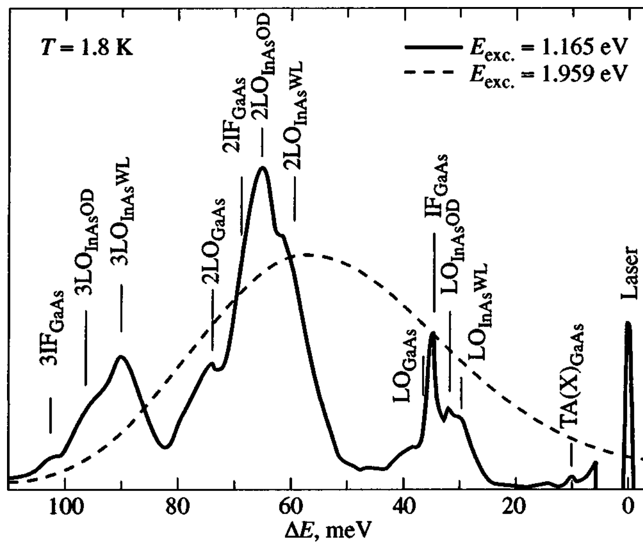


FIG. 14. Photoluminescence spectra of InAs quantum dots in a GaAs matrix for resonant (solid line) and nonresonant (dashed line) excitation. The notation ΔE has been introduced for the quantity 1165 meV$-E$, where E is the energy of the emission in meV.

toluminescence spectrum for nonresonant excitation with photon energy greater than the width of the band gap of the GaAs matrix. It can be seen from the figure that for resonant excitation the photoluminescence spectrum divides into a series of relatively narrow lines, where the spectral distribution of the line intensities corresponds roughly to the photoluminescence spectrum for nonresonant excitation. The energies of the observed peaks agree with a multiple number of GaAs- and InAs-induced *LO* phonons and their combinations, including the *LO* phonons of bulk GaAs (36.6 meV), the interface modes (34 meV), phonons of the InAs wetting layer (29.5 meV), and phonons of the InAs quantum dots (31.9 meV). If the energy of an exciting photon grows, resonant photoluminescence lines follow behind the energy of the exciting photon, broaden, and as a final result the spectrum becomes similar to the photoluminescence spectrum for nonresonant excitation.

An analogous structure is revealed in the quantum-dot photoluminescence excitation spectrum, where the quantum-dot phonons play the dominant role, which accords with the fact that both the ground state and the excited state of the quantum dot are localized mainly inside the quantum dot, and only relatively small wave-function tails protrude into the GaAs matrix. It should also be noted that the energy of the InAs *LO* phonon in a quantum dot (31.9 meV) exceeds the energy of an *LO* phonon in bulk InAs (29.9 meV). This fact is in agreement with theoretical calculations of the strain distributions in coherent InAs quantum dots and the quantum-dot phonon energy calculated with these strains taken into account (32.1 meV). As was already mentioned, modulation of the photoluminescence spectra and photoluminescence excitation spectra with a frequency corresponding to a whole number of *LO* phonons reflects the more efficient relaxation of carriers in the quantum-dot ground state in the case of the corresponding resonance[53] and, consequently, the lesser role of nonradiative recombination in the excited state.

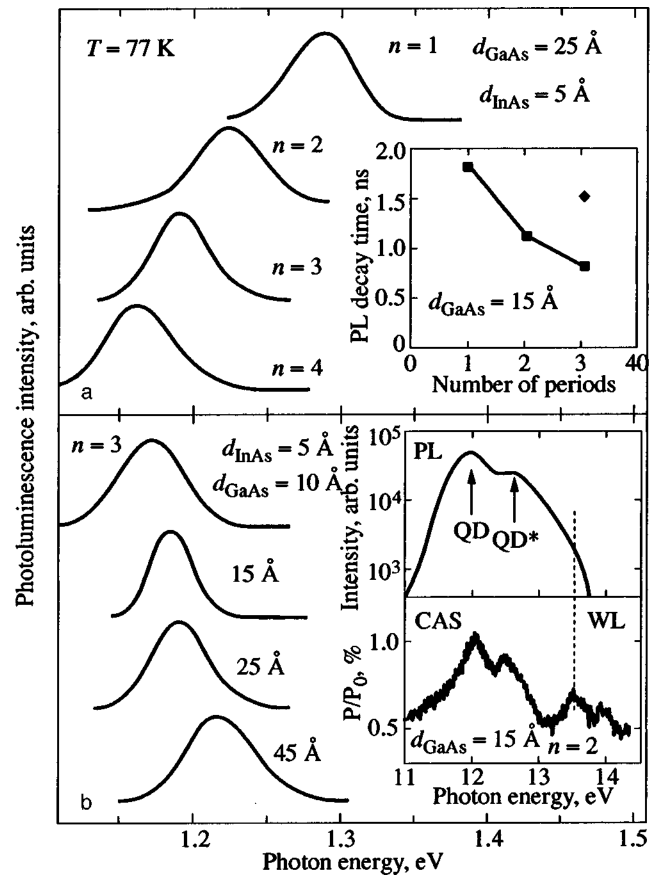


FIG. 15. Photoluminescence spectra of vertically coupled InAs quantum dots (VCQD) with different number of deposition cycles (n) and different thickness of the GaAs spacer. a—spectra for identical InAs and GaAs deposition thicknesses and different number of deposition cycles (n). Photoluminescence decay time as a function of n shown in the inset on the upper right. The rhombus point symbol shows the decay time for $d_{GaAs}=45$ Å. b—photoluminescence spectra for the same number of deposition cycles ($n=3$), the same InAs deposition thickness and different thicknesses of the GaAs spacer. The photoluminescence spectrum of the structure with $d_{GaAs}=15$ Å and two deposition cycles ($n=2$) at excitation power 500 W/cm^2 and the calorimetric absorption spectrum of the structure are shown in the inset on the lower right.

The large width of the phonon lines testifies to the considerable role of multiphonon relaxation mechanisms with participation of acoustic phonons.[6,53,96]

3.7. Luminescence of vertically coupled quantum dots

The luminescence properties of vertically coupled quantum dots was investigated in Refs. 71–75. Increasing the number of deposition cycles (n) led to a long-wavelength shift of the photoluminescence line and the associated resonance peak in the calorimetric absorption spectrum. The effect was most pronounced for small nominal thicknesses of the GaAs spacer (Fig. 15). Increasing the thickness d_{GaAs} for the same number of cycles led to a short-wavelength shift of the quantum-dot line. It was also shown that increasing the number of deposition cycles shortens the lifetime of radiative recombination in the quantum dots. Thus, formation of a new quantum-mechanical object, specifically, tunnel-coupled quantum dots possessing properties substantially different from those of isolated quantum dots, was demonstrated. Le-

dentsov *et al.*[75] used gas-phase epitaxy from metallo-organic compounds (GPE MOC) to fabricate for the first time vertically coupled quantum dots in Ref. 75.

3.8. Luminescence of type-II quantum dots

Type-II quantum dots were first grown by Hatami *et al.*[97] in the system GaSb–GaAs. Previously it had been shown[98,99] that in quantum wells grown in this system the holes are localized in the GaSb regions, whereas for the electrons the GaSb regions present a potential barrier. We grew GaSb–GaAs quantum dots by depositing a GaSb layer with a mean thickness on the order of 3.5 monolayers. Formation of quantum dots led to the appearance, in addition to the wetting-layer luminescence line, of a line shifted toward the longer-wavelength region of the spectrum. A TEM study of the samples enabled us to determine the characteristic size of the GaSb quantum dots, which was $\sim$20–30 nm. In contrast to InAs quantum dots on GaAs substrates, the GaSb dots had a rectangular instead of square base.[97] In a system of type-II quantum dots only the holes are quantized in the GaSb layer, while the electrons are spatially separated from the holes by a potential barrier in the conduction band and are held in place near the holes only by the Coulomb interaction. Increasing the excitation density increases the positive charge of the quantum dots and causes the dipole moment between the localized holes and the electron cloud around the quantum dot to grow. It also leads to a short-wavelength shift of the quantum-dot photoluminescence line.[97]

4. QUANTUM-DOT INJECTION LASERS

4.1. Main advantages of quantum-dot lasers

The advantages of a quantum-dot laser in comparison with a quantum-well laser can be arbitrarily divided into physical and technological. The physical advantages are due mainly to the δ-like spectrum of the density of states and the giant oscillator strength of the optical transitions per unit volume of the quantum dot, which in turn is caused by the effective overlap of the electron and hole wave functions due to their spatial localization. Such advantages include a superhigh temperature stability of the threshold current density,[77,100] giant maximum material gain and maximum differential gain, two to three orders of magnitude larger than the analogous values for quantum-well lasers.[101–104] The advantages of quantum-dot lasers also include a short population time of the ground state and, consequently, high operating frequencies. The technological advantages include the absence or suppression of diffusion of nonequilibrium carriers, which leads to reduced leakage of nonequilibrium carriers from the stripe region, suppression of nonradiative recombination at point and extended defects and, correspondingly, suppression of the effect of dislocation growth, and also suppression of the effect of overheating of the mirrors due to surface recombination. In addition, an ordered quantum-dot array, located in an optical waveguide, can lead to distributed feedback and single-mode generation. In the case of vertically emitting lasers, a laser on one quantum dot can actually be created. This would make it possible to avoid inhomogeneous broadening, which is characteristic of an ensemble of quantum dots, and completely realize the advantages of three-dimensional quantization. The operating characteristics of quantum-dot lasers fabricated by various methods were examined in Refs. 100–113.

4.2. Characteristics of quantum-dot lasers

The characteristics of a laser containing a single plane array of InGaAs quantum dots in the active region were investigated in Ref. 100. At low temperatures lasing begins at energies near the maximum of the photoluminescence peak, indicating that transitions through the ground state of the quantum dots are responsible for lasing. An increase in the temperature has virtually no effect on the threshold current density, which maintaining its value of 80 A/cm^2 up to temperatures on the order of 180 K. If the temperature dependence of the threshold current density is approximated by a dependence of the form $J=J_0\exp(-T/T_0)$, then in this temperature range $T_0=380$ K, which is higher than the theoretical limit for quantum-well lasers. The lasing wavelength in this case is found near the maximum of the peak of the photoluminescence and electroluminescence for a weak level of excitation. As the temperature is raised above 180 K, the threshold current density begins to grow, coinciding with a decrease in the total photoluminescence intensity with activation energy $\sim$80–90 meV. This value is in good agreement with the localization energy of the holes in the quantum wells and indicates that the reason for the growth of the threshold current density is insufficient gain, which is associated with thermal ejection of carriers from the quantum dots. The effect of gain saturation is accompanied by a shift of the lasing wavelength toward shorter wavelengths, which corresponds to the region of emission of the excited states of the quantum dots and the InGaAs wetting layer.

Thus, quantum-dot injection lasers exhibit low threshold current densities and record temperature stability at low temperatures, in agreement with theoretical predictions. However, the insufficient localization energy of the carriers leads to a strong temperature dependence of the threshold current density at temperatures near room temperature.

4.3. Amplification in quantum-dot lasers

The possibility of realizing a quantum-dot injection laser depends to a substantial degree on the relationship between the gain of the optical emission from the quantum dots and the optical losses in the structure. In the case of a quantum-well injection laser, whose typical width is $\sim$100 Å, the optical confinement factor (proportional to the total overlap between the electron wave function and the light wave) is of the order of 0.03. If the well width is decreased to 10–30 Å, this factor is not significantly decreased due to penetration of the wave function into the barriers. Note that the gain in quantum-well lasers is small, and low threshold current densities can be realized only with long cavities, when it is possible to substantially decrease the effect of losses on the output intensity, or in four-cleavage samples.[114–116] In the case of a quantum-dot array the ground-state wave function is completely localized inside the quantum dot. Even in the

case of a dense array of islands only about two InAs monolayers are converted into quantum dots, which corresponds to ~6 Å average exciton volume, averaged over the surface. As a result, the optical confinement factor is small.[6,101] Nevertheless, it has turned out to be possible to realize a quantum-dot injection laser as a result of the giant growth of the "specific gain," in agreement with theoretical predictions.

The specific gain was determined directly from the condition that the gain equal the losses at lasing threshold.[72,101] The internal losses were determined from the dependence of the threshold current density on the cavity length and the dependence of the differential efficiency on the output losses. The optical confinement factor was estimated from available electron-microscope data on the geometrical dimensions of the dots. The value of the maximum "specific gain" obtained in this way is 1.5×10^{-5} cm^{-1} (Ref. 101), which exceeds by more than an order of magnitude the value for quantum-well lasers. As a consequence of the linear relation between the gain and the current, the differential gain in quantum-dot lasers also grows by more than three orders of magnitude[101] in relation to quantum-well lasers and reaches values on the order of 10^{-12} cm^2.

Growth of the gain is a direct consequence of size quantization in the quantum dots. The latter decreases the number of states that need to be filled to achieve a given gain. The main factors lowering the gain for a given current density are thermal ejection of carriers from the quantum dots and leakage through nonradiative recombination in the barrier material.

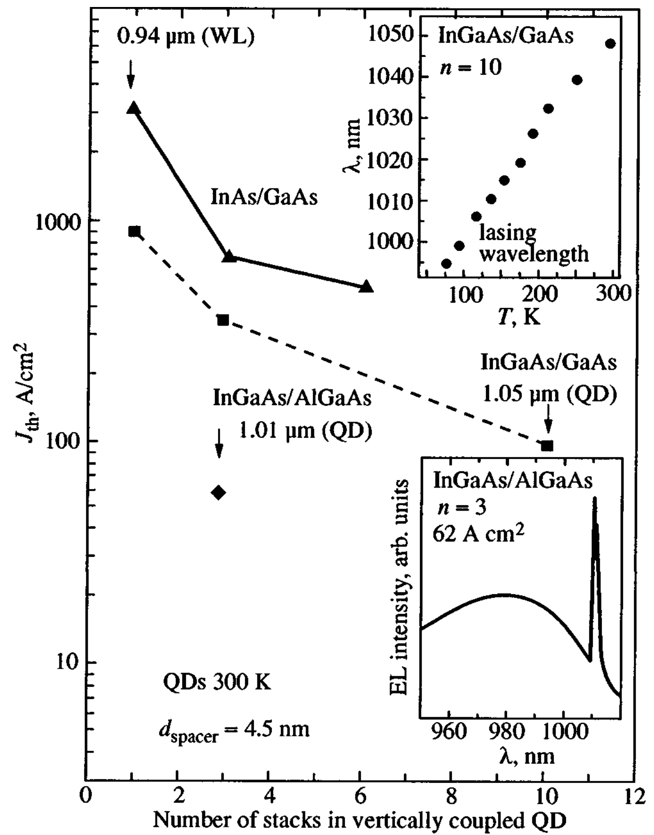


FIG. 16. Threshold current density as a function of the number of deposition cycles of InAs or InGaAs for a vertically coupled quantum-dot injection laser. The lasing wavelength is plotted versus temperature in the inset on the upper right. The emission spectrum is shown in the inset on the lower right.

4.4. Lasers on vertically coupled quantum dots

Degradation of the characteristics of quantum-dot lasers at temperatures above 150–180 K is due to insufficient gain. To increase the gain it is necessary either to increase the homogeneity of the quantum-dot array, which in principle can be done by optimization of the growth regimes, or to increase the dot concentration, which can be done by using vertically coupled quantum dots (VCQD).[71–74] The advantages of vertically coupled quantum dots, besides a large optical confinement factor, are the possibility of faster relaxation of the carriers to the ground state,[102] a shorter radiative recombination time,[73] and the possibility of efficient tunneling of electrons and holes between the dots on neighboring rows, which is absent in the case of rows of isolated quantum dots.

It has been shown experimentally that vertically coupled quantum-dot lasers exhibit much larger optical gain, and gain saturation is absent in them down to short cavity lengths.[71–74] They lase through the ground state of the quantum dots up to room temperature, and the lasing wavelength tracks the temperature dependence of the width of the GaAs band gap. It was also shown that the threshold current density decreases dramatically to values of the order of 90 A/cm^2 (300 K) as the number of stacking cycles is increased to 10 (Fig. 16).[72–74]

This effect is due to an increase in the gain as a result of growth of the optical confinement factor. The differential efficiency also grows with increase of N, reaching 50% at $N=10$. Such rather low values of the differential efficiency are due to the low value of the internal quantum efficiency (0.5), which means that a large fraction of the carriers recombine nonradiatively, most probably in the GaAs layer covering the quantum dot, which is deposited at the low temperature of 480 °C.

4.5. Effect of matrix material on the characteristics of quantum-dot injection lasers

As was already noted, one of the main mechanisms leading to an increase in the threshold current density at elevated temperatures in quantum-dot lasers is thermal ejection of carriers from the quantum-dot states to the states of the wetting layer and the matrix. Further lowering of the threshold current density in lasers was achieved by using AlGaAs instead of GaAs as the matrix material for the vertically coupled quantum dots.[73,74] In this case the barrier height is increased while the wave function of the carriers in the quantum dots undergoes insignificant changes, which leads to an increase in the energy gap between the quantum-dot levels and the levels of the wetting layer and barrier. As a result, a further increase in the gain occurs due to a suppression of the thermal ejection of carriers from the quantum dots due to an increase in the localization energy. The problem of the quality of AlGaAs layers deposited at low temperature during formation of the vertically coupled quantum dots is solved

by efficient annealing of point defects during growth of the upper emitter of the laser structure at 700 °C. Thus, the use of AlGaAs spacers in combination with growth of emitters of the laser structure at 700 °C has made it possible to obtain a room-temperature threshold current density of 60 A/cm^2 (Fig. 16)[73] and a differential efficiency greater than 70%. This improvement of characteristics has made it possible to realize the regime of continuous operation of quantum-dot lasers at room temperature with an output power of 1 W or greater.[117] It has been shown that the regime of lasing through the quantum-dot states is preserved up to pump currents 7–8 times threshold. This result shows that high-power semiconductor lasers based on quantum dots can be built with characteristics at least as good as those of quantum-well lasers.

It was found that in the low-temperature region (77–150 K) in low-threshold lasers based on vertically coupled quantum dots a decrease of the threshold current density is observed with increase of the observation temperature; i.e., the laser has a negative characteristic temperature T_0. Such behavior cannot be explained within the framework of a quasi-equilibrium carrier distribution described by the Fermi function. Zhukov *et al.*[118] proposed a model in which the absence of a quasi-equilibrium population of quantum-dot states at low temperatures is taken as the reason for the appearance of a segment of negative T_0.

In fact, the condition for the setting up of a quasi-equilibrium distribution is that the thermal ejection time of the carriers from the quantum dots be short in comparison with the radiative recombination time in the quantum dots. Carrier ejection is strongly suppressed at low temperatures. Thus, there exists a boundary temperature (T_b) which separates the regions of equilibrium ($T>T_b$) and nonequilibrium ($T<T_b$) population of the quantum-dot states. In the latter case, the population of the quantum-dot states is determined by the probability of electron (hole) capture, which is approximately identical for holes of different sizes. Consequently, the nonequilibrium distribution is characterized by a wider spectrum of populated states and a lower maximum gain in comparison with the Fermi distribution. Thus, when the observation temperature passes through the boundary temperature T_b, the threshold current density and the half-width of the luminescence line should decrease. This is the reason for the observed negative characteristic temperature in quantum-dot lasers at low temperatures.

The boundary temperature T_b increases as the localization energy of the carriers in the quantum dots increases when a wider-band material is used as the matrix. This explains the fact that the region of negative T_0 is more distinct when vertically coupled quantum dots in a AlGaAs matrix are used in comparison with GaAs.

4.6. Dynamics of quantum-dot lasers

Good dynamic characteristics of quantum-dot lasers follow from short relaxation times of the carriers to the ground state and large differential gains.[101,119] Direct measurements of the cutoff frequency of quantum-dot lasers give a value on the order of 10 GHz (Ref. 102).

Another important aspect characterizing the operation of a laser at high frequencies is the spectral broadening factor of the laser line. Any absorption or amplification peak gives rise to a modulation of the refractive index near the energy of the corresponding resonance according to the Kramers–Kronig relations. Thus, the photon wavelength in the crystal can vary during the pump current pulse. This effect is described by the spectral line broadening factor (α). In the case of quantum-well structures the shape of the absorption or gain spectrum is strongly asymmetrical, which accounts for the large value of α (between 1 and 2). In contrast, in the case of quantum-dot structures the absorption or gain spectrum is more symmetrical and has a Gaussian shape. Thus, the energy derivative and, consequently, the variation of the refractive index in the region of an absorption or amplification maximum is equal to zero. The experimentally measured values of α lie in the range $\sim$0.5. This is attributable to the finite contribution of the excited states of the quantum dots to the gain spectrum near the lasing threshold, which gives rise to some asymmetry in the gain profile.[102]

4.7. Gain mechanisms

The condition of transparency in a quantum dot is realized when the quantum dot encompasses one exciton. In this case the probabilities of emitting or absorbing one photon with formation of a biexciton are equal.[20] It should be noted, however, that, in general, the energies of the exciton and biexciton states in a quantum dot are different, and population of the quantum dot by one exciton can lead simultaneously to the appearance of an exciton amplification line and a biexciton absorption line. If the inhomogeneous broadening of the lines is less than the energy difference between the exciton and biexciton energies in a quantum dot, then a purely exciton gain mechanism is possible. The importance of the contribution of charged excitons to the gain spectrum should also be noted.[120]

If the possibility of carrier transport between neighboring quantum dots is absent (which is typical at low temperatures), then the probability of exciton and carrier capture in a quantum dot is temperature independent.[121] At high temperatures thermal ejection of carriers from shallower quantum dots can lead to preferential population of the deeper quantum dots.[118] The behavior of the gain spectra is different in these two cases: In the first case the gain maximum does not change its position with increase of the pump current; in the second case it is shifted toward higher energies.[120]

Effects in quantum-dot structures and lasers based on them were considered theoretically in Refs. 119–124.

4.8. Vertically emitting quantum-dot lasers

Surface emitting quantum-dot lasers operating with optical excitation at low temperatures through the quantum-dot ground state were constructed by Schur *et al.*[125] Injection lasers in vertically stacked, isolated quantum dots were made by Saito *et al.*[126] The threshold current density at room temperature was $\sim$500 A/cm^2, and lasing took place through quantum-dot excited states. Using structures with a 7-μm aperture in aluminum oxide to reduce the current injection

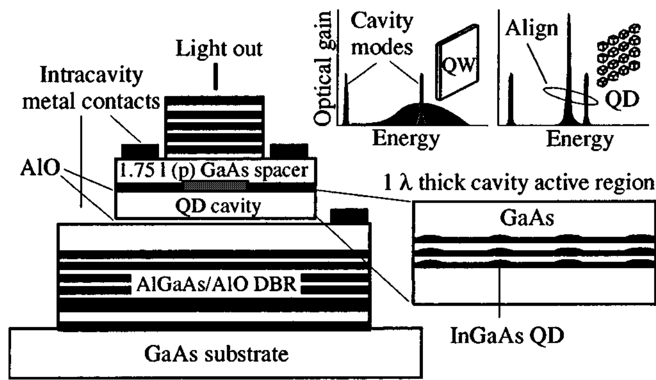


FIG. 17. Diagram of a surface-emitting laser based on InGaAs quantum dots in a GaAs matrix. For comparison, the inset on the upper right shows the optical gain spectra for a quantum-well structure and a quantum-dot structure. The inset on the lower right shows a cross section of the active region.

regions, the regime of lasing through the quantum-dot ground state was realized in Ref. 127 at room temperature (300 K), a current density of ~ 1000 A/cm^2, and threshold current of 0.5 mA. The concept of vertically coupled quantum dots as the active medium was proposed in Refs. 128 and 129 to enhance the operating characteristics of vertical quantum-dot lasers. Such an approach opens the door to the creation of a laser on a single quantum dot, and would allow one to simultaneously increase the spatial confinement factor and preserve the very small width of the quantum-dot amplification line. In structures with a 10-μm mesa (Fig. 17) the regime of continuous lasing at room temperature with a threshold current of 180 μA (180 A/cm^2, 300 K) was realized.[129] Characteristics of the device are shown in Fig. 18. For the case of an optimum number of layers of the upper Bragg reflector the maximum efficiency was 16% and the minimum threshold current was 68 μA for a 1-μm mesa, which corresponds to the best values for vertical quantum-well lasers of such geometry.

CONCLUSION

Laser generation through states of self-organized quantum dots was first observed in 1993 by Ledentsov *et al.*[130] In the years since, impressive progress has been achieved in fabricating high-density arrays of quantum dots of uniform

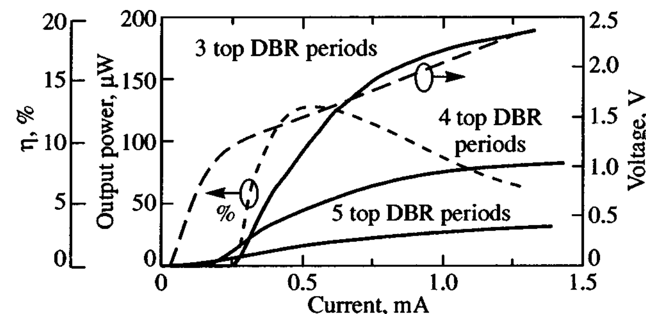


FIG. 18. Output power and quantum yield of a laser based on vertically coupled InGaAs quantum dots in a GaAs matrix plotted as a function of the injection current. Solid lines—output power; short-dashed line—quantum yield, long-dashed lines—current–voltage characteristic.

size, shape, and spacing and suitable for use as the active region of injection heterolasers, and in the study of the properties of quantum-dot lasers. Further progress in the area of quantum-dot heterostructures is tied up with an extension of their range of applications in various devices of micro- and opto-electronics and with fabricating yet more uniform quantum dots, which will make possible a qualitative improvement in the operating characteristics of the majority of present-day devices.

Original works of the authors, mentioned in this review, were supported by the Russian Fund for Fundamental Research, the Volkswagen Foundation, INTAS, the Soros Foundation, and the German and Russian Ministries of Science.

We express our deep gratitude to these organizations. One of the authors (N. N. Ledentsov) also expresses his gratitude to the Alexander von Humboldt Foundation.

[1] Zh. I. Alferov, in *Proceedings of Nobel Symposium 99*, Arild, Sweden, June 4–8, 1996 [Physica Scripta **68,** 32 (1996)]; Fiz. Tech. Poluprovodn. **32**, 1 (1998) [Semiconductors **32**, 3 (1998).

[2] C. M. Sotomayor Torres, F. D. Wang, N. N. Ledentsov, and Y.-S. Tang, in *Proc. SPIE—The International Society for Optical Engineering* (Bellingham, Washington, USA 1994), Vol. 2141, p. 2.

[3] *Optical properties of Low Dimensional Semiconductors*, edited by G. Abstreiter, A. Aydinli, and J.-P. Leburton, NATO ASI Series. Series E: Applied Sciences (Kluwer Academic Publishers, Dordrecht, The Netherlands, 1997), Vol. 344.

[4] A. I. Ekimov and A. A. Onushchenko, JETP Lett. **34**, 345 (1981).

[5] H. Haken, *Synergetics* (Springer, Berlin–Heidelberg, 1997).

[6] N. N. Ledentsov, M. Grundmann, N. Kirstaedter, O. Schmidt, R. Heitz, J. Böhrer, D. Bimberg, V. M. Ustinov, V. A. Shchukin, P. S. Kop'ev, Zh. I. Alferov, S. S. Ruvimov, A. O. Kosogov, P. Werner, U. Richter, U. Gösele, and J. Heydenreich, in *Proc. 7th Int'l. Conf. Modulated Semicond. Struct.*, Madrid, Spain, July, 10–14, 1995 [Solid-State Electron. **40,** 785 (1996)].

[7] V. A. Shchukin, in *Proc. 23rd Int'l. Conf. Phys. Semicond.*, July 22–26, 1996, Berlin, Germany, edited by M. Scheffler, R. Zimmermann (World Scientific, Singapore), Vol. 2, p. 1261.

[8] V. A. Shchukin, N. N. Ledentsov, M. Grundmann, and D. Bimberg, in *Optical Properties of Low Dimensional Semiconductors*, edited by G. Abstreiter, A. Aydinli, and J.-P. Leburton, NATO ASI Series. Series E: Applied Sciences (Kluwer Academic Publishers, Dordrecht, The Netherlands, 1997), Vol. 344, p. 257.

[9] D. Bimberg, I. P. Ipatova, N. N. Ledentsov, P. S. Kop'ev, V. G. Malyshkin, and V. A. Shchukin, Usp. Fiz. Nauk **167**, 552 (1997).

[10] P. S. Kop'ev and N. N. Ledentsov, Fiz. Tekh. Poluprovodn. **22**, 1729 (1988) [Sov. Phys. Semicond. **22**, 1093 (1988)].

[11] S. V. Ivanov, P. S. Kop'ev, and N. N. Ledentsov, J. Cryst. Growth **104**, 345 (1990).

[12] S. V. Ivanov, P. S. Kop'ev, and N. N. Ledentsov, in *Proc. 6th Int'l. Conf. on MBE*, San Diego, USA, 1990 [J. Cryst. Growth **111,** 151 (1991)].

[13] S. V. Ivanov, P. S. Kop'ev, and N. N. Ledentsov, J. Cryst. Growth **108**, 661 (1991).

[14] S. V. Ivanov, P. D. Altukhov, T. S. Argunova, A. A. Bakun, V. V. Chaldyshev, Yu. A. Kovalenko, P. S. Kop'ev, R. N. Kyutt, B. Ya. Meltser, S. S. Ruvimov, S. V. Shaposhnikov, L. M. Sorokin, and V. M. Ustinov, Semicond. Sci. Technol. **8**, 347 (1993).

[15] J. W. Cahn, Trans. Metall. Soc. AIME **242**, 166 (1968).

[16] A. G. Khachaturyan, *Theory of Phase Transitions and Structure of Solid Solutions* [in Russian] (Nauka, Moscow, 1974).

[17] G. B. Stringfellow, J. Cryst. Growth **65**, 454 (1983).

[18] I. P. Ipatova, V. G. Malyshkin, and V. A. Shchukin, J. Appl. Phys. **74**, 7198 (1993).

[19] I. P. Ipatova, V. G. Malyshkin, and V. A. Shchukin, Philos. Mag. B **70**, 557 (1994).

[20] C. Herring, Phys. Rev. **82**, 87 (1951).

[21]V. I. Marchenko and A. Ya. Parshin, Zh. Éksp. Teor. Fiz. **79**, 257 (1980) [Sov. Phys. JETP **52**, 129 (1980)].
[22]A. F. Andreev and Yu. A. Kosevich, Zh. Éksp. Teor. Fiz. **81**, 1435 (1981) [Sov. Phys. JETP **54**, 761 (1981)].
[23]A. F. Andreev, Zh. Éksp. Teor. Fiz. **80**, 2042 (1981) [Sov. Phys. JETP **53**, 1063 (1981)].
[24]V. I. Marchenko, Zh. Éksp. Teor. Fiz. **81**, 1141 (1981) [Sov. Phys. JETP **54**, 605 (1981)].
[25]V. A. Shchukin, A. I. Borovkov, N. N. Ledentsov, and P. S. Kop'ev, Phys. Rev. B **51**, 17 767 (1995).
[26]M. Kasu and N. Kobayashi, Appl. Phys. Lett. **62**, 1262 (1993).
[27]N. N. Ledentsov, G. M. Gur'yanov, G. E. Tsirlin, V. N. Petrov, Yu. B. Samsonenko, A. O. Golubok, S. Ya. Tipisev, Fiz. Tekh. Poluprovodn. **28**, 903 (1994) [Semiconductors **28**, 526 (1994)].
[28]R. Nötzel, N. N. Ledentsov, L. Däweritz, M. Hohenstein, and K. Ploog, Phys. Rev. Lett. **67**, 3812 (1991).
[29]M. Higashiwaki, M. Yamamoto, T. Higuchi, S. Shimomura, A. Adachi, Y. Okamoto, N. Sano, and S. Hiyamizu, Jpn. J. Appl. Phys. **35**, L606 (1996).
[30]M. Kasu and N. Kobayashi, Appl. Phys. Lett. **62**, 1262 (1993).
[31]R. Nötzel, N. N. Ledentsov, L. Däweritz, M. Hohenstein, and K. Ploog, Phys. Rev. B **45**, 3507 (1992).
[32]Zh. I. Alfërov, A. Yu. Egorov, A. E. Zhukov, S. V. Ivanov, P. S. Kop'ev, N. N. Ledentsov, B. Ya. Mel'tser, and V. M. Ustinov, Fiz. Tekh. Poluprovodn. **26**, 1715 (1992) [Sov. Phys. Semicond. **26**, 959 (1992)].
[33]V. A. Shchukin, A. I. Borovkov, N. N. Ledentsov, and D. Bimberg, Phys. Rev. B **51**, 10 104 (1995).
[34]V. A. Shchukin and A. I. Borovkov, in *Abstracts Int'l. Symp. ''Nanostructures: Physics and Technology''*, St. Petersburg, Russia, June 20–24, 1994, p. 227.
[35]A. F. Andreev, Pis'ma Zh. Éksp. Teor. Fiz. **32**, 2654 (1980) [*sic*].
[36]V. I. Marchenko, JETP Lett. **33**, 381 (1981).
[37]Zh. I. Alf'erov, D. Bimberg, A. Yu. Egorov, A. E. Zhukov, P. S. Kop'ev, N. N. Ledentsov, S. S. Ruvimov, V. M. Ustinov, and J. Heydenreich, Usp. Fiz. Nauk **165**, 224 (1995) [*sic*].
[38]P. D. Wang, N. N. Ledentsov, C. M. Sotomayor Torres, P. S. Kop'ev, and V. M. Ustinov, Appl. Phys. Lett. **64**, 1526 (1994).
[39]P. D. Wang, N. N. Ledentsov, C. M. Sotomayor Torres, P. S. Kop'ev, and V. M. Ustinov, Appl. Phys. Lett. **66**, 112 (1995).
[40]V. Bressler-Hill, A. Lorke, S. Varma, P. M. Petroff, K. Pond, and W. H. Weinberg, Phys. Rev. B **50**, 8479 (1994).
[41]N. N. Ledentsov, P. D. Wang, M. Sotomayor Torres, A. Yu. Egorov, M. V. Maximov, V. M. Ustinov, A. E. Zhukov, and P. S. Kop'ev, Phys. Rev. B **50**, 12171 (1994).
[42]C. Ratsch, A. Zangwill, P. Šmilauer, and D. D. Vvedensky, Phys. Rev. Lett. **72**, 3194 (1994).
[43]G. M. Guryanov, G. E. Cirlin, A. O. Golubok, S. Ta. Tipissev, N. N. Ledentsov, V. A. Shchukin, M. Grundmann, D. Bimberg, and Zh. I. Alferov, Surf. Sci. **352–354**, 646 (1996).
[44]L. Goldstein, F. Glas, J. Y. Marzin, M. N. Charasse, and G. Le Roux, Appl. Phys. Lett. **47**, 1099 (1985).
[45]D. E. Eaglesham and M. Cerullo, Phys. Rev. Lett. **64**, 1943 (1990).
[46]J.-W. Mo, D. E. Savage, B. S. Swartzentruber, and M. G. Lagally, Phys. Rev. Lett. **65**, 1020 (1990).
[47]D. Vanderbilt and L. K. Wickham, Mater. Res. Soc. Symp. Proc. **202**, 555 (1991).
[48]C. Ratsch and A. Zangwill, Surf. Sci. **293**, 123 (1993).
[49]I. M. Lifshitz and V. V. Slezov, Zh. Éksp. Teor. Fiz. **35**, 479 (1958) [Sov. Phys. JETP **8**, 331 (1959)].
[50]F. Glas, C. Guille, P. Hénoc and F. Houzay, Inst. Phys. Conf. Ser. **87**, 71 (1987).
[51]J. M. Moison, F. Houzay, F. Barthe, L. Leprince, E. André, and O. Vatel, Appl. Phys. Lett. **64**, 196 (1994).
[52]D. Leonard, M. Krishnamurthy, C. M. Reaves, S. P. Denbaars, and P. M. Petroff, Appl. Phys. Lett. **63**, 3203 (1993).
[53]N. N. Ledentsov, M. Grundmann, N. Kirsraedter, J. Christen, R. Heitz, J. Böhrer, F. Heinrichsdorf, D. Bimberg, S. S. Ruvimov, P. Werner, U. Richter, U. Gösele, J. Heydenreich, V. M. Ustinov, A. Yu. Egorov, M. V. Maximov, P. S. Kop'ev, and Zh. I. Alferov, in *Proc. 22nd Int'l. Conf. Phys. Semicond.*, Vancouver, Canada, August 1994, edited by D. J. Lockwood (World Scientific, Singapore), Vol. 3, p. 1855.
[54]G. Cirlin, G. M. Guryanov, A. O. Golubok, S. Ya. Tipissev, N. N. Ledentsov, P. S. Kop'ev, M. Grundmann, and D. Bimberg, Appl. Phys. Lett. **67**, 97 (1995).
[55]M. Grundmann, J. Christen, N. N. Ledentsov, J. Böhrer, D. Bimberg, S. S. Ruvimov, P. Werner, U. Richter, U. Gösele, J. Heydenreich, V. M. Ustinov, A. Yu. Egorov, A. E. Zhukov, P. S. Kop'ev, and Zh. I. Alferov, Phys. Rev. Lett. **74**, 4043 (1995).
[56]D. Bimberg, M. Grundmann, N. N. Ledentsov, S. S. Ruvimov, P. Werner, U. Richter, J. Heydenreich, V. M. Ustinov, P. S. Kop'ev, and Zh. I. Alferov, Thin Solid Films **267**, 32 (1995).
[57]S. S. Ruvimov, P. Werner, K. Scheerschmidt, U. Richter, U. Gösele, J. Heydenreich, N. N. Ledentsov, M. Grundmann, D. Bimberg, V. M. Ustinov, A. Yu. Egorov, P. S. Kop'ev, and Zh. I. Alferov, Phys. Rev. B **51**, 14766 (1995).
[58]M. Grundmann, N. N. Ledentsov, R. Heitz, L. Eckey, J. Christen, J. Bḧrer, D. Bimberg, S. S. Ruvimov, P. Werner, U. Richter, U. Gösele, J. Heydenreich, J. Heydenreich, V. M. Ustinov, A. Yu. Egorov, A. E. Zhukov, P. S. Kop'ev, and Zh. I. Alferov, Phys. Status Solidi B **188**, 249 (1995).
[59]N. N. Ledentsov, V. M. Ustinov, S. V. Ivanov, B. Ya. Mel'tser, M. V. Maksimov, P. S. Kop'ev, D. Bimberg, Zh. I. Alf'erov, Usp. Fiz. Nauk **166**, 423 (1996) [*sic*].
[60]V. A. Shchukin, N. N. Ledentsov, P. S. Kop'ev, and D. Bimberg, Phys. Rev. Lett. **75**, 2968 (1995).
[61]V. A. Shchukin, N. N. Ledentsov, P. S. Kop'ev, and D. Bimberg, in *Abstracts Int'l. Symp. ''Nanostructures: Physics and Technology''*, St. Petersburg, Russia, June 26–30, 1995, p. 392.
[62]V. A. Shchukin, A. I. Borovkov, N. N. Ledentsov, P. S. Kop'ev, M. Grundmann, and D. Bimberg, Phys. Low-Dimens. Semicond. Struct. **12**, 43 (1995).
[63]V. A. Shchukin, N. N. Ledentsov, P. S. Kop'ev, and D. Bimberg, in *Proc. Int'l. Semicond. Dev. Res. Symp.*, December 1995, Charlottesville, Virginia, USA, Vol. 2, p. 581.
[64]V. A. Shchukin, N. N. Ledentsov, M. Grundmann, P. S. Kop'ev, and D. Bimberg, Surf. Sci. **352–354**, 117 (1996).
[65]V. A. Shchukin, N. N. Ledentsov, V. G. Malyshkin, I. P. Ipatova, P. S. Kop'ev, D. Bimberg, in *Abstracts Int'l. Symp. ''Nanostructures: Physics and Technology''*, St. Petersburg, Russia, June 24–28, 1996, p. 439.
[66]J. D. Eshelby, Proc. R. Soc. London, Ser. A **241**, 376 (1957).
[67]G. X. Qian, R. M. Martin, and D. J. Chadi, Phys. Rev. B **38**, 7649 (1988).
[68]E. Tournié, A. Trampert, and K. Ploog, Europhys. Lett. **25**, 663 (1994).
[69]V. G. Malyshkin and V. A. Shchukin, Fiz. Tekh. Poluprovodn. **27**, 1932 (1993) [Semiconductors **27**, 1062 (1993)].
[70]V. A. Shchukin, in *Abstracts Int'l. Symp. ''Nanostructures: Physics and Technology''*, St. Petersburg, Russia, June 24–28, 1996, p. 175.
[71]Zh. I. Alferov, N. A. Bert, A. Yu. Egorov, A. E. Zhukov, P. S. Kop'ev, I. L. Krestnikov, N. N. Ledentsov, A. V. Lunev, M. V. Maksimov, A. V. Sakharov, V. M. Ustinov, A. F. Tsatsul'nikov, Yu. M. Shernyakov, D. Bimberg, Fiz. Tekh. Poluprovodn. **30**, 351 (1996) [Semiconductors **30**, 346 (1996)].
[72]N. N. Ledentsov, J. Böhrer, D. Bimberg, S. V. Zaitsev, V. M. Ustinov, A. Yu. Egorov, A. E. Zhukov, M. V. Maximov, P. S. Kop'ev, Zh. I. Alferov, O. A. Kosogov, U. Gösele, and S. S. Ruvimov, Mater. Res. Soc. Symp. Proc. **421**, 133 (1996).
[73]N. N. Ledentsov, V. A. Shchukin, M. Grundmann, N. Kirstaedter, J. Böhrer, O. Schmidt, D. Bimberg, S. V. Zaitsev, V. M. Ustinov, A. E. Zhukov, P. S. Kop'ev, Zh. I. Alferov, O. A. Kosogov, S. S. Ruvimov, P. Werner, U. Gösele, and J. Heydenreich, Phys. Rev. B **54**, 8743 (1996).
[74]N. N. Ledentsov, in *Proc. 23rd Int'l. Conf. on the Physics of Semiconductors*, Berlin, Germany, July 21–26, 1996, edited by M. Scheffler and R. Zimmermann (World Scientific, Singapore, 1996), Vol. 1, p. 19.
[75]N. N. Ledentsov, D. Bimberg, I. V. Kochnev, M. V. Maximov, P. S. Kop'ev, Zh. I. Alferov, O. A. Kosogov, S. S. Ruvimov, P. Werner, and U. Gösele, Appl. Phys. Lett. **69**, 1095 (1996).
[76]V. V. Lutskii, Phys. Status Solidi A **1**, 199 (1970).
[77]Y. Arakawa and H. Sakaki, Appl. Phys. Lett. **40**, 939 (1982); M. Asada, M. Miyamoto, and Y. Suematsu, IEEE J. Quantum Electron. **QE-22**, 1915 (1986).
[78]K. J. Vahala, IEEE J. Quantum Electron. **QE-24**, 523 (1988).
[79]J.-Y. Marzin, J. M. Gérard, A. Izraël, D. Barrier, and G. Bastard, Phys. Rev. Lett. **73**, 716 (1994).
[80]H. Benisty, C. M. Sotomayor Torres, and C. Weisbuch, Phys. Rev. B **44**, 10945 (1991).
[81]M. V. Belousov, N. N. Ledentsov, M. V. Maximov, P. D. Wang, I. N.

Yassievich, N. N. Faleev, I. A. Kozin, V. M. Ustinov, P. S. Kop'ev, and C. M. Sotomayor Torres, Phys. Rev. B **51**, 14346 (1995).

[82] P. D. Wang, N. N. Ledentsov, C. M. Sotomayor Torres, I. N. Yassievich, A. Pakhomov, A. Yu. Egorov, P. S. Kop'ev, and V. M. Ustinov, Phys. Rev. B **50**, 1604 (1994).

[83] E. L. Ivchenko, A. V. Kavokin, V. P. Kochereshko, P. S. Kop'ev, and N. N. Ledentsov, Superlattices Microstruct. **12**, 317 (1992).

[84] A. A. Sirenko, T. Ruf, N. N. Ledentsov, A. Yu. Egorov, P. S. Kop'ev, V. M. Ustinov, and A. E. Zhukov, Solid State Commun. **97**, 169 (1996).

[85] N. N. Ledentsov, I. L. Krestnikov, M. V. Maximov, S. V. Ivanov, S. L. Sorokin, P. S. Kop'ev, Zh. I. Alferov, D. Bimberg, and C. M. Sotomayor Torres, Appl. Phys. Lett. **69**, 1343 (1996).

[86] I. L. Krestnikov, M. V. Maximov, S. V. Ivanov, S. L. Sorokin, S. A. Permogorov, A. N. Reznitsky, A. V. Kornievski, N. N. Ledentsov, D. Bimberg, and C. M. Sotomayor Torres, in *Proc. 23rd Int'l. Conf. on the Physics of Semiconductors*, Berlin, Germany, July 21–26, 1996, edited by M. Scheffler and R. Zimmermann (World Scientific, Singapore, 1996), Vol. 4, p. 3187.

[87] E. Gross, S. Permogorov, and A. Razbirin, J. Phys. Chem. Solids **27**, 1647 (1966).

[88] I. L. Krestnikov, N. N. Ledentsov, M. V. Maksimov, A. V. Sakharov, S. V. Ivanov, S. V. Sorokin, L. N. Tenishev, P. S. Kop'ev, and Zh. I. Alfërov, Pis'ma Tekh. Fiz. **23**, 33 (1997) [Tech. Phys. Lett. **23**, 23 (1997)].

[89] A. V. Sakharov, S. V. Ivanov, S. V. Sorokin, B. V. Volovik, I. L. Krestnikov, P. S. Kop'ev, and N. N. Ledentsov, Pis'ma Zh. Tekh. Fiz. **23**, No. 4. 26 (1997) [Tech. Phys. Lett. **23**, 305 (1997)].

[90] N. N. Ledentsov, M. V. Maximov, P. S. Kop'ev, V. M. Ustinov, M. V. Belousov, B. Ya. Meltser, S. V. Ivanov, V. A. Shchukin, Zh. I. Alferov, M. Grundmann, D. Bemberg, S. S. Ruvimov, U. Richter, P. Werner, U. Gösele, U. Heidenreich, P. D. Wang, and C. M. Sotomayor Torres, Microelectron. J. **26**, 871 (1995).

[91] S. S. Ruvimov, P. Werner, K. Scheerschmidt, U. Richter, U. Gösele, J. Heydenreich, N. N. Ledentsov, M. Grundmann, D. Bimberg, V. M. Ustinov, A. Yu. Egorov, P. S. Kop'ev, and Zh. I. Alferov, Inst. Phys. Conf. Ser. **146**, 31 (1995).

[92] S. Ruvimov, Z. Liliental-Weber, N. N. Ledentsov, M. Grundmann, D. Bimberg, V. M. Ustinov, A. Yu. Egorov, P. S. Kop'ev, Zh. I. Alferov, K. Scheerschmidt, and U. Gösele, Mater. Res. Soc. Symp. Proc. **421**, 383 (1996).

[93] M. Grundmann, N. N. Ledentsov, O. Stier, D. Bimberg, V. M. Ustinov, P. S. Kop'ev, and Zh. I. Alferov, Appl. Phys. Lett. **68**, 979 (1996).

[94] M. Grundmann, R. Heitz, N. N. Ledentsov, O. Stier, D. Bimberg, V. M. Ustinov, P. S. Kop'ev, Zh. I. Alferov, S. S. Ruvimov, P. Werner, U. Gösele, and J. Heydenreich, Superlattices Microstruct. **19**, 81 (1996).

[95] M. Grundmann, N. N. Ledentsov, O. Stier, D. Bimberg, V. M. Ustinov, P. S. Kop'ev, and Zh. I. Alferov, Phys. Rev. B **53**, 10 509 (1996).

[96] R. Heitz, M. Grundmann, N. N. Ledentsov, L. Eckey, M. Veit, D. Bimberg, V. M. Ustinov, A. Yu. Egorov, A. E. Zhukov, P. S. Kop'ev, and Zh. I. Alferov, Appl. Phys. Lett. **68**, 361 (1996).

[97] F. Hatami, N. N. Ledentsov, M. Grundmann, F. Heinrichsdorff, D. Bimberg, S. S. Ruvimov, P. Werner, U. Gösele, J. Heydenreich, U. Richter, S. V. Ivanov, B. Ya. Meltser, P. S. Kop'ev, and Zh. I. Alferov, Appl. Phys. Lett. **67**, 656 (1995).

[98] N. N. Ledentsov, J. Böhrer, M. Beer, M. Grundmann, F. Heinrichsdorff, D. Bimberg, S. V. Ivanov, B. Ya. Meltser, I. N. Yassievich, N. N. Faleev, P. S. Kop'ev, and Zh. I. Alferov, in *Proc. 22nd Int'l. Conf. Physics of Semicond.*, Vancouver, Canada, 1994, edited by D. J. Lockwood (World Scientific, Singapore, 1995), Vol. 2, p. 1617.

[99] N. N. Ledentsov, J. Böhrer, M. Beer, F. Heinrichsdorff, M. Grundmann, D. Bimberg, S. V. Ivanov, B. Ya. Meltser, I. N. Yassievich, N. N. Faleev, P. S. Kop'ev, and Zh. I. Alferov, Phys. Rev. B **52**, 14 058 (1995).

[100] N. Kirstaedter, N. N. Ledentsov, M. Grundmann, D. Bimberg, V. M. Ustinov, S. S. Ruvimov, M. V. Maximov, P. S. Kop'ev, Zh. I. Alferov, U. Richter, P. Werner, U. Gösele, and J. Heydenreich, Electron. Lett. **30**, 1416 (1994).

[101] N. Kirstaedter, O. G. Schmidt, N. N. Ledentsov, D. Bimberg, V. M. Ustinov, A. Yu. Egorov, A. E. Zhukov, M. V. Maximov, P. S. Kop'ev, and Zh. I. Alferov, Appl. Phys. Lett. **69**, 1226 (1996).

[102] D. Bimberg, N. Kirstaedter, N. N. Ledentsov, Zh. I. Alferov, P. S. Kop'ev, and V. M. Ustinov, IEEE J. Quantum Electron. **3**, 196 (1997).

[103] D. Bimberg, N. Kirstaedter, N. N. Ledentsov, Zh. I. Alferov, P. S. Kop'ev, V. M. Ustinov, S. V. Zaitsev, and M. V. Maximov, in *Optical Properties of Low Dimensional Semiconductors*, edited by G. Abstreiter, A. Aydinli, and J.-P. Leburton, NATO ASI Series. Series E: Applied Sciences (Kluwer Academic, Dordrecht, The Netherlands. 1997), Vol. 344, p. 315.

[104] D. Bimberg, M. Grundmann, and N. N. Ledentsov, Spectrum der Wissenschaft **63**, 64 (1996).

[105] V. M. Ustinov, A. E. Zhukov, A. Yu. Egorov, N. N. Ledentsov, M. V. Maximov, A. F. Tsatsul'nikov, P. S. Kop'ev, D. Bimberg, and Zh. I. Alferov, in *Devices Based on Low-Dimensional Semiconductor Structures*, edited by M. Balkanski, NATO ASI Series, High Technology (Kluwer Academic, Dordrecht, The Netherlands, 1996), Vol. 14, p. 91.

[106] D. Bimberg, N. N. Ledentsov, M. Grundmann, N. Kirstaedter, O. Schmidt, M. H. Mao, V. M. Ustinov, A. Yu. Egorov, A. E. Zhukov, P. S. Kop'ev, Zh. I. Alferov, S. S. Ruvimov, U. Gösele, and J. Heydenreich, Jpn. J. Appl. Phys. **35**, Pt. 1, 1311 (1996).

[107] O. G. Schmidt, N. Kirstaedter, N. N. Ledentsov, M. H. Mao, D. Bimberg, V. M. Ustinov, A. Yu. Egorov, A. E. Zhukov, M. V. Maximov, P. S. Kop'ev, and Zh. I. Alferov, Electron. Lett. **32**, 1302 (1996).

[108] V. M. Ustinov, A. Yu. Egorov, A. E. Zhukov, N. N. Ledentsov, M. V. Maximov, A. F. Tsatsul'nikov, N. A. Bert, A. O. Kosogov, P. S. Kop'ev, D. Bimberg, and Zh. I. Alferov, Mater. Res. Soc. Symp. Proc. **417**, 141 (1996).

[109] F. Heinrichsdorff, M.-H. Mao, N. Kirstaedter, A. Krost, D. Bimberg, A. O. Kosogov, and P. Werner, Appl. Phys. Lett. **71**, 22 (1997).

[110] Q. Xie, A. Kalburge, P. Chen, and A. Madhukar, IEEE Photonics Technol. Lett. **8**, 965 (1996).

[111] K. Kammath, P. Bhattacharya, T. Sosnowski, and J. Phillios, Electron. Lett. **30**, 1364 (1997).

[112] A. Moritz, R. Wirth, A. Hangleiter, A. Kurtenbach, and K. Eberl, Appl. Phys. Lett. **69**, 212 (1996).

[113] H. Shoji, K. Mukai, N. Ohtsuka, M. Sugawara, T. Uchida, and H. Ishikawa, IEEE Photonics Technol. Lett. **7**, 1385 (1995).

[114] Zh. I. Alfërov, A. M. Vasil'ev, S. V. Ivanov, P. S. Kop'ev, N. N. Ledentsov, M. É. Lutsenko, B. Ya. Mel'tser, V. M. Ustinov, Pis'ma Zh. Tekh. Fiz. **14**, No. 19, 1803 (1988) [Sov. Tech. Phys. Lett. **14**, 782 (1988)].

[115] S. V. Ivanov, P. S. Kop'ev, N. N. Ledentsov, B. Ya. Meltser, M. Yu. Nadtochi, V. M. Ustinov, and A. M. Vasilev, Phys. Status Solidi A **118**, 169 (1990).

[116] Zh. I. Alfërov, S. V. Ivanov, P. S. Kop'ev, N. N. Ledentsov, M. É. Lutsenko, Yu. Ya. Mel'tsep, M. I. Nemenov, V. M. Ustinov, and S. V. Shaposhnikov, Fiz. Tekh. Poluprovodn. **24**, 152 (1990) [Sov. Phys. Semicond. **24**, 92 (1990)].

[117] Yu. M. Shernyakov, A. Yu. Egorov, A. E. Zhukov, S. V. Zaĭtsev, A. R. Kovsh, I. L. Krestnikov, A. V. Lunev, N. N. Ledentsov, M. V. Maksimov, A. V. Sakharov, V. M. Ustinov, Zhao Zhen', P. S. Kop'ev, Zh. I. Alfërov, and D. Bimberg, Pis'ma Zh. Tekh. Fiz. **23**, No. 1, 51 (1997) [Tech. Phys. Lett. **23**, 70 (1997)].

[118] A. E. Zhukov, V. M. Ustinov, A. Yu. Egorov, A. R. Kovsh, A. F. Tsatsul'nikov, N. N. Ledentsov, S. V. Zaitsev, N. Yu. Gordeev, P. S. Kop'ev, and Zh. I. Alferov, in *Proc. Int'l. Symp. on the Formation, Physics, and Device Application of Quantum Dot Structures*, November 4–7, 1996, Hokkaido University Conference Hall, Sapporo, Japan [Jpn. J. Appl. Phys. **36**, Pt. 1, 4216 (1997)].

[119] Y. Arakawa and A. Yariv, IEEE J. Quantum Electron. **QE22**, 1887 (1986).

[120] M. Grundmann and D. Bimberg, Jpn. J. Appl. Phys. **36**, Pt. 1, 4181 (1997).

[121] M. Grundmann and D. Bimberg, Phys. Rev. B **55**, 9740 (1997).

[122] L. V. Asryan and R. A. Suris, Semicond. Sci. Technol. **11**, 554 (1996).

[123] L. V. Asryan and R. A. Suris, IEEE J. Sel. Top. Quantum Electron. **3**, 148 (1997).

[124] R. A. Suris, in *Future Trends in Microelectronics*, edited by S. Luryi, J. Xu, and A. Zaslavsky (Kluwer Academic, Dordrecht, The Netherlands, 1996), p. 197.

[125] R. Schur, F. Sogawa, M. Nishioka, S. Ishida, and Y. Arakawa, Jpn. J. Appl. Phys. **35**, L357 (1997).

[126] H. Saito, K. Nishi, I. Ogura, S. Sugou, and Y. Sugimoto, Appl. Phys. Lett. **69**, 3140 (1996).

[127] D. L. Huffaker, O. Baklenov, L. A. Graham, B. G. Streetman, and D. G. Deppe, Appl. Phys. Lett. **70**, 2356 (1997).

[128] N. N. Ledentsov, V. M. Ustinov, A. Yu. Egorov, A. R. Kovsh, M. V. Maximov, P. S. Kop'ev, Zh. I. Alferov, J. A. Lott, and D. Bimberg, in *Proc. Int'l. Symp. ''Nanostructures: Physics and Technology''*, St. Petersburg, 1997, p. 195.

[129] J. A. Lott, N. N. Ledentsov, V. M. Ustinov, A. Yu. Egorov, A. E. Zhukov, P. S. Kop'ev, Zh. I. Alferov, and D. Bimberg, Electron. Lett. **33**, 1150 (1997).
[130] N. N. Ledentsov, V. M. Ustinov, A. Yu. Egorov, A. E. Zhukov, M. V. Maksimov, I. G. Tatabadze, and P. S. Kop'ev, Fiz. Tekh. Poluprovodn. **28**, 1281 (1994) [*sic*].

Translated by Paul F. Schippnick

ATOMIC STRUCTURE AND NON-ELECTRONIC PROPERTIES OF SEMICONDUCTORS

Effect of nonuniform distribution of radiation defects in GaAs on the DLTS spectra

V. A. Novikov and V. V. Peshev

V. D. Kuznetsov Siberian Physicotechnical Institute, 634050 Tomsk, Russia
(Submitted June 21, 1996; accepted for publication September 15, 1997)
Fiz. Tekh. Poluprovodn. **32**, 411–416 (April 1998)

An attempt has been made to relate the origin of the U band in the DLTS spectrum of neutron-irradiated GaAs with the known $P2$ and $P3$ defects localized near disordered regions. The shape and temperature position of the $P2$ and $P3$ peaks in the DLTS spectra are assumed to vary as a result of the influence of the electric field of these disordered regions on the rate of electron emission from the defect levels. The DLTS spectra for $P2$ and $P3$ centers located in regions with internal electric fields have been calculated. A comparison of the total calculated spectrum for $P2$ and $P3$ centers with a U band reveals satisfactory agreement with experiment. © *1998 American Institute of Physics.* [S1063-7826(98)00204-X]

It is well known that when n-type GaAs is bombarded by α particles, protons (with energy $E>10$ MeV), and neutrons a wide band, called the U band, appears on the deep-level transient spectroscopy (DLTS) curves. The nature of this band has not been determined unambiguously, although there are some conjectures about its origin.[1]

In the present work we have made an effort to link the appearance of the U band with the known radiation defects $P2$ and $P3$ (Ref. 2). Here we have assumed that the defects are localized in the electric fields of the disordered regions (DR) and the electric field of the disordered regions alters the shape of the $P2$ and $P3$ peaks and their position in the DLTS spectrum. That the U band may be due to $P2$ and $P3$ defects is supported by the following data in the literature. In Ref. 3 it was shown that upon annealing at $T=770$ K the U band splits into two peaks. It follows from Ref. 4 that the amplitude of the U band varies slightly upon annealing at $T=500$ K. Therefore, the defects responsible for the formation of the U band should have high thermal stability. $E1-E5$ defects are annealed out at $T\simeq 500$ K (Ref. 2). The $P2$ and $P3$ traps that were observed in electron and γ-irradiated samples have significantly greater thermal stability, but a low incorporation rate.[2] They are observed after the $E4$ and $E5$ centers masking them have been annealed out or after high-temperature electron irradiation.[5] However, there are grounds to assume that the fraction of $P2$ and $P3$ traps in the total number of radiation defects grows abruptly with increase in the recoil energy of the atoms.[6] Increased recoil energy of the atoms leads to the formation of disordered regions. The presence of an electric field in the vicinity of the disordered regions leads to an increase in the electron emission rate from the deep levels. Therefore, the superposition of $P2$ and $P3$ peaks in the DLTS spectra is shifted toward lower temperatures. The larger the recoil energy of the atoms, the larger the degree of compensation in the damaged regions and the larger the contact potential difference between these regions and the crystal matrix (there is a limiting value of the potential difference, which is associated with the limiting position of the Fermi level[7]). It is therefore possible that a tendency toward an increase in the shift and half-width of the peaks with increase in the mass of the bombarding high-energy particle was observed in Ref. 8.

Our goal in this study was to calculate the DLTS signal for the $P2$ and $P3$ deep centers localized in the electric fields of the disordered regions and to compare it with the U band in n-type GaAs irradiated with neutrons in a pulsed reactor. The neutron irradiation dose was $D_n=6.3\times 10^{13}$ cm^{-2}. For samples we used Schottky-barrier diodes. The Schottky barrier was created in a GaAs layer grown by vapor-phase epitaxy on a heavily doped substrate. The free electron concentration in the layer was 5.5×10^{15} cm^{-3}.

We used the following model picture. The disordered regions consist of a central part with radius r_1 with a high defect density $N_1=10^{19}-10^{20}$ cm^{-3} (Ref. 9), in which the Fermi level is pinned at the limiting position, and a spherical layer in which the defect concentration $N_T(r)$ falls off with distance r from the center of the disordered region according to a Gaussian law

$$N_T(r)=N_{T0}\exp(-r^2/2\sigma^2), \tag{1}$$

where σ is the variance of the distribution. The spherical defect layer is found in a semiconductor matrix in which the free carrier concentration is equal to N_2. Between the matrix and the central part of the disordered region there is a contact potential difference φ_c, and defects located in the spherical layer are found in the electric field (Fig. 1). Traps located a distance r from the center of the disordered region $r_1<r<r_0$ are destroyed even without the application of an external electric field and do not participate in the formation of the DLTS spectrum. Here r_0 is the radius of the spherical surface on which the Fermi level E_F intersects the deep trap level. The shallower the trap level E_T, the less the number of traps of the given sort will participate in the formation of the corresponding DLTS peak. The internal electric field in-

 © 1998 American Institute of Physics

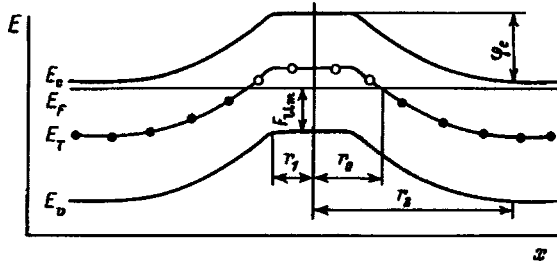


FIG. 1. Band diagram of the disordered region.

creases the relative electron emission rate α from the levels due to the Poole–Frenkel' effect and tunneling in which phonons participate. Since the electric field intensity F is a function of the distance r from the center of the disordered region, the emission rate also depends on r. Therefore, the larger is E_T, the wider will be the interval of values of α for the traps participating in the recharging and the greater will be the influence of the internal field on the half-width of the corresponding DLTS peak.

The expression for the DLTS signal has the form[10]

$$R(T)=\frac{C}{W_1^2N_2}\int_{W_0}^{W_1}xN_T[\exp(-\alpha t_1)-\exp(-\alpha t_2)]dx, \tag{2}$$

where t_1 and t_2 are the gate times of relaxation of the capacitance, T is the temperature, W_0 is the length of the space charge region (SCR) of the Schottky barrier at the time of action of the filling pulse, and W_1 and C are the length of the space charge region and the capacitance of the Schottky barrier in the absence of the filling pulse. The quantities W_0, W_1, and C are related by well-known relations to N_2, to the voltage U applied to the diode, and to the contact potential difference φ_{Sch} of the Schottky barrier.[11] Calculation of the integral in Eq. (2) presents certain difficulties since the distance x is reckoned from the surface of the semiconductor, while the distance r in the distribution of deep centers and electric fields $F(r)$ is reckoned from the centers of the disordered regions. However, expression (2) can be simplified, bearing in mind the following. In the vicinity of the points x at which the disordered regions are found, changes in the emission rate due to the field take place at short intervals $[r_0,3\sigma]$, reckoned from the centers of the disordered regions. The lengths of such intervals (~ 600 Å) are negligible in comparison with the distance x over essentially the entire integration interval $[W_0,W_1]$. Therefore, for a sufficiently large concentration of disordered regions it may be assumed that at any point x the macroscopic concentration of traps $\bar{N}_T$ does not depend on x and the discrete value of the emission rate (for $T=\mathrm{const}$) is split into a band of emission rates with boundaries $\alpha(r_0)$ and $\alpha(3\sigma)$. The band of emission rates is characterized by a density of states in the band $\partial\bar{N}_T/\partial\alpha$. Here the density of states is known if we know the microscopic distribution of traps $N_T(r)$ and the dependence of the electric field on distance reckoned from the center of the disordered region, $F(r)$:

$$\frac{\partial\bar{N}_T}{\partial\alpha}\sim\frac{\partial N_T}{\partial r}\frac{\partial r}{\partial\alpha}. \tag{3}$$

Taking these points into account, and also Eqs. (1) and (3), we recast expression (2) in the form

$$R(T)\simeq\frac{C}{N_2W_1^2}\int_{W_{00}}^{W_1}xdx\int_{\alpha(r_0)}^{\alpha(3\sigma)}\frac{\partial\bar{N}_T}{\partial\alpha}[\exp(-\alpha t_1)-\exp(-\alpha t_2)]\partial\alpha=\frac{\bar{N}_TC(W_1^2-W_0^2)}{2N_2W_1^2\sigma^2}\times\int_{r_0}^{3\sigma}r\exp(-r^2/2\sigma^2)[\exp(-\alpha t_1)-\exp(-\alpha t_2)]dr. \tag{4}$$

In the present paper we take into account only the Poole–Frenkel' effect. In this case the expression for the emission rate in the presence of an electric field has the form[12]

$$\alpha(r,T)=\alpha_0Y(E_p,T), \tag{5}$$

where

$$\alpha_0=b\sigma_{na}T^2\exp(-E_{na}/kT), \tag{6}$$

$$Y(E_p,T)=(kT/E_p)^2\{1+[(E_p/kT)-1]\times\exp(E_p/kT)\}+0.5, \tag{7}$$

$$E_p=\beta\sqrt{|F(r)|}. \tag{8}$$

Here α_0 is the relative emission rate in the absence of an electric field, b is a constant characterizing the semiconductor,[11] σ_{na} is the apparent cross section for electron capture at the trap level, $E_{na}=(E_\sigma+E_T)$ is the ionization energy of a trap in the absence of an electric field (where E_σ is the electron capture barrier of the trap), Y is a function that allows for the influence of the electric field on the emission rate, and β is a coefficient that is equal to $\beta=\beta_0=2e\sqrt{e/s}$ in the case of capture of an electron at a singly positively charged center. Here e is the charge of the electron, and ε is the absolute dielectric constant of the semiconductor.

Let us briefly analyze expression (4). It can be seen from expression (4) that the height of the DLTS peak is directly proportional to the trap concentration [the same as in the case $N_T(x)=\mathrm{const}$ and $F=0$]. However, for the same traps, keeping their concentration fixed, the height of the peak in the case of formation of a disordered region is always less than in the case of a uniform defect distribution. This is so for two reasons. The first is that traps located in the region $r<r_0$ in all the disordered regions are excluded from the recharging process. The second consists in peak broadening due to the dependence $\alpha=f(r)$. It can also be seen from expression (4) that the proportionality constant between peak height and the trap concentration depends in a complicated way on the doping level N_2, the trap parameters E_{na} and σ_{na} and the variance of the distribution σ. Therefore, the ratio of peak heights for different traps is not equal to the ratio of their concentrations.

TABLE I. Parameters of radiation defects.

Center	$E2$	$E3$	$P2$	$P3$
E_{na}, eV	0.16	0.38	0.5	0.72
σ_{na}, cm^2	5×10^{-13}	6.2×10^{-15}	1.4×10^{-15}	1.4×10^{-13}
$\bar{N}_T$, cm^{-3}	5.7×10^{14}	3.6×10^{13}	3.3×10^{13}	2.4×10^{14}

To use expression (4) for calculations, it is necessary to know r_0 and the dependence $E_p(r)$. These quantities can be obtained by using the potential obtained in the Gossick approximation for the disordered regions.[13] The potential (in the SI system) for the region $r_1<r<r_2$ is

$$\varphi(r)=\frac{N_2e}{6\varepsilon}\frac{(3r_2^2r-r^3-2r_2^3)}{r}, \qquad (9)$$

where r_2 is the boundary of the space charge region surrounding the cluster. In the region $0<r<r_1$ $|\varphi|=\varphi_c$. The quantity r_2 is determined from the boundary condition that the potential difference between spherical surfaces with radii r_1 and r_2 is equal to φ_c, i.e., r_2 is a root of the cubic equation

$$|\varphi(r_1)|=\varphi_c=\frac{N_2e}{6\varepsilon}\frac{2r_2^3+r_1^3-3r_2^2r_1}{r_1}. \qquad (10)$$

In Ref. 7 it was shown that $F_{\rm lim}$ (see Fig. 1) in GaAs is located 0.6 eV above the top of the valence band. In this light, the expression for φ_c can be written as

$$e\varphi_c=E_g-(0.6\text{ eV}+E_F), \qquad (11)$$

where E_g is the width of the band gap.

The radius r_0 of the spherical boundary separating the filled from the unfilled traps is found from the condition that the quantity $(E_T-E_F)/e$ and the potential difference between the points r_0 and r_2 should be equal, i.e., r_0 is a root of the cubic equation

$$|\varphi(r_0)|=(E_T-E_F)/e=\frac{N_2e}{6\varepsilon}\frac{(2r_2^3+r_0^3-3r_2^2r_0)}{r_0}. \qquad (12)$$

Note that r_2 and r_0 depend on T since φ_c, E_g, and E_F in Eqs. (10)–(12) are functions of temperature. The electric field intensity in the vicinity of the disordered region are found by differentiating expression (9):

$$|F(r)|=\frac{N_2e}{3\varepsilon}\left[\left(\frac{r_2}{r}\right)^3-1\right]r. \qquad (13)$$

Using Eqs. (13) and (5), we obtain the dependence of the relative emission rate on the distance from the center of the disordered region. In order to find r_2 from Eq. (10), it is necessary to know the size of a cluster r_1. We used the mean cluster size $r_1=130$ Å, obtained from electrical measurements in Ref. 9 for neutron-irradiated GaAs. The characteristics E_{na} and σ_{na} of radiation defects in GaAs are well known[2,8] and are given in Table I.

To determine the radius r_0, it is necessary to know the depth E_T of the trap level. In Ref. 14 the height of the barrier to electron capture to the $E3$ level was determined to be $E_\sigma=0.08$ eV, which gives $E_T=0.3$ eV. For other centers information about E_σ is absent, and we have simply assumed $E_T=E_{na}$. Thus, all the necessary parameters and dependences used in Eq. (4) are known. The variable parameters are only the variance of the distribution and the trap concentration. The main fitting parameter here is the variance since its variation leads to a change in the height, shape, and temperature of the peak maximum, and variation of $\bar{N}_T$ changes only the height.

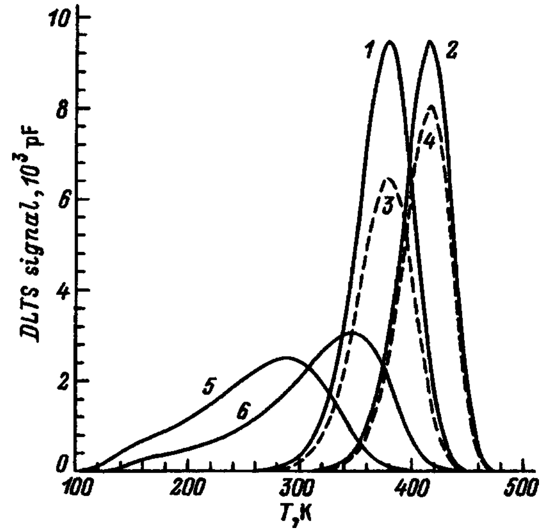


FIG. 2. Calculated $P2$ (*1,3,5*) and $P3$ peaks (*2,4,6*). $N_2=1\times10^{15}$ cm^{-3}; $t_1/t_2=4\times10^{-5}$ s/2×10^{-4} s. $\bar{N}_T=10^{14}$ cm^{-3} for $P2$ and $P3$. *1,2*—calculated according to Eq. (2), $N_T(x)=$const, $F=0$; *3,4*—calculated according to Eq. (4), $\sigma=250$ Å, $\beta\simeq0$; *5,6*—calculated according to Eq. (4), $\sigma=250$ Å, $\beta=\beta_0$.

Figure 2 shows the effect on the DLTS curves of the electric fields of the disordered regions for the case of $P2$ and $P3$ centers, whose concentrations are the same, $\bar{N}_T=1\times10^{14}$ cm^{-3}, and which are created in GaAs with doping level $N_2=1\times10^{15}$ cm^{-3}. Curves *1* and *2* were calculated according to the classical formula obtained from Eq. (2) when $N_T(x)=$const and $F=0$. In this case the half-width of the peaks and their positions in the spectrum correspond to the experimentally observed $P2$ and $P3$ peaks in electron-irradiated GaAs. Curves *3* and *4* were obtained from Eq. (4) for the hypothetical case in which $\sigma=250$ Å, $\beta\simeq0$ [for exact equality $\beta=0$ expression (7) is undefined]. In this case the $P2$ and $P3$ centers are localized in regions with electric fields, but the field does not have any effect on the emission rate, only varying the number of traps that participate in the measurement. It can be seen from Fig. 2 that the half-width and temperature position of the peaks are the same as before, but the height of the peaks has changed. As can be seen, the change in the height of the peak is greater for the shallower center ($P2$). Curves *5* and *6* were also calculated from Eq. (4) for $\sigma=250$ Å and $\beta=\beta_0$. In this case the built-in fields increase the electron emission rate from the $P2$ and $P3$ levels as they do from the deep donor levels. It can be seen that the shape and temperature position of the peaks have changed. The height of the peaks has also changed due to broadening. Thus, formula (4) indeed gives the results that were expected above on the basis of physical arguments.

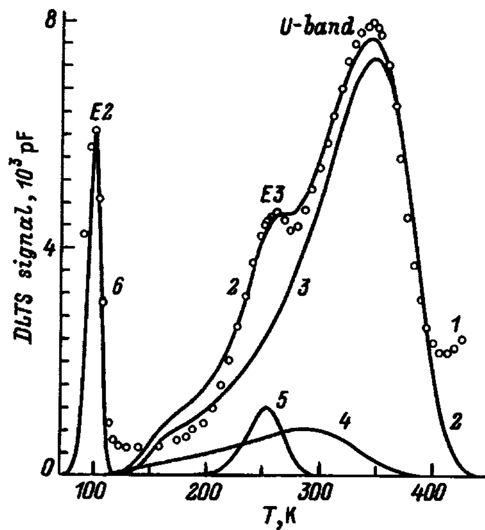


FIG. 3. Experimental (*1*) and calculated (*2–6*) DLTS spectra. $N_2=5.5\times10^{15}$ cm^{-3}. $t_1/t_2=4\times10^{-5}$s/2×10^{-4} s. $\sigma=250$ Å. *1*—$D_n=6.3\times10^{13}$ cm^{-2}; *2*—total spectrum of the calculated $E2$, $E3$, $P2$, and $P3$ peaks; *3*—$P3$, $\bar{N}_T=2.4\times10^{14}$ cm^{-3}, $\beta=\beta_0$; *4*—$P2$, $\bar{N}_T=3.3\times10^{13}$ cm^{-3}, $\beta=\beta_0$; *5*—$E3$, $\bar{N}_T=3.6\times10^{13}$ cm^{-3}, $\beta\simeq0$; *6*—$E2$, $\bar{N}_T=5.7\times10^{14}$ cm^{-3}, $\beta\simeq0$.

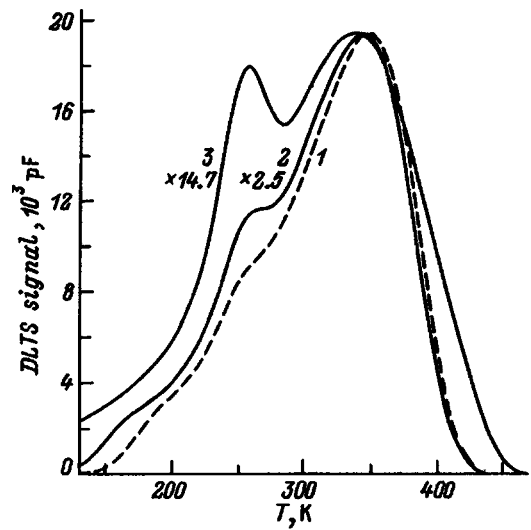


FIG. 4. Total spectra of the calculated $E3$, $P2$, and $P3$ peaks in GaAs with different doping levels. $t_1/t_2=4\times10^{-5}$ s/2×10^{-4} s. N_2, cm^{-3}: *1*—10^{15}, *2*—5.5×10^{15}, *3*—10^{17} cm^{-3}. $\bar{N}_T$, cm^{-3}: $E3$—3.6×10^{13}; $P2$—3.3×10^{13}; $P3$—2.4×10^{14}. $\sigma=250$ Å.

Figure 3 shows an experimentally obtained (curve *1*) and a calculated DLTS spectrum (curve *2*) for GaAs irradiated by fast neutrons. The experimental spectrum was obtained for amplitude of the filling pulse $U_p=6$ V, reverse bias $U_b=6$ V, rate window $t_1/t_2=40/200$ [μs/μs] and duration of the filling pulse $t_p=50$ μs. The contact diameter of the diode was $d=180$ μm, the contact potential difference of the Ti/GaAs Schottky barrier was $\varphi_{\text{Sch}}=0.7$ V, and the free electron concentration was 5.5×10^{15} cm^{-3}. From Fig. 3 (curve *1*) it can be seen that the peaks associated with the $E2$ and $E3$ centers, and also the U band, are observed in the spectrum. To calculate the spectrum according to Eq. (4), we calculated the $E2$, $E3$, $P2$, and $P3$ peaks (curves *3–6*), and then summed them up (curve *2*). We assumed that $P2$ and $P3$ are donors, i.e., $\beta=\beta_0$. In the fitting of curve *2* to the experimental spectrum (curve *1*), we varied two parameters: $\bar{N}_T$ and σ. The variance was assumed to be the same for all the centers. This equality of the variance is neither necessary nor obvious. However, the best fit of curve *2* to curve *1* in the region of the U band is obtained when the values of σ in the distributions of the $P2$ and $P3$ centers coincide: $\sigma=250$ Å. The magnitude of the Poole–Frenkel' effect depends on what kind of trap is present in the given conductivity type of material, i.e., on whether it is a donor or acceptor. Coincidence of curves *1* and *2* in the region of the $E2$ and $E3$ peaks is achieved only for $\beta\simeq0$. The weak effect of the electric field is possibly due to the acceptor nature of the $E2$ and $E3$ centers. For $\beta\simeq0$ variation of σ affects only the height of the peaks, which leads to an indeterminacy in the choice of the value of the variance. Therefore, for the $E2$ and $E3$ centers the value $\sigma=250$ Å was chosen only by analogy with the $P2$ and $P3$ centers. The values of the concentrations of the traps, which form spectrum *2*, are given in Table I. It can be seen from Fig. 3 that the calculated spectrum is in good agreement with the experimental spectrum.

It is interesting to compare the total average concentration $\Sigma\bar{N}_T\simeq10^{15}$ cm^{-3} (see Table I) of all the detected deep electron traps in a sample irradiated with a dose equal to 6.3×10^{13} cm^{-2} with the data in Ref. 9. It follows from Ref. 9 that for $D_n=4.2\times10^{17}$ cm^{-2} the concentration of disordered regions is $N_{dr}=5.6\times10^{16}$ cm^{-3}, and each disordered region contains roughly 1000 displaced atoms. Thus, for linear kinetics of accumulation of disordered regions the average concentration of displaced atoms in our samples should be $\sim6.3\times10^{15}$ cm^{-3}. If we take into account that the displaced atoms along with the deep electron traps can form other types of defects invisible to DLTS in n-type GaAs with a Schottky barrier, in particular, hole traps, then the agreement of our data with those of Ref. 9 is seen to be satisfactory.

It was shown in Ref. 15 that as the doping level of neutron-irradiated GaAs is increased, the half-width of the U band increases. Here the increase in the half-width takes place mainly on account of the low-temperature edge of the U band. Figure 4 shows the superposition of the calculated $E3$, $P2$, and $P3$ peaks in GaAs with $N_2=10^{15}$, 5.5×10^{15}, and 10^{17} cm^{-3}. To obtain these spectra, we used values of the trap concentrations and variance of the distribution obtained above by comparing the spectra in Fig. 3. It can be seen from the figure that the nature of the change in the half-width of the calculated spectrum in response to changes in the doping level correlates with the experimental data of Ref. 15.

Note that the derivation of expression (4) disregarded the electric field created during the DLTS measurements. Below we present estimates of the consequence of ignoring the magnitude of the external field. Since the value of the field intensity is important from the viewpoint of increasing the emission rate α, it is necessary to compare the values of the internal and external fields in those space charge regions of

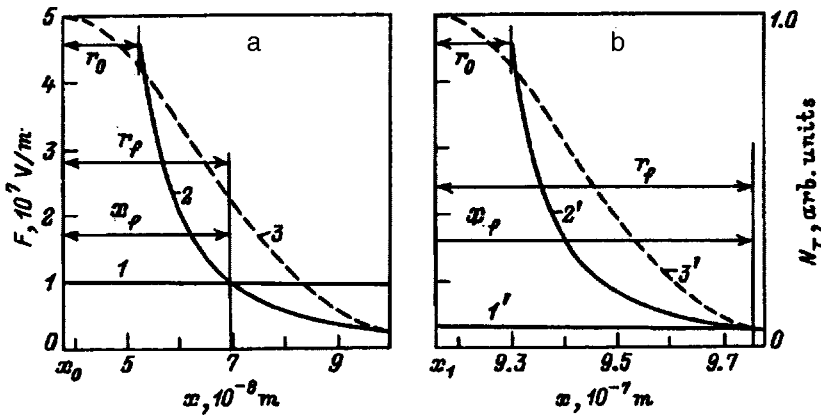


FIG. 5. Dependence of the external electric field intensity (*1,1′*) in the space charge region of the Schottky barrier, and of the field intensities (*2,2′*) and defect concentration N_T of two disordered regions (*3,3′*) on distance. *a*—Center of the disordered region at x_0; *b*—center of the disordered region at x_1.

the Schottky barrier where 1) defects are located and 2) defects can be recharged during the measurement. In the absence of disordered regions in the DLTS measurements only those traps can participate that are located in the space charge region of the Schottky barrier between the points $x_0 = W_0 - \lambda$ and $x_1 = W_1 - \lambda$. Here x_0 is the coordinate of the intersection point of the defect level with the Fermi level in a diode without reverse bias, and x_1 is the coordinate of the intersection point with the Fermi quasilevel in a reverse-biased diode. In the presence of built-in fields in the disordered regions not all traps located in this interval participate in the measurement. At the point x_0 the intensity of the external field having an effect on α at the time the traps are being destroyed has the most significance. Therefore, we should first compare the fields in the vicinity of the disordered region whose center is located at the point x_0 (Fig. 5a). Since it was shown above that the $P3$ peak dominates in the makeup of the U band, we used the characteristics of the $P3$ center in our calculations. We calculated for $N_2 = 5.5 \times 10^{15}$ cm^{-3} since quantitative values of the variance and defect concentration were obtained from the experimentally obtained spectra, and for the values r_1, φ_{Sch}, U_b, σ, and the temperature of the maximum of the $P3$ peak, which were used or found above in the calculations of the spectra. We obtained the following values: $r_2 = 1.553 \times 10^{-7}$, $r_0 = 1.472 \times 10^{-8}$, $x_0 = 3.784 \times 10^{-8}$, $x_1 = 9.156 \times 10^{-7}$, $W_0 = 4.192 \times 10^{-7}$, $W_1 = 1.297 \times 10^{-6}$ m.

It turned out that for $x = x_0 + r_0$, starting at which defects of the given disordered region can be recharged, the internal field is ~ 4.62 times larger than the external field and is equal to 4.58×10^7 V/m. The field of the disordered region (curve *2*) falls off faster than the external field (curve *1*) and at $x_1 = 6.94 \times 10^{-8}$ m the fields become equal. It can be seen from Fig. 5 that the coordinate $x = x_{dr}$ of the center of the disordered region is linked by the relation $x_f - x_{dr} = r_f$, from which we can find r_f, which is equal to the distance from the field intersection point to the center of the disordered region. For $x > x_f$ the external field dominates; however, with increase of the distance from the center of the disordered region the defect concentration (curve *3*) decreases. For a specific disordered region we write the ratio q of the number of defects being recharged in that part of the disordered region where the field of the disordered region (the internal field) dominates, to the total number of defects in this region taking part in the recharging

$$q = \int_{r_0}^{r_f} 4\pi r^2 N_T(r)\,dr \Big/ \int_{r_0}^{3\sigma} 4\pi r^2 N_T(r)\,dr$$
$$= \int_{r_0}^{r_f} r^2 \exp\left(-\frac{r^2}{2\sigma^2}\right) dr \Big/ \int_{r_0}^{3\sigma} r^2 \exp\left(-\frac{r^2}{2\sigma^2}\right) dr. \tag{14}$$

From Eq. (14) we find that when $x_{dr} = x_0$ (Fig. 5a), $q = 0.315$, and for $x_{dr} = x_1$ (Fig. 5b) $q = 0.91$. Thus, q depends on the coordinate of the center of the disordered region. We write for the space charge region of the Schottky barrier the ratio Q of the number of defects being recharged in the dominant fields of all the disordered regions, to the total number of defects participating in the recharging

$$Q = \int_{x_0}^{x_1} S N_{dr} q(x_{dr})\,dx_{dr} \Big/ \int_{x_0}^{x_1} S N_{dr}\,dx_{dr}$$
$$= \int_{x_0}^{x_1} q(x_{dr})\,dx_{dr} \Big/ x_1 - x_0, \tag{15}$$

where S is the area of the metal–semiconductor contact. We used numerical methods to obtain the dependences $q = f(x_{dr})$ on the intervals $[x_0, x_1]$ for different values of N_2. Substituting these dependences into Eq. (15), we obtained the following values of Q: 0.85, 0.55, 0.45, and 0.16 for $N_2 = 1 \times 10^{15}$, 5.5×10^{15}, 1×10^{16}, 1×10^{17} cm^{-3}, respectively. Thus, Eq. (4) (from the viewpoint of neglecting the external field) gives a good qualitative description of the spectrum for $N_2 = 10^{15}$ cm^{-3}, a satisfactory description for $N_2 = (5.5-10) \times 10^{15}$ cm^{-3}, and only a qualitative description for $N_2 = 10^{17}$ cm^{-3}, with the necessary inclusion, in the last case, of peculiarities arising as a result of a nonuniform defect distribution. Note that the quantity Q is not a direct measure of the presence of defects, although it is related to it. For example, $Q = 55\%$ means that $100\% - Q = 45\%$ of the defects actually form the spectrum for values of the external electric field corresponding to curves *1* and *1′* for $x > x_f$ in Fig. 5, but Eq. (4) treats these defects as forming the spectrum in the closer field of the disordered region (curves *2* and *2′* for $x > x_f$ in Fig. 5).

In summary, we have proposed an approximate analytical expression to describe the DLTS spectra of semiconductors bombarded by particles with a large mass and energy. We have shown that the U band in the spectrum of neutron-bombarded n-type GaAs can be satisfactorily described assuming it is formed by $P2$ and $P3$ centers localized in the electric fields of the disordered regions.

[1] G. M. Martin, E. Estève, P. Langlade, and S. Makram-Ebeid, J. Appl. Phys. **56**, 2655 (1984).
[2] D. Pons and J. C. Bourgoin, J. Phys. C: Solid State Phys. **18**, 3839 (1985).
[3] R. Magno, M. Spencer, J. G. Giessner, and E. R. Weber, in *13th Int'l. Conf. on Defects in Semiconductors* (Coronado, California, 1984), p. 981.
[4] V. N. Brudnyĭ, N. G. Kolin, and A. I. Potapov, Fiz. Tekh. Poluprovodn. **27**, 260 (1993) [Semiconductors **27**, 145 (1993)].
[5] V. N. Brudnyĭ and V. V. Peshev, Phys. Status Solidi A **105**, K57 (1990).
[6] V. A. Ivanyukovich, V. I. Karas', and V. M. Lomako, Fiz. Tekh. Poluprovodn. **24**, 1427 (1990) [Sov. Phys. Semicond. **24**, 893 (1990)].
[7] V. N. Brudnyĭ, S. N. Grinyaev, and V. E. Stepanov, Physica B: Condens. Matter. **212**, 429 (1995).
[8] D. Lang, in *Point Defects in Solids* [Russian translation] (Mir, Moscow, 1979), p. 187.
[9] R. Coates and E. W. J. Mitchell, Adv. Phys. **24**, 593 (1975).
[10] J. C. Bourgoin and M. Lanoo, *Point Defects in Semiconductors* (Springer-Verlag, New York, 1980).
[11] L. S. Berman, *Capacitance Methods for Studying Semiconductors* [in Russian] (Nauka, Leningrad, 1972).
[12] J. L. Hartke, J. Appl. Phys. **39**, 4871 (1968).
[13] B. R. Gossick, J. Appl. Phys. **30**, 1214 (1959).
[14] D. Stievenard, J. C. Bourgoin, and M. Lanoo, J. Appl. Phys. **55**, 1477 (1984).
[15] C. E. Barnes, T. E. Zipperian, and L. R. Dawson, J. Electron. Mater. **14**, 95 (1985).

Translated by Paul F. Schippnick

Influence of the initial boron doping level on the boron atom distribution arising as a result of heat treatment in silicon implanted with boron ions

V. I. Obodnikov and E. G. Tishkovskiĭ

Institute of Semiconductor Physics, Siberian Branch of the Russian Academy of Sciences, 630090 Novosibirsk, Russia

(Submitted January 28, 1997; accepted for publication September 15, 1997)
Fiz. Tekh. Poluprovodn. **32**, 417–420 (April 1998)

The dependence of the boron distribution on the initial boron concentration in the range $(1-9)\times 10^{19}$ cm^{-3} was investigated by secondary-ion mass-spectrometry (SIMS) after heat treatment of boron-ion implanted silicon at 900 °C. It was found that when the initial boron concentration exceeds the solubility limit at the annealing temperature used, two additional peaks arise in the boron concentration profiles at the boundaries of the ion-implantation disordered region. It is suggested that their appearance in these regions at high doping levels is due to clustering of excess interstitial impurity atoms not built into the lattice sites following displacement of boron atoms from the lattice sites by intrinsic interstitials that leave the disordered region. © *1998 American Institute of Physics.* [S1063-7826(98)00304-4]

INTRODUCTION

In Refs. 1 and 2 it was found that the boron concentration profiles obtained as a result of annealing at 900 °C samples of silicon initially doped with boron to a concentration of 2×10^{20} cm^{-3} and then bombarded with $^{10}B^{+}$ ions at a dose of 10^{16} cm^{-2} with energies of 180 or 400 eV are nonmonotonic and have five maxima. Processes that can lead to the formation of such spatial distributions for such high initial impurity concentrations are discussed in Ref. 2.

Since such an effect is not observed in samples with a low initial doping level (see, e.g., Ref. 3), it might be expected that one of the main reasons for its appearance is precisely a high initial doping level.

The aim of the present work is to investigate the influence of the initial level of boron doping of silicon on the nature of the impurity depth profiles arising as a result of heat treatment of samples bombarded with boron ions.

EXPERIMENT

As our starting samples, we used samples of silicon with (100) surface orientation, boron doped in the concentration interval from 10^{19} to 9×10^{19} cm^{-3}. Samples with boron concentration 10^{19} cm^{-3} were prepared from KDB-0.006 silicon wafers and contained both isotopes of boron—^{10}B and ^{11}B in a concentration ratio of $\sim$1:4 corresponding to their natural abundance. Samples with boron concentrations $N_B \simeq 3\times 10^{19}$, 6×10^{19} and 9×10^{19} cm^{-3} were prepared by ion doping. The $^{10}B^{+}$ ions with energy $E=300$ keV were implanted in KDB-10 silicon wafers at dose levels $D=5\times 10^{15}$, 10^{16}, and 1.5×10^{16} cm^{-2} on an HVEE-400 setup. The bombarded wafers were annealed at 1075 °C for 100 min (10 min in dry oxygen +90 min in dry nitrogen). As a result, regions of thickness ~ 1 μm were created in surface regions of the wafers, doped to the indicated concentrations with only one isotope of boron.

The samples were then bombarded by 180-keV $^{10}B^{+}$ ions at a dose of 10^{16} cm^{-2}. In every case the thickness of the region affected by the implantation process was smaller than the thickness of the pre-doped (to the above-indicated boron concentrations) surface regions. After bombardment, the samples were annealed at 900 °C for 1 h in an inert medium.

The boron isotope distribution with depth x in the samples was studied by secondary-ion mass-spectroscopy (SIMS) on an MIQ-256 setup (CAMECA-RIBER), which measured the yields of $^{10}B^{+}$ and $^{11}B^{+}$ secondary ions. The O_2^{+} primary beam ions had an energy of 13 keV. The scanning area of the focused primary ion beam was $\sim 0.4\times 0.6$ mm^2. The etch rate was ~ 0.5 nm/s, and its constancy during the measurements was monitored by the $^{30}Si^{+}$ secondary ion yield. To eliminate the contribution of secondary boron ions knocked out of the walls of the crater, the analyzed signal was taken from the central part of the crater making up 10% of the scanning area of the primary ion beam.

Control measurements in which a primary beam of cesium ions was used and the yield of $(^{28}Si^{10}B)^{-}$ and $(^{28}Si^{11}B)^{-}$ secondary ions was recorded gave the same results as the measurements with a primary beam of oxygen ions.

EXPERIMENTAL RESULTS

The measurements on the samples with the lowest initial boron doping level $N_B=10^{19}$ cm^{-3} showed (Fig. 1) that the shape of the depth profile of the ^{10}B isotope, formed as a result of annealing, faithfully tracked the boron profiles obtained in Ref. 3 under similar conditions of bombardment ($E=70$ keV, $D=10^{16}$ cm^{-2}) and annealing ($T=900$ °C, $t=35$, 70, and 105 min) in silicon that had not been pre-doped with boron. The impurity attached to a segment in the region of the distribution maximum of the implanted boron,

 © 1998 American Institute of Physics

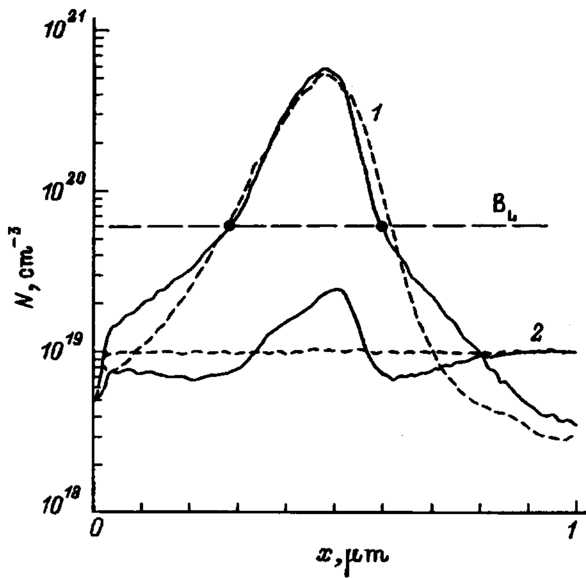


FIG. 1. Concentration profiles (N) of boron isotopes in silicon with initial boron concentration $N_B = 10^{19}$ cm^{-3}, resulting from ^{10}B ion implantation ($E = 180$ keV, $D = 10^{16}$ cm^{-2}) and annealing at 900 °C for 1 h. *1*—^{10}B, *2*—^{11}B. Dashed curves—before annealing, solid curves—after annealing. $B_L = 6 \times 10^{19}$ cm^{-3}—maximum solubility of boron at substitution sites at $T = 900$ °C.

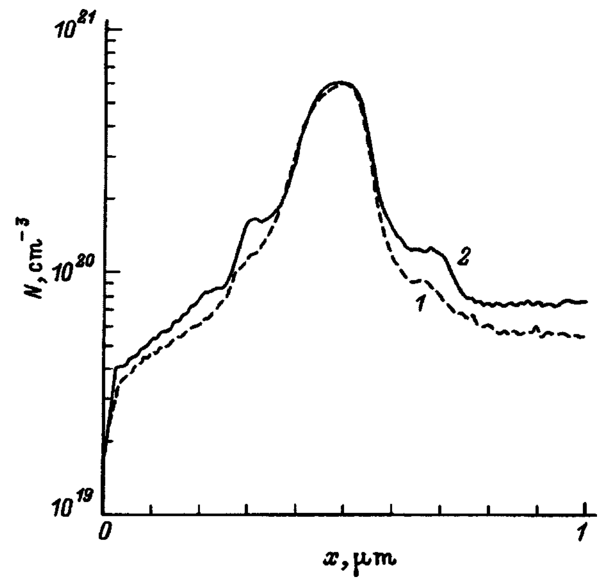


FIG. 2. Concentration profiles (N) of the isotope ^{10}B in silicon with different initial dopant levels N_B, resulting from ^{10}B ion implantation ($E = 180$ keV, $D = 10^{16}$ cm^{-2}) and annealing at 900 °C for 1 h. N_B, cm^{-3}: *1*—6×10^{19}, *2*—9×10^{19}.

where its concentration is greater than $\sim 6 \times 10^{19}$ cm^{-3}, the limiting solubility of boron atoms at the lattice sites (B_L) at $T = 900$ °C (Refs. 4 and 5); a broadening of the profile beyond the bounds of this region (the inflection points of the profile, indicated in the figure) is also observed. Such an attachment of the boron atoms was also experimentally observed earlier under other conditions of bombardment and annealing,[3,4,6] but only in those cases in which the amount of implanted impurity created an excess boron atom concentration relative to its limiting solubility at substitution sites for the temperatures used. In Refs. 3 and 4 the relative immobility of the boron atoms in this segment was tied to the attachment to immobile boron-containing clusters of an excess of impurity which was not built into the sites due to the limitation of the solubility limit.

Changes in the ^{11}B isotope profile, observed after annealing at 900 °C, show (Fig. 1) that on the segment in the region of the maximum of the implanted boron distribution there takes place not only an attachment but also an accumulation of impurity, distinctly revealed in the appearance of a maximum on the initially flat profile of the isotope ^{11}B. Consequently, if the observed attachment of boron atoms is to be tied to the formation of boron-containing clusters,[3,4] then the latter can serve as sinks for the mobile boron component which maintain the impurity flux in the region of the maximum of the implanted boron distribution.

As the initial doping level is increased, the boron atom depth profiles resulting from $^{10}B^+$ ion implantation and subsequent annealing change their shape.

Whereas for an initial boron concentration of $\sim 3 \times 10^{19}$ cm^{-3} the resulting profile mimics the ^{10}B profile obtained in samples with an initial concentration of 10^{19} cm^{-3}, already at boron concentrations of $\sim 6 \times 10^{19}$ cm^{-3}, along with a gathering of boron into the region adjacent to the maximum of the implanted impurity distribution, two inflection points are observed in the wings of the profile (Fig. 2, curve *1*). As the initial doping level is increased to $\sim 9 \times 10^{19}$ cm^{-3}, additional maxima appear in place of these inflection points (Fig. 2, curve *2*).

Thus, features appear in the impurity concentration profiles only when the initial doping level roughly coincides with or exceeds the limiting boron solubility for the annealing temperature used.

DISCUSSION

The results obtained by us can be explained within the framework of the approach proposed in Ref. 2, which is based on the following assumptions:

—during implantation of boron ions in the bombarded region, along with the accumulation of boron atoms, a significant fraction of which do not occupy regular lattice sites, there also takes place an accumulation of defect complexes containing interstitial silicon atoms (I) and vacancies (V);

—the region where the main fraction of implanted boron is located and the region where the main fraction of primary radiation defects are generated essentially coincide; as a result, the boron concentration at substitution positions (B_S) in this common region is lower than the initial doping level.

—at the initial stage of annealing the defect complexes decay, which leads to a liberation of interstitial silicon atoms and vacancies; boron atoms are also liberated from defect–impurity associates and then show up at interstitial sites (B_I);

—the liberated vacancies participate in reactions associated with incorporation of boron atoms into lattice sites, and in annihilation reactions with intrinsic interstitials; the intrin-

sic interstitials also take part in reactions displacing the boron atoms from the lattice sites (the Watkins reaction).[7,8]

Consequently, at the initial stage of annealing the most probable reactions are

$$B_S + I \rightarrow B_I, \quad (1)$$

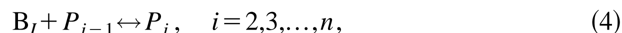
$$B_I + V \rightarrow B_S, \quad (2)$$

$$V + I \rightarrow 0. \quad (3)$$

Relaxation of nonequilibrium elementary defects inside the implantation-affected region due to reactions (1)–(3) is accompanied by their diffusive spreading. Since in the ion-implantation-disturbed region the B_I and I concentrations are large, the propagation of vacancies is initially limited by reactions (2) and (3), and the vacancies are localized inside this region. At the same time, the propagation of intrinsic interstitials in this region is limited only by the annihilation reaction (3) since the B_S concentration is small and the efficiency of reaction (1) is reduced.

As a result of this circumstance, the unreacted intrinsic interstitials, in concentration significantly exceeding that of the vacancies, diffuse out from the implantation-disturbed region. They pass beyond the nominal boundaries of this region and penetrate into the silicon, where the boron concentration at the substitution sites is close to the initial doping level, where they displace boron atoms from the lattice sites.

In the case where the initial doping level exceeds the boron solubility limit, not all the boron atoms displaced from lattice sites can again occupy positions in the silicon lattice; i.e., the excess of mobile boron atoms at interstitial sites cannot be removed by just the one reaction (2). Thus it can be assumed that it is simultaneously removed by a chain of clustering reactions:



$$B_I + P_{i-1} \leftrightarrow P_i, \quad i = 2,3,\ldots,n, \quad (4)$$

where P_i is an immobile cluster that contains i boron atoms.

The boron-containing clusters forming at the boundaries of the bombardment-damaged region serve as sinks for the free boron component. Upon subsequent heating this also ensures the accumulation of impurity in these newly formed objects.

Thus, the appearance of side maxima may be due to processes completely analogous to those discussed in Ref. 3, where attachment of the impurity in the region of the maximum of the implanted-boron distribution was linked with an impurity excess that cannot dissolve into the sites of the silicon lattice and accumulates in clusters. The difference is that at the center of the implantation-disturbed region the high concentration of boron atoms not occupying regular sites in the silicon lattice is created during heating as a result of the decay of defect–impurity complexes formed during ion implantation, and upon the formation of side maxima—due to reaction (1) displacing boron atoms from the sites of interstitials that have escaped the implantation-disturbed region.

In the case where the initial doping level is lower than the boron solubility limit in the substitution sites, the mobile boron component, as it propagates beyond the limits of the implantation-disturbed region, takes part in reaction (2) and occupies lattice sites where under the conditions of the experiment it is relatively immobile. This course of events manifests itself in the formation of ''wings'' which escape, as it were, outward from under the limiting solubility level (the ^{10}B profile after annealing in Fig. 1, and also the data in Refs. 3 and 4).

Thus, the experimentally observed array of boron atom profiles in silicon, depending as it does on the initial doping level, can be consistently explained, in our opinion, on the basis of the well-known reactions (2)–(4) without invoking any additional mechanisms.

CONCLUSIONS

1. We have shown that the distribution of boron atoms with depth in silicon bombarded with 180-keV boron ions at a dose of 10^{16} cm^{-2} and annealed at 900 °C for 1 h depends on the initial boron concentration. At the initial doping level $N_B > 6 \times 10^{19}$ cm^{-3}, which exceeds the limiting solubility of boron at the annealing temperature, along with aggregation and attachment of boron atoms in the region near the maximum of the implanted-impurity distribution, additional maxima appear in the impurity concentration profiles.

2. We believe that the formation of additional maximum in the region near the boundaries of the implantation-disturbed region is connected with clustering of the excess of interstitial boron atoms and the flow-off to these newly formed clusters of the mobile boron component. The excess concentration of interstitial boron atoms is created in this segments due to the reaction which displaces boron atoms from the sites of intrinsic interstitial atoms that escape the implantation-disturbed region.

We wish to thank V. G. Seryapin and B. I. Fomin for assistance in the preparation of the samples. We also thank E. I. Cherepov, L. I. Fedinaya, and V. V. Kalinin for interest in this work and for discussion of the results.

[1] A. M. Myasnikov, V. I. Obodnikov, V. G. Seryapin, E. G. Tishkovskiĭ, B. I. Fomin, and E. I. Cherepov, JETP Lett. **60**, 102 (1994).
[2] A. M. Myasnikov, V. I. Obodnikov, V. G. Seryapin, E. G. Tishkovskiĭ, B. I. Fomin, and E. I. Cherepov, Fiz. Tekh. Poluprovodn. **31**, 338 (1997) [Semiconductors **31**, 279 (1997)].
[3] W. K. Hofker, H. W. Werner, D. P. Oosthoek, and H. A. M. de Grefte, Appl. Phys. Lett. **2**, 265 (1973).
[4] H. Ryssel, K. Muller, K. Haberger, R. Henkelmann, and F. Jahnel, Appl. Phys. Lett. **22**, 35 (1980).
[5] V. E. Borisenko and S. G. Yudin, Phys. Status Solidi A **101**, 123 (1987).
[6] M. G. Dowsett, E. A. Clark, and M. N. Lewis, in *Proceedings of the Sixth International Conference on Secondary-Ion Mass Spectrometry SIMS-VI* (1988), p. 725.
[7] G. D. Watkins, *Radiation Damage in Semiconductors* (Dunod, Paris, 1964), p. 97.
[8] G. Watkins, in *Lattice Defects in Semiconductors* (Conf. Ser. No. 23., Inst. Phys. London–Bristol, 1975), p. 1.

Translated by Paul F. Schippnick

Observation of low-temperature diffusion of aluminum impurity atoms in hydrogen-implanted silicon

Yu. V. Gorelkinskiĭ, B. N. Mukashev, and Kh. A. Abdullin

Physicotechnical Institute of the Ministry of Science, Academy of Sciences of the Republic of Kazakhstan, 480082 Almaty, Kazakhstan

(Submitted March 20, 1997; accepted for publication October 2, 1997)
Fiz. Tekh. Poluprovodn. **32**, 421–428 (April 1998)

Single-crystal samples of aluminum-doped silicon implanted with hydrogen at ~80 K were examined by electron spin resonance (ESR). Two new ESR spectra labeled Si-AA15 and Si-AA16 were observed. The hyperfine structure of the spectra unambiguously reveals that two aluminum atoms and one aluminum atom are incorporated in the AA15 and AA16 defects, respectively. Observation of Al–Al pairs (the AA15 defect) is evidence in favor of long-range migration of aluminum atoms. The migration occurring in the experiment at $T \leqslant 200$ K cannot be normal or recombination-enhanced (injection-enhanced) diffusion. Tentative models including Al–Al interstitial pairs as well as a hydrogen–enhanced migration mechanism are discussed. © *1998 American Institute of Physics.* [S1063-7826(98)00404-9]

INTRODUCTION

It is well known[1] that the migration rate of interstitial atoms can be radically increased as the result of various physical processes. For example, isolated interstitial aluminum atoms Al_i, which correspond to the ESR spectrum designated Si-G18 (Refs. 1 and 2), are stable at 500 K; however, under conditions of nonequilibrium free carrier injection they anneal out at 300 K and the activation energy of the diffusion process decreases from 1.2 to 0.3 eV (Ref. 3). The annealing temperature of interstitial boron B_i (Ref. 4) and carbon atoms C_i (Refs. 5 and 6) is also reduced by more than 100 K under carrier injection conditions. Intrinsic interstitial atoms are mobile under conditions of electron bombardment even at 4.2 K.[1,2] In all these cases the energy necessary for activation of the migration process is transferred to the interstitial atom upon recombination of the nonequilibrium electrons and holes by means of different electron–phonon interaction mechanisms.[3]

Another example of enhanced diffusion is hydrogen-enhanced diffusion. When samples were subjected to hydrogen-plasma treatment at temperatures ~350 K more rapid formation of oxygen thermal donors was observed[7] than for thermal annealing at the same temperature. An analogous effect was observed in samples annealed in a hydrogen atmosphere at $T \sim 900-1300$ °C with subsequent quenching at room temperature.[8] These results are evidence of the catalytic action of hydrogen atoms on the diffusion of interstitial oxygen. Theoretical models have been proposed to explain the microscopic mechanism of hydrogen-enhanced diffusion of an oxygen impurity.[9,10]

In the present paper we report the observation of enhanced diffusion of aluminum atoms in hydrogen-doped silicon. Migration of aluminum atoms was inferred from the appearance of a previously unknown ESR spectrum (designated as Si-AA15) arising from a defect whose structure contains two almost equivalent aluminum atoms. Of course, the formation of an Al–Al pair can take place only in the presence of diffusion of aluminum atoms. Such migration, which takes place in our experiments at least at temperatures $T < 200$ K, cannot be recombination-enhanced diffusion. The necessity of implanting hydrogen atoms in the samples to form Al–Al pairs is evidence of hydrogen-enhanced diffusion.

EXPERIMENT

Samples of float-zone silicon (FZ-Si) doped with aluminum to a concentration of 5×10^{16} cm^{-3} were studied. Samples with dimensions $15\times1\times0.5$ mm were cut from a silicon wafer oriented in the {111} plane, where the long side of the samples corresponded to the ⟨011⟩ axis. A hydrogen impurity was implanted in the samples in two ways: by ion-implantation of protons and by annealing in water vapor. In the first case the samples were bombarded by 30-MeV protons. The ion beam was directed perpendicular to the plane of the sample. Since the mean free path of 30-MeV protons in silicon is ~5 mm, to dope the samples with hydrogen atoms an aluminum absorber of thickness ~4 mm was placed in front of the sample. To obtain a more uniform proton distribution over the sample, additional aluminum foils of different thickness (50–300 μm) were placed in front of the sample during bombardment. Allowing for the straggling width of 30-MeV protons in silicon (~150 μm) and for the typical total dose 3×10^{15} cm^{-2}, the mean volume concentration of hydrogen after implantation is found to be $\sim5\times10^{17}$ cm^{-3}. Implantation was performed at a temperature of ~80 K, after which the samples were transported without intermediate heating into a cryostat for the ESR measurements. In the second method of implanting hydrogen atoms in the samples the starting samples were placed in a sealed quartz cell containing a few milligrams of water and were annealed at ~1200 °C for roughly an hour at room temperature. Such a heat treatment leads to doping of the

 © 1998 American Institute of Physics

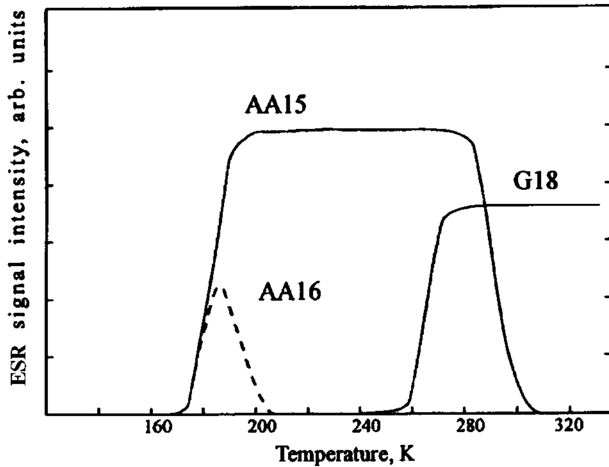


FIG. 1. Temperature region of formation and annealing of AA15 and AA16 centers in a *FZ*-Si(Al) sample implanted at 80 K with a dose of protons corresponding to a volume concentration of hydrogen equal to $5\times10^{17}\ \mathrm{cm}^{-3}$.

samples with hydrogen to a concentration of $\sim10^{16}\ \mathrm{cm}^{-2}$ (Ref. 11). These samples together with control samples were bombarded at ~80 K by protons without the absorber. Since the mean free path of 30-MeV protons is almost an order of magnitude greater than the thickness of the sample, all the protons braked beyond the sample and bombardment introduced only radiation defects uniformly distributed throughout the volume of the sample. We also used samples bombarded with ~40-MeV α particles as controls. All ESR measurements were performed at a sample temperature of 77 K at a frequency of ~37 GHz in the absorption regime.

EXPERIMENTAL RESULTS

After hydrogen implantation at 80 K and before annealing, only known centers are present in the ESR spectra. In contrast to the case of electron bombardment,[1,2] ESR spectra of interstitial aluminum Si-G18 were not observed immediately after hydrogen bombardment or implantation in either the hydrogen-doped samples or the control samples. Isochronous annealing in hydrogen-doped samples in the temperature range 180–200 K led to the appearance of previously unknown spectra, denoted as Si–AA15 and Si–AA16 (Fig. 1). The intensities of these spectra were higher in samples in which hydrogen had been introduced by ion implantation to a concentration of $\sim(1-5)\times10^{17}\ \mathrm{cm}^{-3}$, and increased in proportion to the hydrogen concentration. In samples saturated with hydrogen at high temperatures (with hydrogen concentration $\sim10^{16}\ \mathrm{cm}^{-3}$), the intensity of the AA15 and AA16 spectra was roughly 3–5 times lower, did not depend on the quenching regime, and was the same for both the air-cooled and water-quenched samples. The AA15 and AA16 spectra are absent in samples that had not been subjected to a preliminary heat treatment in water vapor at ~1200 °C and were bombarded with α particles or protons without an absorber.

The AA15 spectrum consists of 11 groups of lines. Figure 2a shows two groups of lines: on the low-field side and on the high-field side. The observed structure of the spectrum indicates that two nuclei (α and β) with spin $I=5/2$, natural abundance 100%, and similar hyperfine interaction (HFI) parameters enter into the makeup of the AA15 center. Such nuclei in our samples can belong only to ^{27}Al atoms. Each group of lines corresponds to a certain value of the projection m_I of the total nuclear spin $I=I_\alpha+I_\beta$. The spectrum AA16 can be described as arising from the anisotropic symmetry defect C_{1h} by the spin Hamiltonian

$$H=\mu_B\mathbf{HgS}+\sum_{j=\alpha,\beta}\mathbf{SA}_j\mathbf{I}_j, \tag{1}$$

where the first term describes the electron Zeeman interaction, and the second term describes the hyperfine interaction. As was noted, the hyperfine interaction parameters of the nuclei are similar but not equal. The nonequivalence of the α and β nuclei is clearly visible from the structure of the internal groups of the lines corresponding to $m_I=4$, 3, and 2. In the case of equivalent nuclei, splitting within each group would arise due to removal of degeneracy for lines with different values of the total nuclear spin I, and the lines within each group would not be equivalent and the shift of each line would be proportional to $I(I+1)-m^2$. In our case the positions of the lines are satisfactorily described by a formula which takes second-order corrections into account separately[12] for the α and β nuclei:

$$g\beta H\cong h\nu_0-\sum_{j=\alpha,\beta}\{A_jm_j+(\langle A_j\rangle)^2/2g\beta H)[I_j(I_j+1)-m_j^2]\} \tag{2}$$

(where β is the Bohr magneton and ν_0 is the frequency of the microwave field); therefore, splitting inside the groups is due to the difference in the hyperfine interaction tensors—A_α and A_β. From an analysis of the experimental angular dependence of the AA15 spectrum we determined the principal values and axes of the g, A_α, and A_β tensors, shown in Fig. 3 and in Table I. All three tensors have symmetry C_{1h}. The maximum discrepancy between the calculated (Fig. 2b) and experimental angular dependence of the AA15 spectrum is observed for $\mathbf{H}\|\langle111\rangle$ and amounts to ~0.3 mT. The angular dependence of the g-factor and the hyperfine splitting for different equivalent positions of the AA15 center, calculated using the constants of the g and A tensors, is shown in Fig. 4. The maximum values of A_α and A_β lie on the $\langle011\rangle$ axis; i.e., the corresponding molecular orbitals are oriented perpendicular to the symmetry plane of the defect.

The hyperfine interaction data were analyzed in the same way as was done for other defects in silicon,[13] in the approximation of linear combination of the $(3s, 3p)$ orbitals for the wave function of the paramagnetic electron of the AA15 center. Noting the parameters of the wave function of the aluminum atom: $|\psi_{3s}(0)|^2=20.4\times10^{24}\ \mathrm{cm}^{-3}$ and $\langle r_{3p}^{-3}\rangle=8.95\times10^{24}\ \mathrm{cm}^{-3}$ (Ref. 13), we calculated that the degree of localization of the wave function at the α and β nuclei is roughly 18 and 10%, respectively, and the contribution of the $3s$ state is ~50% (see Table I).

The ESR spectrum of the AA16 center consists of six groups of lines, which indicates the presence of one aluminum atom in the structure of this defect. Figure 5 shows two

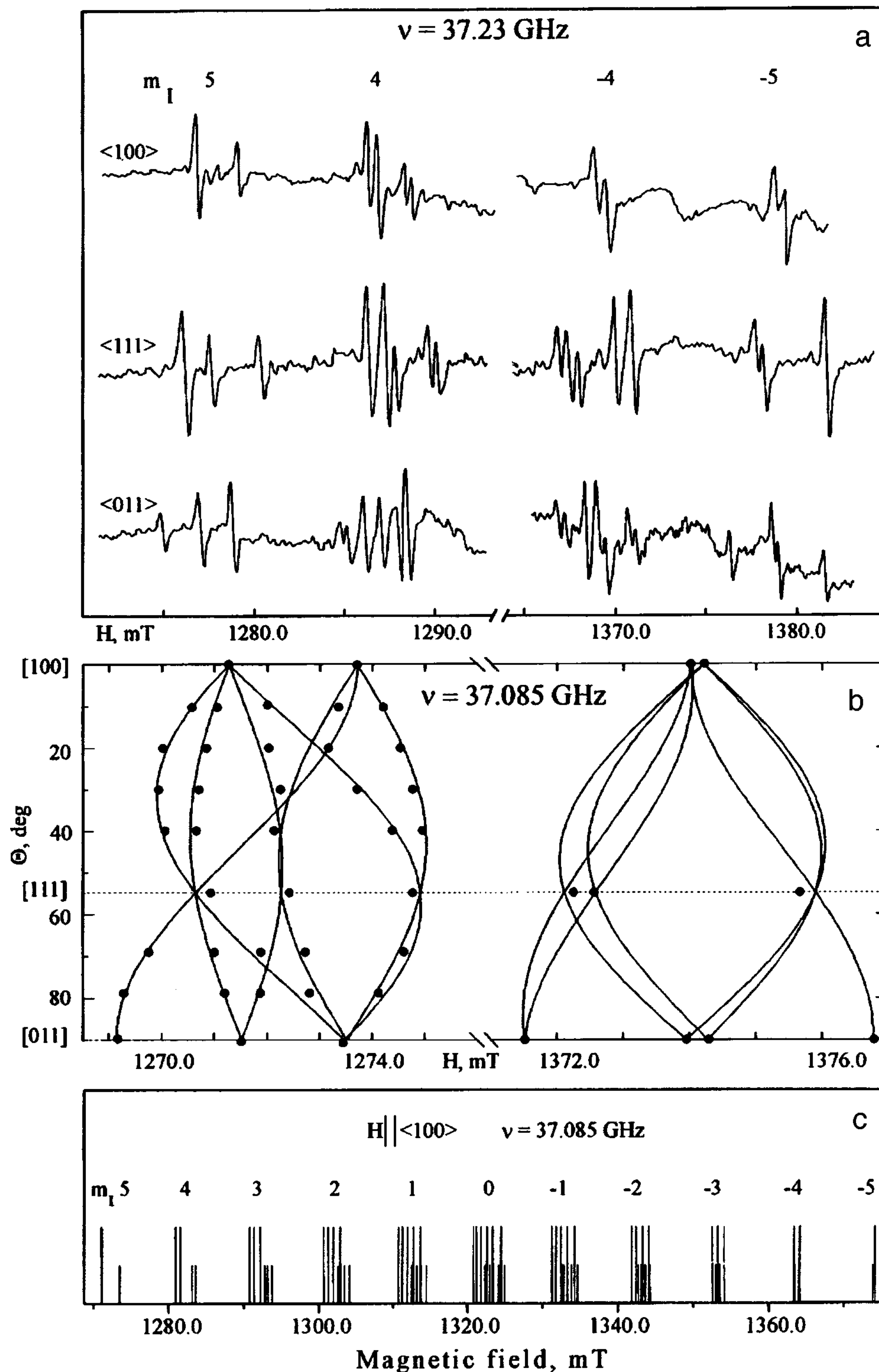


FIG. 2. a—ESR spectra of Si-AA15 in a FZ-Si(Al) sample implanted with protons at ~80 K and annealed at 200 K for 10 min. Groups of lines with total nuclear spin projection $m_i=5$ and 4 (lower-magnetic-field side of the spectrum) and -5, -4 (higher-magnetic-field side of the spectrum) for three orientations of the sample. Measurement temperature 77 K. The nuclear spin projection values indicated assume that the hyperfine interaction constant is positive. b—experimental values (points) and calculated angular dependence (solid curves) of line positions of the AA15 spectrum with $m_i=5$ and -5 (calculated with the constants of the g- and A-tensors from Table I). c—general form of the spectrum for the orientation $H\|\langle 100\rangle$, obtained using the parameters from Table I.

of the six groups of lines on the low-field and high-field sides. For the AA16 spectrum it was found that the g-factor has a weak anisotropy with mean value $g=2.0035$; the hyperfine interaction tensor has trigonal symmetry (see Table I). The degree of localization of the wave function at the Al nucleus is ~30%, and the contribution of the $3s$ state is ~80%.

The signal-to-noise ratio for the AA15 and AA16 spectra was insufficient to record forbidden transitions and to account for the contribution of the quadrupole interaction for

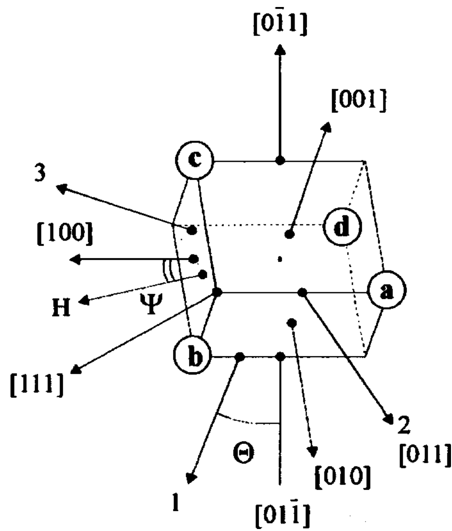


FIG. 3. Principal values and orientations of the axes of the g-, A_α-, and A_β tensors for one of the equivalent positions (cb in Fig. 4) of the AA15 defect.

the ^{27}Al nuclei. Ignoring this contribution to the Hamiltonian probably is the cause of the noticeable ($\sim$0.3 mT) discrepancy between the calculated and experimental angular dependences of the AA15 and AA16 spectra. However, if we take into account that for the magnetic field oriented along, say, the [111] axis in the (011) crystal plane the AA15 spectrum breaks up into 108 lines, then it turns out that the intensity of the AA15 spectrum is dominant in comparison with the other defects. The amplitude of the spectrum does not vary up to the annealing temperature 300 K. After annealing at room temperature the AA15 spectrum disappears. Note that in both the hydrogen-doped and the control samples the Si-G18 (Al_i) spectrum appears after annealing at

TABLE I. Parameters of the ESR spectra of the AA15 and AA16 centers and coefficients of the molecular wave function (α^2—isotropic part, β^2—anisotropic part of the hyperfine interaction, η^2—degree of localization of the wave function).

ESR Spectrum	$g(\pm 0.0003)$		A, MHz	α^2	β^2	η^2
		$^{27}Al_\alpha$	$A_1 = 276.8$			
	$g_1 = 2.0008$		$A_2 = 318.0$	0.35	0.65	0.18
AA15	$g_2 = 2.0025$		$A_3 = 290.1$			
	$g_3 = 2.0035$		$\Theta = 20 \pm 1°$			
	$\Theta = 70 \pm 1°$					
		$^{27}Al_\beta$	$A_1 = 267.3$			
			$A_2 = 285.5$	0.57	0.43	0.10
			$A_3 = 277.5$			
			$\Theta = 33 \pm 1°$			
AA16	$g = 2.0035$	^{27}Al	$A_\parallel = 823$	0.80	0.20	0.29
			$A_\perp = 836$			

260–280 K, and no correlation is observed in the behavior of the Si-AA15 and Si-G18 centers. As was noted above, the AA16 spectrum appears after annealing at 200 K simultaneously with the AA15 spectrum. However, repeated heating of the sample to 200 K leads to irreversible disappearance of the spectrum. Such a narrow temperature region of existence of the defect may be an indication that the AA16 center is an intermediate or metastable state of the defect in the formation of the Al–Al pairs (the AA15 center).

DISCUSSION

Formation of the AA15 defect, which includes two Al impurity atoms in its structure, is proof that migration of the Al atoms by a distance greater than 100 lattice constants takes place at temperatures below 200 K. The experimental data provide evidence of the important role of the hydrogen

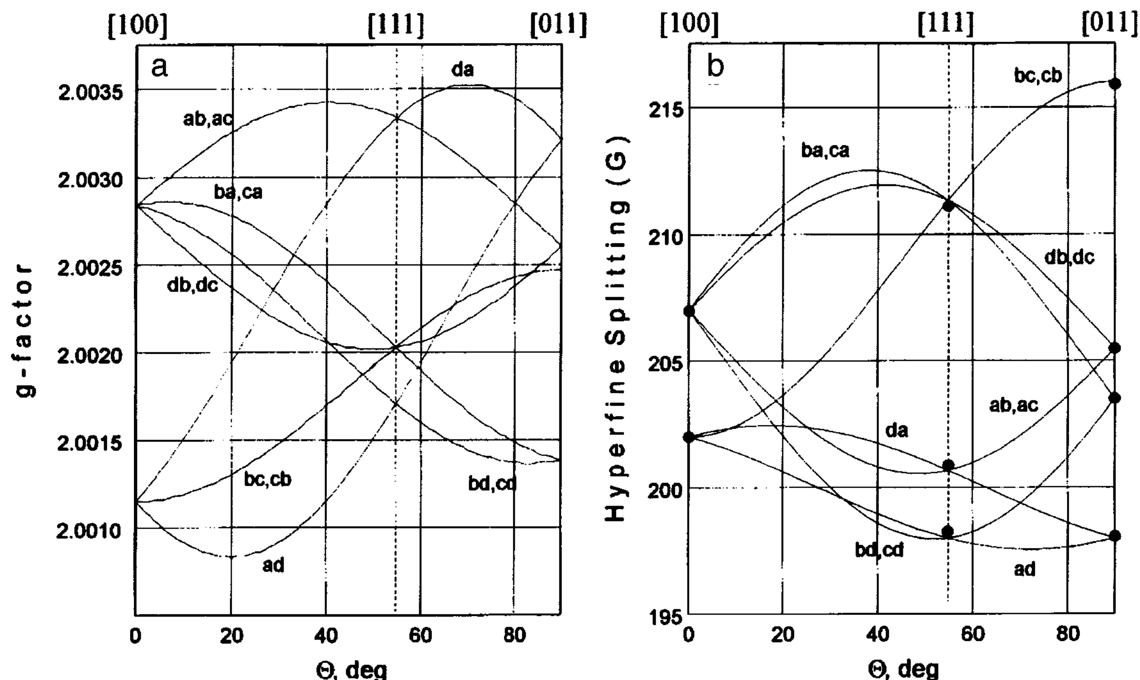


FIG. 4. Calculated curves of the angular dependence of the g-factor (a) and hyperfine splitting ($A = A_\alpha + A_\beta$) (b) for the AA15 center.

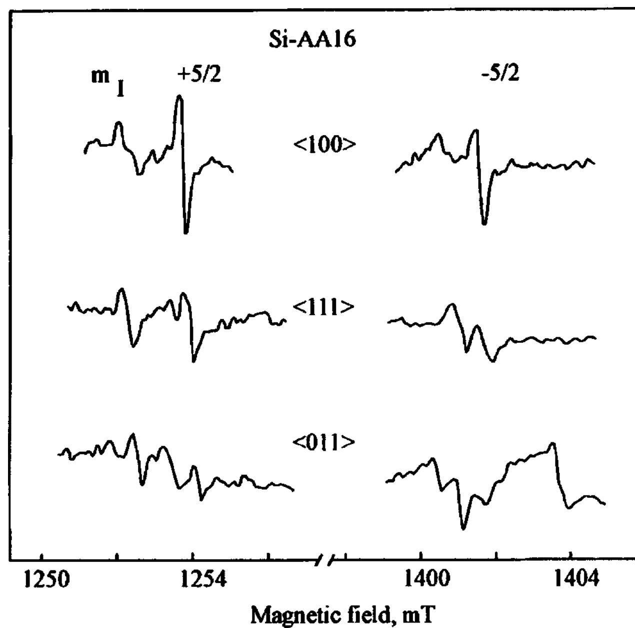


FIG. 5. High-field and low-field parts of the Si-AA16 spectrum for an *FZ*-Si(Al) sample implanted with protons at 80 K and annealed at 190 K. The nuclear spin projection values assume that the hyperfine interaction constant is positive.

atoms in this process. Although bombardment of the sample with high-energy particles generates nonequilibrium electrons and holes, the long mean-free-path migration of aluminum atoms which we have observed cannot be injection-enhanced. Even assuming that during bombardment the conditions of saturation of the annealing rate of the Al_i centers for the given injection rate are achieved, the characteristic anneal time for the interstitial aluminum atoms, according to the data of Ref. 2, is $\sim 10^9$ h at ~ 80 K. On the other hand, the characteristic time during which injection-enhanced diffusion can take place cannot exceed the bombardment time, i.e., in our case ~ 1 h. Consequently, the injection-enhanced mechanism cannot be the reason for diffusion of aluminum atoms. Diffusion also cannot be bombardment-enhanced since the AA15 center was not observed in the control samples bombarded with α particles or protons. Which mechanism is then responsible for the enhanced diffusion of the aluminum impurity?

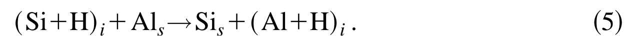
As is well known, Al_i-Al_s impurity pairs (G19 and G20 centers[1]) are formed as a result of displacement by intrinsic interstitials (Si_i) of the Al_s lattice-site atoms into the T_d interstices according to the substitution reaction[6]

$$Si_i + Al_s \rightarrow Si_s + Al_i \quad (3)$$

and subsequent migration and capture of Al_i centers by Al lattice-site atoms at temperatures ~ 500 K. In our case, the temperature at which Al–Al impurity pairs (the AA15 defect) are formed is substantially lower (Fig. 1). The structure of the AA15 defect also differs radically from the familiar G19 and G20 centers since it includes almost-equivalent aluminum atoms. Al_i defects, which always precede the appearance of the familiar Al_i-Al_s impurity pairs, are not observed. The necessity for the presence of hydrogen atoms is evidence of a hydrogen-enhanced mechanism of aluminum atom migration. Diffusion probably occurs in the form of the Al+H complex. It is well known[14] that hydrogen can be captured by acceptors of a third group (B, Al, Ga) situated at lattice sites, and that it can passivate their electrical activity, forming Al_s+H complexes. However, there are no reports in the literature of the capacity of Al_s+H complexes for diffusion. Therefore, it is obvious that aluminum diffusion proceeds via interstitial positions. In this case, hydrogen is probably bound to the interstitial aluminum atom, forming the complex Al_i+H. The absence of interstitial Al_i atoms (the G18 spectrum) leads us to the necessity of assuming that the AA15 complex is formed not in two steps, i.e., after reaction (3) and

$$Al_i + H \rightarrow (Al-H)_i , \quad (4)$$

but directly as a result of the reaction



$$(Si+H)_i + Al_s \rightarrow Si_s + (Al+H)_i . \quad (5)$$

As is well known, interstitial Si_i atoms possess a very high mobility in p-Si under conditions of electron bombardment,[1] and the interstitial impurity defects Al_i, B_i, and C_i are observed directly after low-temperature (4.2 K) bombardment. However, when bombarding with protons or α-particles the Al_i and C_i centers, monitored by ESR and deep-level relaxation spectroscopy (deep-level transient spectroscopy, DLTS), appear only after annealing the p-Si samples at ~ 280 K.[15,16] This does not mean that Si_i atoms are immobile under conditions of proton or α-particle bombardment since, as we have shown,[16] immediately after proton bombardment at ~ 80 K $(Si-O)_i$ centers are observed in n-type as well as p-type samples, which is evidence of migration of Si_i defects. Of course, for the formation of interstitial impurity defects it is necessary that two conditions be satisfied: not only the possibility of migration of intrinsic interstitials, but also the overcoming of an energy barrier to force an impurity atom into an interstitial position. The striking difference in the behavior of interstitial impurity centers may be a consequence of the energy barrier for a substitution reaction. The energy barrier depends on the chemical nature of the impurity and on its charge state, i.e., on the Fermi level. As confirmation of this assertion we may cite the results of studies of *FZ*-Si samples bombarded with electrons, where interstitial carbon atoms (C_i) appear during low-temperature bombardment in p-type samples, whereas in n-type samples they appear only after annealing at 160 K.[15,17] Note that the injection level, which is different for different types of bombardment (e.g., for electron and proton bombardment) and which depends on the energy and intensity of the beam, can also influence the defect distribution via the charge states and thereby the efficiency of formation of the interstitial impurity centers.

Thus, in the absence of hydrogen atoms the substitution reaction (3) takes place at 260–280 K under conditions of proton and α-particle bombardment, but hydrogen catalyzes the reaction and substantially lowers the temperature at which the Al_i centers are formed. However, as a result of reaction (5), the Al_i centers are bound up with hydrogen in the $(Al+H)_i$ complex. This reaction can occur during im-

plantation at ~80 K. However, it is more probable that it proceeds during thermal annealing at ~200 K. As is well known,[18] at these temperatures the hydrogen atoms occupying positions at the bond (*BC*) become mobile and, consequently, can form complexes with defects and impurities. To explain the fact of the formation of Al–Al pairs (the AA15 center), it is necessary to assume that $(Si-H)_i$ and $(Al-H)_i$ complexes are mobile at these temperatures.

Note that the AA16 center is a good candidate for the role of the $(Al-H_i)$ defect. The different width of the lines of the AA16 spectrum for different equivalent positions of the defect (Fig. 5) may be a consequence of a weak hyperfine interaction at the hydrogen nucleus. The nature of the hyperfine interaction at the ^{27}Al nucleus for the AA16 center shows that the degree of localization of the wave function of the paramagnetic electron on the Al atom is similar to the case of an Al_i center at a T_d interstice (Si–G18) and the wave function has a preferentially $3s$ character. Thus, the Al atom in the AA16 defect is found in the Al^{++} charge state as in the defects G18–G20 (Ref. 1). The insignificant contribution of the $3p$ state indicates a perturbation shifts the Al atom away from the T_d interstice. Such a perturbation can create, for example, a negative hydrogen ion (H^-), which is located at a neighboring T_d interstice. Note that the $(Al-H)_i$ defect can be efficiently formed as a result of the long-range Coulomb attraction between the components forming this complex. The absence of a strong hyperfine interaction at the H nucleus can be caused by weak localization of the wave function at the hydrogen atom. In the given model it is necessary to explain the capability of this defect for migration. Theoretical studies of the process of enhanced migration of Al_i atoms[19] shows that the energy barrier for diffusion is lowered from 1.3 eV for the charge state Al^+ to 0.5 eV for Al^-. Therefore, electron injection leads to recharging of the Al_i centers and a substantial enhancement of the process of migration. Of course, the given mechanism is inapplicable in our case. Further studies will be directed at a more careful examination of the AA16 defect and the process of diffusion enhancement with hydrogen participation.

Let us discuss a possible model for the AA15 center. The familiar ESR centers G19 and G20, identified as Al_i-Al_s pairs, have a significantly different degree of localization of the wave function on the Al_i and Al_s atoms. At an aluminum atom located at a T_d interstice, the localization of the wave function is ~38%, whereas at a lattice-site aluminum atom the degree of localization of the wave function is ~3–4%. Thus, at G19 and G20 centers substantially different positions of the Al atoms entrained in the Al–Al pair are clearly manifested. In the case of the AA15 center the Al atoms are almost equivalent and have the same symmetry of the hyperfine interaction (C_{1h}). It can be assumed that the aluminum atoms are found at positions that are equivalent in the undamaged lattice, and the lowering of the symmetry from C_{2v} to C_{1h} is due either to the Jahn–Teller effect or to the presence of some additional defect located on the periphery of the structure of the AA15 center. (Although hyperfine splitting from hydrogen was not observed in the AA15 spectrum, we cannot completely rule out the participation of the hydrogen atom in the structure of this defect.) In addition,

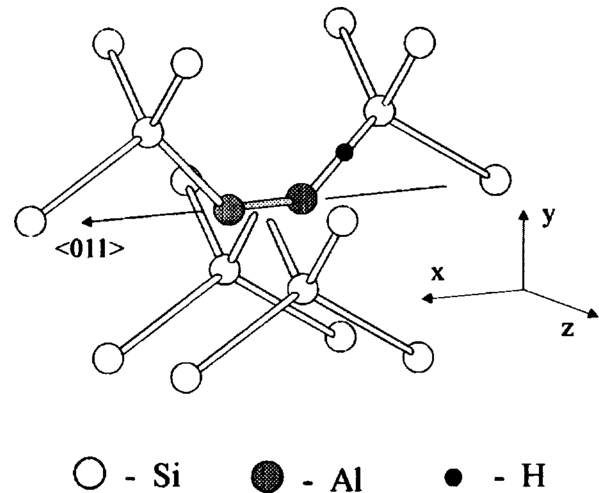


FIG. 6. Tentative model of an AA15 center–⟨011⟩-split interstitial.

the two Al atoms should lie symmetrically about the C_{1h} {011} plane. Therefore, the possible interstitial variants of the model of the AA15 complex are restricted to centers based at two types of split interstices: ⟨011⟩ and ⟨001⟩, whose symmetry is lowered to C_{1h} by the additional perturbation.

In the case of a ⟨001⟩-split interstice it may be expected that the paramagnetic electron will be localized mainly on the $3p$ orbital. This follows from a comparison with the C_i (Si-G12), $C_s-Si_i-C_s$ (Si-G11), and C_i-O_i (Si-G15) centers,[20–22] in which the paramagnetic electron is situated at a molecular orbital, ~95% of which comes from the $2p$ orbital of the carbon atom. Therefore, the model of a ⟨001⟩-split interstice is in poor agreement with the HFI data for the AA15 center. In contrast, in the case of the ⟨011⟩-split interstice model (Fig. 6) it is easy to construct a molecular orbital from a mixture of the $3s$ and $3p_z$ orbitals of the aluminum atom. Here the s and p_x orbitals form the Al–Al bond, the p_x and p_y orbitals form the bonds with the neighboring silicon atoms, and the E orbital, formed from the p_z and s states, is occupied by the paramagnetic electron. Therefore, the hyperfine interaction in the direction perpendicular to the symmetry plane of the defect is maximum, in agreement with experiment. Figure 6 shows one possible model of the AA15 center, whose structure contains one pair of Al atoms. The symmetry (C_{2v}) of such a center can be lowered to C_{1h} by the presence of a hydrogen atom on the periphery of the defect or by the Jahn–Teller distortion (if hydrogen is not entrained in the structure of the defect). Note that the symmetry of the hyperfine interaction with two almost equivalent aluminum atoms observed for the AA15 center can also be obtained in the model of a complex, whose nucleus is a vacancy with two lattice-site aluminum atoms or a pair of Al atoms in a T_d interstice. The symmetry of such complexes (C_{2v}) can be lowered to C_{1h} by an additional perturbation. However, the formation of such a complex requires the interaction of more than one radiation defect with each other, which renders its formation improbable. Further studies are required to establish the exact structure of the AA15 center.

Note that theoretical calculations[23] give similar values of the defect energy in the case of trigonal[1,2] Al_i-Al_s pairs and

orthorhombic pairs which are not observed experimentally. On the other hand, two types of trigonal pairs are observed; namely, G19 and G20. These results may be evidence of peculiarities of the electronic structure of defects in which aluminum atoms participate and of the presence of several energy minima of similar depth in the configuration diagram, which accounts for the variety of observed structures for Al–Al impurity pairs.

CONCLUSIONS

The phenomenon of low-temperature hydrogen-enhanced diffusion of Al impurity atoms has been observed. Such diffusion is evidenced by the formation of Al–Al impurity pairs in hydrogen-doped silicon, into which radiation defects have been introduced at ~80 K. The formation of Al–Al pairs takes place at low temperatures (<200 K) and does not depend on the means of introducing hydrogen into the samples, whether it is ion doping or high-temperature (~1200 °C) annealing in water vapor. Presumably, migration of aluminum impurity atoms takes place in the form of a the $(Al–H)_i$ complex (ESR spectrum AA16). For the ESR center AA15 the best agreement with experiment is achieved for the model of the ⟨110⟩-split interstice $(Al–Al)_i$.

[1] G. D. Watkins, Mater. Sci. Forum **143–147**, 9 (1994).
[2] G. D. Watkins, *Radiation Damage in Semiconductors* (Dunod, Paris, 1964) p. 97.
[3] J. R. Troxell, A. P. Chatterjee, G. D. Watkins, and L. C. Kimerling, Phys. Rev. B **19**, 5336 (1979).
[4] J. R. Troxell and G. D. Watkins, Phys. Rev. B **22**, 921 (1980).
[5] L. C. Kimerling, P. Blood, and W. M. Gibson, Inst. Phys. Conf. Ser. **N 46**, 273 (1978).
[6] Kh. A. Abdullin, B. N. Mukashev, and M. F. Tamendarov, Fiz. Tekh. Poluprovodn. **25**, 684 (1991) [Sov. Phys. Semicond. **25**, 412 (1991)].
[7] R. C. Newman, in *20th Int. Conf. Phys. Semicond.*, edited by E. M. Anastassakis, J. D. Joannopoulos (World Scientific, Singapore. 1991), Vol. 1, p. 332.
[8] S. A. McQuaid, C. A. Londos, M. J. Binns, R. C. Newman, and J. H. Tucker, Mater. Sci. Forum **143–147**, 963 (1994).
[9] R. Jones, S. Öberg, and A. Umerski, Mater. Sci. Forum **83–87**, 551 (1992).
[10] S. Estreicher, Phys. Rev. B **41**, 9886 (1990–I).
[11] M. J. Binns, S. A. McQuaid, R. C. Newman, and E. C. Lightowlers, Semicond. Sci. Technol. **8**, 1908 (1993).
[12] E. L. Elkin and G. D. Watkins, Phys. Rev. **174**, 881 (1968).
[13] G. D. Watkins and J. W. Corbett, Phys. Rev. **134**, A1359 (1964).
[14] S. J. Pearton, J. W. Corbett, and M. Stavola, Springer Ser. Mater. Sci. **16** (1992).
[15] Kh. A. Abdullin and B. N. Mukashev, Fiz. Tekh. Poluprovodn. **28**, 1831 (1994) [Semiconductors **28**, 1012 (1994)].
[16] Kh. A. Abdullin, B. N. Mukashev, Yu. V. Gorelkinskiĭ, Semicond. Sci. Technol. **11**, 1696 (1996).
[17] R. G. Harris and G. D. Watkins, in *13th Int. Conf. Def. in Semicond.* (Coronado, California, 1985) p. 799.
[18] Yu. V. Gorelkinskiĭ and N. N. Nevinnyi, Physica B **170**, 155 (1991).
[19] G. A. Baraff and M. Schluter, Phys. Rev. B **30**, 3460 (1984).
[20] G. D. Watkins and K. L. Brower, Phys. Rev. Lett. **36**, 1329 (1976).
[21] K. L. Brower, Phys. Rev. B **9**, 2607 (1974).
[22] J. M. Trombetta and G. D. Watkins, Appl. Phys. Lett. **51**, 1103 (1987).
[23] M. Hagen and H. Overhof, Mater. Sci. Forum **143–147**, 1197 (1994).

Translated by Paul F. Schippnick

Thermal expansion and characteristic features of the strength of interatomic bonds in melts of III–V compounds (AlSb, GaSb, InSb, GaAs, InAs)

V. M. Glazov and O. D. Shchelikov

Moscow Institute of Electronic Engineering, 103498 Moscow, Russia

(Submitted April 2, 1997; accepted for publication October 9, 1997)
Fiz. Tekh. Poluprovodn. **32**, 429–431 (April 1998)

The temperature dependence of the specific volume of melts of III–V compounds has been studied thermometrically and with penetrating γ radiation. The thermal expansion of these melts has been estimated at various temperatures. Based on the similarity of the structure of the melts and the elastic continuum, the characteristic Debye temperatures and rms dynamic displacements of the atoms in the close-range-order structure of these melts have been calculated from the estimated thermal expansion. A considerable change in the indicated characteristics is observed at the transition from the solid state to the liquid state, indicating significant changes in the vibrational spectrum of III–V compounds upon melting. The thermal expansion of the melts is observed to increase upon heating, indicating a further loss of strength of the interatomic bonds in the melts with growth of temperature. © *1998 American Institute of Physics.* [S1063-7826(98)00504-3]

The electrical and magnetic properties indicate that the melting of III–V semiconductor compounds is accompanied by an abrupt increase in the charge-carrier concentration. As a result, the electrical conductivity reaches a value that is typical of molten metals and according to Regel's classification, III–V compounds melt according to the type semiconductor–metal.[1–4]

The high-temperature properties of III–V compounds have not been studied in adequate detail. At the same time, the results of such studies are of interest in connection with the development of technology and instrumentation for obtaining samples of compounds in the form of single crystals and epitaxial structures.

Specifically, the results of studies of the bulk properties and thermal expansion can be used in thermal-physical calculations of the design characteristic of reactors in view of the anomalous (of water type) bulk changes of III–V compounds upon their transition from the liquid state to the solid state.[5]

The density of III–V compounds in a wide temperature range, including the liquid phase, have been studied extensively.[1–5] However, the high-temperature region, especially that for gallium and indium arsenides, received inadequate attention. To obtain a deeper understanding of the physical-chemical nature of melts of these compounds, a detailed study was undertaken of the temperature dependence of the specific volume of their melts with subsequent calculation of such characteristics of the strength of the interatomic bond as the thermal expansion coefficient, the characteristic Debye temperature, and the rms dynamic displacements of atoms from their equilibrium positions.

To solve this problem, we synthesized the compounds InSb, GaSb, AlSb, InAs, and GaAs. In our experiments we used macrocrystalline (GaAs, InAs) or single-crystal (InSb, GaSb, AlSb) samples of these compounds.

The density of the melts (with the exception of GaAs) was determined thermometrically, by tracking the level of the melt in a calibrated, specially shaped quartz cell. In the case of AlSb the cell was coated on the inside by a mirror layer of pyrocarbon. The density of GaAs was determined by penetration γ radiation. The design of the devices and the technique of the experiments are described in more detail in Refs. 4 and 6.

Figure 1 plots the temperature dependence of the specific volume (V_{sp}) for the investigated compounds.

The measurement results were reduced by the least-squares method, which made it possible to represent them in the form of first-degree interpolation formulas characterizing the temperature dependence of the density over the investigated temperature interval

$$\frac{1}{V_{sp}}=d=A+B(T-T_m). \tag{1}$$

Here T_m is the melting temperature. Table I lists the values of the coefficients A and B in equations such as (1) for all the investigated III–V compounds.

On the basis of the obtained results we calculated the thermal expansion coefficients from the equation

$$\alpha=\frac{1}{3V_{sp}}\left(\frac{\partial V_{sp}}{\partial T}\right)_P. \tag{2}$$

The temperature dependence of the thermal expansion of melts of the investigated compounds is plotted in Fig. 2.

From the thermal expansion data we calculated the characteristic Debye temperatures, using for this purpose a relation obtained by Sirota,[7] which combines the Lindeman–Borellius and Grüneisen equations:

$$\theta=19.37(\bar{A}V_M^{2/3}\alpha)^{-1/2}. \tag{3}$$

 © 1998 American Institute of Physics

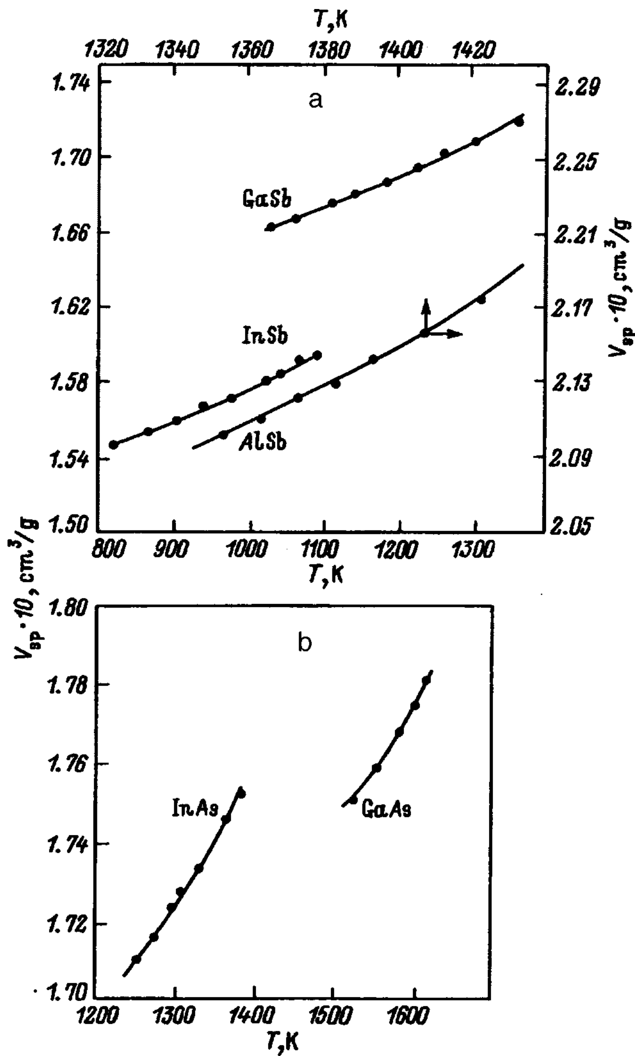


FIG. 1. Temperature dependence of the specific volume of III–V alloys: a—AlSb, GaSb, and InSb; b—GaAs and InAs.

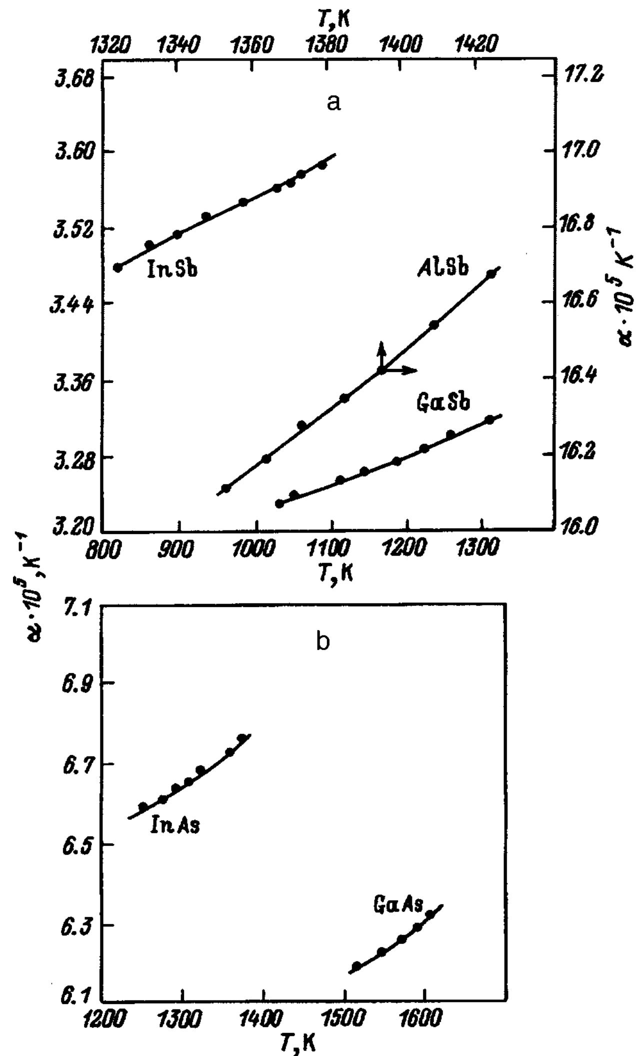


FIG. 2. Temperature dependence of the thermal expansion of III–V alloys: a—AlSb, GaSb, and InSb; b—GaAs and InAs.

Here $\bar{A}$ is the rms atomic weight, and V_M is the molecular volume.

It should be noted that the calculation of the Debye temperatures using Eq. (3), according to Refs. 8 and 9, is completely applicable to melts since the Lindeman–Borellius and Grüneisen relations, like the concept of a characteristic temperature, were derived for an elastic isotropic continuum, which has much more in common with a liquid than with a crystal even with cubic structure. Knowing the values of the Debye temperatures allowed us to calculate the rms dynamic displacements of the atoms from their equilibrium positions in the melt. For these calculations we used the Debye–Waller relation

$$\bar{u}^2 = 4.3\times 10^{-14}\left[\frac{D(\theta/T)}{\theta/T} + \frac{1}{4}\right] \bigg/ \bar{A}\cdot\theta, \tag{4}$$

where $D(\theta/T)$ is the Debye function, whose values are given in Ref. 10.

It should be noted that in a rigorous approach expression (4) can be used for comparatively low temperatures ($T < \theta/8$). For higher temperatures ($T > 1.6\theta$) the expression[10]

$$\bar{u}^2 = 1.52\times 10^{-20}\cdot T(V^{1/3}E)^{-1} \tag{5}$$

is recommended to calculate the rms dynamic displacements of the atoms from their equilibrium positions. Here V is the molar (atomic) volume, and E is the elasticity modulus, whose determination for a melt is problematic.

At the same time, it is well known[10] that Young's modulus is related to the Debye temperature by the relation

TABLE I. Values of the coefficients in equations of type (1) for the temperature dependence of the density of alloys of III–V compounds.

Compound	$A \pm \Delta A$	$-(B \pm \Delta B)\times 10^3$
InSb	6.466±0.0212	0.675±0.128
InAs	5.880±0.024	1.155±0.251
GaSb	6.033±0.016	0.582±0.082
GaAs	5.71 ±0.0016	1.06 ±0.039
AlSb	4.75 ±0.081	2.27 ±0.643

TABLE II. Values of the thermal expansion coefficients, Debye temperatures, and rms dynamic displacements of atoms from their equilibrium positions in alloys of III–V compounds.

Compound	$\alpha_L \times 10^4$, K^{-1}	θ_L, K	$\alpha_s \cdot 10^6$, K^{-1}	θ_s, K	$\sqrt{\bar{u}^2}$, Å	$\bar{m}\theta^2 \cdot 10^{18}$, r·$K^2$
InSb	0.348	115	5.9	250	0.471	2.57
InAs	0.650	96	5.2	322	0.765	1.430
GaSb	0.320	135	2.4	274	0.484	2.90
GaAs	0.620	124	5.6	476	0.764	1.84
AlSb	1.59	62	4.9	341	1.319	0.484

$$\theta = \frac{1.68\times 10^3 \cdot \sqrt{E}}{\bar{A}^{1/3} d^{1/6}}, \tag{6}$$

where $\bar{A}$ is the rms atomic weight, and d is the density. Determining the value of Young's modulus from expression (6) and substituting it in Eq. (5) gives

$$\bar{u}^2 = \frac{4.29\times 10^{-14} T}{\bar{A}\theta^2}. \tag{7}$$

Note that Eqs. (4) and (7) yield the rms value $\sqrt{\bar{u}^2}$; obviously it should be different for the subsystems A and B. However, noting the small differences in the atomic masses of the atoms in the investigated compounds, we ignore possible differences in the values of $\sqrt{\bar{u}^2}$.

We also estimated the values of the quantity $\bar{m}\theta^2$, which characterizes the energetics of the close-range-order structure of the corresponding melts.

Results of our calculations using Eqs. (2)–(7) at the melting temperature of the corresponding III–V compounds are presented in Table II. Note that the calculations of $\sqrt{\bar{u}^2}$ using Eqs. (4) and (7) give practically identical results.

Comparison of these data with the corresponding quantities for the crystalline III–V compounds at room temperature shows that the thermal expansion in the liquid phase is significantly greater and the Debye temperature is significantly smaller than in the solid phase. The rms dynamic displacements of the atoms from their equilibrium positions in the melts also grow considerably. All these facts indicate a substantial weakening of the cohesive forces between the particles upon transition of the considered materials from the solid state to the liquid state. The thermal expansion increases considerably with increasing temperature (Fig. 2). This indicates a further decrease in the strength of the interatomic bonds in the melts of these compounds upon heating.

The substantial differences in the characteristics of the strength of the interatomic bond in the solid and liquid phases indicate that melting produces very serious changes in the nature of the vibrational spectrum of the atoms of the considered materials.

In summary, our studies of thermal expansion of melts of III–V compounds have allowed us to estimate the characteristics of the strength of the interatomic bond and to obtain values of the average Debye temperatures, the thermal expansion, and the rms dynamic displacements of the atoms from their equilibrium positions in the close-range-order structure. These results, especially when compared with the corresponding values for the crystalline materials at low temperatures, shed some light on the physical nature of the changes taking place during the crystal–melt phase transition.

[1] A. R. Regel', in *Structure and Physical Properties of Materials in the Liquid State* [in Russian] (Kiev State Univ., Kiev, 1954), p. 117.
[2] A. R. Regel' and V. M. Glazov, *Periodic Law and Physical Properties of Electronic Alloys* [in Russian] (Nauka, Moscow, 1978).
[3] A. R. Regel' and N. P. Mokrovskiĭ, Zh. Tekh. Fiz. **22**, 1281 (1952).
[4] V. M. Glazov, S. N. Chizhevskaya, and N. N. Glagoleva, *Liquid Semiconductors* [in Russian] (Nauka, Moscow, 1967).
[5] V. M. Glazov, S. N. Chizhevskaya, and S. B. Evgen'ev, Zh. Fiz. Khim. **43**, 378 (1969).
[6] V. M. Glazov, M. Vobst, and V. I. Timoshenko, *Methods of Studying the Properties of Liquid Metals and Semiconductors* [in Russian] (Metallurgiya, Moscow, 1989).
[7] N. N. Sirota and S. N. Chizhevskaya, *Physics and Physical-Chemical Analysis* [in Russian] (Gosstroĭizdat, Moscow, 1957), p. 138.
[8] V. M. Glazov, M. Kasymova, and A. R. Regel', Fiz. Tekh. Poluprovodn. **13**, 2049 (1979) [Sov. Phys. Semicond. **13**, 1196 (1979)].
[9] V. M. Glazov and O. D. Shchelikov, Fiz. Tekh. Poluprovodn. **18**, 662 (1984) [Sov. Phys. Semicond. **18**, 411 (1984)].
[10] I. N. Frantsevich, in *Questions of Powder Metallurgy and Strength of Materials* [in Russian] (Izdat. Akad. Nauk SSSR, Kiev, 1956), Issue No. III, p. 3.

Translated by Paul F. Schippnick

ELECTRONIC AND OPTICAL PROPERTIES OF SEMICONDUCTORS

Optical absorption and photosensitivity of $CuIn_xGa_{1-x}Se_2$ thin film structures

Yu. V. Rud'

State Technical University, 195251 St. Petersburg, Russia

V. Yu. Rud'

A. I. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia

I. V. Bodnar' and V. F. Gremenok

State University of Informatics and Radio Electronics, 220027 Minsk, Belarus
(Submitted May 26, 1997; accepted for publication May 27, 1997)
Fiz. Tekh. Poluprovodn. **32**, 432–435 (April 1998)

Thin polycrystalline films of $CuIn_xGa_{1-x}Se_2$ $(0 \leq x \leq 1)$ were fabricated by pulsed laser evaporation. Results of measurements of the optical properties, photocurrent polarization indicatrices, and spectral dependence of the photoconversion quantum yield of In–p-$CuIn_xGa_{1-x}Se_2$ structures are discussed. A window effect in the photosensitivity has been observed, and it is concluded that it is possible to use $CuIn_xGa_{1-x}Se_2$ thin films as photoconverters of solar radiation. © *1998 American Institute of Physics.*
[S1063-7826(98)00604-8]

Ternary chalcogenide compounds and solid solutions based on them are finding ever wider use in the construction of high-efficiency solar photoconverters.[1–4] On the basis of such compounds it has been possible to fabricate such structures with efficiencies as high as 17% (Ref. 4), and further progress in this area lies on the path of expanding studies of the fundamental properties of these materials closely linked with variations in the conditions of their fabrication. We report here the results of such a study of thin films in the system $CuIn_xGa_{1-x}Se_2$, which by way of controlling its atomic composition offers the possibility of a smooth ''fit'' of its parameters to the requirements of specific applications.

Polycrystalline films of this system were fabricated by pulsed laser sputtering of a target consisting of a single-phase material with the required ratio of indium and gallium concentrations. The films were deposited on a heated glass substrate, whose temperature could be varied within the range[5] 300–500 °C.

Let us examine the results of studies of the optical properties of films of several compositions. The films had an area of 1×3 cm^2 and possessed good adhesion to the glass surface.

1. Studies of the optical transmittance T and reflectance R were carried out in natural and linearly polarized radiation.

Figure 1 shows typical spectral curves of the optical transmittance of films of several compositions; their parameters are given in Table I. For films with thicknesses in the range $d\simeq1$ μm a rather high transmittance $T\simeq30$–50% in the transparency region is typical. Variations in the value of T for fixed values of x in the range 25–30% is common for these films and does not have a direct link with the temperature of the deposition process T_S (see Table I). Most likely, these variations are due to variations in the optical quality and homogeneity of the films. For higher-quality films $R\simeq30\%$ for all values of x from 0 to 1. Therefore, we took this value as our basis in the analysis of the absorption data as it turned out to be close to the known value of R for bulk crystals of similar atomic composition.[1]

As can be seen from Fig. 1, in the spectral range shown there is no clear manifestation of interference in the curves, while at photon energies $\hbar\omega>1$ eV in the films with $x=1$ and 0.8, and also for $\hbar\omega>1.25$ eV for films with $x=0.6$ a sudden exponential decrease of T appears. This decrease is probably due to the appearance of interband optical transitions. The shift of the exponential edge T toward shorter wavelengths as the composition is varied from $x=1$ and 0.8 to $x=0.6$ can be attributed to an increase in the width of the band gap of the solid solutions.[1,3] Comparison of the spectral curves for the two films with $x=0.8$ (Fig. 1, curves *2* and *3*), which were deposited on substrates with different temperature T_S allows us to conclude that the optical transmittance is sensitive to this parameter. On the basis of these spectral curves (T), taking multiple reflection into account, we calculated the optical absorption of $CuIn_xGa_{1-x}Se_2$ films using the standard formula[6,7]

$$\alpha=\frac{1}{d}\ln\left\{\frac{(1-R)^2}{2T}+\sqrt{\left[\frac{(1-R)^2}{2T}\right]+R^2}\right\}.$$

Spectral curves of α for the investigated films are shown in Fig. 2. They are typical for bulk and film samples of ternary chalcogenides[1,2] and are characteristic of semiconductors with direct optical transitions; they are straight lines in the coordinates $(\alpha\hbar\omega)^2-\hbar\omega$ (Fig. 3). By extrapolating

 © 1998 American Institute of Physics

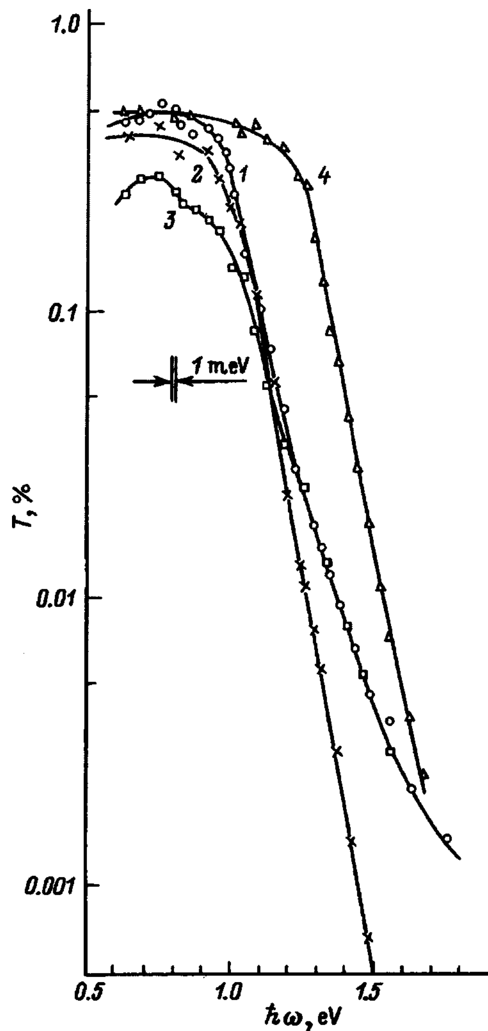


FIG. 1. Spectral dependence of the optical transmittance of thin films of the system $CuIn_xGa_{1-x}Se_2$ at $T=300$ K in natural light (*1*—sample 2, *2*—sample 3, *3*—sample 4, *4*—sample 5).

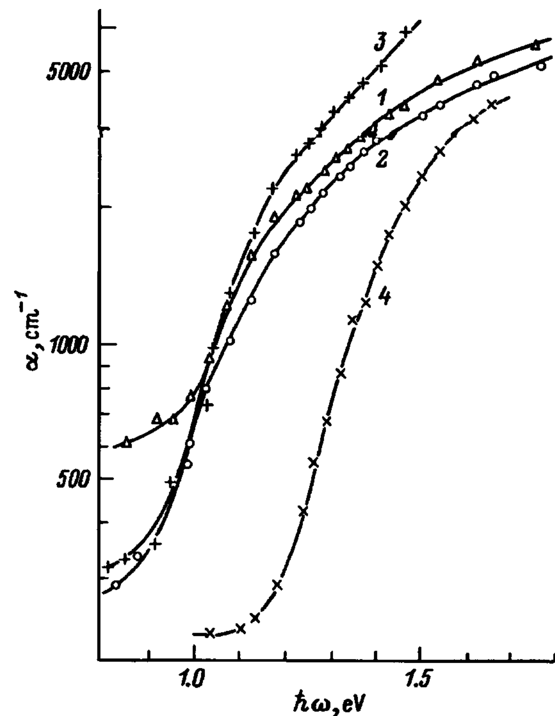


FIG. 2. Spectral dependence of the optical absorption of $CuIn_xGa_{1-x}Se_2$ at $T=300$ K (notation the same as in Fig. 1).

the quantity $(\alpha\hbar\omega)^2$ to zero we determined the width of the band gap E of the investigated films from the intercept on the energy axis (see Table I). These values are similar to known values for films based on ternary chalcogenides,[1,2] and the observed overshoot of the known value of the width of the band gap $E_G=1.03$ eV for $CuInSe_2$ reflects the dependence of this parameter on the temperature and other conditions of deposition, established by many authors.[1] This is probably the reason why the width of the band gap was similar in the films with $x=1$ and 0.8. Our main conclusion, which follows from the data of Fig. 3, is that changes in x do not affect the structure of the band spectrum and the interband transitions remain direct, whereas the width of the band gap in the solid-solution films depends on the conditions of deposition and grows as x is lowered below 0.8. The observed decrease of E_G for the films with $x=0.8$ relative to the films with $x=1$ (see Fig. 3 and Table I) may also be evidence of fluctuations in the atomic composition of the film relative to the starting bulk material.

TABLE I. Optical properties of $CuIn_xGa_{1-x}Se_2$ films.

Sample No.	x	T_s, °C	d_l, μm	T, % ($\hbar\omega=0.8$ eV)	E_G, eV $(\alpha\hbar\omega)^2\to 0$
2	1	380	1.3	49	1.14
3	0.8	380	1.2	37	1.09
4	0.8	490	1.2	27	1.10
5	0.6	490	1.4	49	1.36

Since reports of achieving a high degree of texturization can be found in a number of works on laser-deposited films,[8,9] we made an effort to observe the optical absorption anisotropy commonly encountered in chalcopyrite semiconductors. If the films are textured in such a way that the crystals in the film are oriented preferentially in the (112) crystallographic plane, then on the basis of Ref. 10 one should expect anisotropy in their edge absorption. Figure 4 shows typical polarization indicatrices T for one of the investigated films for four values of the photon energy. Such dependences are observed in the transparency region and deep in the fundamental absorption region and are characteristic of all the examined compositions. From these results we may assume that texturing is not manifested in our films, and that the absence of a polarization effect on the optical absorption may be a consequence of the polycrystalline structure of $CuIn_xGa_{1-x}Se_2$ thin films obtained by laser evaporation.

2. In addition to the optical absorption of these thin films, we also examined the photosensitivity of structures based on them. With this goal in mind, we deposited semitransparent layers of indium ($\simeq 1$ μm) through a template. A diagram of the structures obtained in this way is shown in the inset in Fig. 5. To make it possible to determine the homogeneity of the photosensitivity, we created several

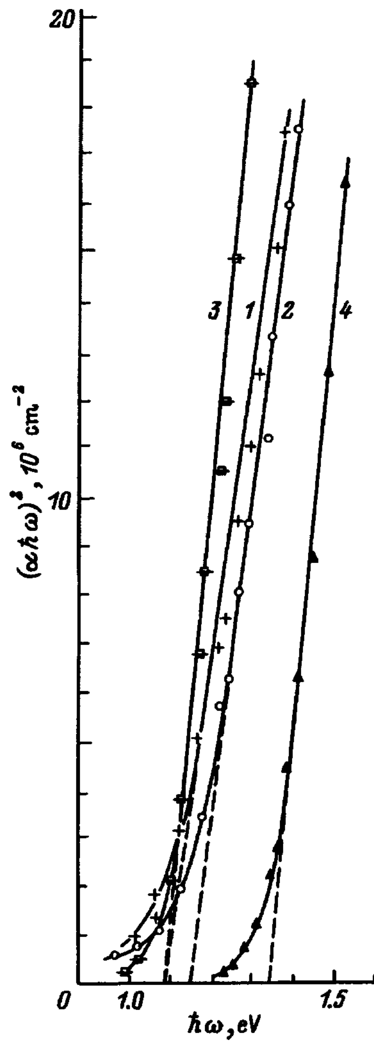


FIG. 3. Spectral dependence of $(\alpha\hbar\omega)^2$ for films of the system $CuIn_xGa_{1-x}Se_2$ at $T=300$ K (x: *1*—1, *2,3*—0.8, *4*—0.6. T_S, °C: *1,2*—380, *3,4*—490).

$In/CuIn_xGa_{1-x}Se_2$ structures on the film each having an area of 2×2 mm.

The structures were illuminated both with natural and with linearly polarized light normal to their photodetector plane from the side with the indium layer. The photosensitivity was measured in the short-circuit photocurrent (i) regime, where the photoresponse is proportional to the flux density of the incident light. The photocurrent spectra were scaled to the number of incident photons and are therefore equal to the relative quantum efficiency of photoconversion η. For the better structures the maximum absolute photosensitivity reached 3 mA/W at $T=300$ K and mainly for the films with $x=0.6$.

Typical spectral curves of η are shown in Fig. 5 and their main features can be described as follows.

The long-wavelength edge of the photosensitivity for the fabricated surface-barrier structures is exponential and is formed in the interband absorption region of the film material. Its slope $s=\delta(\ln\eta)/\delta(\hbar\omega)$ turns out to be quite high, reaching 25–60 eV^{-1}, and reflects the direct nature of the interband optical transitions. Spectral curves of η for struc-

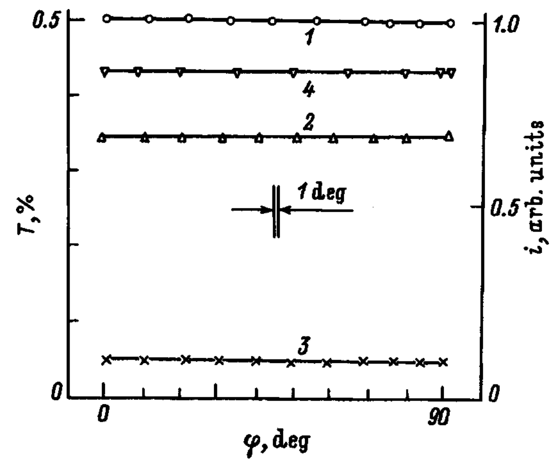


FIG. 4. Polarization indicatrices of the optical transmittance of the $CuIn_{0.6}Ga_{0.4}Se_2$ film (*1–3*) and short-circuit photocurrent (*4*) of the structure In/p-$CuIn_{0.6}Ga_{0.4}Se_2$ at $T=300$ K for illumination by linearly polarized light. (Sample 5. $\hbar\omega$, eV: *1*—0.7, *2*—1.2, *3,4*—1.4. Illumination of the structure $In/CuIn_{0.6}Ga_{0.4}Se_2$ from the side of the barrier contact along the normal to the photodetector plane.)

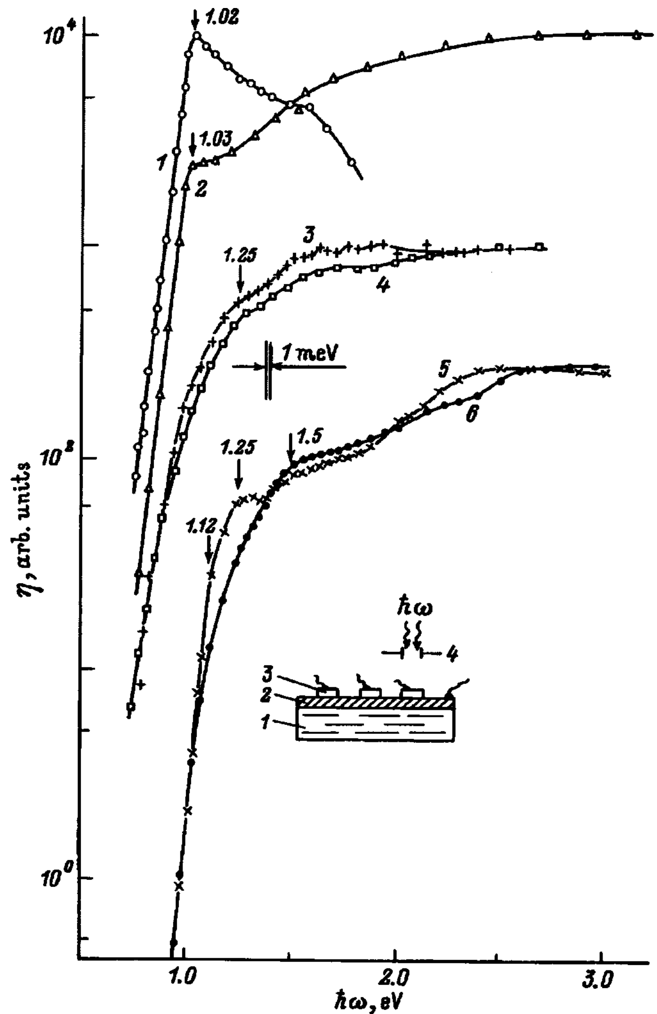


FIG. 5. Spectral dependence of the relative quantum yield η of In/p-$CuIn_xGa_{1-x}Se_2$ surface-barrier structures in natural light at $T=300$ K. *1,2*—structures on different segments of film 2 ($x=1$); *3*—structure on film 4 ($x=0.8$); *5,6*—structures on different segments of film 5 ($x=0.6$). T_s, °C: *1–3*—380, *4–6*—490. Inset: diagram of illumination and design of structures: *1*—substrate, *2*—film, *3*—energy barrier, *4*—movable diaphragm.

tures fabricated on different segments of the $CuIn_xGa_{1-x}Se_2$ films usually reveal some differences in their spectral shape (Fig. 5, curves *1* and *2* and also curves *5* and *6*), although the energy position of their features are usually well reproduced. From a comparison of the spectral curves of η for films deposited at different substrate temperatures T_S (Fig. 5, curves *3* and *4*) it is possible to conclude that the observed differences are not large and of the same order as those encountered in scanning different segments of the film. On the whole, the obtained spectra seem to indicate that the photosensitivity of the surface-barrier structures is determined by the energy spectrum of $CuIn_xGa_{1-x}Se_2$. The absence of a short-wavelength falloff in the spectral curves of all the $In/CuIn_xGa_{1-x}Se_2$ structures (Fig. 5, curves *2–6*) indicates that laser-pulsed deposition yields solid-solution films of the given system of adequately high quality.

The photosensitivity of the surface-barrier structures is in complete agreement with the optical transmittance polarization data and is also isotropic. The polarization indicatrices of the photocurrent i_φ of such structures for normally incident, linearly polarized light (Fig. 4, curve *4*) are straight lines. The established insensitivity of the photocurrent to the orientation of the polarization plane, defined by the azimuthal angle φ, accords with the above-stated conclusion of the absence of texture in the laser-deposited films.

In summary, the method of pulsed-laser deposition[5] allows one to obtain homogeneous polycrystalline $CuIn_xGa_{1-x}Se_2$ films which can be used to build high-efficient solar photoconverters, where the photoactive absorption region can be controlled by the atomic composition of the film.

This work was supported by INTAS (Grant No. 94-3998).

[1] T. J. Coutts, L. L. Kamerskiik, and S. Wagner, [Eds.], *Copper Indium Diselenide for Photovoltaic Applications* (Amsterdam, 1986).
[2] Y. Bougnot, S. Bochemin, and M. Savelli, Sol. Cells **16**, 221 (1986).
[3] H.-W. Schock, Appl. Surf. Sci. **92**, 606 (1996).
[4] V. Nadenau, D. Braunger, D. Hariskos, M. Kaiser, Ch. Koble, A. Oberacker, M. Ruckh, U. Ruhle, R. Schaffler, D. Schmid, T. Walter, Z. Zweigart, and H.-W. Schock, in *Progress in Photovoltaic Research and Applications*, Vol. 3, p. 363 (1995).
[5] V. F. Gremenok, E. P. Zaretskaya, I. V. Bodnar, Yu. V. Rud, and M. A. Magomedov, Thin Solid Films **232**, 139 (1993).
[6] Yu. I. Ukhanov, *Optical Properties of Semiconductors* [in Russian] (Nauka, Moscow, 1977).
[7] I. S. Gorban', *Optics* [in Russian] (Vishcha Shkola, Kiev, 1979).
[8] E. Armed, A. E. Hill, J. Leppavouti, R. D. Pilkington, R. D. Tomlinson, J. Levoska, and O. Kusmartseva, Adv. Mater. Opt. Electron. **4**, 423 (1994).
[9] Y. Yamamoto, T. Yamaguchi, Y. Demiru, T. Tanaka, A. Ganjao, and A. Yoshida, in *Proc. 10th Int'l. Conf. on Ternary and Multinary Compounds* (Stuttgart, 1995).
[10] I. V. Bodnar', A. A. Vaĭpolin, V. Yu. Rud', and Yu. V. Rud', Fiz. Tekh. Poluprovodn. **28**, 1332 (1994) [Semiconductors **28**, 748 (1994)].

Translated by Paul F. Schippnick

On trapping of minority current carriers in *n*-type $Cd_xHg_{1-x}Te$ at low temperatures

S. G. Gasan-zade and G. A. Shepel'skiĭ

Institute of Semiconductor Physics, Ukrainian National Academy of Sciences, 252650 Kiev, Ukraine
(Submitted June 2, 1997; accepted for publication July 14, 1997)
Fiz. Tekh. Poluprovodn. **32**, 436–438 (April 1998)

The photoconductivity and photoelectromagnetic effect was measured under conditions of uniaxial elastic deformation in *n*-type $Cd_xHg_{1-x}Te$ crystals. The increase in the current-carrier lifetime in the low-temperature range ($T<40-50$ K) is shown to be due to trapping of minority current carriers (holes) in shallow acceptor-type attachment levels rather than interband Auger recombination. © *1998 American Institute of Physics.* [S1063-7826(98)00704-2]

Recombination mechanisms in the narrow-gap semiconductors $Cd_xHg_{1-x}Te$ ($x=0.2-0.3$) — materials of infrared (IR) photoelectronics for the range 3–5 and 10–12 μm — have now been well studied (see, for example, the review article[1]). Specifically, this refers to *n*-type material and the conventional working temperature range of photodetectors (77–300 K). The results are not so clear for low temperatures (4.2–50 K). In particular, there exist different points of view concerning the reasons for the rapid increase of the current-carrier lifetime τ with decreasing temperature $T<40$ K in *n*-type crystals. The increase in τ can be explained (just as in the intrinsic conductivity region) by predomination by interband Auger recombination.[2,3] This stems from the fact that $\tau_A(T)$ contains an exponential factor $\exp((b/1+b)\cdot(\varepsilon_g/kT))$, which ensures that the laws of conservation of energy and momentum are satisfied in recombination transitions. Here $b=m_v/m_c$ is the ratio of the effective masses of heavy holes and conduction-band electrons.

On the other hand, trapping of minority current carriers (holes) in shallow acceptor-type attachment levels can also lead to rapid growth of τ.[4]

Simultaneous measurement of the photoconductivity (PC) and photomagnetic effect (PME) in the same sample could be helpful for determining the recombination mechanism in the case described. Indeed, in *n*-type material τ_{PM} is close to the lifetime of minority current carriers (holes), while τ obtained from measurements of the PC (τ_{PC}) corresponds to the electron lifetime. The large difference between the values of τ_{PC} and τ_{PM} ($\tau_{PM}<\tau_{PC}$ always being satisfied) rules out interband Auger recombination as the main mechanism of relaxation of the nonequilibrium current carrier density. However, the fact that the values of τ_{PM} and τ_{PC} are not the same is still not an indication of the existence of attachment of minority carriers, since the relation $\tau_{PM}<\tau_{PC}$ can also be explained by the fact that the electron and hole trapping cross sections of a deep recombination level are different in the Shockley–Read model.

To obtain an unequivocal result it may be helpful in this situation to apply to the crystal a mechanical uniaxial stress. This suggestion is based on the selective character of the effect of a directed elastic deformation on the electronic states in the conduction band and donor levels, on the one hand, and the valence band and acceptor levels, associated with it, on the other, in crystals with cubic symmetry. Indeed, a uniaxial stress has virtually no effect on the conduction band, if the small anisotropy introduced is neglected. Donors associated with the conduction band are likewise virtually unaffected by deformation. The situation is completely different in the case of the valence band, which in the absence of deformation is degenerate at the point $k=0$ — the point of tangency of the heavy- and light-hole bands. A uniaxial stress lifts the degeneracy of the valence band by lowering the symmetry of the crystal. As a result, a fourfold degenerate valence band, which takes into account the spin, splits into two subbands with anisotropic hole effective masses.[5] The magnitude of the splitting between the bands at the point $k=0$ is

$$\varepsilon_p=\frac{1}{4}|b|(s_{11}-s_{12})\cdot\chi \quad \text{for deformation direction } \chi\|[100],$$

$$\varepsilon_p=\frac{1}{8\sqrt{3}}|d|s_{44}\chi \quad \text{for deformation direction } \chi[111].$$

Here b and d are constants in the deformation potential, s_{ij} are the components of the elastic compliance tensor, and χ is the magnitude of the uniaxial stress.

For weak deformations, when $E_P<E_A$, a fourfold degenerate acceptor level likewise splits into two levels, and the magnitude of the splitting E_A is of the order of the splitting of the valence band. In ordinary semiconductors, for example, Ge and Si, because of the small ratio m_h/m_l of the heavy- and light-hole effective masses, a uniaxial deformation has only a very small effect on the parameters of the valence band and acceptor levels. For $Cd_xHg_{1-x}Te$ ($x=0.20$) $m_h/m_l>50$. For this reason, the splitting of the subbands, the change in the hole effective masses in the bands, and the change in the depth of the acceptors for the same values of the deformation are much stronger.[6] Indeed, even for $\chi=1$ kbar the splitting of the bands is $E_P=12$ meV for deformation along direction [100]. This is greater than the depth of the shallow acceptors in the initial crystal. In this case, at low temperatures the carriers occupy only the

 © 1998 American Institute of Physics

upper subband of the split valence subbands, while the ground state of the acceptor is formed by the energy states at the top of this band. The energy of the acceptor can be sought as the energy of a Coulomb center in a simple anisotropic band with effective masses

$$m_{\parallel}=(\gamma_1+2\bar{\gamma})m_0 \quad \text{and} \quad m_{\perp}=(\gamma_1-\bar{\gamma})m_0 .$$

Here γ_1, $\bar{\gamma}=\gamma_2=\gamma_3$ are the Luttinger parameters. The heavy-hole mass in the initial crystal is $m_h=(\gamma_1-2\bar{\gamma})^{-1}m_0$. Since the average effective mass of the upper subband of the split subbands is much less than m_h in the initial crystal, this increases the Bohr radius a of the acceptor center and decreases the ionization energy of the center. For zero deformation $a_0=\hbar(2m_h\cdot\varepsilon_A)^{-1/2}$. In the limit of infinite deformation $a_\infty=\hbar[(2m_\perp^2 m_\parallel)^{1/3}\cdot\varepsilon_A]^{-1/2}$.

If the acceptor denstiy in $Cd_xHg_{1-x}Te$ is sufficiently high ($N_A>10^{15}$ cm^{-3}), then uniaxial compression $\sim$4 kbar can induce a Mott transition (provided that $N_a^3=0.02$).[6] At the same time the ground state energy of the acceptor vanishes. For us it is important that a uniaxial stress results in a rapid decrease in the ionization energy of the acceptor. As a result, the acceptor level is no longer an effective hole trapping center, since the thermal exchange of carriers between the level and the valence band accelerates sharply. It was shown earlier that application of a uniaxial stress is a convenient method for distinguishing interband recombination mechanisms — radiative and Auger.[7] As a result of a decrease in the hole effective mass in the upper subband, the Auger recombination rate decreases, while the radiative rate, conversely, increases. As one can see from the estimates presented above, a uniaxial stress can reveal trapping centers for minority current carriers — holes in n-CdHgTe, and in the limit it can eliminate the trapping mechanism itself.

Experimental measurements of the PC and PME in a stationary regime under conditions of uniaxial elastic deformation were performed on a number of n-type samples. The experimental procedures for performing the measurements and preparing the samples are described in Ref. 8.

A characteristic feature of the experimental samples was the rapid growth of τ_{PC} in the temperature range $T<30-40$ K. The complete temperature dependence of τ_{PC} was obtained for n-$Cd_xHg_{1-x}Te$ samples in the range 10–300 K (Fig. 1). The temperature dependences of the PME were also recorded for some samples. As one can see from the figure, the increase in the PC and the simultaneous decrease of the PME at low temperatures can reliably attest to an impurity and not an interband character of recombination in the indicated temperature range.

Figures 2 and 3 show the deformation dependences of the PC and PME at different temperatures. A LG-126 laser (λ=0.63, 1.15, and 3.39 μm) was used as the radiation source. Qualitatively similar dependences are observed for all experimental n-$Cd_xHg_{1-x}Te$ samples at least at $T=10-30$ K: As voltage increases, the PC signal decreases and the PME increases. Changing the wavelength of the exciting radiation does not change the form of these dependences. In our view, the deformation dependences presented attest unequivocally to the presence of trapping of minority

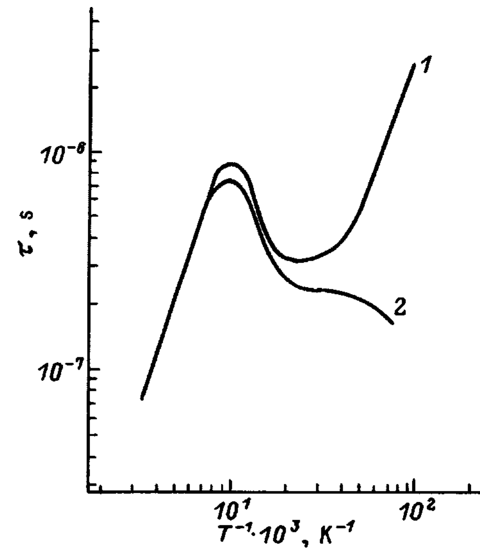


FIG. 1. Temperature dependences of the lifetime τ determined from PC (*1*) and PME (*2*) for a n-$Cd_xHg_{1-x}Te$ sample.

current carriers (holes) in n-type crystals at low temperatures. The residence time of minority carriers in an attachment level starts to decrease as the elastic stress increases. This is due to the decrease in the depth E_A of the corresponding acceptor.

In summary, the trapping of holes on shallow acceptors, rather than Auger recombination, leads to the observed growth of τ with decreasing temperature in the range $T<30-40$ K. On the basis of the temperature dependences of the PC in the indicated range, the ionization energy of these acceptors in the initial state for the experimental samples lies

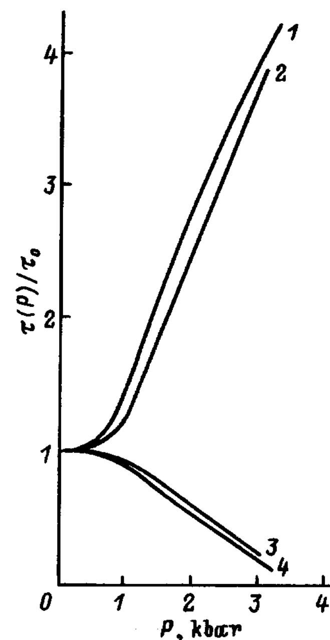


FIG. 2. Deformation dependences $\tau(P)/\tau_0$ for n-$Cd_xHg_{1-x}Te$ samples determined from PME (*1*, *2*) and PC (*3*, *4*). $T=15$ K. $\chi \parallel [100]$. *1*, *4* — $x=0.220$, $n=3.16\times10^{14}$ cm^{-3}, $\mu_n=1.13\times10^5$ cm^2/(V·s), *2*, *3* — same sample as in Fig. 1.

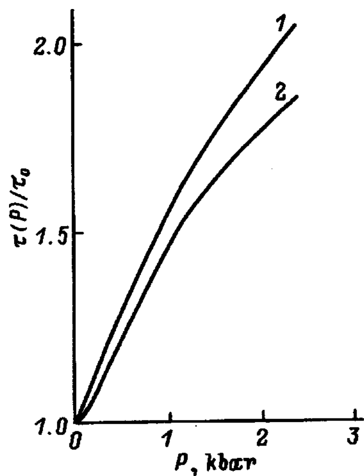


FIG. 3. $\tau(P)/\tau_0$ versus the applied compression for a n-$Cd_xHg_{1-x}Te$ sample. $T=100$ K. $\chi \parallel [100]$. *1* — Determined from the PME, *2* — PC. $x=0.220$, $n=3.16\times10^{14}$ cm^{-3}, $\mu_n=1.13\times10^5$ cm^2/(V·s).

in the range 7–10 meV, while the acceptor density N_A is less than 10^{15} cm^{-3}.

It is interesting that as the temperature increases to $T>100$ K, the sign of the changes in the PC and PME with deformation becomes the same, as expected. Indeed, in this temperature range the lifetime of the current carriers is determined by interband impact recombination (Auger recombination), and $\tau_{PC}\simeq\tau_{PM}$. In this case, τ increases because the hole effective mass in the upper band of the split valence bands decreases. As a result, the thresholds for Auger processes increase.

[1] N. S. Baryshev, B. L. Gel'mont, and M. I. Ibragimov, Fiz. Tekh. Poluprovodn. **24**, 209 (1990) [Sov. Phys. Semicond. **24**, 127 (1990)].
[2] I. Calas and I. Allegre, Phys. Status Solidi B **112**, 179 (1982).
[3] R. R. Gerhards, R. Dornhaus, and G. Nimtz, Solid-State Electron. **21**, 1467 (1978).
[4] N. L. Bazhenov, B. L. Gel'mont, V. I. Ivanov-Omskiĭ, A. A. Mal'kova, V. K. Ogorodnikov, and T. U. Totieva, Fiz. Tekh. Poluprovodn. **16**, 109 (1982) [Sov. Phys. Semicond. **16**, 64 (1982)].
[5] G. L. Bir and G. R. Pikus, *Symmetry and Strain-Induced Effects in Semiconductors*, Israel Program for Scientific Translations, Jerusalem; Wiley, N. Y., 1975 [Russian original, Moscow, 1972].
[6] A. V. Germanenko, G. M. Min'kov, V. L. Rumyantsev, O. É. Rut, and O. V. Inisheva, Fiz. Tekh. Poluprovodn. **23**, 796 (1989) [Sov. Phys. Semicond. **23**, 500 (1989)].
[7] F. T. Vas'ko, S. G. Gasan-zade, M. V. Strikha, and G. A. Shepel'skiĭ, JETP Lett. **50**, 318 (1989).
[8] S. G. Gasan-zade and G. A. Shepel'skiĭ, Fiz. Tekh. Poluprovodn. **27**, 1326 (1993) [Semiconductors **27**, 733 (1993)].

Translated by M. E. Alferieff

Short-wavelength photoluminescence of SiO_2 layers implanted with high doses of Si^+, Ge^+, and Ar^+ ions

G. A. Kachurin and I. E. Tyschenko

Institute of Semiconductor Physics, Siberian Branch of the Russian Academy of Sciences, 630090 Novosibirsk, Russia

L. Rebohle, W. Skorupa, R. A. Yankov, and H. Froeb

Rossendorf Research Center, Dresden, Germany

T. Boehme and K. Leo

Technical University of Dresden, Dresden, Germany

(Submitted January 28, 1997; accepted for publication July 16, 1997)
Fiz. Tekh. Poluprovodn. **32**, 439–444 (April 1998)

The short-wavelength (400–700 nm) photoluminescence (PL) spectra of SiO_2 layers implanted with Si^+, Ge^+, and Ar^+ ions in the dose range $3.2\times10^{16}-1.2\times10^{17}$ cm^{-2} are compared. After Ar^+ implantation an extremely weak luminescence, which vanishes completely after annealing for 30 min at 400 °C or 20 ms at 1050 °C, was observed. After implantation of group-IV elements the luminescence intensities were 1 to 2 orders of magnitude higher, and the luminescence remained not only with annealings but it could also increase. The dose and heating dependences of the luminescence show that it is due to the formation of impurity clusters and this process is more likely to be of a percolation than a diffusion character. For both group-IV impurities an intense blue band and a weaker band in the orange part of the spectrum were observed immediately after implantation. The ratio of the excitation and emission energies of the blue luminescence is characteristic of oxygen vacancies in SiO_2; its properties are determined by the direct interaction of group-IV atoms. On this basis it is believed that the centers of blue PL are chains of Si (or Ge) atoms embedded in SiO_2. The orange luminescence remained after annealings only in the case of Si^+ implantation. This is attributed directly to the nonphase precipitates of Si in the form of strongly developed nanometer-size clusters.
© *1998 American Institute of Physics.* [S1063-7826(98)00804-7]

1. INTRODUCTION

The recently discovered capability of structures based on porous Si to give strong visible-range emission has attracted extraordinary attention because of its great scientific and practical value. Different methods of constructing light-emitting nanostructures of indirect-gap Si and Ge, which differ advantageously from procedures used to fabricate porous Si by the absence of ''wet'' processes and by their compatibility with modern semiconductor technology, have recently been developed and are now being successfully refined.[1–10] Ordinarily, they are based on the decomposition of solid solutions of Si and Ge in SiO_2. At the same time, the physical nature of the sources of visible-range luminescence is still not entirely clear and continues to be a subject of debate.[6–13] Quantum-well Si and Ge crystals, defects in the surrounding matrix, interfacial states, and foreign compounds have been named as sources.

Two types of radiation sources can now be quite definitely distinguished. The sources of the first type emit predominantly in the long-wavelength part of the spectrum ($\lambda>700$ nm), they are formed after high-temperature ($T>900$ °C) annealings, and they have relaxation times up to 10^{-3} s. Their luminescence is similar to the red radiation of porous Si and is most likely due to size quantization in the nanocrystals formed after annealing. Sources of the second type emit shorter-wavelength light ($\lambda<700$ nm) of lower intensity, their formation does not require annealing, and the relaxation times lie in the submicrosecond range. The nature of centers of this radiation is less certain, since it is observed in SiO_2[14,15] and after bombardment of SiO_2 with inert-gas atoms and γ-rays,[9,16] and as a result of introducing excess Si or Ge into the oxide.[8,10,17] Depending on temperature and duration of the annealings, it can both intensify and weaken.[6–8,12,18,19]

The difficulties in identifying the sources of short-wavelength radiation largely stem from the fact that their nature is most often judged according to the appearance of luminescence bands, without regard for the intensity and excitation conditions of the bands. It is virtually impossible to compare data from different studies, since the emission spectra are recorded under different conditions, while the intensities are given in relative units. To obtain more complete and objective information about the short-wavelength luminescence from implanted SiO_2 layers we bombarded these layers by Si^+, Ge^+, and Ar^+ ions under comparable conditions and then subjected them to identical heat treatment,

 © 1998 American Institute of Physics

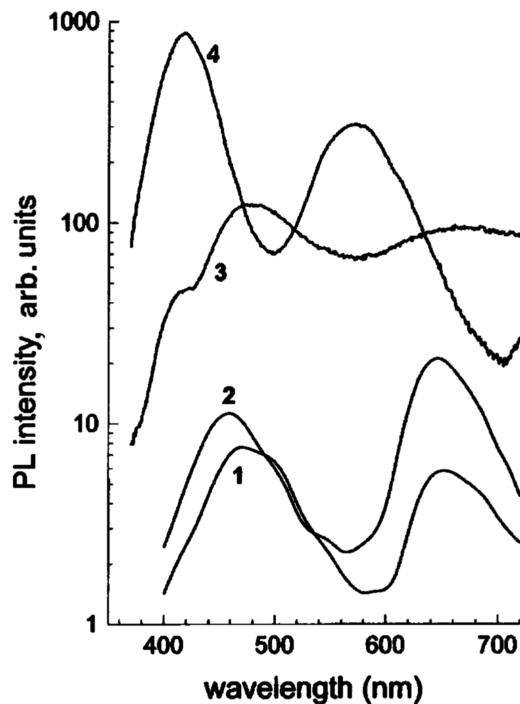


FIG. 1. Photoluminescence spectra of the initial SiO_2 layers (*1*) and of the same layers after Ar^+ (*2*), Si^+ (*3*), and Ge^+ (*4*) implantation. Total doses, 10^{16} cm^{-2}: *2* — 4.5, *3* — 4.8, *4* — 6.6.

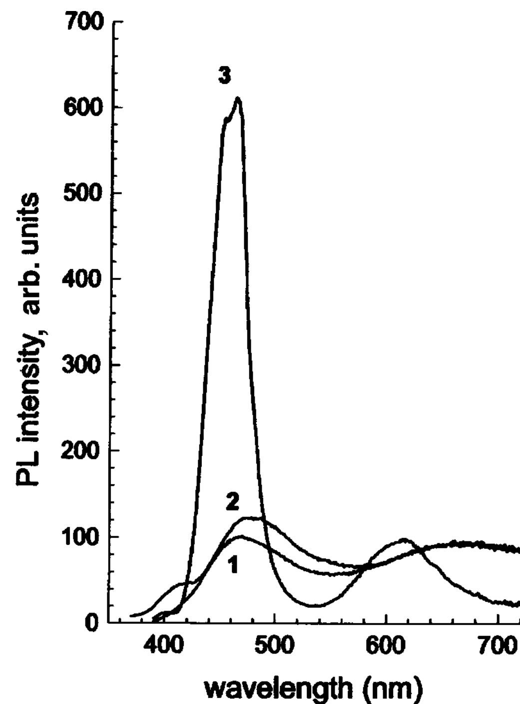


FIG. 2. Si^+ dose dependence of the photoluminescence spectra. Total doses, 10^{16} cm^{-2}: *1* — 4.8, *2* — 6.7, *3* — 12.

while the spectra were recorded on the same apparatus in the entire range of interest at a fixed excitation level.

2. PROCEDURE

Si^+, Ge^+, and Ar^+ ions were implanted in 500-nm-thick SiO_2 layers grown thermally on Si(100). Implantation was performed with two energies in order to produce a region with a relatively uniform impurity distribution at the center of the SiO_2 layer. The energies were chosen on the basis of range calculations: 200 and 100 keV for Si^+, 250 and 170 keV for Ar^+, and 350 and 200 keV for Ge^+. The doses at the high and low energies were in the ratio 1:0.6. The total Si^+ and Ge^+ doses varied from 3.2×10^{16} cm^{-2} up to 1.2×10^{17} cm^{-2}, ensuring approximately the same volume densities for both elements. The total Ar^+ dose was taken to be 4.5×10^{16} cm^{-2} in order that the number of atomic displacements produced in SiO_2 would be comparable to the effect of Si^+ and Ge^+ ions. Implantations were performed at low temperatures (from -150 °C to -100 °C), with the exception of the cases mentioned. The response of the irradiated samples to heating was studied either by 30-min anneals in a furnace at 400 °C or by exposing the samples to short heat pulses of 900 and 1050 °C for 20 ms. The photoluminescence (PL) spectra were recorded on a Spex Fluoromax apparatus in the wavelength range 400–700 nm at 20 °C. The source of exciting radiation was an ultraviolet lamp, whose required line was filtered out with a monochromator. A photomultiplier was used as a detector. The fact that the spectra were recorded under identical conditions makes it possible to compare quantitatively the PL intensities in Figs. 1–4, where the scales are graduated in the same units, and 250 nm radiation is used for excitation.

3. RESULTS

Figure 1 shows on a logarithmic scale the PL spectra before and after implantation of Si^+, Ge^+, and Ar^+ doses in SiO_2, which are comparable with respect to the doping level and defect formation. One should note first and foremost that there is a large difference in the PL intensities between the results of implantation of Si^+ and Ge^+, on the one hand, and Ar^+, on the other. For example, the peak in the blue band after Ge^+ implantation is 100 times higher than after Ar^+ implantation. In general, implantation of 4.5×10^{16} cm^{-2} Ar^+ ions gives a very small increase in PL; in addition, the bands near 460 and 650 nm, which are present in the grown SiO_2 layers initially, grow (Fig. 1, curves *1* and *2*). Second, there is an appreciable difference between the results of Ge^+ and Si^+ implantation. In the first case a very bright band at 420 nm dominates. The other band (580 nm) is almost three times weaker than the first band. Such a strong domination of one band does not occur after implantation of Si^+. Luminescence with approximately an order of magnitude higher intensity than after implantation of Ar^+ is observed in the entire visible region. It clearly consists of several lines, the strongest of which lies near 480 nm. We underscore that the differences between Si^+, Ge^+, and Ar^+ are observed after low-temperature implantation and without heating the samples above 20 °C.

The Si^+ dose dependence of PL is shown in Fig. 2. The luminescence is observed for doses above 10^{16} cm^{-2}, and additional annealing is not required for it to appear. After total doses of 4.8×10^{16} cm^{-2} and 6.7×10^{16} cm^{-2} the luminescence spectrum is broadened over the entire experimental

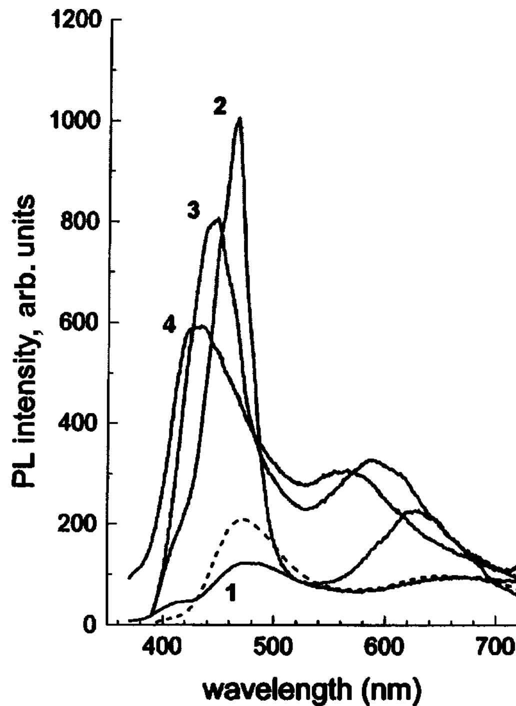


FIG. 3. Effect of annealings and implantation temperature on photoluminescence with implantation of 4.8×10^{16} cm^{-2} Si^+. (*1*) — Immediately after implantation. Annealings: 400 °C, 30 min (*2*); 900 °C, 20 ms (*3*); and 1050 °C, 20 ms (*4*). Dashed line — implantation at 20 °C.

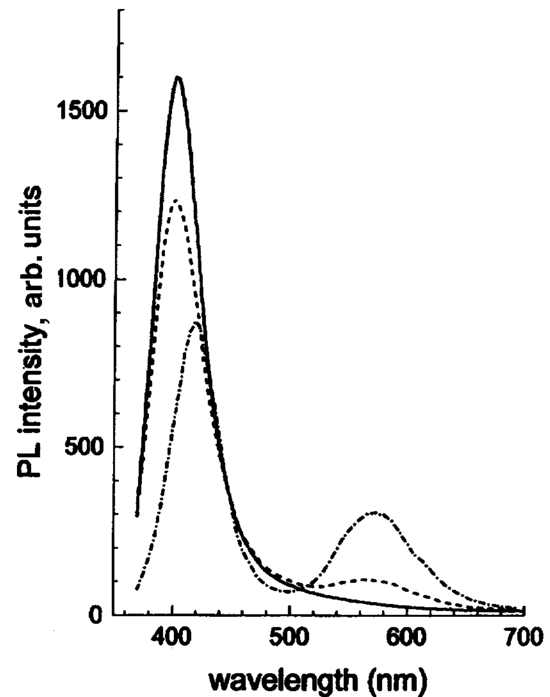


FIG. 4. Effect of annealings on the photoluminescence of SiO_2 layers after implantation of 6.6×10^{16} cm^{-2} Ge^+. Dot-dashed curve — immediately after implantation; dashed and solid curves — 900 and 1050 °C for 20 ms, respectively.

range. When the dose is increased to 1.2×10^{17} cm^{-2}, band near 450 nm and the weaker orange band acquire a pronounced intensity.

A similar transformation of the PL spectrum also occurs if instead of increasing the Si^+ dose, annealing for 30 min at 400 °C is performed (Fig. 3, curves *1* and *2*). A narrow intense peak of blue PL and a weaker and broadened orange band appear. With further annealing, the ratio of the intensities between the blue and orange PL starts to change in favor of the orange PL and both bands shift in the direction of short wavelengths. One can also see from Fig. 3 that a simple transition from low-temperature to room-temperature implantation intensifies the blue emission without any additional heat treatments.

The transformation of the spectra accompanying annealings after Ge^+ implantation is shown in Fig. 4. The intense blue band seen immediately after implantation is intensified even more by annealings. The second band, however, decreases rapidly with heating and vanishes completely after annealing for 20 ms at 1050 °C. The PL centers introduced by Ar^+ ions also vanish completely with annealings at 400 °C for 30 min and at 1050 °C for 20 ms. The Ar^+ data are not shown in the annealing plots (Figs. 3 and 4) because of the initial weakness of the signal (Fig. 1).

In addition to the PL spectra, we also investigated the PL excitation spectra. Figure 5 shows such spectra for the blue emission lines after Si^+ and Ge^+ implantation. In both cases the processes develop similarly. Before annealings the emission intensities start to increase rapidly when the wavelength of the exciting line becomes shorter — 230 nm. The behavior of these curves was virtually independent of the type of ion implanted and was of an edge character. As the annealing temperature increases, sharp peaks appear instead of ''edge'' excitation, and the spectrum depends on the impurity. For Ge this was a single band with a maximum at 240 nm and at least three bands could be distinguished after Si implantation: a dominant band with $\lambda\cong250$ nm and two weaker bands near 215 and 280 nm.

4. DISCUSSION

Comparing the PL intensities of layers in SiO_2 after implantation of Si^+, Ge^+, and Ar^+ ions shows unequivocally that implantation of group-IV elements does not simply reduce to the formation of radiation damage. According to Ref. 19, an increase of the density of PL centers introduced in SiO_2 by bombardment with Ar^+ ions is observed in the dose range $10^{12}-10^{14}$ cm^{-2}, after which the process saturates. The Ar^+ doses which we investigated gave a fortyfold displacement of the matrix atoms, but the increment to the PL was nonetheless extremely small compared with the effect produced by Si and Ge implantation. Moreover, the centers introduced by inert-gas atoms are easily annealed, while after implantation of group-IV impurity growth of the intensity and transformation of the PL spectra are also possible with heating (Figs. 1, 3, and 4).

The dose dependence of the PL and the response of the emission to annealing make it possible to understand better the mechanism leading to the appearance of centers. Indeed, centers form at doses that give an impurity density greater than 10^{21} cm^{-3}, which corresponds to an average interatomic distance of less than 1 nm. If the PL centers were individual Si or Ge atoms, then emission would have grown gradually

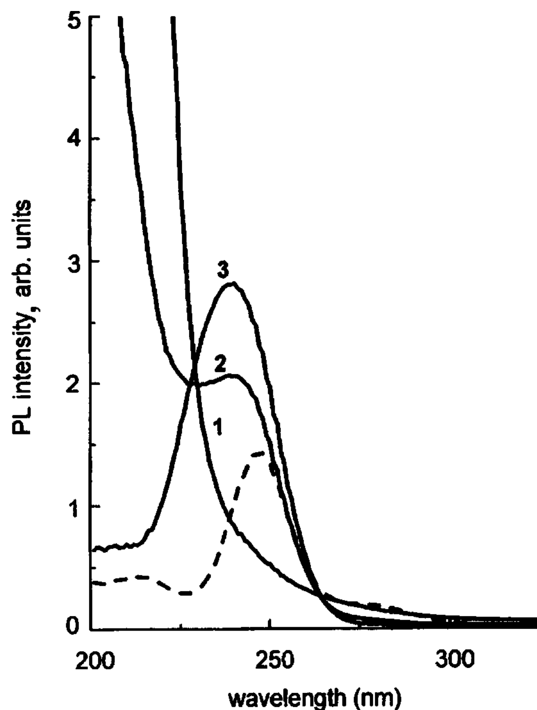


FIG. 5. Excitation spectra of blue photoluminescence after implantation of 6.6×10^{16} cm^{-2} Ge^+ (solid lines) and 4.8×10^{16} cm^{-2} Si^+ (dashed line — implantation at 20 °C). *1* — Immediately after implantation; *2* — annealing at 900 °C for 20 ms; *3* and dashed curve — annealing at 1050 °C for 20 ms.

with increasing dose. In our case, however, a rapid growth of the PL intensity occurred when interaction between impurity atoms became possible. Therefore, the formation of PL centers is due to clustering of impurities. The sharp intensification of the blue band in the preannealing spectra as the Si^+ dose is increased by less than a factor of 2 supports this argument (Fig. 2).

The number and size of impurity clusters clearly increases not as a result of the diffusion of impurity atoms toward sinks. Our experiments showed that the processes occur even in the case of low-temperature implantation without additional annealing, if a definite doping level has been reached. Nonetheless, increasing the implantation temperature to room temperature transforms the spectrum, intensifying the blue band (Fig. 3). Therefore, thermal activation occurs, as is also indicated by data from annealings (Figs. 3 and 4). If annealings are compared with respect to the computed diffusion displacement $l=(Dt)^{1/2}$ of excess Si atoms in SiO_2, then using the value of D from Ref. 20, we obtain the minimum distance $l=10^{-3}$ nm for 400 °C and 30 min and the maximum distance $l=10^{-2}$ nm for 1050 °C and 20 ms. In the case where diffusion displacement is much shorter than the single atomic hopping lengths and the distance between the impurity atoms is of the order of this length, the formation of impurity complexes and clusters will be most likely of a percolation character.

We assume that in the case of Si^+ and Ge^+ implantation the shortest-wavelength Pl bands which arise are due to the formation of extremely small percolation clusters of impurities (point complexes, atomic chains, and their combinations) as a result of the displacement of excess atoms from stoichiometric SiO_2. The segregation process is thermally activatable, but at low temperatures it is not a diffusion process. It occurs by means of single hops of impurity atoms to the nearest atoms of the growing percolation cluster. After the layers implanted with Si^+ are annealed the direct sources of the blue PL are centers which absorb ~5-eV photons and emit in the region ~2.7 eV (Fig. 5). Such centers have been observed many times in SiO_2, and ordinarily they were attributed to an oxygen deficiency and the formation of oxygen vacancies.[7,14,15] Similar centers also appear in SiO_2 containing Ge. They absorb and emit, respectively, at ~5 and ~3.1 eV,[21,22] and their nature is attributed to the presence of Ge and a deficiency of oxygen. In SiO_2 with a O deficiency the oxygen vacancies materialize in the form of directly interacting (no intermediate O) nearest-neighbor Si atoms with formation of Si–Si bonds. In the presence of Ge Si–Ge (vacancy–impurity complex) and Ge–Ge pairs are possible. The sharp intensification of the short-wavelength luminescence accompanying implantation of high doses of Si and Ge and the obvious similarity in transformation of the PL excitation spectra (Fig. 5) allow us to conclude that in both cases known vacancy-type defects are formed in SiO_2. This is essentially a direct closure of bonds between excess atoms with formation of impurity chains inside SiO_2. According to calculations,[23] the formation of such bonds introduces in the band gap of SiO_2 local levels whose position is determined mainly by the interacting pair with a small contribution from the six nearest peripheral oxygen atoms.

Of the three types of ions, only implantation of Si^+ preserved and even increased emission in the orange region of the spectra after annealing (Fig. 3 and 4). It is natural to attribute it directly to developed clusters of segregated Si. We note that there are greater possibilities for development of silicon clusters in SiO_2, since one third of the matrix itself consists of Si. This can explain the diffuseness of the PL spectra after Si implantation and the presence of a large number of lines. In the case of Ge implantation orange luminescence is observed only before annealing and is apparently due to mixed clusters of Ge atoms and silicon displaced from SiO_2. Heating restores the SiO_2 network, while for germanium atoms the percolation radii are shorter and, correspondingly, the clusters will be smaller and of more uniform size. This corresponds to the fact that when the orange band in Ge-implanted layers vanishes as a result of annealing, the intensity of the blue band increases at the same time (Fig. 4).

The luminescence in the orange and adjoining regions of the spectrum (500 nm $\leqslant\lambda\leqslant$700 nm) has been attributed in many studies to the formation of different nanosize objects associated with excess Si.[6,16,17] In keeping with this conclusion, our results make it possible to specify to a high degree the nature and behavior of the silicon precipitates. Being developed percolation clusters, they consist of nonphase precipitates of Si, since they do not possess sharp phase boundaries. The number and sizes of such clusters at first grow with annealing; this promotes an increase in the PL intensity. However, the branching nonphase precipitates then contract into nuclei of a Si phase in SiO_2.[13] The smallest phase precipitates ($\lesssim$2 nm) cannot preserve the crystal structure because of the destabilizing influence of the surface.[24] In amor-

phous silicon, however, PL is suppressed by nonradiative recombination centers in the form of dangling bonds. Such ideas make it possible to understand the reasons why the PL decreases after increasing, as shown in Fig. 3. Stable Si nanocrystals form only as a result of diffusion-limited growth with appropriate temperatures and durations of annealing. Because of quantum-well limitations, they emit a very strong radiation similar to the red PL of porous silicon.[2,6,13,25]

5. CONCLUSIONS

The appearance of a rather intense short-wavelength PL ($\lambda \leqslant 700$ nm) immediately after low-temperature implantation of high doses of Si^+ and Ge^+ ions in SiO_2 does not reduce simply to radiation damage in the SiO_2 network. Secondary damage after implantation of comparable doses of Ar^+ give only a very weak increase in the PL intensity, and moreover this increase is completely removed by annealings at 400 °C for 30 min or 1050 °C for 20 ms. The effects produced by implantation of group-IV elements are manifested at doping levels above 10^{21} cm^{-3}, when the average distance between the impurity atoms decreases to <1 nm. This circumstance, as well as the intensification or transformation of PL spectra accompanying heat treatments, which give a computed diffusion distance of impurity atoms of only $10^{-3}-10^{-2}$ nm, indicate that impurity clustering plays a decisive role, and this process is apparently of a percolation character. The intense blue PL band is due to radiative recombination centers with a ratio of excitation and emission energies that is characteristic of oxygen vacancies in SiO_2. Since O vacancies in SiO_2 indicate a direct interaction of the type Si–Si (or Si–Ge), the centers of blue PL are not phase precipitates but rather chains of group-IV atoms embedded in SiO_2. The weaker emission in the orange part of the spectrum is preserved after moderate annealings only in the case of Si^+ implantation. Therefore, it is due to the Si precipitates themselves. It is unlikely that these precipitates are silicon nanocrystals, considering the low thermal budget of annealings employed. It makes more sense to interpret them as nonphase nanoprecipitates of Si, segregated from SiO_2 in the form of strongly developed or merged clusters. Such precipitates are capable of subsequently contracting into nuclei of an amorphous Si phase, and after definite sizes are reached they can form stable silicon nanocrystals.

[1] H. Tamura, M. Ruckschloss, T. Wirschem, and S. Veprek, Appl. Phys. Lett. **65**, 1357 (1994).
[2] P. Mutti, G. Ghislotti, S. Bertoni, L. Bonoldi, G. F. Cerofolini, L. Meda, E. Grilli, and M. Guzzi, Appl. Phys. Lett. **66**, 851 (1995).
[3] J. G. Zhu, C. W. White, J. D. Budai, S. P. Withraw, and Y. Chen, J. Appl. Phys. **78**, 4386 (1995).
[4] Q. Zhang, S. C. Bayliss, and D. A. Hutt, Appl. Phys. Lett. **66**, 1977 (1995).
[5] H. Morisaki, H. Hashimoto, F. W. Ping, H. Nozava, and H. Ono, J. Appl. Phys. **74**, 2977 (1993).
[6] T. Shimizu-Iwayama, Y. Terao, A. Kamiya, M. Takeda, S. Nakao, and K. Saitoh, Nucl. Instrum. Methods Phys. Res. B **112**, 214 (1996).
[7] L.-S. Liao, X.-M. Bao, N.-S. Li, X.-Q. Zheng, and N.-B. Min, J. Lumin. **68**, 199 (1996).
[8] A. K. Dutta, Appl. Phys. Lett. **68**, 1189 (1996).
[9] K. S. Min, K. V. Shcheglov, C. M. Yang, H. A. Atwater, M. L. Brongersma, and A. Polman, Appl. Phys. Lett. **68**, 2511 (1996).
[10] Y. Kanemitsu, H. Uto, Y. Masumoto, and Y. Maeda, Appl. Phys. Lett. **61**, 2187 (1992).
[11] H. M. Cheong, W. Paul, S. P. Withrow, J. G. Zhu, J. D. Budai, S. W. White, and D. M. Hembree, Appl. Phys. Lett. **68**, 87 (1996).
[12] W. Skorupa, R. A. Yankov, L. Rebohle, H. Frob, T. Bohme, K. Leo, I. E. Tyschenko, and G. A. Kachurin, Nucl. Instrum. Methods Phys. Res. B **119**, 106 (1996).
[13] G. A. Kachurin, I. E. Tyschenko, K. S. Zhuravlev, N. A. Pazdnikov, V. A. Volodin, A. K. Gutakovsky, A. F. Leier, W. Skorupa, and R. A. Yankov, Nucl. Instrum. Methods Phys. Res. B **122**, 571 (1997).
[14] R. Tohmon, Y. Shimogaichi, H. Mizuno, and Y. Ohki, Phys. Rev. Lett. **62**, 1388 (1989).
[15] H. Nishikawa, T. Shiroyama, R. Nakamura, Y. Ohki, K. Nagasawa, and Y. Hama, Phys. Rev. B **45**, 586 (1992).
[16] H. Nishikawa, E. Watanabe, D. Ito, Y. Sakurai, K. Nagasawa, and Y. Ohki, J. Appl. Phys. **80**, 3513 (1996).
[17] A. J. Kenyon, P. F. Towoga, C. W. Pitt, and G. Rehm, J. Appl. Phys. **79**, 9291 (1996).
[18] G. Ghislotti, B. Nielsen, P. Asoka-Kumar, K. G. Lynn, A. Gambhir, L. E. Di Mauro, and C. E. Bottani, J. Appl. Phys. **79**, 8660 (1996).
[19] S. Bota, B. Garrido, J. R. Morante, A. Baraban, and P. P. Konorov, Solid-State Electron. **34**, 355 (1996).
[20] L. A. Nesbit, Appl. Phys. Lett. **46**, 38 (1985).
[21] H. Hosono, Y. Abe, D. L. Kinser, R. A. Weeks, K. Muta, and H. Kawazoe, Phys. Rev. B **46**, 11 445 (1982).
[22] M. Gallagher and U. Osterberg, Appl. Phys. Lett. **63**, 2987 (1993).
[23] V. B. Sulimov, V. O. Sokolov, J. Non-Cryst. Solids **191**, 260 (1995).
[24] A. N. Goldstein, Appl. Phys. A **62**, 33 (1996).
[25] G. A. Kachurin, I. E. Tyschenko, V. Skorupa, R. A. Yankov, K. S. Zhuravlev, N. A. Pazdnikov, V. A. Volodin, A. K. Gutakovskiĭ, and A. F. Leĭer, Fiz. Tekh. Poluprovodn. **31**, 730 (1997) [Semiconductors **31**, 627 (1997)].

Translated by M. E. Alferieff

Characteristic features of the defect formation process in $Pb_{1-x}Sn_xSe(x\leqslant 0.06)$

A. N. Veĭs and N. A. Suvorova

St. Petersburg State Technical University, 195251 St. Petersburg, Russia
(Submitted October 6, 1997; accepted for publication October 9, 1997)
Fiz. Tekh. Poluprovodn. **32**, 445–449 (April 1998)

The spectral dependences of the absorption coefficient in undoped p-$Pb_{1-x}Sn_xSe(x=0.00-0.06)$ at $T=300$ K were investigated. A quasilocal level associated with a chalcogen vacancy was found in all experimental samples. It is shown that the chalcogen vacancy density in samples with tin is appreciably different from that in samples without tin, and it increases with x. The nature of this phenomenon is discussed. © *1998 American Institute of Physics.*
[S1063-7826(98)00904-1]

One of the most unusual properties of the dilute solid solutions $Pb_{1-x}Sn_xSe(x\leqslant 0.02)$ is undoubtedly the pinning of the Fermi level E_F, which is observed when sodium impurity is introduced, and the accompanying transfer of tin from the state Sn^{2+} into the state Sn^{4+} (Refs. 1–5). These data were used as the basis of models according to which the pinning of E_F in $Pb_{1-x}Sn_xSe$:Na is due either to the charge transfer on tin $Sn^{2+}\leftrightarrows Sn^{4+}$ (Refs. 4 and 5) or the binding of tin in neutral complexes $Na_2Se+SnSe_2$ (Refs. 2 and 3). Another mechanism of stabilization of E_F in $Pb_{1-x}Sn_xSe$:Na ($x\leqslant 0.02$) was proposed by Veĭs and Suvorova.[6] The experiments performed in that study did not permit observing any localized or quasilocalized states associated with tin or with complexes which include tin. Just as in Ref. 7, only the anionic vacancies V_{ch}, which, however, are in different charge states, were manifested in the optical absorption spectra α of dilute solid solutions $Pb_{1-x}Sn_xSe$:Na. Comparing the results obtained in Refs. 6 and 7 with data from electro investigations[2] (see also Ref. 8) suggests[6] that the pinning of E_F in the solid solutions studied is due to charge transfer on chalcogen vacancies $V^0_{ch}\leftrightarrows V^{2-}_{ch}$ (the superscript indicates, just as in Ref. 6, the number of electrons localized in a vacancy). Later,[9] the results of Ref. 6 were confirmed in experimental investigations of the spectra $\alpha(\hbar\omega)$ in PbSe doped with acceptor impurities Na or Tl and superstoichiometric lead. According to Ref. 10, anionic vacancies are the dominant type of intrinsic defect in such samples.

It appears, however, that the investigation of the characteristic features of defect formation in the dilute solid solutions $Pb_{1-x}Sn_xSe$ should be continued. This is due to the fact that the model proposed in Ref. 6 does not reflect the presence of unusual features, found in Refs. 4 and 5, in the Mössbauer spectra of $Pb_{1-x}Sn_xSe$:Na, Tl. In addition, the model does not permit explaining the reason for the high V_{ch} densities in the solid solution studied. The experimentally observed increase in the density of anionic vacancies in $Pb_{1-x}Sn_xSe$:Na could be due not only to self-compensation of the acceptor action of sodium[10] but also to a specific effect of tin. For this reason, it is first necessary to establish whether or not there exists any relation between the density of V_{ch} and the amount of tin in the solid solutions.

We have accordingly investigated the spectra $\alpha(\hbar\omega)$ in $Pb_{1-x}Sn_xSe$ with variable tin concentration $x\leqslant 0.06$ in the charge, which, however, did not contain any acceptor impurities. All samples were p-type. The aim of the experiments was to reveal in the spectra $\alpha(\hbar\omega)$ of the samples investigated additional absorption bands associated with V_{ch} and to compare the characteristics of these bands with the corresponding data for "pure" p-type PbSe with no acceptor impurities.

We note that the spectra $\alpha(\hbar\omega)$ in "pure" p-type PbSe were studied earlier.[11,12] However, no features associated with V_{ch} were found in these spectra. For this reason, the photon energy range in which the optical absorption in pure p-PbSe was investigated, was much wider than in Refs. 11 and 12. All experiments were performed at $T=300$ K.

The objects of investigation were single crystals and polycrystalline samples. The single crystals were grown by the Bridgman–Stockbarger method and the polycrystalline samples were formed by hot pressing. All samples were subjected to homogenizing annealing for 100 h at $T=650$ °C. The composition of the polycrystals is given in terms of the load in the charge and that of the single crystals is given with allowance for the data from x-ray crystallographic analysis.

The experimental results for the single crystals and pressed samples were found to be similar. Some of them are shown in Fig. 1. One can see from Fig. 1 that the spectra $\alpha(\hbar\omega)$ for samples with and without tin are qualitatively different. Characteristic features attesting to the existence of localized and quasilocal states are clearly manifested in the optical absorption spectra of the dilute solutions $Pb_{1-x}Sn_xSe$. Intense bands of additional absorption α_1, which have asymmetric spectral curves and a sharp red boundary, are present in the long-wavelength region of the experimental spectra of this group of samples. The shape of the frequency dependences $\alpha_1(\hbar\omega)$ attests to the fact that the bands studied are associated with optical transitions of electrons from allowed states in the valence band into localized states E_1 above the valence-band top (see energy scheme in Fig. 1). Bell-shaped features α_2 are clearly seen against their background. These features are due to optical transitions of electrons between a quasilocal level E_2 in the

 © 1998 American Institute of Physics

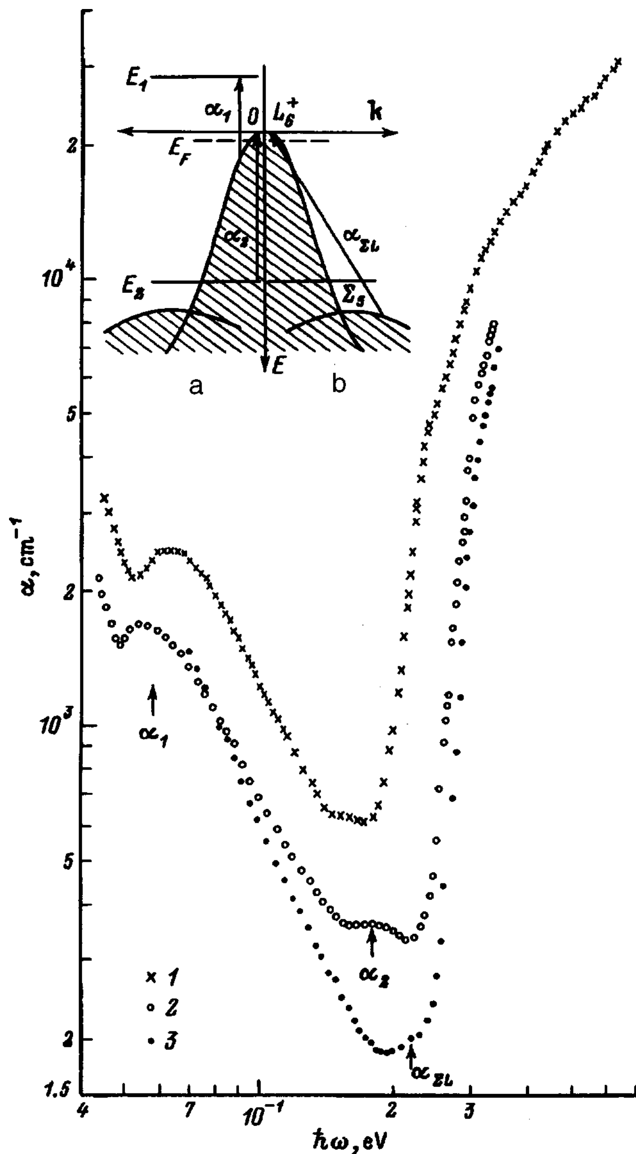


FIG. 1. Spectral dependences of the optical absorption coefficient α in $Pb_{1-x}Sn_xSe$ at $T=300$ K. *1* — Pressed sample, *2*, *3* — single crystals. $p_H \times 10^{-18}$, cm^{-3} (x): *1* — 3.0 (0.06), *2* — 2.7 (0.015), *3* — 2.8 (0). Sample thickness, μm: *1* — 9.4 and 2.4, *2* — 9.4, *3* — 9.1. Inset: Energy schemes of p-$Pb_{1-x}Sn_xSe$ (a) and p-PbSe (b). The arrows mark the observed optical transitions.

valence band and allowed states of the valence band near its top.

The nature of these bands has been investigated in detail in Refs. 6, 7, and 9. It was shown there that the bands α_1 and α_2 are associated with complex and anionic vacancies, respectively.

The optical charge-transfer energies of localized ($E_1^{\rm opt}$) and quasilocalized ($E_2^{\rm opt}$) states observed in $Pb_{1-x}Sn_xSe$ were determined by calculating the spectral dependences of the additional absorption coefficient α_{ad}. Accordingly, we separated individual components of the additional absorption coefficient from the experimental spectra $\alpha(\hbar\omega)$ by the procedure described in detail in Refs. 7 and 13. Some results obtained in this manner are shown in Fig. 2. The frequency dependences $\alpha_1(\hbar\omega)$ were calculated, just as in Ref. 7, from

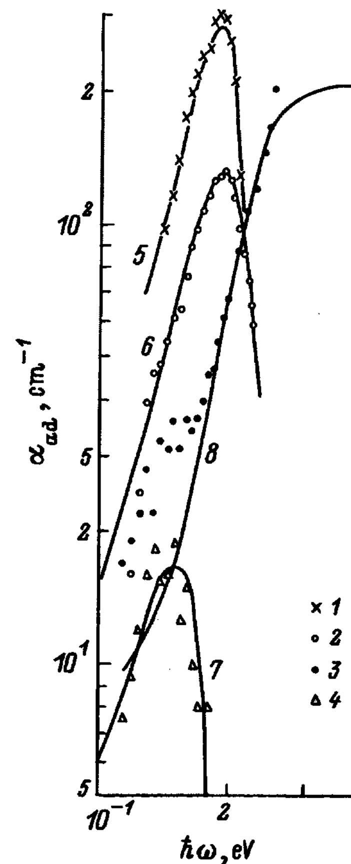


FIG. 2. Spectral dependences of the additional absorption coefficient α_{ad} in $Pb_{1-x}Sn_xSe$ at $T=300$ K. Dots — experiment: *1–3* — correspond to the sample designations in Fig. 1. *4* — $\alpha_2(\hbar\omega)$ in sample *3* in Fig. 1, lines — calculations: *5–7* — $\alpha_2(\hbar\omega)$, *8* — $\alpha_{\Sigma L}(\hbar\omega)$.

the expression $\alpha_1 \sim (\hbar\omega)^{-3}(\hbar\omega - E_1^{\rm opt})^{1/2}$ and the expression (1) from Ref. 14 was used to calculate $\alpha_2(\hbar\omega)$. The computed dependences $\alpha_2(\hbar\omega)$ are presented in Fig. 2 (solid lines). The calculation showed that the quantities $E_1^{\rm opt}$ and $E_2^{\rm opt}$ do not vary appreciably in the range $0.01 \leqslant x \leqslant 0.06$ and equal 0.054 ± 0.006 and 0.205 ± 0.020 eV, respectively. The half-widths Γ_2 of the quasilocal bands in the experimental samples were found to be the same within the limits of the experimental error and equal to 0.010 ± 0.005 eV. We note that the values obtained for the energy parameters of localized and quasilocal states in $Pb_{1-x}Sn_xSe$ agree well with the corresponding data from Refs. 6, 7, and 9.

The spectral dependences $\alpha(\hbar\omega)$ for pure p-PbSe, not doped with acceptor impurities, appear at first glance to be much less rich. Besides the fundamental band edge and absorption by free current carriers, the only component of $\alpha_{\Sigma L}$ clearly manifested in them is due to optical transitions of electrons between nonequivalent extrema of the valence band ($\Sigma_5 \rightarrow L_6^+$) separated by an energy gap ΔE_v. However, detailed analysis of the results shows that the dependences $\alpha_{ad}(\hbar\omega)$ in pure p-PbSe are in fact more complex. This conclusion follows from a comparison of the experimental spectra α_{ad} and the computed dependences $\alpha_{\Sigma L}(\hbar\omega)$, constructed, just as in Ref. 12, on the basis of the Hagi–Kimura theory, according to Eq. (4) of Ref. 15. The corresponding data for one of the experimental samples are shown in Fig. 2.

The dots (*3*) in this figure represent the dependence $\alpha_{ad}(\hbar\omega)$ separated from the experimental spectrum $\alpha(\hbar\omega)$ by subtracting the spectrum extrapolated into the short-wavelength region of absorption by free current carriers. The line (8) was obtained by calculating the component of $\alpha_{\Sigma L}$. In constructing the curves $\alpha_{\Sigma L}(\hbar\omega)$ in all experimental samples, the value $\Delta E_v = 0.23$ eV was used, in accordance with Ref. 12. As one can see from Fig. 2, good agreement between the experimental dots and the computed curve is observed in a wide spectral interval. However, in the long-wavelength region of the spectrum α_{ad} the computed curve $\alpha_{\Sigma L}(\hbar\omega)$ lies substantially below the experimental points. Similar results were also obtained in the other samples of pure p-PbSe. This shows that besides the components of $\alpha_{\Sigma L}$ there is another component in the spectral dependences of the additional absorption in pure p-PbSe.

This component was separated from the experimental curves by subtracting the computed values of $\alpha_{\Sigma L}$. One result obtained in this manner is shown in Fig. 2 (curve *4*). It was found that this component of the spectra α_{ad} of pure p-PbSe is bell-shaped with sharply defined red and short-wavelength boundaries. This makes it possible to attribute it to optical transitions of electrons from a quasilocal level in the valence band into allowed states of the valence band near its top. The agreement between the energy position of this feature and the data for $Pb_{1-x}Sn_xSe$ (Fig. 2) indicates that the component of the spectra of pure p-PbSe, which we are discussing, is attributable to anionic vacancies. This is also indicated by the results of calculations of the frequency dependences of the absorption coefficient in the bands studied, which was performed, just as in the case of $Pb_{1-x}Sn_xSe$, using Eq. (1) of Ref. 14. The calculation showed that the energy parameters of the center responsible for the appearance of the bell-shaped features in the spectra $\alpha(\hbar\omega)$ of pure p-PbSe ($E^{\mathrm{opt}} = 0.190 \pm 0.025$ eV and $\Gamma = 0.010 \pm 0.005$ eV) correspond to the data of Refs. 6, 7, and 9 for V_{ch} in lead selenide and dilute lead-selenide-based solid solutions with different composition of the master alloy.

To compare the experimental data for different samples the ratio of the total absorption cross section S_2 in the band α_2 ($S_2 \sim \int_0^\infty \alpha_2 d(\hbar\omega)$) to the Hall hole density p_H was determined for each of them. The function $S_2(x)/p_H$ obtained in this manner is shown in Fig. 3. One can see from the figure that the ratio S_2/p_H in samples with tin differs substantially from that in samples without tin and it tends to increase with the tin content in the charge.

The decrease in the band gap E_g in the experimental samples attests to the increase in the tin density in the solid solutions with increasing amount of tin in the charge. In accordance with the results of Ref. 16, the quantities E_g were determined according to the intercept of the curves $\alpha^2(\hbar\omega)$ on the abscissa. It was found that as x increases, the band gaps E_g decrease from 0.29 ± 0.01 eV in PbSe to 0.245 ± 0.010 eV in $Pb_{0.94}Sn_{0.06}Se$, as should happen in solid solutions of lead and tin selenides.[17]

The function $S_2(x)/p_H$ shown in Fig. 3 indicates that in $Pb_{1-x}Sn_xSe$ ($x \leqslant 0.06$) there does indeed exist a relation between the amount of tin in the charge and the density of anionic vacancies. This suggests in turn that the rather high

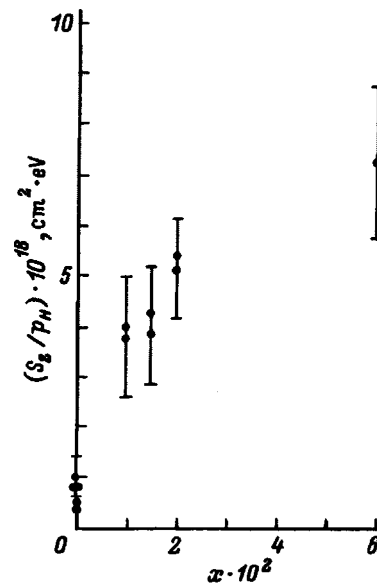


FIG. 3. Ratio of the total absorption cross sections S_2 in the bands α_2 to the Hall hole density p_H in $Pb_{1-x}Sn_xSe$ versus the tin concentration x at $T = 300$ K.

densities of V_{ch} in $Pb_{1-x}Sn_xSe$:Na could be due to not only self-compensation of the acceptor action of sodium but also a restructuring in the subsystem of isoelectronic impurity atoms. According to Refs. 2 and 3, such a restructuring occurs in the subsystem of tin atoms as a result of the formation of the complexes Na_2Se:$SnSe_2$, whose concentration can be comparable to x. However, the samples investigated in the present work do not contain sodium impurity, and complexes with the indicated composition cannot be formed in them. As shown earlier in Ref. 7, other complexes which could contain intrinsic defects arise in such samples. But, according to the data in Ref. 7, the concentration of these complexes does not exceed $10^{17} - 10^{18}$ cm^{-3}. For this reason, the appearance of moderately high concentrations (10^{19} cm^{-3} and higher) of anionic vacancies in $Pb_{1-x}Sn_xSe$ with no electrically active impurities cannot be attributed to their formation. This shows that some other rearrangement mechanism, which gives rise to the appearance of moderately high concentrations of V_{ch}, materializes in the subsystem of tin atoms in the solid solution studied.

Specifically, it cannot be ruled out that in $Pb_{1-x}Sn_xSe$ tin can occupy two nonequivalent crystallographic positions: sites of the metal sublattice and tetrahedral voids of the chalcogen sublattice. On the basis of this assumption it is possible not only to explain the characteristic features of the Mössbauer spectra of $Pb_{1-x}Sn_xSe$, but also to indicate the likely reason why moderately high V_{ch} concentrations arise in these compounds.

Indeed, some sites in the metal sublattice should be free as a result of such a restructuring in the subsystem of tin atoms. Since the electrical activity of the cationic vacancies (doubly charged acceptors[18–21]) is different from that of interstitial metal atoms (singly charged donors[18]), one would expect that the hole density will increase substantially as a result of the transfer of some tin atoms into interstices. In

reality, however, this does not happen. According to the data of Refs. 2 and 7, the densities p_H in $Pb_{1-x}Sn_xSe$ ($x \leqslant 0.02$), which do not contain acceptor impurities, do not exceed $\approx 5 \times 10^{18}$ cm^{-3}. Such a situation can occur if in the process of change in the structural position of tin there arise, together with cationic vacancies, anionic vacancies as well, which act as donors in lead chalcogenides[18–21] and compensate for the acceptor action of metal vacancies. Thus, one possible reason for the appearance of the high concentrations of V_{ch} in the dilute solid solutions $Pb_{1-x}Sn_xSe$ could be self-compensation.

The assumption that some tin atoms transfer into interstices also agrees with Mössbauer spectroscopy data.[4,5] The experimentally observed change in the isomeric shift of the ^{119}Sn line could be due to a difference in the type of chemical bond of tin occupying a lattice site (p-orbitals) and interstitial positions (sp^3 hybrid orbitals). We also note that this assumption is consistent with the basic premises of the crystal chemistry of IV–VI compounds (see, for example, Ref. 22). This circumstance should be viewed as an argument in support of this assumption.

The results obtained in $Pb_{1-x}Sn_xSe$ attest to the fact that self-compensation can have a strong effect on the properties of IV–VI compounds even in the absence of electrically active impurities. This is also indicated by the data for pure p-PbSe, which were studied above. According to Ref. 22, in such samples the dominant type of intrinsic defects are metal vacancies. Conversely, the density of intrinsic defects exhibiting a donor effect should be low in such samples. Despite this circumstance, characteristic features associated with optical charge transfer on anionic vacancies were found in the optical absorption spectra of the pure p-PbSe samples investigated in this study. Comparing the data obtained from them (Fig. 3) with the corresponding results of Ref. 9 shows that the V_{ch} density in the pure p-PbSe samples that we investigated is not low and can reach $(2-3) \times 10^{18}$ cm^{-3}. This signifies that all pure p-PbSe samples investigated in the present work are self-compensated, and the values of p_H in them are determined by the difference between the cationic and anionic vacancy densities. This idea was first stated in Ref. 23. Apparently, the results of the present work should be viewed as direct experimental confirmation of this idea.

We wish to thank L. V. Prokof'eva for providing the samples for the investigations and the x-ray crystallographic data. We also thank O. E. Kvyatkovskiĭ for a helpful discussion.

[1] L. V. Prokof'eva, M. N. Vinogradova, and S. V. Zarubo, Fiz. Tekh. Poluprovodn. **14**, 2201 (1980) [Sov. Phys. Semicond. **14**, 1304 (1980)].
[2] L. V. Prokof'eva, E. A. Gurieva, K. G. Gartsman, Sh. M. Zhumaksanov, S. V. Zarubo, Kh. R. Maĭlina, and K. T. Urazbaeva, Preprint No. 1049, A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, Leningrad (1986).
[3] Yu. A. Degtyarev, P. P. Konstantinov, Kh. R. Maĭlina, and L. V. Prokof'eva, Fiz. Tekh. Poluprovodn. **23**, 1576 (1989) [Sov. Phys. Semicond. **23**, 977 (1989)].
[4] L. V. Prokof'eva, S. V. Zarubo, F. S. Nasredinov, and P. P. Seregin, JETP Lett. **33**, 12 (1981).
[5] F. S. Nasredinov, L. V. Prokof'eva, and P. P. Seregin, Zh. Éksp. Teor. Fiz. **87**, 951 (1984) [Sov. Phys. JETP **60**, 542 (1984)].
[6] A. N. Veĭs and N. A. Suvorova, Fiz. Tekh. Poluprovodn. **29**, 278 (1995) [Semiconductors **29**, 138 (1995)].
[7] A. N. Veĭs, E. A. Gurieva, O. G. Nefedov, and L. V. Prokof'eva, Fiz. Tekh. Poluprovodn. **18**, 1723 (1984) [Sov. Phys. Semicond. **18**, 1080 (1984)].
[8] G. T. Alekseeva, V. A. Gurieva, P. P. Konstantinov, N. V. Maksimova, and L. V. Prokof'eva, Fiz. Tekh. Poluprovodn. **29**, 1388 (1995) [Semiconductors **29**, 719 (1995)].
[9] A. N. Veĭs and N. A. Suvorova, Fiz. Tekh. Poluprovodn. **30**, 2089 (1996) [Semiconductors **30**, 1089 (1996)].
[10] L. I. Bytenskiĭ, V. I. Kaĭdanov, R. B. Mel'nik, S. A. Nemov, and Yu. I. Ravich, Fiz. Tekh. Poluprovodn. **14**, 74 (1980) [Sov. Phys. Semicond. **14**, 40 (1980)].
[11] T. G. Potashevskaya and N. S. Tselishcheva, Tr. LPI **325**, 111 (1971).
[12] A. N. Veĭs, R. F. Kuteĭnikov, S. A. Kumzerov, and Yu. I. Ukhanov, Fiz. Tekh. Poluprovodn. **10**, 2218 (1976) [Sov. Phys. Semicond. **10**, 1320 (1976)].
[13] A. N. Veĭs, A. Yu. Rydanov, and N. A. Suvorova, Fiz. Tekh. Poluprovodn. **27**, 701 (1993) [Semiconductors **27**, 386 (1993)].
[14] A. N. Veĭs, Dokl. Akad. Nauk SSSR **289**, 1355 (1986) [Sov. Phys. Dokl. **31**, 654 (1986)].
[15] A. N. Veĭs and Yu. I. Ukhanov, Fiz. Tekh. Poluprovodn. **10**, 1315 (1976) [Sov. Phys. Semicond. **10**, 780 (1976)].
[16] I. A. Drabkin, Yu. Ya. Eliseeva, G. F. Zakharyugina, I. V. Nel'son, and Yu. I. Ravich, Fiz. Tekh. Poluprovodn. **8**, 1947 (1974) [Sov. Phys. Semicond. **8**, 1261 (1974)].
[17] G. Nimtz and B. Schliht, *Narrow Gap Semiconductors*, Springer-Verlag, Berlin, 1983, Vol. 98, p. 1.
[18] H. Heinrich, in *Proceedings of the Int. Summer School on Narrow Gap Semicond. Physics and Applications,* Nimes, France, 1979; Springer-Verlag, Berlin, 1980, p. 407.
[19] G. A. Kalyuzhnaya and K. V. Kiseleva, Tr. FIAN **177**, 5 (1987).
[20] N. J. Panda and G. W. Pratt, Phys. Rev. Lett. **22**, 180 (1969).
[21] B. A. Volkov and O. A. Pankratov, Zh. Éksp. Teor. Fiz. **88**, 280 (1985) [Sov. Phys. JETP **61**, 164 (1985)].
[22] N. Kh. Abrikosov and L. E. Shelimova, *Semiconductor Materials Based on IV–VI Compounds* [in Russian], Nauka, Moscow, 1975.
[23] R. F. Brebrick, J. Phys. Chem. Solids **18**, 116 (1961).

Translated by M. E. Alferieff

Energy spectrum of acceptors in the semimagnetic semiconductors p-$Hg_{1-x}Mn_xTe$ in the spin-glass region

E. I. Georgitsé, V. I. Ivanov-Omskiĭ, and D. I. Tsypishka

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia

(Submitted October 7, 1997; accepted for publication October 9, 1997)
Fiz. Tekh. Poluprovodn. **32**, 450–452 (April 1998)

The far-IR transmission and photoconductivity in the semimagnetic alloys p-$Hg_{1-x}Mn_xTe$ ($x=0.20-0.22$) at temperatures 2–7 K were investigated at fixed frequencies of optically pumped molecular lasers in the region 49–311 μm. We report the observation of photoexcitation of acceptors from the ground into excited states under conditions when the direct interaction of the magnetic moments of the manganese ions becomes substantial and results in the formation of a spin-glass phase. It was found that the internal field produced by the spontaneous and external polarizations of the magnetic moments of the Mn^{2+} ions in the spin-glass temperature range influences the energy spectrum of acceptors in a magnetic field. © *1998 American Institute of Physics.* [S1063-7826(98)01004-7]

1. INTRODUCTION

In semimagnetic semiconductors, the magnetic moments of Mn^{2+} ions produce an internal magnetic field. As a result of this field, the ground state energy of acceptors decreases with increasing external magnetic field.[1] In Ref. 2 it is shown that the internal field has the effect that if the ground state energy of an acceptor in the absence of an external magnetic field is of the order of the heavy-hole Bohr energy, then in a sufficiently strong magnetic field it is of the order of the light-hole Bohr energy.

In Refs. 3 and 4 we reported on laser magnetooptic spectroscopy investigations of the energy spectrum of acceptors in the narrow-gap semiconductors $Hg_{1-x-y}Cd_xMn_yTe$ in a magnetic field. In the present paper we report the results of similar investigations for the solid solutions p-$Hg_{1-x}Mn_xTe$ ($x=0.20-0.22$). In these materials a spin-glass phase, consisting of a disordered magnet in which the exchange interaction energy varies randomly over the volume of the sample, is observed at temperatures $T<5$ K. We show how this circumstance gives rise to characteristic features in the photoexcitation spectrum of acceptors as a result of restructuring of their energy spectrum in a magnetic field at a transition into a spin-glass phase.

2. EXPERIMENTAL METHODS AND RESULTS

The far-IR photoconductivity and transmission of p-$Hg_{1-x}Mn_xTe$ ($x=0.20-0.22$) samples in the range of photon energies 4–27 meV at temperatures 2–7 K were investigated. Bulk crystals were grown by the Bridgman method. The experimental samples were oriented in the direction of the crystallographic axis (111). The density of acceptors of unidentified nature was $N_A-N_D=(1-6)\times10^{15}$ cm^{-3}. These values were obtained from Hall coefficient measurements at $T=77$ K. The parameters of the experimental samples are presented in Table I.

The radiation source were submillimeter CH_3OH and CH_3OD vapor gas lasers with optical pumping by a tunable CO_2 laser. The measurements were performed in the Faraday geometry ($\mathbf{q}\perp\mathbf{B}$, where $\mathbf{q}$ is the wave vector of the radiation incident on the sample and $\mathbf{B}$ is the external magnetic field) and the Voigt geometry ($\mathbf{q}\parallel\mathbf{B}$). The spectra were recorded by scanning the magnetic field up to 6.5 T with fixed laser wavelength. The samples were cooled down to the measurement temperature in magnetic fields up to 1 T and without a field.

Figures 1 and 2 show the transmission and photoconductivity spectra of a $Hg_{0.78}Mn_{0.22}Te$ sample recorded at a temperature of 2 K for the cases where the sample was cooled in a magnetic field $B_{cool}=0.6$ T (curve a) and without a field (curve b). One can see by comparing the spectra that in the case where the sample is cooled to the measurement temperature in a magnetic field the lines become sparser and are shifted in the direction of weaker magnetic fields. The displacement of the lines in the spectra depends on B_{cool} and reaches a maximum value for $B_{cool}\geqslant0.6$ T.

Figure 3 shows the photoconductivity spectra of the same sample recorded at temperature $T=6$ K. At this temperature the shape of the lines in the spectra does not depend on the cooling method. Shifting of the lines likewise is not observed when the sample is cooled in a magnetic field.

To analyze the spectra and identify the lines we constructed plots of the energies of the observed transitions versus the magnetic field. were constructed. Figure 4 shows the magnetic field dependence of the energies of the transmission peaks at $T=2$ K. The dashed lines show the dependences for a p-$Hg_{0.78}Mn_{0.22}Te$ sample cooled to temperature 2 K in a magnetic field, while the solid lines are for the case of zero field cooling. When these curves are extrapolated to zero magnetic field, the energy of the observed optical transitions does not vanish. This makes it possible to attribute them to intracenter excitation of the acceptors, since the experimental samples were p-type and the spectra were re-

TABLE I. Parameters of the experimental samples.

Sample	E_g, eV	$E_A(0)$, meV (calculation)	$E_A(0)$, meV (measurements)	$N_0\beta$, eV	B_{int}, T
$Hg_{0.8}Mn_{0.2}Te$	0.72	11.0	11.2	0.85 ± 0.05	0.52 ± 0.03
$Hg_{0.78}Mn_{0.22}Te$	0.75	11.1	11.4	0.85 ± 0.05	0.52 ± 0.03

corded in the absence of interband illumination. Since two curves were obtained, and since it is assumed that at the measurement temperature the acceptor ground state is populated by holes, the observed magnetic resonances can be attributed to transitions from the ground state into two excited states.

As follows from the plot of the magnetic field dependences of the transition energies, the maximum shift of the lines in a magnetic field is $\Delta B=0.50\pm0.05$ T. The energy difference between the corresponding curves obtained with the sample cooled in the absence and presence of a magnetic field is $\Delta E=0.4\pm0.05$ meV.

In contrast to narrow-gap semimagnetic semiconductors, the curves of the transition energies versus the magnetic field do not show anticrossing of the energy levels of acceptors in a magnetic field.[4] This is explained by the fact that the depth of the acceptor is greater and the Zeeman splitting of its ground state is substantially smaller in the experimental samples as a result of the larger band gap. The calculations of the ground state energies of the acceptor assumed that the static permittivity $\varepsilon=18$, while the effective heavy-hole mass is $m_n^*=0.4m_0$, where m_0 is the mass of a free electron. Analysis of the magnetic field dependences of the transition energies showed that the ground state energy of the acceptor in the experimental samples in a field $B=0$ is ~ 11 meV, in good agreement with the calculations.

To determine the exchange integral $N_0\beta$ for valence band states the magnetic field dependences of the transition energies were calculated for temperatures of 2 and 6 K. The method for calculating $N_0\beta$ is described in Ref. 4. The computational results for the experimental samples are presented in Table I. We note that in the calculations the following values were used for the parameters appearing in the matrix of the modified Pidgeon–Brown Hamiltonian:[5] $E_p=18.3$ eV, $\Delta=1$ eV, $\gamma_1=0.5$, $K=0.1$, and $N_0\alpha=-0.45$ eV.

3. DISCUSSION

Of special interest among the results which we obtained is the shift of the transmission and photoconductivity lines in the spectra in the direction of weaker magnetic fields when the sample is cooled to the measurement temperature in a magnetic field. Analysis of the spectra of all experimental samples showed that the magnitude of the shift depends on the Mn^{2+} ion density and on the magnitude of the applied field during cooling. For each sample, the shift of the lines reached the maximum magnitude at a definite value of the external field. For example, for the $Hg_{0.78}Mn_{0.22}Te$ sample this shifts is $\Delta B=0.50\pm0.05$ T. As the magnetic field increased further during cooling of the sample, no changes were observed in the spectra. Moreover, it is interesting that when the samples are cooled in the absence of an external magnetic field, the lines in the spectra are much wider and are not Gaussian.

These results can be explained, in our opinion, as follows. When an external magnetic field is switched on and the temperature of the sample is higher than that required for the formation of a spin-glass phase, the magnetic moments of the Mn^{2+} ions are partially oriented along the external field. As the external field increases, the number of ''oriented'' moments of the magnetic ions also increases until all ions are oriented.

As the sample is cooled, when the spin-glass phase appears, the magnetic moments of the Mn^{2+} ions remain ori-

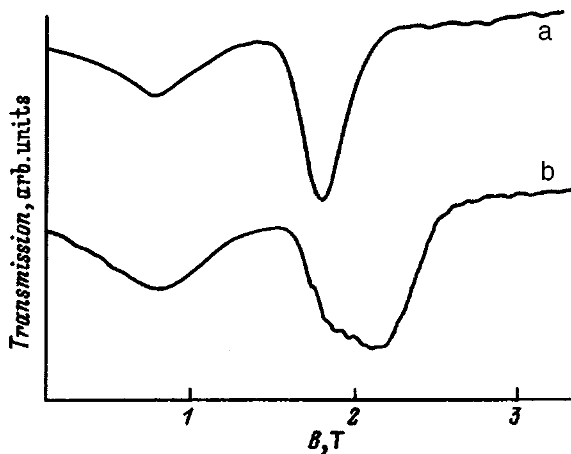


FIG. 1. p-$Hg_{0.78}Mn_{0.22}Te$ transmission spectra for photon energy 7.27 meV at $T=2$ K: a — Cooling in a magnetic field $B_{cool}=0.6$ T; b — $B_{cool}=0$.

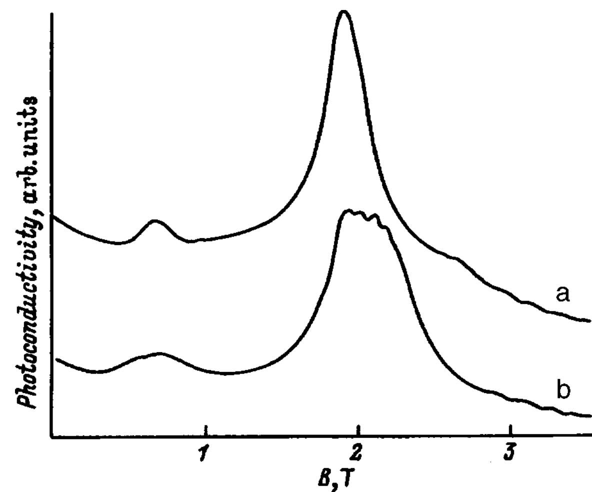


FIG. 2. p-$Hg_{0.78}Mn_{0.22}Te$ photoconductivity spectra with excitation by 7.27-meV photons at $T=2$ K: a — Cooling in a magnetic field $B_{cool}=0.6$ T; b — $B_{cool}=0$.

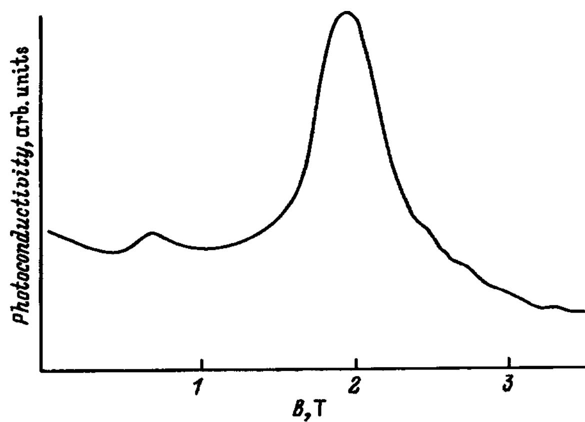


FIG. 3. p-$Hg_{0.78}Mn_{0.22}Te$ photoconductivity spectra with excitation by 7.27-meV photons at $T=6$ K.

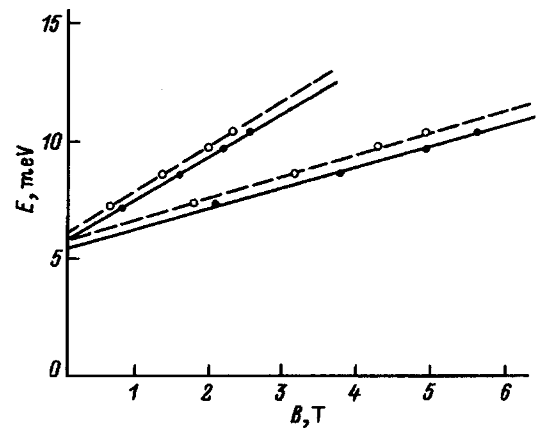


FIG. 4. Position of transmission minima in a magnetic field versus photon energy in the case of cooling in zero magnetic field (solid line) and in a magnetic field $B=0.6$ T (dashed line).

ented even after the external field is removed. An internal magnetic field is thereby produced. This field restructures somewhat the energy spectrum of the acceptor. When the transmission or photoconductivity spectra are recorded in a magnetic field oriented in the same direction as the internal field (as in the present experiments), the lines corresponding to transitions of an acceptor from the ground state to the excited states undergo a shift in the direction of decreasing applied field by an amount of the internal field. As expected, the magnitude of the internal field depends on the magnetic ion density (see Table I).

The narrowing of lines in the spectra of the samples cooled in a magnetic field is just as interesting. This result can be interpreted as follows. When the sample is cooled in a zero field, the magnetic moments of the Mn^{2+} ions freeze in random directions. For this reason, the local fields produced by the Mn^{2+}ions around individual acceptors have different magnitudes. Therefore, the energy spectra of photoexcited acceptors in an external magnetic field differ somewhat from one another. In the spectra this corresponds to wide lines with a diffuse peak. However, when the sample is cooled in a magnetic field and the magnetic moments of almost all Mn^{2+} ions are oriented in the direction of the field, the internal fields around the acceptors all have almost the same magnitude (sample inhomogeneity can cause some deviations from the average value). In this case the photoexcitation lines in the spectra become narrower.

Returning to the experimentally obtained magnetic field dependences of the transition energies (Fig. 4), we call attention to the fact that these dependences are almost linear for both methods of cooling. This shows that in the present wide-gap semimagnetic semiconductors the binding energy of a magnetopolaron on an acceptor should be very small. If it is assumed for a given sample that this energy can be determined as $V=2\Delta E$ (Ref. 4), where ΔE is the energy splitting between the curves shown in Fig. 4, then we obtain $V=0.8$ meV. For comparison, we note that in the samples with a lower manganese concentration the binding energy, paradoxically, is appreciably larger [$V=2.8$ meV (Ref. 4)].

[1] A. Mycielsky and J. Micielski, J. Phys. Soc. Jpn. **49** (Suppl.), 807 (1980).
[2] A. D. Bykhovskiĭ, É. M. Vakhabova, B. L. Gel'mont, and I. A. Merkulov, Zh. Éksp. Teor. Fiz. **94**, 183 (1988).
[3] E. J. Georgitse, V. I. Ivanov-Omskii, D. I. Tsypishka, and V. A. Kharchenko, Int. J. Infrared Millim. Waves **13**, 155 (1992).
[4] V. I. Ivanov-Omskiĭ, V. A. Kharchenko, and D. I. Tsypishka, Fiz. Tekh. Poluprovodn. **26**, 1728 (1992) [Sov. Phys. Semicond. **26**, 967 (1992)].
[5] G. Bastard, C. Rigaux, J. Guidner, and A. Mycielski, J. de Pysique **39**, 87 (1978).

Translated by M. E. Alferieff

Indirect electronic transitions in semiconductors occurring as a result of scattering of charge carriers by dislocations in a quantizing magnetic field

É. M. Kazaryan and K. A. Mkhoyan

Erevan State University, 375049 Erevan, Armenia

(Submitted July 17, 1997; accepted for publication October 13, 1997)
Fiz. Tekh. Poluprovodn. **32**, 453–454 (April 1998)

The light absorption due to indirect electronic transitions in a semiconductor in a quantizing magnetic field is calculated under the assumption that an edge dislocation plays the role of a third body. The characteristic light frequency and magnetic field dependences of the absorption coefficient are determined for the mechanism considered. © *1998 American Institute of Physics.*
[S1063-7826(98)01104-1]

Indirect interband transitions in a semiconductor with a dislocation mechanism for scattering of charge carriers (CCs) under certain conditions imposed on the temperature and dislocation density can compete with indirect transitions due to the phonon scattering mechanism.[1,2]

In the present paper we study indirect interband transitions of CCs in a semiconductor in a uniform magnetic field with edge dislocations playing the role of a third body.

The light absorption (AC) associated with indirect electronic transitions in an external magnetic field can be calculated in second-order perturbation theory as

$$\alpha(\omega,H)=\frac{2\pi\hbar c}{n_r\omega}\frac{n_D S}{|A_0|^2}\frac{2\pi}{\hbar V}\times\sum_{N,\mathbf{k}}\sum_{N'',\mathbf{k}''}\sum_{N',\mathbf{k}'}\frac{|M^{\rm PHOT}|^2|M^{\rm DIS}|^2}{[E_g^0+\varepsilon_{N'}^c(k_z')-\varepsilon_N^v(k_z)-\hbar\omega]^2}\times\delta(E_g+\varepsilon_{N''}^c(k_z'')-\varepsilon_N^v(k_z)-\hbar\omega),\tag{1}$$

where ω and A_0 are the frequency and amplitude of the incident light wave, n_r is the refractive index of the medium, n_D is the dislocation density on an area S, V is the volume of the sample, E_g is the band gap, and E_g^0 is the band gap at the center of the Brillouin zone ($\mathbf{k}=0$):[1)] Since the z axis of the coordinate system is chosen in the direction of the magnetic field $\mathbf{H}$, while the vector potential $\mathbf{A}=(0, Hx, 0)$, the state of an electron in the band l (c or ν) is characterized by the set $\{N, k_y, k_z, l\}$.

In interactions of an edge dislocation with CCs the electrostatic part of the potential of the dislocation always predominates over the deformation part. In calculating the matrix element of the electron-dislocation interaction $M^{\rm DIS}$ we can therefore restrict the discussion to the electrostatic part. This potential in the dangling-bond approximation has the form[3]

$$V(\boldsymbol{\rho})=v_0K_0(\lambda\rho),\tag{2}$$

where $\boldsymbol{\rho}$ is the two-dimensional radius vector perpendicular to the line of the dislocation, $K_0(x)$ is the modified Bessel function of order 0, λ is the reciprocal of the Debye screening length, and

$$v_0=\frac{2e^2}{\varepsilon_S a}f.\tag{3}$$

Here a is the distance between the dangling bonds, f is the bond filling factor, and ε_S is the permittivity of the semiconductor.

Calculation of the matrix element $M^{\rm DIS}$ with the potential (2) shows that when the line of the dislocation is perpendicular to the magnetic field, we obtain[2)] the following expression for $M^{\rm DIS}$:

$$M^{\rm DIS}=\frac{2\pi}{L_yL_z}\frac{v_0}{\lambda^2+|\mathbf{k}''-\mathbf{k}'|^2}\delta_{N',N''},\tag{4}$$

where L_y and L_z are the linear dimensions of the sample in the corresponding directions.

Using the expression (4) and the well-known expression for the optical matrix element $M^{\rm PHOT}$, and switching in Eq. (1) from summation over $\mathbf{k}''$ and $\mathbf{k}$ to integration, we finally obtain the following expression for the AC:

$$\alpha(\omega,h)=\frac{2\pi^2n_D}{n_r\hbar\omega}\left(\frac{e}{m_0c}\right)^2(\mathbf{ep}_{cv}(0))^2\times v_0^2\frac{(m_L^cm^V)^{1/2}}{\hbar^2}\frac{eH}{(\lambda^2+k^2)^{3/2}}\times\sum_N\frac{\Theta\left(\hbar\omega-E_g-\left(N+\frac{1}{2}\right)\hbar\omega_H^*\right)}{\left[E_g^0+\left(N+\frac{1}{2}\right)\hbar\omega_H-\hbar\omega\right]^2},\tag{5}$$

where m_L^c is the electron effective mass at the point $\mathbf{k}_L$ of the main minimum,

$$\omega_H=\frac{eH}{c}\left(\frac{1}{m^c}+\frac{1}{m^v}\right)\quad\text{and}\quad\omega_H^*=\frac{eH}{c}\left(\frac{1}{m_L^c}+\frac{1}{m^v}\right)$$

are reduced cyclotron frequencies.

As one can see from Eq. (5), the dependence of $\alpha(\omega, H)$ on $\hbar\omega-E_g$ for an indirect-gap conductor is of a step char-

 © 1998 American Institute of Physics

acter at the fundamental absorption edge. The jumps correspond to values of $\hbar\omega$, for which transitions between energetically more distant levels are excited.

In contrast to the generally accepted phonon mechanism,[4] the main characteristic features of the light absorption mechanism, studied here, in a semiconductor in an external magnetic field are as follows:

1. $\alpha(\omega,H)\sim H$, i.e., the absorption coefficient for light is directly proportional to the magnetic field, in contrast to the phonon mechanism where the dependence is quadratic.

2. The absorption depends strongly on the momentum transfer $\hbar\mathbf{k}_L(\alpha\sim 1/k_L^3)$, which can give an additional possibility of determining experimentally the displacement of the conduction-band bottom relative to the valence-band top.

3. The absorption coefficient is proportional to the density n_D of scattering dislocations.

The numerical estimates presented here show that, for example, for GaP with density $n_D\simeq 3\times 10^7$ cm^{-2} and magnetic field $H\approx 10^2$ G, the dislocation mechanism can compete with the phonon mechanism even at temperature $T\approx 100$ K.

[1)]The expression (1) takes into account the fact that the valence band is completely filled, while the conduction band is empty.

[2)]It is easy to show that dislocations oriented parallel to **H** do not contribute to indirect transitions.

[1] A. A. Kirakosyan, M. K. Kumashyan, K. A. Mkhoyan, and A. A. Sarkisyan, Izv. Nats. Akad. Nauk Armenii, Fiz. **30**, 208 (1995).
[2] E. M. Kazaryan, K. A. Mkhoyan, and H. A. Sarkisyan, Thin Solid Films **7**, 302 (1997).
[3] V. L. Bonch-Bruevich and V. B. Glasko, Fiz. Tverd. Tela (Leningrad) **3**, 36 (1961) [Sov. Phys. Solid State **3**, 26 (1961)].
[4] B. Lax and S. Zwerdling, Prog. Semicond. **5**, 221 (1960).

Translated by M. E. Alferieff

Recombination of donor nickel excitons in the solid solutions $ZnSe_{1-y}S_y$:Ni

V. I. Sokolov[a)] and O. V. Dolzhenkov

Institute of Metal Physics, Ural Branch of the Russian Academy of Sciences, 620219 Ekaterinburg, Russia

(Submitted July 14, 1997; accepted for publication October 26, 1997)
Fiz. Tekh. Poluprovodn. **32**, 455–457 (April 1998)

We have found that the zero-phonon line of the donor exciton Ni $[d^7e]$ and its vibrational repetitions in the electroabsorption spectrum depend strongly on the composition of the solid solution $ZnSe_{1-y}S_y$:Ni ($y\leqslant 20\%$). Using the idea of an intermediate virtual $\{d^8\}^*$ state, we interpret this on the basis of the adiabatic potential model as an interaction of the d^7 configuration with the nearest-neighbor environment and nonradiative tunneling recombination $[d^7e]\rightarrow(d^8)^*$. © *1998 American Institute of Physics.* [S1063-7826(98)01204-6]

In II–VI compounds Ni atoms isovalently replace group-II atoms and possess a d^8 configuration. Transitions between multielectronic energy states in this configuration are manifested in the absorption and photoluminescence spectra. For ZnSe:Ni hydrogen-like Ni states are also observed at the edge of the photoionization band of Ni $d^8\rightarrow[d^9h]$ (acceptor exciton — AE) and in the photoionization band of Ni $d^8\rightarrow[d^7e]$ (donor exciton — DE).[1] Phonon-free transitions associated with Ni AE and DE in II–VI compounds depend very strongly on the temperature. This has been noted by many authors and has been investigated in greater detail in the electroabsorption spectra (EA) for CdS:Ni (Ref. 2) and in the excitation spectra of intracenter photoluminescence of ZnS:Ni.[3] In the present paper we report that disordering of the solid solution $ZnSe_{1-y}S_y$:Ni strongly influences the zero-phonon transitions associated with Ni DE.

The EA spectra were ordinarily recorded[1] for samples of the solid solutions $ZnSe_{1-y}S_y$:Ni obtained from melt in the composition range corresponding to $0.01\leqslant y\leqslant 20\%$ at temperatures 4.2 and 77 K in electric fields 5–20 kV/cm. The Ni density was equal to approximately $(1-3)\times 10^{17}$ cm^{-3}. Figures 1 and 2 show the EA spectra for three samples from the indicated range of compositions. The EA spectra for these samples are due to the effect of an electric field on the zero-phonon Ni DE line (2.64 eV), its vibrational repetitions, and the free-exciton line, as follows from the fact that these spectra undergo a definite transformation with increasing sulfur content in the solid solution, temperature, and amplitude E_M of the ac electric field. The EA spectrum for the sample with $y=0.05\%$ at 4.2 K is very similar to the EA spectrum of ZnSe:Ni, but the peaks are smaller and slightly broadened. The spectrum consists of a zero-phonon Ni DE line (2.64 eV), its vibrational repetitions, and a large positive peak due to the effect of an electric field on the free-exciton line. For the sample with $y\simeq 1\%$ at 4.2 K the peaks are broadened even more and slightly shifted in the direction of high energies. At 77 K only a large positive peak, due to the effect of an electric field on the free-exciton line, is recorded. The EA spectrum for the sample with $y\simeq 12\%$ at 4.2 K changes very radically. Only a large positive peak, shifted in the direction of high energies in accordance with the increase in the band gap E_g, is observed. The decrease in the amplitude of this peak attests to broadening of the free-exciton line for this composition of the solid solution. The zero-phonon Ni DE line and its first vibrational repetitions are not observed. The weak feature on the low-energy side of the large peak can be explained as the manifestation of the most intense repetitions of the zero-phonon line of the Ni DE, since it vanishes at 77 K. For this sample, the free-exciton line is observed well in the reflection spectrum at 77 K. Its position ($\hbar\omega=2.845$ eV) makes it possible to estimate the composition of the solid solution quite reliably. The EA spectrum for the sample with $y=20\%$ consists of a positive peak with the energy of the maximum determined by the influence of the field on the free-exciton line whose position is determined according to the character of the dropoff in the reflection spectrum at 77 K.

Of the diverse changes occurring in the EA spectrum with increasing y, we note only the obvious weakening of the zero-phonon Ni DE line. It is due to broadening of the zero-phonon Ni DE line. As a result, the line and its first vibrational repetitions become unobservable for samples with y near 10%. This sharp dependence of the line intensity on the composition of the solid solution contrasts with the weak dependence of the free-exciton line on the composition of the solid solution $ZnSe_{1-y}S_y$:Ni in our electroabsorption and reflection spectra, as well as in the photoluminescence spectra[4] and of the N AE line (A line) in the solid solutions $GaP_{1-y}As_y$:N.[5] Thus, we can state that the hydrogen-like excited states of Ni in II–VI solid solutions depend much more strongly on the composition of the solid solution than free-exciton states and hydrogen-like excited states of simple isoelectronic impurities. In many ways this is reminiscent of the sharp difference between the effect of temperature on the no-phonon Ni DE line, on the one hand, and on the free-exciton line and the A line in GaP:N, on the other.

The broadening of the no-phonon Ni DE line could be due to the effect of the disordering of the solid solution both on the hydrogen-like states and on the d shell. In our view, the fluctuation relief of the conduction-band bottom does not explain the strong effect of the disordering of the solid solu-

 © 1998 American Institute of Physics

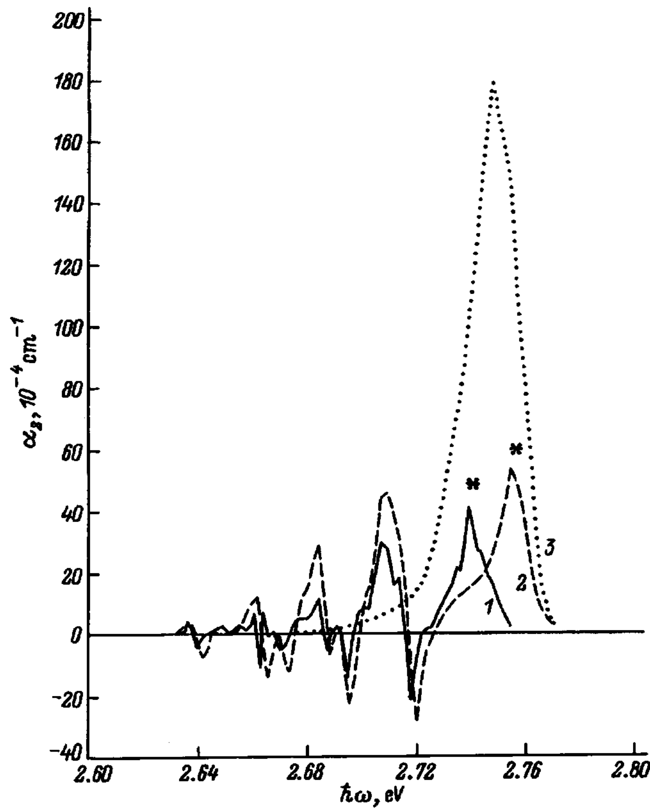


FIG. 1. Spectral dependence of the amplitude α_2 of the second harmonic of electroabsorption for samples with $y=0.05\%$ (*1*) and $y\simeq 1\%$ (*2, 3*). Measurement regimes: *1* — $T=4.2$ K, $F_m=10$ kV/cm; *2* — $T=4.2$ K, $F_m=14$ kV/cm; *3* — $T=77$ K, $F_m=16.1$ kV/cm. The positive peaks marked by the asterisk are decreased by a factor of 10.

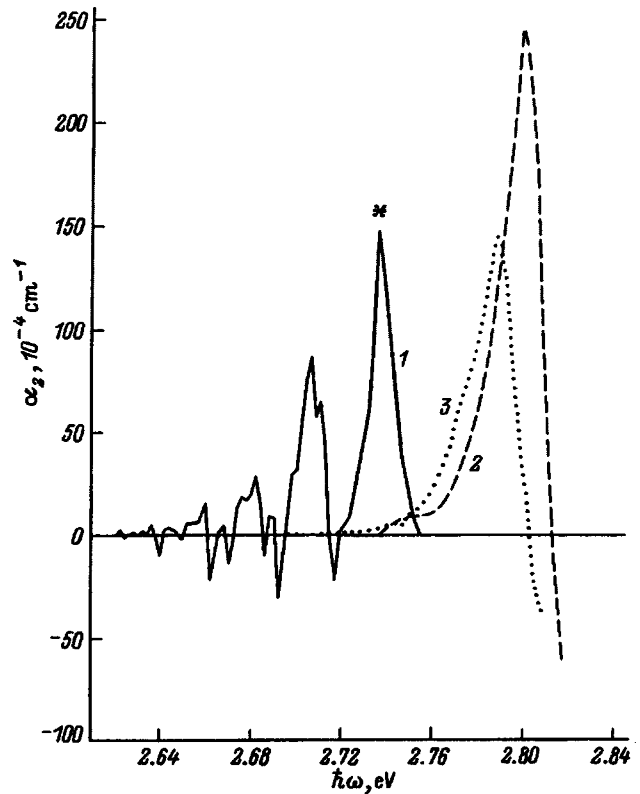


FIG. 2. Spectral dependence of the amplitude α_2 of the second harmonic of electroabsorption for samples with $y=0.05\%$ (*1*) and $y\simeq 12\%$ (*2, 3*). Measurement regimes: *1* — $T=4.2$ K, $F_m=15$ kV/cm; *2* — $T=4.2$ K, $F_m=21.4$ kV/cm; *3* — $T=77$ K, $F_m=21$ kV/cm. The positive peak marked by the asterisk is decreased by a factor of 10; the vertical bar on the right-hand side shows the position of the free-exciton line for the sample with $y\simeq 12\%$ at 77 K.

tion on the hydrogen-like excited states of Ni impurity atoms. The free-exciton line, clearly observed in the reflection and electroabsorption spectra of solid solutions with $y=12$ and 20%, should be broadened as a result of fluctuations of both the valence-band top (ΔE_v) and the conduction-band bottom (ΔE_c). The Ni DE line broadens only as a result of fluctuations of the conduction-band bottom, which are much weaker, since in the solid solution $ZnSe_{1-y}S_y$:Ni the change ΔE_c equals 1/3 of the change ΔE_v, while the electron mass is less than the heavy-hole mass ($0.17m_0$ and $0.70m_0$).[6] Broadening of the (0/+) level, i.e., a shift of the d^8 and d^7 configuration states as a result of large-scale fluctuations, is unlikely. The formation of an appreciable fraction of $NiSe_4S_1$ clusters should result in the appearance of a new zero-phonon Ni DE line similarly to the appearance of several *A* lines in accordance with the different number of As atoms in the second coordination sphere for the solid solutions $GaP_{1-y}As_y$:N.[5] New no-phonon Ni DE lines did not appear in the solid solutions $ZnSe_{1-y}S_y$:Ni.

The sharp decrease in the amplitude of the Ni DE peaks with increasing *y* for small values of *y* can be understood as the result of a decrease in the Ni DE lifetime, which is described in the adiabatic-potential model, similarly to the strong temperature dependence.[1,2] Figure 3 shows the configuration diagrams for two excited states of the d^8 configuration and for the Ni DE state $[d^7e]$. The no-phonon Ni DE line is due to transitions occurring under the influence of a photon $\hbar\omega_d$ into the zeroth electronic-vibrational state at the bottom of the adiabatic potential for Ni DE (the transition $d^8+\hbar\omega_d\rightarrow[d^7e]$).[1] In the solid solution $ZnSe_{1-y}S_y$:Ni with small *y* all Ni atoms are surrounded mainly by four Se atoms and the relative fraction of $NiSe_3S_1$ clusters is very small. As *y* increases, the lattice constant *a* decreases systematically in accordance with Vegard's law $a=a_0-ky$, which results in an additional change in the impurity–nearest-neighbor distance for all $NiSe_4$ clusters. The configuration curve for the $[d^7e]$ state will undergo a shift relative to the curves for the excited states of the d^8 configuration (Fig. 3), since the d^7 configuration, which corresponds to the Ni^{3+} ion, is more sensitive to a displacement of the nearest-neighbor ions. We believe that the process which we term tunneling annihilation of a Ni DE intensifies with the shift. At the point *B* a transformation $[d^7e]\rightarrow\{d^8\}^*$ into one of the excited states of the d^8 configuration without a change in the coordinate *Q* occurs. This transformation could be caused by nonadiabaticity terms containing derivatives of the electronic wave function with respect to the normal coordinate. As *y* increases, the slope of the curve for $[d^7e]$ and the corresponding derivatives increase. After the transformation the nearest-neighbor environment undergoes restructuring and an intermediate virtual state $\{d^8\}^*$ transforms into a normal excited state $(d^8)^*$ of the d^8 configuration. The resulting transition is shown by the arrow in Fig. 3. After the transition

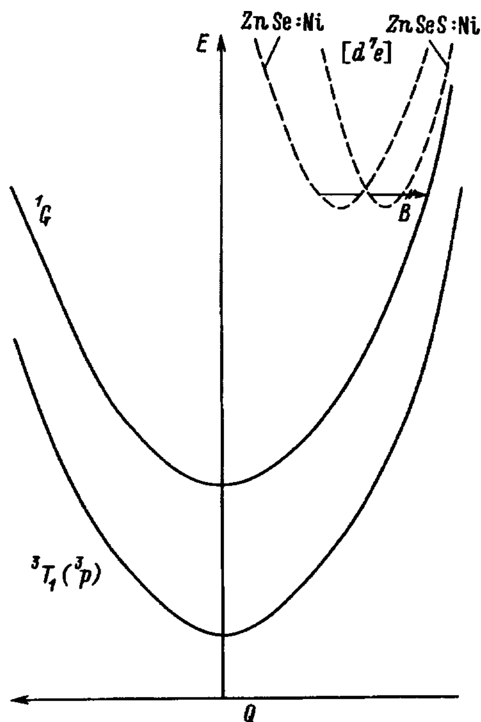


FIG. 3. Configuration curves for the d^8 states of the Ni^{2+} ion in $ZnSe_{1-y}S_y$:Ni and for the Ni DE states $[d^7e]$ in ZnSe:Ni and $ZnSe_{1-y}S_y$:Ni. The rightward displacement of the Ni DE curve with increasing y is shown schematically. The arrow at the point B shows the tunneling transition from the Ni DE state into the state of a d^8 configuration.

vibrational relaxation starts along the configuration curve for the $(d^8)^*$ state. The intermediate virtual state $\{d^8\}^*$ can be interpreted similarly to an intermediate state in the conduction band for describing indirect transitions in Ge and Si.

In summary, we have shown that the disordering of the solid solution $ZnSe_{1-y}S_y$:Ni strongly influences the no-phonon Ni DE line $[d^7e]$. This occurs through the d^7 configuration accompanying a change in the lattice constant with increasing y of the solid solution and not via the electron in a hydrogen-like orbit, which corresponds to the ideas of Refs. 7 and 8. This influence gives rise to a tunneling transition $[d^7e]\rightarrow(d^8)$ in the configuration-curve model. The novelty here as compared with Ref. 2 lies in the application of the idea of an intermediate virtual state $\{d^8\}^*$. The very strong dependence of the d^7 configuration on the nearest-neighbor environment potentially could be helpful for investigating local nonuniformities of Ni-doped II–VI solid solutions.[1]

We thank A. Tsunger for a helpful discussion of the problem of $3d$ impurities in solid solutions of semiconductors.

[a]Fax: (3432) 745-244; e-mail: Optics@ifm.e-burg.su

[1] V. I. Sokolov, Fiz. Tekh. Poluprovodn. **28**, 545 (1994) [Semiconductors **28**, 329 (1994)].
[2] V. I. Sokolov and A. N. Mamedov, JETP Lett. **43**, 237 (1986).
[3] R. Heitz, A. Hoffman, and I. Broser, Phys. Rev. **48**, 8672 (1993).
[4] P. N. Newbury, K. Shahzad, J. Petruzello, and D. A. Cammack, J. Appl. Phys. **66**, 4950 (1989).
[5] G. F. Glinskiĭ, M. V. Lupal, I. I. Parfenov, and A. N. Pikhtin, Fiz. Tekh. Poluprovodn. **26**, 644 (1992) [Sov. Phys. Semicond. **26**, 364 (1992)].
[6] N. N. Ablyazov, M. E. Raĭkh, and A. L. Éfros, Fiz. Tverd. Tela (Leningrad) **25**, 353 (1983) [Sov. Phys. Solid State **25**, 199 (1983)].
[7] K. A. Kikoin, V. I. Sokolov, V. N. Flerov, and V. V. Chernyaev, Zh. Éksp. Teor. Fiz. **83**, 2335 (1982) [Sov. Phys. JETP **56**, 1354 (1982)].
[8] V. I. Sokolov and T. P. Surkova, Fiz. Tverd. Tela (Leningrad) **29**, 2938 (1987) [Sov. Phys. Solid State **29**, 1689 (1987)].

Translated by M. E. Alferieff

SEMICONDUCTOR STRUCTURES, INTERFACES, AND SURFACES

Photosensitivity of thin-film structures based on laser-deposited $CuIn(Te_xSe_{1-x})_2$ layers

I. V. Bodnar'[a)] and V. F. Gremenok

State University of Informatics and Radioelectronics, 220027 Minsk, Belarus

V. Yu. Rud'

St. Petersburg State Technical University, 195251 St. Petersburg, Russia

Yu. V. Rud'

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia

(Submitted May 27, 1997; accepted for publication May 29, 1997)
Fiz. Tekh. Poluprovodn. **32**, 458–460 (April 1998)

Thin films of the solid solutions $CuIn(Te_xSe_{1-x})_2$ $(0<x<1)$ exhibiting chalcopyrite structure were obtained by the method of laser deposition. Using half-transmitting indium layers, Schottky diodes were prepared on the basis of the films obtained. The spectral dependence of the sensitivity as a function of the ratio between Te and Se was investigated by illuminating the structures through the In contact. Analysis of the experimental results showed that the region of spectral sensitivity of such thin-film structures depends on the tellurium content in the $CuIn(Te_xSe_{1-x})_2$ layers. © *1998 American Institute of Physics.* [S1063-7826(98)01304-0]

1. INTRODUCTION

Ternary chalcopyrite semiconductors ABC_2 (A — Cu, Ag; B — Al, Ga, In; C — S, Se, Te) are now being closely scrutinized by investigators as promising materials for producing different optoelectronic devices.[1–3] The most interesting compound in this class is copper selenoindiate, which has been used to develop and fabricate solar cells which have an efficiency above 16% (Ref. 4) and which are more stable and show a higher radiation resistance than similar cells based on silicon and gallium arsenide.[3]

A number of investigators have also shown that films based on the solid solutions $CuIn(Te_xSe_{1-x})_2$ can be successfully used as active layers in thermophotovoltaic systems for converting radiation in the spectral range 0.5–1.0 eV.[5–8]

In the present paper we report the results of investigations of the photosensitivity of thin-film structures based on laser-deposited layers of the solid solutions $CuIn(Te_xSe_{1-x})_2$.

2. FILM PREPARATION

A GOS-1001 laser operating in the free-lasing mode at wavelength $\lambda=1.06$ μm with pulse duration $\simeq 10^{-3}$ s was used to obtain the films. Following the procedure described earlier,[9–11] the films were deposited on glass substrates with $T_p=430-470$ °C and pressure in the chamber $(2-4)\times 10^{-5}$ Torr by a series of 180 to 200-J pulses with energy flux density at the target $(3-5)\times 10^5$ J/cm^2. The thickness of the films on the active area $\simeq 2$ cm^2 was equal to 0.4–1.2 μm, depending on the number of pulses. Large-block, 12–16 mm in diameter, $CuIn(Te_xSe_{1-x})_2$ ingots were used as the targets for sputtering.

3. RESULTS AND DISCUSSION

The homogeneity of the films and the structures obtained was monitored by the x-ray method using a DRON-3M diffractometer with CuKα radiation and a nickel filter. The diffraction patterns of the solid solutions contained a system of lines that corresponds to the chalcopyrite structure characteristic of the ternary compounds $CuInSe_2$ and $CuInTe_2$ forming this system. The observed resolution of the high-angle lines indicates that the films obtained were quite homogeneous.

The atomic composition of the films was determined with a JEOL scanning electron microscope. The results of these measurements are presented in Table I. One can see that within the limits of accuracy attained in these measurements ($\simeq 4\%$) the composition of the films corresponds satisfactorily to that of the source material.

Uniform films of the solid solutions $CuIn(Te_xSe_{1-x})_2$ were used as the initial films for producing photosensitive structures of several types. One consisted of an optical contact which was produced by clamping the surfaces of the films of the solid solutions to the natural cleavage surfaces of the layered semiconductors *n*-InSe and *n*-GaSe. Such structures exhibited sufficiently high photosensitivity ($\simeq 10^3$ V/W) under illumination on the III–VI semiconductor side. The photosensitivity of these anisotypic structures dominated in the fundamental absorption region of the

 © 1998 American Institute of Physics

TABLE I. Results of analysis of the composition of the compounds $CuInSe_2$ and $CuInTe_2$ and the solid solutions $CuIn(Se_{1-x}Te_x)_2$.

Copper, at. %			Indium, at. %			Tellurium, at. %			Selenium, at. %		
calc.	exp.		calc.	exp.		calc.	exp.		calc.	exp.	
	crystal	film		crystal	film		crystal	film		crystal	film
25.00	24.85	25.42	25.00	25.07	24.62	50.00	50.33	50.82	—	—	—
25.00	25.32	26.84	25.00	25.61	25.07	42.51	42.00	41.22	7.50	7.18	6.77
25.00	24.88	25.02	25.00	25.00	25.23	24.54	25.12	24.70	25.00	24.30	26.02
25.00	25.29	24.73	25.00	24.63	25.08	15.00	14.99	15.20	35.00	35.08	35.80
25.00	24.65	25.42	25.00	25.20	24.98	—	—	—	50.00	50.10	49.77

III–VI crystals. This is due to their higher resistance with respect to films of the solid solutions $CuIn(Te_xSe_{1-x})_2$ and, correspondingly, localization of the space charge layer mainly in the volume of the InSe and GaSe crystals. For this reason, such structures did not make it possible to observe photosensitivity in the fundamental absorption region of $CuIn(Te_xSe_{1-x})_2$.

The structures of the second type were surface-barrier structures obtained on the exterior 2×2-mm surface of the half-transmitting indium layers by thermal deposition through a mask. When such structures were illuminated on the metal/semiconductor barrier side as well as on the solid-solution film side, the barrier contact always became negatively charged and the sign of the charge remained unchanged under all possible illumination geometries in the entire region of photosensitivity of the structures obtained. This shows that the observed photovoltaic effect is determined by separation of photogenerated pairs in a single active region arising on the indium/semiconductor boundary. As a rule, the photosensitivity dominates when these structures are illuminated on the barrier contact side. For the best structures, which we were able to obtain on films with a tellurium content in the range 30–50 mole%, the maximum current photosensitivity reaches 1 mA/W at $T=300$ K. It should also be noted that the obtained spectral dependences of the short-circuit photocurrent, which was proportional to the flux density of the incident radiation, was converted into an equal number of incident photons, so that they actually reflect the relative quantum efficiency of the photoconversion of natural radiation by means of such elements. When the surface of the barrier contact was scanned, the photoelectric parameters were reproduced well from point to point. This shows that the films of the solid solutions obtained are quite uniform (the diameter of the light probe is 0.2–0.4 mm).

Figure 1 shows for the $In/CuIn(Te_xSe_{1-x})_2$ structures obtained the typical spectral dependences of the relative quantum efficiency η of photoconversion at room temperature and under illumination by natural light in a direction normal to the photoreceiving plane. It is evident from the figure that despite the changes in the atomic composition of the solid solutions, the photosensitivity curves remain similar in the entire concentration range. As photon energy increases above 0.8 eV, the photosensitivity follows an exponential law with slope $S=\delta(\ln\eta)/\delta(\hbar\omega)\simeq30-50$ eV^{-1}, which is consistent with the assumption that the bands in the film studied possess a simple structure. As photon energy is increased, a step (curves *1* and *2*) or kink (curves *3*–*5*) forms in the spectral curves $\eta(\hbar\omega)$ at a definite value of $\hbar\omega$ for each composition of the solid solution. The inset in Fig. 1 shows the position of these features on the energy axis versus the composition of the films in comparison with the dashed line connecting the values of the band gaps of the ternary compounds $CuInSe_2$ and $CuInTe_2$.[2,12] In the composition range $x>0.3$ the features in the quantum efficiency spectra fall on the dashed line, making it possible to relate the long-wavelength exponential edge of $\eta(\hbar\omega)$ with interband absorption in the semiconductor film with the corresponding value of x. As x is lowered, $x<0.3$, these features in the spectral curve $\eta(\hbar\omega)$ start to shift relative to the dashed line into the low-energy region, which could indicate that the long-wavelength photosensitivity edge is due to an increase in the contribution of shallow defect levels to photoactive absorption. It should also be noted that despite such a strong

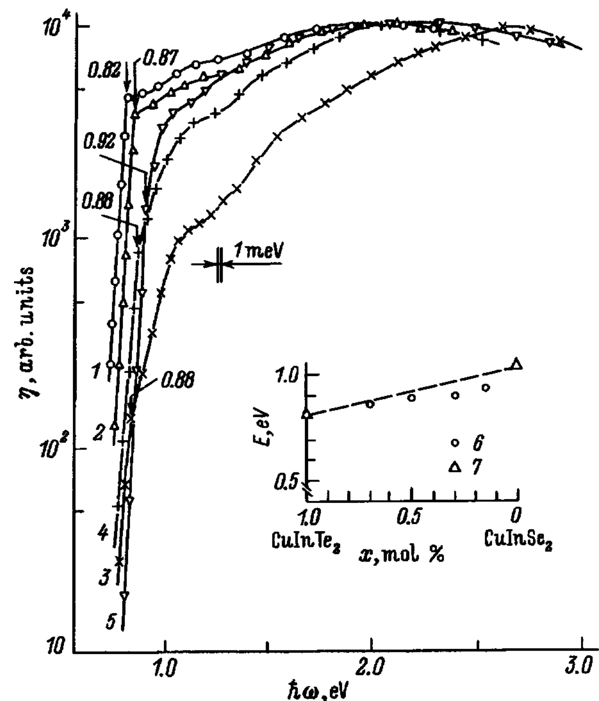


FIG. 1. Spectral dependences of the relative quantum efficiency η of photoconversion in $In/p\text{-}CuIn(Te_xSe_{1-x})_2$ structures at $T=300$ K; x, mol%: *1* — 0, *2* — 0.7, *3* — 0.5, *4* — 0.3, *5* — 0.15. Inset: Band gap of the solid solutions versus composition (*6* — position of the features of the long-wavelength edge on the energy axis in the relative quantum efficiency spectra, *7* — band gaps of ternary compounds according to Refs. 2 and 13).

variation in composition of the films ($x=0-1$), the long-wavelength edge of η shifts relatively little, which is attributable, correspondingly, to the small effect of the positional substitution of selenium for tellurium atoms on the band gap of the solid solutions which we obtained.[2,12] As one can see from the figure, for all film compositions the photosensitivity increases as photon energy increases further. Therefore, recombination at the surface of the solid-solution films can be suppressed in the surface-barrier structures. This fact also shows at the same time that the proposed laser-deposition regime makes it possible to obtain films of solid solutions with a quite perfect (from the standpoint of carrier recombination) exterior surface without resorting to any additional treatment. The increase of photosensitivity in the fundamental absorption region of the solid solutions $CuIn(Te_xSe_{1-x})_2$ is in qualitative agreement with the spectral dependence of the optical absorption coefficient in $CuInSe_2$ thin films.[2,13,14] Therefore, there are grounds for judging on the basis of the observed short-wavelength growth of the photosensitivity of the surface-barrier structures the optical absorption spectrum of thin films obtained by pulsed laser deposition.[9]

On the whole, it can be concluded on the basis of the results presented that the technology for producing $CuIn(Te_xSe_{1-x})_2$ films could find application for construction of high-efficiency, thin-film solar cells.

It should also be specially noted that an attempt to observe natural photopleochroism in the structures did not lead to observation of any dependence of the photosensitivity on the polarization of the incident radiation.[13,14] This result could be due to the polycrystalline structure of the films, indicating at the same time the absence of pronounced texture in them.

[a]Fax: (375)-(172)-310914; e-mail: cit@micro.rei.minsk.by

1. J. L. Shay and J. H. Wernick, *Ternary Chalcopyrite Semiconductors: Growth, Electronic Properties and Applications*, Pergamon Press, N. Y., 1975.
2. T. J. Coutts, L. L. Kamerskii, S. Wagner [Eds.], *Copper Indium Diselenide for Photovoltaic Applications*, Amsterdam, 1985.
3. T. J. Coutts and J. O. Meakin [Eds.], *Current Topics in Semiconductor Photovoltaics*, Academic Press, N. Y., Vols. 1–3, 1875–1988 [Russian translation, Mir, Moscow, 1988].
4. H.-W. Schock, Sol. Energy Mater. Sol. Cells **34**, 19 (1994).
5. E. S. Vera, J. J. Laferskii, and N. Spitzer, in *Proceedings of the 4th European International Conference on Photovoltaic Solar Energy Conversion*, Italy, 1996, p. 659.
6. M. Quitero and J. C. Wooley, J. Appl. Phys. **55**, 15 (1984).
7. R. Diaz, M. Leonard, and F. Ruenda, J. Vac. Sci. Technol. A **10**, 3295 (1992).
8. M. Leonard, R. Diaz, and F. Ruenda, J. Vac. Sci. Technol. A **12**, 3082 (1994).
9. V. F. Gremenok, E. P. Zaretskaya, I. V. Badnar, Yu. V. Rud', and M. A. Magomedov, Thin Solid Films **232**, 139 (1993).
10. E. P. Zaretskaya, V. F. Gremenok, I. A. Viktorov, and I. V. Bodnar', Pis'ma Zh. Tekh. Fiz. **22**, 53 (1996) [Sov. Tech. Phys. Lett. **23**, 578 (1996)].
11. V. F. Gremenok, I. A. Viktorov, M. V. Yakushev, and V. V. Kindyak, Pis'ma Zh. Tekh. Fiz. **22**, 9 (1996) [Tech. Phys. Lett. **22**, 954 (1996)].
12. *Reference Data on the Physical and Chemical Properties of Semiconductors* [in Russian], Nauka, Moscow, 1979.
13. N. N. Konstantinova, M. A. Magomedov, V. Yu. Rud', and Yu. V. Rud', Fiz. Tekh. Poluprovodn. **25**, 2047 (1991) [Sov. Phys. Semicond. **25**, 1233 (1991)].
14. N. N. Konstantinova, M. A. Magomedov, V. Yu. Rud', and Yu. V. Rud', Fiz. Tekh. Poluprovodn. **26**, 558 (1992) [Sov. Phys. Semicond. **26**, 317 (1992)].

Translated by M. E. Alferieff

Electric and photoelectric properties of n-$Ga_xIn_{1-x}N/p$-Si anisotypic heterojunctions

S. E. Aleksandrov, V. A. Zykov, T. A. Gavrikova, and D. M. Krasovitskiĭ

St. Petersburg State Technical University, 195251 St. Petersburg, Russia
(Submitted October 1, 1997; accepted for publication October 30, 1997)
Fiz. Tekh. Poluprovodn. **32**, 461–465 (April 1998)

We have developed a technology for producing n-type $Ga_xIn_{1-x}N/p$-Si heterostructures by combined pyrolysis of indium and gallium monoammoniate chlorides, making it possible to obtain heterolayers with composition varying over wide limits (from GaN up to InN). The composition and basic electric and optical characteristics of nitride films were determined. The electric and photoelectric properties of the heterostructures with $Ga_xIn_{1-x}N$ films of different composition were investigated. It was shown that the anisotypic heterojunction n-$Ga_xIn_{1-x}N/p$-Si is a promising photosensitive element for detecting visible-range radiation. The maximum values of the specific detectivity were $D^* = 1.2\times 10^{11}\ \mathrm{Hz}^{1/2}\cdot \mathrm{W}^{-1}$ at 290 K. A band diagram of the heterojunction was constructed. © *1998 American Institute of Physics.*
[S1063-7826(98)01404-5]

The solid solutions of the semiconductor compounds $Ga_xIn_{1-x}N$ possess unique properties, which are of great interest for applications in optoelectronic devices.[1] In these direct-gap materials, depending on the composition, the band gap E_g varies continuously from 1.9 to 3.4 eV, covering the entire visible and part of the UV range of photon energies. Together with the progress made in recent years in their use in light-emitting devices[2,3] nitrides hold promise for applications in photoreceiving devices, for example, in solar cells, photodiodes, and phototransistors. Heterostructures formed by deposition of a material with a large band gap on a semiconductor substrate are usually characterized by a higher photosensitivity because of the so-called ''window effect.''[4] Moreover, in the case where a wide-gap semiconductor exhibits comparatively high electrical conductivity, as in the case of gallium and indium nitrides as a result of deviations of their composition from stoichiometry in the direction of excess metal, a higher photosensitivity can be expected on account of a decrease in the series resistance of the structure as a whole. If the heterolayer is formed by chemically stable and radiation-resistant material, which is also characteristic of nitrides, the photosensitive structure does not require formation of additional protective coatings on the surface.

On this basis it can be assumed that $Ga_xIn_{1-x}N$/Si structures, being cheap, should exhibit characteristics comparable and even exceeding those of existing photodetectors. Unfortunately, such heterostructures have not been adequately investigated, apparently because of technological problems associated with the formation of perfect layers of nitride solid solutions.[5] Molecular-beam epitaxy (MBE) and chemical vapor deposition using metalorganic compounds (MOCVD) are successfully used to obtain high-quality films of GaN and nitride solid solutions $Ga_xIn_{1-x}N$ ($x \geqslant 0.7$). The technology for obtaining InN and solid solutions with a high indium content is much less well developed.[5–11] At the same time, chemical vapor deposition (CVD) methods using halogen-containing initial reagents continue to be used and give fairly good results.[11–14]

Our objective in the present work is to develop a technology for producing heterostructures of the type $Ga_xIn_{1-x}N$/Si that permits varying the composition of the heterolayer over wide limits (from GaN to InN) as well as studying the basic electrical and optical properties of films and the photoelectric characteristics of the heterostructures.

The nitride films were obtained by chemical vapor deposition, based on combined pyrolysis of gallium monoammonia chloride (GAC) and indium monoammonia chloride (IAC). The process was conducted in a nitrogen–ammonia atmosphere in a continuous-flow system. The deposition process was conducted at atmospheric pressure and substrate temperature 550 °C. The gallium and indium ratio in the growing film was fixed by varying the GAC and IAC ratio in the initial charge. The ratio of the molar flow rate of ammonia to the total molar flow rate of the metal containing reagents (V/III ratio) at the reactor entrance was varied within the limits 10–100. The growth rate of the films was in the range 0.8–1 μm/h and their thickness was equal to 1 μm. KU-1 optical quality quartz and KDB-10 (100) Si wafers were used as substrates. The reactor layout and the film deposition procedure are described in Ref. 15.

The films obtained consist of polycrystalline deposits with good adhesion and continuity. The color varies with increasing indium content in the load from almost transparent (pure GaN) to reddish-brown. The results of an analysis of the transmission and reflection spectra of the films in the wavelength range 300–1200 nm, in the course of which the optical band gap $E_g^{\rm opt}$ was determined and the values of the optical constants (absorption coefficient k and refractive index n) were calculated by the method of Ref. 16, show that the layers were single-phase and uniform. The composition of the $Ga_xIn_{1-x}N$ films was estimated from Auger-spectroscopy data. The correlation established in the present work between the values of $E_g^{\rm opt}$ and x agree satisfactorily

 © 1998 American Institute of Physics

TABLE I. Technological production parameters and properties of $Ga_xIn_{1-x}N$ films deposited on quartz substrates.

Sample No.	Molar fraction of GAC in the charge	V/III ratio	x	E_g^{opt}, eV	k $\lambda=0.8$ μm	n
1	1.0	10	1.00	3.4	0.018	1.92
2	0.37	70	0.85	3.05	0.044	2.03
3	0.23	100	0.64	2.25	0.050	2.08

TABLE II. Basic parameters of the heterojunctions n-$Ga_xIn_{1-x}N/p$-Si at $T=290$ K.

Sample No.	Film composition	E_g^{opt}, eV	R_0, kΩ	V_D, eV	V_{Dn}, eV	τ, μm	E_τ, eV
1	GaN	3.40	20	1.0	0.44	300	0.29
2	$Ga_{0.85}In_{0.15}N$	3.05	5	1.2	0.36	50	0.28
3	$Ga_{0.64}In_{0.36}N$	2.25	2	0.8	0.13	20	0.26

with the results of a study of $E_g^{opt}(x)$ for solid solutions in the system GaN–InN.[6,17]

All films obtained are n-type. The current carrier density determined from Hall effect measurements at 300 K was equal to $10^{17}-10^{19}$ cm^{-3}, increasing with In concentration in the solution. The Hall electron mobility in the films did not exceed 10 cm^2/(V·s) at 300 K. The electrical resistance of the films depended largely on their indium content and varied at room temperature from ~ 100 kΩ, characteristic for pure GaN, down to values ~ 500 Ω measured for $Ga_{0.64}In_{0.36}N$ films. The basic parameters of the films grown on quartz substrates are presented in Table I.

Anisotropic heterojunctions n-$Ga_xIn_{1-x}N/p$-Si were prepared in the above-described experiments on growing nitride films. Deposition was done both on naturally oxidized silicon surfaces and on surfaces from which oxide was removed by gas etching in a HCl flow. It should be noted that the precleaning of the silicon surface had no effect on any of the parameters of the heterojunctions investigated. This fact apparently reflects a specific feature of the film deposition method: The feature associated with the presence of hydrogen chloride in the reaction gas mixture in the deposition zone, resulting in the removal of oxide at the initial stages of the process.

The heterojunctions obtained were subjected to a battery of investigations, including measurement of the current-voltage characteristics, the photoresponse signal (in the photo-emf regime) due to blackbody radiation at 900 K, the spectral distribution of the sensitivity, and the photoresponse time. The measurements were performed at temperatures $T=80-290$ K. A MDR-23 monochromator was used to measure the spectral characteristics. The photoresponse time was determined from the relaxation curves of the photoresponse signal. In this case a GaAs LED served as the source of radiation. In all experiments the radiation was incident in a direction perpendicular to the surface of the heterojunction through a wide-gap material — a $Ga_xIn_{1-x}N$ film. The samples had an area of 5×5 mm^2.

Chromium was used as the contact material for investigating the electric and photoelectric properties of the heterojunctions. It was deposited in the form of a 0.3-μm-thick 0.7×5-mm^2 strips, as well as in the form of a continuous layer on the backside of the silicon by magnetron sputtering onto a wide-gap material. The linearity of the I-V characteristics of such contacts shows that they are nearly ohmic.

The basic parameters of the heterojunctions investigated at $T=290$ K are presented in Table II. Here R_0 is the differential resistance of the junction under zero bias, V_D is the total contact potential difference, V_{Dn} is the diffusion potential on the $Ga_xIn_{1-x}N$ film side, τ is the photoresponse time, and E_τ is the activation energy of the photoresponse time.

All the heterojunctions which we investigated in this study are photosensitive in the temperature range $T=80-290$ K. At each temperature the magnitude of the photoresponse signal depends on the composition of the nitride film. The highest photo-emf was recorded for heterojunctions with a film of the material with the largest band gap — GaN. For this sample the specific detectivity at 290 K is $D^*=1.2\times10^{11}$ Hz$^{1/2}$·W^{-1}·cm. As the indium content in the film increases, the photosensitivity decreases by almost an order of magnitude (Table II, sample 3). The temperature dependences of the photoresponse signal S are nonmonotonic. As the temperature is lowered from 290 K, the photoresponse signal in all the heterojunctions increases, reaches maximum values in the temperature range $200-250$ K, and then decreases. For some heterojunctions the sensitivity at the maximum of the temperature dependence is more than an order of magnitude higher than at 290 K: For example, for the heterojunction n-$Ga_{0.85}In_{0.15}N/p$-Si $D^*=2\times10^{12}$ Hz$^{1/2}$· W^{-1}· cm at $T=200$ K. The values of the photo-emf at 80 and 290 K are approximately the same for all heterojunctions.

The spectral characteristics of the photosensitivity measured at $T=290$ K (Fig. 1) have the typical form for heterojunctions. The red edge of the spectral characteristic is determined by the band gap of the narrow-gap semiconductor — Si. The short-wavelength decrease of the photosensitivity coincides in energy with the band gap of the wide-gap semiconductor. As E_g in the nitride film increases, the short-wavelength part of the spectral characteristic shifts into the region of shorter wavelengths. It should be noted that for all wavelengths of the incident radiation the sensitivity of a heterojunction increases with increasing E_g of the wide-gap layer.

The maximum photosensitivity for all heterojunctions falls in the absorption region of silicon. However, the sensitivity distribution, which depends on the film composition, is observed to be nonmonotonic in this region of the spectral characteristic. This feature is most clearly seen in the spectral characteristics of the photoresponse signal $S/S_{max}=f(\lambda)$ (Fig. 2), where S_{max} is the maximum value of the photoresponse on the spectral characteristic of the given sample. As the indium content in the nitride film increases, an additional long-wavelength sensitivity peak appears at $\lambda\simeq1.1$ μm, which for the heterojunction n-$Ga_{0.64}In_{0.36}N/p$-Si is comparable in magnitude to the maximum at $\lambda\simeq0.8$ μm present in all heterojunctions (Fig. 2, curve *3*). In the sample with a lower indium content in the film this peak decreases in mag-

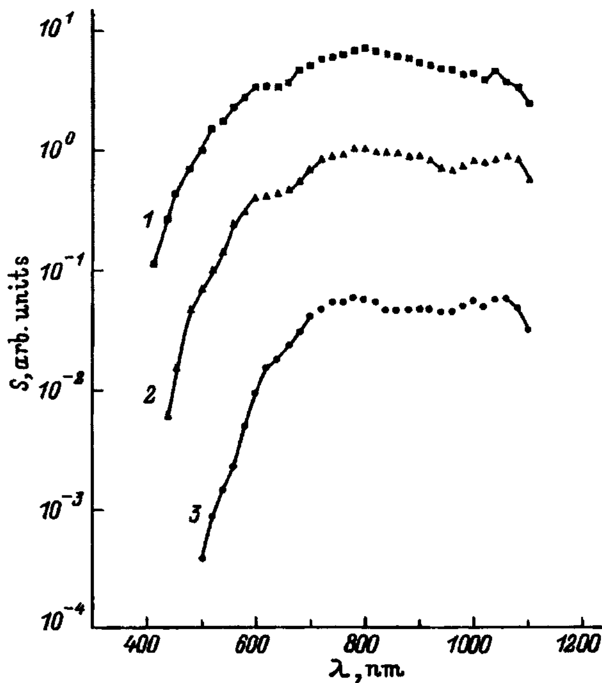


FIG. 1. Spectral photosensitivity characteristics of heterojunctions: *1* — GaN/Si, *2* — $Ga_{0.85}In_{0.15}N/Si$, *3* — $Ga_{0.64}In_{0.36}N/Si$.

nitude (Fig. 2, curve *2*), while it is virtually absent in the heterojunction *n*-GaN/*p*-Si (Fig. 2, curve *1*). The described nonmonotonic variations of the photoresponse signal of heterojunctions and their explicit dependence on the composition of the wide-gap film could be due to a characteristic feature of the transmission of light by the films in this wave-

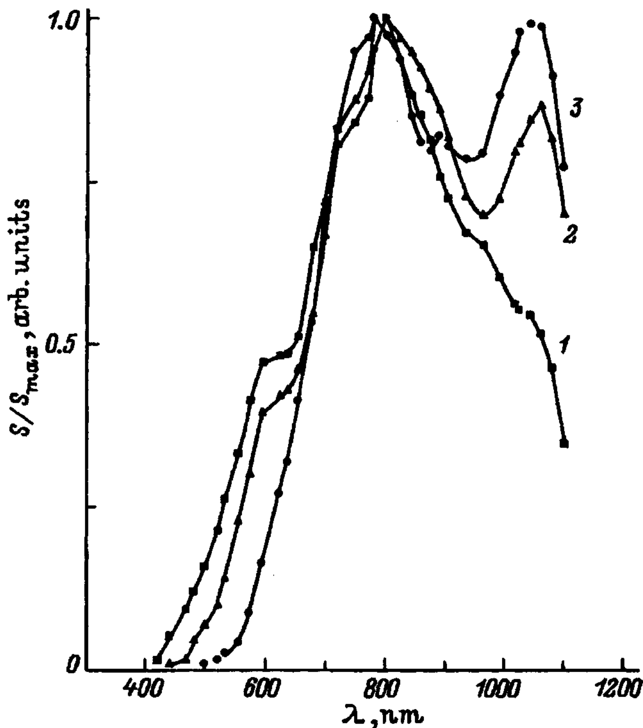


FIG. 2. Normalized spectral photosensitivity characteristics of heterojunctions: *1* — GaN/Si, *2* — $Ga_{0.85}In_{0.15}N/Si$, *3* — $Ga_{0.64}In_{0.36}N/Si$.

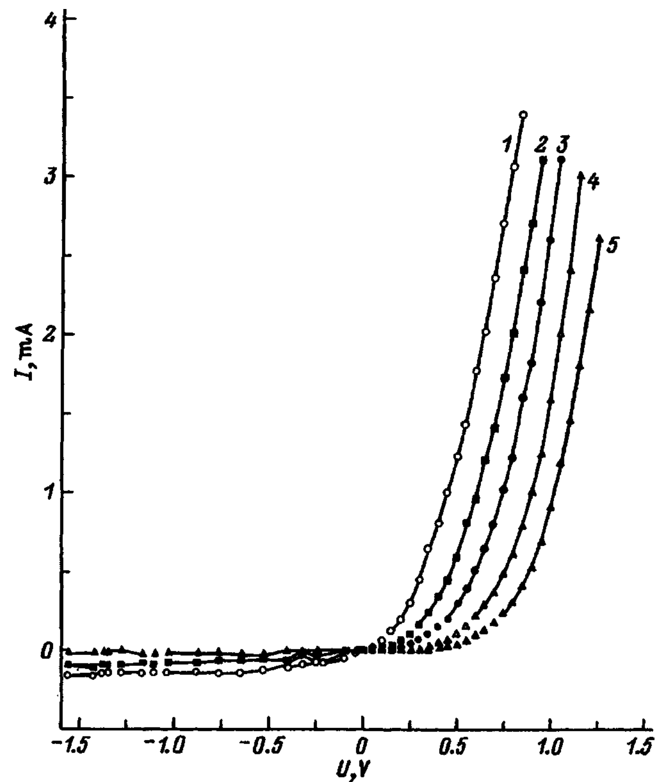


FIG. 3. Current-voltage characteristics of the heterojunction $Ga_{0.64}In_{0.36}N/Si$, *T*, K: *1* — 290, *2* — 250, *3* — 200, *4* — 150, *5* — 80.

length range. This assumption is entirely justified, since pure indium nitride has a maximum in the transmission spectra at wavelengths close to 1 μm (Ref. 18), while the transmission of gallium nitride in this part of the spectrum is much lower and for $\lambda > 0.8$ μm the transmission coefficient varies monotonically.

The photoresponse time at $T = 290$ K is 20–300 μs in all structures investigated (Table II). The maximum value $\tau = 300$ μs was observed in the heterojunction GaN/Si. As we have already noted, this heterojunction exhibits the highest photosensitivity. As the temperature is lowered to 200–250 K, for all heterojunctions the photoresponse time grows exponentially and the activation energy is $E_\tau = 0.26 - 0.29$ eV. The exponential growth of τ with the same activation for all HJs investigated can be attributed to the existence of a level that determines the lifetime of the minority current carriers. The nature of this level is unclear at present. However, since the values presented for the photoresponse time refer to the absorption range of silicon, this level can be tentatively associated with energy states (defect, impurity) either in silicon or on the interface of the heterojunction.

All HJs investigated exhibit rectifying properties. Figure 3 shows typical current-voltage characteristics measured at different temperatures. The direct branches of the I–V characteristics for all HJs investigated have two exponential sections (Fig. 4). The first section (for low biases) is described satisfactorily by the relation $I = I_1 \exp(eU/\beta_1 kT)$ with $\beta_1 = 3 - 6$ at 290 K. The sections of the I–V characteristics with higher biases correspond to $I = I_2 \exp(AU)$ with $\beta_2 = 1/A = 1 - 12$ at $T = 290$ K. As the temperature is lowered, both sec-

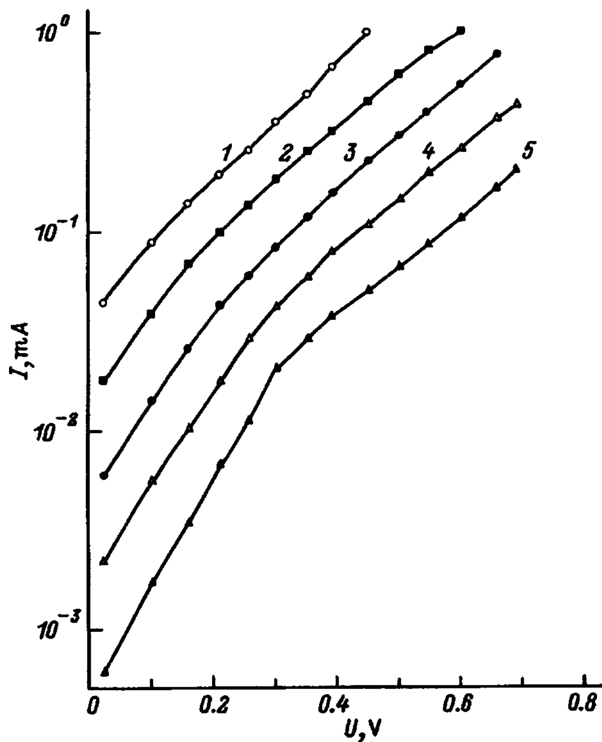


FIG. 4. Direct branches of the I–V characteristics of the heterojunction $Ga_{0.64}In_{0.36}N/Si$. T, K: *1* — 290, *2* — 250, *3* — 200, *4* — 150, *5* — 80.

tions remain but β_1 increases and β_2 is virtually temperature-independent. The voltage corresponding to a transition between these two regions decreases with increasing temperature. Analysis of the current-voltage characteristics makes it possible to characterize the current flow in the experimental heterostructures as follows. At low bias voltage the total current is limited by recombination at the interface. For high direct biases current passage is determined by tunneling. The reverse branches of the I–V characteristics with low voltages (except initial) have a nearly linear power-law dependence.

The differential resistance at zero bias (R_0) at $T=290$ K varies, depending on the composition of the wide-gap layer, from 2 to 20 kΩ (Table II). As temperature decreases to ~ 150 K, R_0 increases exponentially. The activation energy of this process for different structures varied from 0.13 to 0.44 eV. At lower temperatures the differential resistance is temperature-independent. The exponential section of the temperature dependence of R_0 is described well in the recombination–tunneling model with barrier height $V=_{Dn}=0.13-0.44$ eV for different heterojunctions (Table II) for diffusion or thermal-emission currents with current-carrier recombination at the interface. The weak temperature dependence of R_0 at low temperatures could be due to a tunneling mechanism of current transport.

An energy diagram of the anisotypic heterojunction n-$Ga_{0.64}In_{0.36}N/p$-Si was constructed on the basis of the investigations performed and the published data. The diagram was constructed in the Anderson model approximation taking into account the interfacial states at the metallurgical boundary of the heterojunction.[4] These states are due to the mismatch between the lattice parameters of the film and substrate and their density for the heterojunction under study is $\sim 10^{14}$ cm^{-2}. The current carrier density in the $Ga_{0.64}In_{0.36}N$ film was determined from Hall effect measurements to be $n=2\times 10^{18}$ cm^{-3}. The carrier density in the silicon was $p=1.5\times 10^{15}$ cm^{-3}. The total contact potential difference was determined from the current-voltage characteristics by extrapolating its linear section to the voltage axis. For the present heterojunction $V_D=0.8$ eV. The diffusion potential on the $Ga_{0.64}In_{0.36}N$ solid solution side was determined from the temperature dependence of the differential resistance to be $V_{Dn}=0.13$ eV. Such a high value of V_{Dn} with current carrier density 2×10^{18} cm^{-3} in the solid solution could be due to the high density of interfacial states. The band diagram of the heterojunction is presented in Fig. 5, where $V_{Dp}=0.67$ eV, $\Delta E_c=0.053$ eV, $\Delta E_v=1.24$ eV, $E_{g1}=1.12$ eV, and $E_{g2}=2.25$ eV.

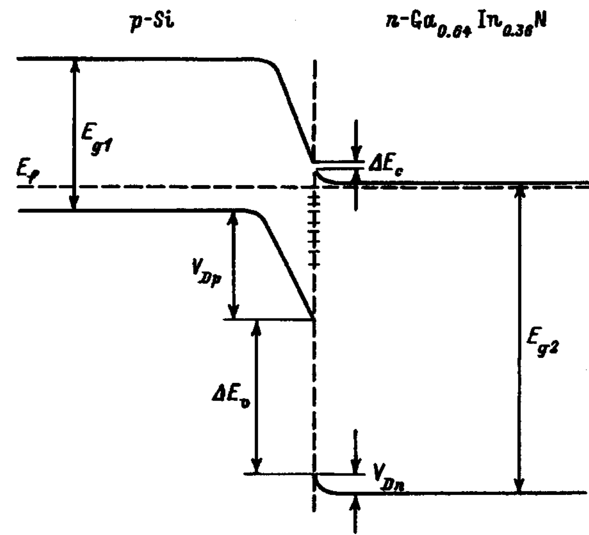


FIG. 5. Band diagram of the heterojunction $Ga_{0.64}In_{0.36}N/Si$.

In summary, our studies have shown that the anisotypic heterojunction n-$Ga_xIn_{1-x}N/p$-Si is a promising photosensitive element for detecting visible-range radiation. Improvement of the device characteristics depends directly on the technology for producing films of solid solutions and, first and foremost, on decreasing the current carrier density in them.

[1] S. Strite, M. E. Lin, and H. Markoç, Thin Solid Films **231**, 197 (1993).
[2] S. Nakamura, T. Mukai, and M. Senoh, Appl. Phys. Lett. **64**, 1687 (1994).
[3] X. H. Yang, T. J. Schmidt, W. Shan, J. J. Song, and B. Goldenberg, Appl. Phys. Lett. **66**, 1 (1995).
[4] A. G. Milnes and D. L. Feucht, *Heterjunctions and Metal–Semiconductor Junctions*, Academic Press, N. Y., 1972 [Russian translation, Moscow, 1975].
[5] A. Yamamoto, M. Tsujino, M. Ohkubo, and A. Hashimoto, J. Cryst. Growth **137**, 415 (1994).
[6] T. Matsuoka, T. Sasaki, and A. Katsui, Optoelectron., Devices Technol. **5**, 53 (1990).
[7] A. Wakahara and T. Tsuchiya, A. Yoshida, J. Cryst. Growth **99**, 385 (1990).
[8] W. E. Hoke, P. J. Lemonias, D. G. Weir, J. Cryst. Growth **111**, 1024 (1991).
[9] S. Strite, D. Chandrasekhar, D. J. Smith, J. Sariel, H. Chen, N. Teraguchi, and H. Markoç, J. Cryst. Growth **127**, 204 (1993).

[10] T. Maruyama and T. Morishita, J. Appl. Phys. **76**, 5809 (1994).
[11] Q. Guo, H. Ogawa, H. Yamano, and A. Yoshida, Appl. Phys. Lett. **66**, 715 (1995).
[12] K. Naniwae, Sh. Itoh, H. Amano, K. Hiramatsu, and I. Akasaki, J. Cryst. Growth **99**, 381 (1990).
[13] O. Igarashi, Jpn. J. Appl. Phys. , Pt. 1 **31**, 2665 (1992).
[14] Y. Sato and S. Sato, J. Cryst. Growth **144**, 15 (1994).
[15] S. E. Aleksandrov and V. A. Kryakin, Inventor's Certificate SSSR No. 1436762; Patent No. 4210081, March 11, 1987.
[16] A. V. Rakov, *Spectrophotometry of Thin-Film Semiconductor Structures* [in Russian], Moscow, 1975.
[17] T. Nagamoto and O. Omoto, J. Phisique IV **C5**, 1173 (1995).
[18] T. L. Tansley and C. P. Foley, J. Appl. Phys. **60**, 2092 (1986).

Translated by M. E. Alferieff

LOW-DIMENSIONAL SYSTEMS

Intraband absorption of light in quantum wells induced by electron-electron collisions

G. G. Zegrya and V. E. Perlin

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia
(Submitted July 22, 1997; accepted for publication July 27, 1997)
Fiz. Tekh. Poluprovodn. **32**, 466–471 (April 1998)

It is shown that in a semiconductor with quantum wells intraband absorption of long-wavelength radiation due to Coulomb interaction of the electrons is possible as a result of band nonparabolicity. Analytical expressions for the absorption are found for the limiting cases of nondegenerate and strongly degenerate, two-dimensional electron, gas. At high carrier densities the absorption due to electron-electron interaction can be much stronger than absorption due to electron-phonon interaction. © *1998 American Institute of Physics.* [S1063-7826(98)01504-X]

1. INTRODUCTION

The physical processes occurring in semiconductor heterostructures with quantum wells are being widely studied. Numerous experiments show the presence of strong absorption of radiation with photon energy ($\hbar\omega\sim 1-5$ meV) smaller than the splitting between the quantum-well subbands in such systems.[1] Such radiation can be absorbed only as a result of intraband and therefore indirect transitions. In order for the laws of conservation of energy and momentum to be satisfied simultaneously in such junctions, an electron which has absorbed a photon must be scattered by a third particle. Mainly intraband transitions with scattering by photons or with the participation of impurities are studied in the literature.[2,3] In the present paper we study a new mechanism of intrasubband absorption due to electron-electron scattering in a free electron gas. This absorption mechanism is not realized in the case of a quadratic law governing dispersion in a subband. In order to calculate this mechanism it is necessary to take into account the nonparabolicity arising in the electron spectrum as a result of interaction of the conduction band and the valence band. At high charge carrier densities the absorption mechanism due to electron-electron collisions can become more effective than the phonon or impurity mechanism.

2. TRANSITION MATRIX ELEMENTS

Let the semiconductor structure be unbounded and uniform in the (y,z) plane and let it contain an infinitely deep quantum well with potential (in the conduction band) $U(x)=0$ in the interval $x\in[0, a]$ and $U(x)=\infty$ outside this interval (a is the well width). The energy of an electron in the conduction band, measured from the center of the band gap E_g, can be calculated in the Kane model[4]

$$E_c=(k,\mathbf{q})=\sqrt{(E_g/2)^2+\gamma^2(k^2+q^2)},$$

where k is the wave number of the motion along the x axis, $\mathbf{q}=(q_y,q_z)$ is the wave vector of free motion in the (y,z) plane, and γ is the parameter, proportional to the matrix element of the momentum operator between valence- and conduction-band states, in the Kane model. We assume that only the lower quantum-well subband is filled, so that $k=\pi/a$. Expanding the energy $E_c(\mathbf{q})$ in powers of $\mathbf{q}$ for small q and replacing $q_{y,z}$ by $p_{y,z}/\hbar$, we write the electron kinetic energy operator in the (y,z) plane in the form

$$\hat{T}=\frac{\hat{\mathbf{p}}^2}{2m_c}-\left(\frac{\gamma}{E_g\hbar}\right)^4 E_g\hat{\mathbf{p}}^4. \quad (1)$$

Here m_c is the electron effective mass in the conduction band, which can be easily expressed in terms of γ, E_g, and k. The term proportional to $\hat{p}^4$ gives a correction to the quadratic dispersion relation. Its smallness is determined by the smallness of the parameter $q_T\gamma/E_g\ll 1$, where $q_T=\sqrt{2m_cT/\hbar^2}$ is the characteristic wave number of an electron in the case of nondegenerate electron gas. If the gas is degenerate, then here q_T must be replaced by the Fermi momentum q_F. We shall solve the problem of finding the absorption coefficient to first order in the nonparabolicity. It is easy show that in this case it should be taken into account only in the operator $\hat{V}_{\mathrm{opt}}$ describing the interaction of electrons with the photon field, while the effect of the nonparabolicity on the wave functions can be disregarded.

The wave function of an electron in the ground state in the well is $\Psi_{\mathbf{q}}(\mathbf{r})=(1/\sqrt{S})e^{i\mathbf{q}\cdot\boldsymbol{\rho}}\Psi_x(x)$, where ρ is the radius vector in the plane of the well, $\mathbf{r}=(x, \rho)$, and $\Psi_x(x)=\sqrt{2/a}\sin kx$ in the well and $\Psi_x(x)=0$ outside the well is the wave function of motion along the x axis. Here normalization ''in a box'' with area S is used. The wave function of a system of two electrons is $\Psi_{\mathbf{q}_1,\mathbf{q}_2}(\mathbf{r}_1,\mathbf{r}_2)=(1/\sqrt{2})\times[\Psi_{\mathbf{q}_1}(\mathbf{r}_1)\Psi_{\mathbf{q}_2}(\mathbf{r}_2)-\Psi_{\mathbf{q}_1}(\mathbf{r}_2)\Psi_{\mathbf{q}_2}(\mathbf{r}_1)]$. We find the matrix element of the interaction of this system with the photon field $\mathbf{A}(\mathbf{r}, t)$. Accordingly, we replace in the Hamiltonian $\hat{\mathbf{p}}$ by $\hat{\mathbf{p}}+(e/c)\mathbf{A}$ and, taking into account only terms linear in the field (corresponding to single-photon absorption), we obtain the interaction operator

 © 1998 American Institute of Physics

$$\hat{V}_{\text{opt}}=\frac{e}{cm_c}[\mathbf{A}(\mathbf{r}_1,t)\cdot\hat{\mathbf{p}}_1+\mathbf{A}(\mathbf{r}_2,t)\cdot\hat{\mathbf{p}}_2]-\frac{4e\gamma^4}{cE_g^3\hbar^4}[\mathbf{A}(\mathbf{r}_1,t)\cdot\hat{\mathbf{p}}_1^3+\mathbf{A}(\mathbf{r}_2,t)\cdot\hat{\mathbf{p}}_2^3]. \quad (2)$$

For a monochromatic plane wave the vector potential is $\mathbf{A}(\mathbf{r},t)=A_0\mathbf{e}\exp[i(\varkappa\cdot\mathbf{r}-\omega t)]$, where A_0 is the amplitude associated with the photon volume density N by the relation $A_0^2=2\pi N\hbar c^2/\omega n_\epsilon^2$ (n_ϵ is the refractive index of the medium at frequency ω), ω and $\varkappa=(\varkappa_x,\varkappa_{\mathbf{p}})$ are the frequency and wave vector of the light, and $\mathbf{e}=(e_x,\mathbf{e}_{\mathbf{p}})$ is the unit polarization vector. In deriving Eq. (2) we employed the Landau gauge $\phi=0$, $\nabla\cdot\mathbf{A}=0$. We shall calculate the matrix element between the initial state $\Psi_i=\Psi_{\mathbf{q}_1,\mathbf{q}_2}$ and the final state $\Psi_f=\Psi_{\mathbf{q}_3,\mathbf{q}_4}$. An integral of the form $\int_{-\infty}^{\infty}e^{i\varkappa_x x}|\Psi_x(x)|^2dx$ arises in the integration over x. It can be assumed that this integral equals 1, since the wave vector of the photon $\varkappa_x\sim10^4\ \text{cm}^{-1}$ is much less than $1/a\sim10^6\ \text{cm}^{-1}$, and the exponential in the integral equals 1. Integration in the (y,z) plane gives a Kronecker δ function. Finally, the matrix element of the electron–photon interaction is

$$V_{\text{opt},if}=(-i)e^{-i\omega t}\frac{e\hbar}{cm_c}A_0\{[\mathbf{e}_{\mathbf{p}}\cdot(\mathbf{q}_3+\mathbf{q}_4)]-b_0(\mathbf{e}_{\mathbf{p}}\cdot\mathbf{q}_3)q_3^2-b_0(\mathbf{e}_{\mathbf{p}}\cdot\mathbf{q}_4)q_4^2\}\times(\delta_{\mathbf{q}_1\mathbf{q}_3}\delta_{\mathbf{q}_2\mathbf{q}_4}-\delta_{\mathbf{q}_2\mathbf{q}_3}\delta_{\mathbf{q}_1\mathbf{q}_4}), \quad (3)$$

where $b_0=4m_c\gamma^4/\hbar^2E_g^3$.

Let us now find the matrix element of the electron-electron interaction. The interaction operator is $\hat{V}_{\text{ee}}=e^2(\epsilon|\mathbf{r}_1-\mathbf{r}_2|)$, where ϵ is the permittivity of the medium. The matrix element between the states $\Psi_i=\Psi_{\mathbf{q}_1,\mathbf{q}_2}$ and $\Psi_f=\Psi_{\mathbf{q}_3,\mathbf{q}_4}$, which are described by antisymmetric wave functions, has the form

$$V_{\text{ee},\,if}=V_2-V_1,$$

where

$$V_1=\int\Psi_1^*(\mathbf{r}_1)\Psi_4(\mathbf{r}_1)\frac{e^2}{\epsilon|\mathbf{r}_1-\mathbf{r}_2|}\Psi_2^*(\mathbf{r}_2)\Psi_3(\mathbf{r}_2)d^3\mathbf{r}_1d^3\mathbf{r}_2,$$

$$V_2=\int\Psi_1^*(\mathbf{r}_1)\Psi_3(\mathbf{r}_1)\frac{e^2}{\epsilon|\mathbf{r}_1-\mathbf{r}_2|}\Psi_2^*(\mathbf{r}_2)\Psi_4(\mathbf{r}_2)d^3\mathbf{r}_1d^3\mathbf{r}_2. \quad (4)$$

Here, for simplicity, we write $\Psi_{q_i}=\Psi_i$. To calculate these integrals we switch to the Fourier representation. We then obtain the following expression for V_1:

$$V_1=\frac{4\pi e^2}{(2\pi)^3\epsilon}\int\frac{I_{14}(\mathbf{q})I_{23}(-\mathbf{q})}{q^2}d^3\mathbf{q}, \quad (5)$$

where $I_{ij}(\mathbf{q})=\int\Psi_i^*(\mathbf{r})\Psi_j(\mathbf{r})e^{i\mathbf{q}\cdot\mathbf{r}}d^3r$. Substituting here the explicit expressions for the wave functions, we obtain

$$V_1=\frac{4\pi e^2}{(2\pi)^3\epsilon}\frac{(2\pi)^4}{S^2}\delta(\mathbf{q}_1+\mathbf{q}_2-\mathbf{q}_3-\mathbf{q}_4)J(|\mathbf{q}_p|), \quad (6)$$

where $\mathbf{q}_p=\mathbf{q}_1-\mathbf{q}_4=\mathbf{q}_3-\mathbf{q}_2$ is the momentum transfer due to Coulomb scattering, and

$$J(q_p)=\int_{-\infty}^{\infty}dq\left[\frac{1}{q_p^2+q^2}\left|\int_{-\infty}^{\infty}dx|\Psi_x(x)|^2e^{iqx}\right|^2\right]=\frac{1}{2a^2}\left[\frac{4\pi}{q_p^3}(q_pa-1+e^{-q_pa})+\frac{2\pi a}{4k^2+q_p^2}+\frac{2\pi(4k^2-q_p^2)(1-e^{-q_pa})}{(4k^2+q_p^2)^2q_p}+\frac{6\pi(1-e^{-q_pa})}{(4k^2+q_p^2)q_p}\right]. \quad (7)$$

The integral V_2 can be found by analogy with V_1. The factor $[J(|\mathbf{q}_1-\mathbf{q}_3|-J(|\mathbf{q}_2-\mathbf{q}_3|)]$ in the expression for the Coulomb matrix element, equal to V_2-V_1, can be substantially simplified. The momentum transfer q_p equals in order of magnitude the thermal momentum q_T (in the case of Boltzmann statistics). At room temperature $q_T\sim10^6\ \text{cm}^{-1}$. In the case of degenerate statistics all momenta are comparable to the Fermi momentum q_F, which is of the same order of magnitude. For a sufficiently narrow quantum well $a\sim100$ Å the quantity q_pa is less than or of the order of 1 and can be approximated by the difference $\pi(|\mathbf{q}_1-\mathbf{q}_3|^{-1}-|\mathbf{q}_2-\mathbf{q}_3|^{-1})$ to a high degree of accuracy, which increases with decreasing a. We can then write the Coulomb matrix element can be in the form

$$V_{\text{ee},if}=\frac{(2\pi)^3e^2}{\epsilon S^2}\delta(\mathbf{q}_1+\mathbf{q}_2-\mathbf{q}_3-\mathbf{q}_4)\times\left(\frac{1}{|\mathbf{q}_1-\mathbf{q}_3|}-\frac{1}{|\mathbf{q}_1-\mathbf{q}_4|}\right). \quad (8)$$

Here the δ function expresses the law of conservation of momentum in a collision of two electrons.

3. CALCULATION OF THE ABSORPTION COEFFICIENT

The absorption process due to electron-electron collisions can be calculated in second-order perturbation theory. The number of transitions of a system of two electrons per unit time per unit area is

$$W=\frac{1}{S}\frac{2\pi}{\hbar}\sum_{i,f}\left|\sum_m\left(\frac{V_{\text{opt},i\,m}V_{\text{ee},m\,f}}{\varepsilon_m-\varepsilon_i-\hbar\omega}+\frac{V_{\text{ee},i\,m}V_{\text{opt},m\,f}}{\varepsilon_m-\varepsilon_i}\right)\right|^2\times\delta(\varepsilon_f-\varepsilon_i-\hbar\omega)\{f(\mathbf{q}_1)f(\mathbf{q}_2)[1-f(\mathbf{q}_3)]\times[1-f(\mathbf{q}_4)]-f(\mathbf{q}_3)f(\mathbf{q}_4)[1-f(\mathbf{q}_1)][1-f(\mathbf{q}_2)]\}. \quad (9)$$

Here the summation extends over all initial, final, and intermediate states of the particles — $\Psi_i=\Psi_{\mathbf{q}_1,\mathbf{q}_2}$, $\Psi_f=\Psi_{\mathbf{q}_3,\mathbf{q}_4}$, and $\Psi_m=\Psi_{\mathbf{q}m1,\mathbf{q}m2}$, S is the area of the heterojunction, ε_i, ε_f, and ε_m are the energies, and $f(\mathbf{q}_j)$, $j=1, 2, 3, 4$ are the probabilities that the corresponding states are occupied by electrons.

It follows from the form of the matrix element $V_{\text{opt},im(mf)}$ (3) that the matrix element assumes nonzero values only for four intermediate states, which to within a transposition of

the indices are identical to the initial or final state. It is easy to see that on summing over these indices the terms not containing the coefficient b_0, which describes the nonparabolicity of the spectrum, cancel as a result of the law of conservation of momentum $\mathbf{q}_1+\mathbf{q}_2=\mathbf{q}_3+\mathbf{q}_4$. This means that the absorption mechanism we are studying is not realized in the case of a quadratic dispersion relation. The composite matrix element is proportional to the coefficient b_0:

$$|M|^2=\left|\sum_m (\ldots)\right|^2=4A_0^2b_0^2\frac{(2\pi)^2e^4}{\epsilon^2S^4}\left(\frac{e\hbar}{m_cc}\right)^2\frac{1}{(\hbar\omega)^2}\times\left(\frac{1}{|\mathbf{q}_1-\mathbf{q}_3|}-\frac{1}{|\mathbf{q}_2-\mathbf{q}_3|}\right)^2\delta(\mathbf{q}_1+\mathbf{q}_2-\mathbf{q}_3-\mathbf{q}_4)^2\times[(\mathbf{e}_p\mathbf{q}_3)q_3^2+(\mathbf{e}_p\mathbf{q}_4)q_4^2-(\mathbf{e}_p\mathbf{q}_1)q_1^2-(\mathbf{e}_p\mathbf{q}_2)q_2^2]^2. \quad (10)$$

We shall examine two limiting cases of the electron distribution over the states.

1. *Nondegenerate electron gas.* In this case we assume that the occupation numbers are small, $f(\mathbf{q}_j)=(2\pi n\hbar^2/m_cT)\exp(-q_j^2/2m_cT)\ll 1$, and we ignore in Eq. (9) the products that contain no more than two of them. The corresponding term in parentheses in Eq. (9) can be simplified further by making use of the energy conservation law $q_3^2/2m_c+q_4^2/2m_c-q_1^2/2m_c-q_2^2/2m_c=\hbar\omega$:

$$\{\ldots\}=f(\mathbf{q}_1)f(\mathbf{q}_2)-f(\mathbf{q}_3)f(\mathbf{q}_4)=\left(\frac{4\pi n\hbar^e2}{2m_cT}\right)^2\times\exp\left[\frac{\hbar^2(q_1^2+q_2^2)}{2m_cT}\right]\left[1-\exp\left(\frac{-\hbar\omega}{T}\right)\right]. \quad (11)$$

We now switch in the expression (9) from summation over the initial and final momenta to integration $\Sigma_{\mathbf{q}_j}\to[S/(2\pi)^2]\int d^2\mathbf{q}_j$. The integration over $\mathbf{q}_4$ can then be performed using one of the δ functions from Eq. (10), and we replace the second integral by a Kronecker δ function: $\delta(\mathbf{q}_1+\mathbf{q}_2-\mathbf{q}_3-\mathbf{q}_4)=[S/(2\pi)^2]\delta_{\mathbf{q}_1+\mathbf{q}_2,\mathbf{q}_3-\mathbf{q}_4}=S/(2\pi)^2$. This gives a sixfold integral that can be calculated analytically (see Appendix 1).

The absorption coefficient G equals the number of transitions per unit volume of the material, divided by the photon flux density $N\nu=Nc/n_\epsilon$. The concept of an absorption coefficient is meaningless for a single quantum well, since the experimentally determined transmission coefficient $[1-\exp(-aG)]$ is virtually equal to 1. To observe absorption in two-dimensional systems it therefore makes sense to use as a sample a multilayer structure consisting of a large number of absorbing layers (quantum wells), alternating with transparent layers. If the period of such a structure equals L, then the number of transitions per unit volume can be obtained from the number of transitions, found above, per unit area by dividing by L. Thus we obtain for the absorption coefficient

$$G=\frac{Wn_\epsilon}{LNc}=\frac{2^{10}\pi^3}{n_\epsilon}\frac{n^2\lambda_g^4}{L}\frac{e^2}{\hbar c}\frac{T^2}{(\hbar\omega)^2}\frac{E_B}{(\hbar\omega)}\times\left(1+\frac{5\hbar\omega}{2T}\right)\left[1-\exp\left(-\frac{\hbar\omega}{T}\right)\right], \quad (12)$$

where $\lambda_g=\hbar/\sqrt{2m_cE_g}$, and $E_B=m_ce^4/2\hbar^2\epsilon^2$ is the Bohr energy for an electron in a crystal. Here we took into account that $\gamma\simeq\hbar\sqrt{E_g/2m_c}$, and $b_0=2\lambda_g^2$ [see Eq. (3)]. To estimate the magnitude of absorption we use typical parameters of standard semiconductors, for example, GaAs: refractive index $n_\epsilon=3.5$, static permittivity $\epsilon=10$, electron effective mass in the conduction band $m_c=0.1m_0=10^{-28}$ g, room temperature $T=0.026$ eV, photon energy $\hbar\omega=3\times10^{-3}$ eV, two-dimensional electron density in the quantum well $n=10^{11}$ cm^{-2}, band gap $E_g\simeq1.6$ eV, and interwell distance is $L=10^{-4}$ cm. Substituting these values into Eq. (12), we obtain the estimate $G\simeq1$ cm^{-1}.

2. *Strongly degenerate electron gas.* If the photon energy $\hbar\omega$ is much less than the Fermi energy $\varepsilon_F=2\pi\hbar^2n/m_c$, then all electron transitions due to the absorption of this light must occur near the Fermi level. If the "width" $\sim T$ of the Fermi–Dirac distribution is much less than the transition energy $\hbar\omega$, then to study these transitions this distribution can be replaced by a step function: $f(\mathbf{q})=1$ for $q<q_F=2\sqrt{\pi n}$ and $f(\mathbf{q})=0$ for $q>q_F$. The conditions $T\ll\hbar\omega\ll\varepsilon_F$ for photons with energy of the order of several meV hold, for example, for electron density $n=10^{12}$ cm^{-2} at liquid-helium temperatures.

We shall solve the problem of finding the absorption coefficient in lowest order in the small parameter $\xi=\hbar\omega/\varepsilon_F$. On absorbing a photon two electrons, which scatter each other, pass from the states 1 and 2 in a narrow layer of thickness of the order of ξq_F inside the Fermi surface into the states 3 and 4, respectively, in a similar narrow layer outside the Fermi surface. Two cases are possible here: 1) the momentum transfer $\mathbf{q}_5=\mathbf{q}_3-\mathbf{q}_1=\mathbf{q}_2-\mathbf{q}_4$ is small compared with q_F ($q_5\sim\xi q_F$) and 2) the momentum transfer $\mathbf{q}_5$ is not small compared with q_F, where the law of conservation of momentum requires that the momenta $\mathbf{q}_1$ and $\mathbf{q}_2$ be almost opposite to one another $|\mathbf{q}_1+\mathbf{q}_2|\sim\xi q_F$. In the first case q_5 is a small quantity of first order in ξ; all other momenta $q_{1,2,3,4}$, $q_6=|\mathbf{q}_3-\mathbf{q}_2|$ are comparable to q_F. In the second case q_5 is also comparable to q_F. The expression for the transition probability (9) contains the squared Coulomb matrix element (8)

$$V_{\mathrm{ee},if}^2\sim\left(\frac{1}{|\mathbf{q}_3-\mathbf{q}_1|}-\frac{1}{|\mathbf{q}_3-\mathbf{q}_2|}\right)^2=\frac{1}{q_5^2}-\frac{2}{q_5q_6}+\frac{1}{q_6^2}. \quad (13)$$

In the first case the second and third terms in Eq. (13), which are proportional to ξ^{-1} and ξ^0, respectively, are small compared with the first term, which is proportional to ξ^{-2}, and can be ignored. In the second case all three terms are of order ξ^0, and it is easy to show that the contributions of transitions with a large momentum transfer to the total number of transitions is small (second order in ξ). Therefore, to lowest order in ξ only the first term from the first case should be taken into account, and the number of transitions (9) is

$$W=\frac{4A_0^2b_0^2e^4}{(2\pi)^3\hbar\epsilon^2(\hbar\omega)^2}\left(\frac{e\hbar}{m_cc}\right)^2\int d^2\mathbf{q}_1\int d^2\mathbf{q}_2\int d^2\mathbf{q}_3$$
$$\times\frac{1}{|\mathbf{q}_3-\mathbf{q}_1|^2}[q_3^2(\mathbf{q}_3\mathbf{e})+q_4^2(\mathbf{q}_4\mathbf{e})-q_1^2(\mathbf{q}_1\mathbf{e})$$
$$-q_2^2(\mathbf{q}_2\mathbf{e})]^2\delta(\varepsilon_f-\varepsilon_i-\hbar\omega). \quad (14)$$

The switch from summation in Eq. (9) to integration and the integration over $\mathbf{q}_4$ are performed just as in the nondegenerate case. In Eq. (14) the integration extends over all possible configurations with a small momentum transfer due to Coulomb interaction (see Appendix B). Finally, similarly to the nondegenerate case, we obtain for the absorption coefficient

$$G=\frac{Wn_\epsilon}{NcL}=\frac{4\chi}{n_\epsilon}\frac{e^2}{\hbar c}\frac{E_B}{E_g}\frac{n\lambda_g^2}{L}, \quad (15)$$

where $\chi=1.333$ is a numerical factor. We note that absorption does not depend on the frequency of the incident light. An estimate of the absorption with the same parameters as for the nondegenerate case (only $n=10^{12}$ cm^{-2}) gives $G\simeq10^{-1}$ cm^{-1};, i.e., the effect for degenerate statistics under the condition $\hbar\omega\ll\varepsilon_F$ is much weaker than for the Boltzmann statistics with $\hbar\omega\ll T$.

4. DISCUSSION

Light absorption by a free electron gas with a nonparabolic electron dispersion relation (1) can be easily explained on the basis of qualitative considerations. The equations of motion of electrons in the field $\mathbf{E}$ of an electromagnetic wave (in the dipole approximation) have the form

$$\frac{d\mathbf{p}_i}{dt}=e\mathbf{E}(t)-\nabla_i\sum_j V_{ij}, \quad i=1..N, \quad (16)$$

where V_{ij} is the interaction potential energy of the ith and jth electrons. The energy V does not appear in the equation for the total momentum $\mathbf{P}=\Sigma\mathbf{p}_i$:

$$\frac{d\mathbf{P}}{dt}=eN\mathbf{E}(t). \quad (17)$$

Let us switch to the Fourier representation

$$\mathbf{P}=\frac{eN\mathbf{E}}{i\omega}. \quad (18)$$

The total electron current is $\mathbf{j}=en\Sigma\mathbf{v}_i=\sigma\mathbf{E}$. In the case of a parabolic band $\Sigma\mathbf{v}_i=\mathbf{P}/m_c$ and the conductivity σ is a purely imaginary quantity; this means that there is no absorption. In the case of the dispersion law (1) we have $\Sigma\mathbf{v}_i=\mathbf{P}/m_c+(4\gamma^4/E_g^3\hbar^4)\Sigma\mathbf{p}_i^3$. The presence of a term proportional to the sum of the cubed momenta makes absorption possible.

In the case of a Boltzmann electron distribution, absorption due to electron-electron interaction is quite strong and somewhat stronger than the electron-phonon absorption (see Ref. 2). The fact that it is 10 times weaker in the strongly degenerate gas is explained by the fact that for $\hbar\omega\ll E_F$ only a small fraction of the electrons can participate in absorption. The electron density dependence of the absorption changes from quadratic for the nondegenerate gas to linear for the strongly degenerate gas. In a real situation, none of the limiting cases examined can be realized and the absorption has intermediate values.

The calculation presented above was performed for the case of an infinitely deep quantum well. The effect of the finitenesss of the depth of a real well can be taken into account in the form of a small correction to the absorption. As one can easily see (see Appendix C), all results obtained above remain valid, both qualitatively and quantitatively.

This work was supported in part by the Russian Fund for Fundamental Research (Grants Nos. 96-02-17952 and 97-02-18151) and the Russian State Program ''Physics of solid-state nanostructures'' (Grants Nos. 97-0003, 97-1035, and 97-2014).

5. APPENDIX

A. Calculation of the integral for the absorption in the case of nondegenerate electron statistics.

The desired integral can be written in the form [see Eqs. (9)–(11)]

$$W_1=\int d^2q_1\int d^2q_2\int d^2q_3\exp[-\alpha(q_1^2+q_2^2)]$$
$$\times\left(\frac{1}{|\mathbf{q}_1-\mathbf{q}_3|^2}-\frac{1}{|\mathbf{q}_1-\mathbf{q}_4|}\right)^2$$
$$\times[(\mathbf{e}_\mathrm{p}\mathbf{q}_3)q_3^2+(\mathbf{e}_\mathrm{p}\mathbf{q}_4)q_4^2-(\mathbf{e}_\mathrm{p}\mathbf{q}_1)q_1^2-(\mathbf{e}_\mathrm{p}\mathbf{q}_2)q_2^2]^2$$
$$\times\delta[(\mathbf{q}_1-\mathbf{q}_3)(\mathbf{q}_2-\mathbf{q}_3)-C],$$

where $\alpha=\hbar^2/2m_cT$, $C=m_c\omega/\hbar$, and the $\mathbf{q}_1$, $\mathbf{q}_2$, and $\mathbf{q}_3$ integrals extend over the entire plane. First we make the substitution of variables $\mathbf{q}_5=\mathbf{q}_3-\mathbf{q}_1$, $\mathbf{q}_6=\mathbf{q}_3-\mathbf{q}_2$, and $\mathbf{q}_7=\mathbf{q}_1+\mathbf{q}_2$ with the Jacobian of the transformation equal to 1/2. We integrate in the polar coordinates $\mathbf{q}_i=(q_i,\theta_i)$, $i=5$, 6, 7, where θ_i is the angle measured from the projection $\mathbf{e}_\mathrm{p}$ of the unit polarization vector $\mathbf{e}$ of the light onto the (y,z) plane. The integrand is symmetric with respect to q_5 and q_6, so that the factor $(q_5^{-1}-q_6^{-1})^2$ can be replaced by $2(q_5^{-2}-q_5^{-1}q_6^{-1})$. Next, after switching to the angular variables θ_5, $\theta=\theta_6-\theta_5$, and θ_7 the integration over q_5 (with the aid of a δ function) and over q_7,θ_5, and θ_7 is elementary. Introducing the variable $x=q_6^2$, we obtain the double integral

$$W_1=\frac{2\pi^2C\,\exp(\alpha C)}{\alpha^2}\int_0^\infty dx\int_0^{2\pi}d\theta$$
$$\times\exp\left[-\frac{\alpha}{2}\left(\frac{C^2}{x\cos^2\theta}+x\right)\right]\left(4+\frac{1}{\cos^2\theta}\right).$$

The integration over θ is performed by the method of differentiation with respect to a parameter: The function $f_1(\beta)\equiv\int_0^{2\pi}\exp(-\beta/\cos^2\theta)/\cos^2\theta d\theta=2\sqrt{\pi/\beta}\exp(-\beta)$ (here the integral is calculated by making the substitution $t=\tan\theta$) equals to within a sign the β derivative of the function $f_2(\beta)\equiv\int_0^{2\pi}\exp(-\beta/\cos^2\theta)d\theta$. Integrating $f_1(\beta)$, we find $f_2(\beta)=2\pi\mathrm{erfc}(\sqrt{\beta})$. After integrating over x with the aid of the relations

$$\int_0^\infty \exp\left(\nu - y - \frac{\nu^2}{4y}\right)\sqrt{y}dy = \frac{\sqrt{\pi}}{2}(1+\nu),$$

$$\int_0^\infty \exp(\nu - y)\ \mathrm{erfc}\left(\frac{\nu}{2\sqrt{y}}\right)dy = 1,$$

we obtain the final result

$$W_1 = \frac{2^7\pi^3 m_c^4 T^4}{\hbar^8}\left(1 + \frac{5\hbar\omega}{2T}\right).$$

B. Calculation of the integral for the absorption in the case of degenerate electron statistics

We make the substitution of variables $\mathbf{q}_1$, $\mathbf{q}_2$, $\mathbf{q}_3 \to \mathbf{q}_1$, $\mathbf{q}_2$, $\mathbf{q}_5 = \mathbf{q}_3 - \mathbf{q}_1$ in the integral from Eq. (14) (which we designate as W_2). We shall calculate the integral using polar coordinates $\mathbf{q}_i = (q_i, \theta_i)$, $i = 1, 2, 5$, with the polar axis directed along the unit vector $\mathbf{e}_\mathrm{p}$. In the region of integration the quantities q_1 and q_2 are close to q_F, while q_5 is much smaller than q_F. We shall perform the calculations in first order with respect to $q_5/q_F \sim \xi$. The angle θ_5 assumes all values from 0 to 2π. For a fixed small momentum transfer $\mathbf{q}_5$, it is easy to obtain from the conditions for the momenta $\mathbf{q}_1$ and $\mathbf{q}_2$ to lie inside and the momenta $\mathbf{q}_3 = \mathbf{q}_1 + \mathbf{q}_5$ and $\mathbf{q}_4 = \mathbf{q}_2 - \mathbf{q}_5$ to lie outside the Fermi circle the range of integration over $\mathbf{q}_1$ and $\mathbf{q}_2$. The angle θ_1 varies from $\theta_5 - \pi/2$ to $\theta_5 + \pi/2$ and the angle θ_2 varies from $\theta_5 + \pi/2$ to $\theta_5 + 3\pi/2$. In first order in ξ, the integrals over q_1 and q_2 can be replaced by the expression obtained by substituting in the integrand q_F for q_1 and q_2 and multiplying by the length of the integration segments $q_5\cos(\theta_1 - \theta_5)$ and $-q_5\cos(\theta_2 - \theta_5)$. Ignoring terms that contain q_5^2, we can transform the δ function in Eq. (14) to

$$\delta(\dots) = \frac{m}{\hbar^2}\delta[\mathbf{q}_5(\mathbf{q}_2 - \mathbf{q}_1 - \mathbf{q}_5) - C]$$

$$= \frac{m}{\hbar^2}\delta[2q_5 q_F \sin[(\theta_1 - \theta_2)/2]\cos\theta_{125} - C],$$

where $\theta_{125} = (\theta_2 + \theta_1 - \pi)/2 - \theta_5$ is the angle between the vectors $\mathbf{q}_2 - \mathbf{q}_1$ and $\mathbf{q}_5$. In the same approximation we repalce the expression in parentheses in Eq. (14) by

$$[\dots] = q_5 q_F^2[\cos(\theta_5 - 2\theta_1) - \cos(\theta_5 - 2\theta_1)].$$

In the expression which we obtain

$$W_2 = -\frac{m_c q_F^6}{\hbar^2}\int dq_5 \int_0^{2\pi} d\theta_5 \int_{\theta_5 - \pi/2}^{\theta_5 + \pi/2} d\theta_1 \int_{\theta_5 + \pi/2}^{\theta_5 + 3\pi/2} d\theta_2 q_5^3$$
$$\times \cos(\theta_1 - \theta_5)\cos(\theta_2 - \theta_5)$$
$$\times [\cos(\theta_5 - 2\theta_1) - \cos(\theta_5 - 2\theta_2)]^2$$
$$\times \delta\{2q_5 q_F \sin[(\theta_1 - \theta_2)/2]\cos\ \theta_{125} - C\}$$

we perform the integration over q_5 with the aid of a δ function, switch to the new angular variables θ_5, $\theta_8 - \theta_1 - \theta_5$, and $\theta_9 = \theta_2 - \theta_5$, perform an elementary integration over θ_5, and obtain

$$W_2 = \frac{\chi m_c q_F^2 C^3}{\hbar^2},$$

where we have introduced the numerical constant

$$\chi = -2\pi \int_{-\pi/2}^{\pi/2} d\theta_8 \int_{\pi/2}^{3\pi/2} d\theta_9$$
$$\times \frac{\cos\ \theta_8\ \cos\ \theta_9[1 - \cos 2(\theta_8 - \theta_9)]}{(\cos\ \theta_8 - \cos\ \theta_9)^4} = 1.333.$$

C. Allowance for the finiteness of the depth of the quantum well

Let the quantum well have a finite depth U. Then the wave function of an electron in the bottom level in this well will have the form

$$\Psi(x) = \begin{cases} C\ \exp(\varkappa x), & x < 0, \\ B\ \sin(kx + \delta), & 0 < x < a, \\ C\ \exp[\varkappa(a - x)], & x > a. \end{cases}$$

In the case of an infinite well $B = \sqrt{2/a}$, $C = 0$, $k = \pi/a$, and $\delta = 0$. We shall find the corrections to these quantities in the lowest orders in the small parameter $\gamma_0 = \sqrt{2\hbar^2/m_c U a^2}$ (for ordinary semiconductor heterostructures $a \sim 10^{-6}$ cm, $U \sim 0.3$ eV, and $m_c \sim 10^{-28}$ g the parameter $\gamma_0 \sim 0.2$ and can be assumed to be small) with the aid of standard equations for a square quantum well $ka = \pi n - 2\sin^{-1}\sqrt{\hbar^2 k^2/2m_c U}$, $n = 1$, and $\sin\delta = \sqrt{\hbar k/2m_c U}$, the boundary conditions at the points $x = 0$, a, and the normalization conditions for the wave function. Ignoring all the terms of the expansion in γ_0 beginning with γ_0^3, we obtain $k = \pi(1 - \gamma_0 + \gamma_0^2)/a$, $\delta = \pi(\gamma_0 - \gamma_0^2)/2$, $B^2 = 2(1 - \gamma_0 + \gamma_0^2)/a$, and $C = \gamma_0 B$. In the present paper the explicit form of the wave function for the motion of an electron along the x axis is used only to calculate the integral $J(q_p)$ (7). Performing the tedious integration and simplifying the answer similarly to the manner in which this was done with Eq. (7), we obtain that the Coulomb matrix element (8) in lowest order in γ_0 is multiplied by $(1 - \gamma_0^2)$. Therefore, the effect of the finiteness of the well on the absorption is described simply by multiplying the absorption by $(1 - \gamma_0^2)^2 = (1 - 2\gamma_0^2)$. We can now rewrite Eqs. (12) and (15) in the form

$$G = \frac{Wn_\epsilon}{LNc} = \frac{2^{10}\pi^3}{n_\epsilon}\frac{n^2\lambda_g^4}{L}\frac{e^2}{\hbar c}\frac{T^2}{(\hbar\omega)^2}\frac{E_B}{(\hbar\omega)}$$
$$\times\left(1 + \frac{5\hbar\omega}{2T}\right)\left[1 - \exp\left(-\frac{\hbar\omega}{T}\right)\right]\left(1 - \frac{4\hbar^2}{m_c U a^2}\right)$$

for the nondegenerate case and

$$G = \frac{4\chi}{n_\epsilon}\frac{e^2}{\hbar c}\frac{E_B}{E_g}\frac{n\lambda_g^2}{L}\left(1 - \frac{4\hbar^2}{m_c U a^2}\right)$$

for the degenerate case.

[1] L. E. Vorobiev and S. N. Danilov *et al.*, in *23rd International Conference on the Physics of Semiconductors*, Berlin, Germany, 1996, Vol. 3, p. 1887.
[2] V. L. Gurevich, D. A. Parshin, and K. É. Shtengel', Fiz. Tverd. Tela (Leningrad) **30**, 1466 (1988) [Sov. Phys. Solid State **30**, 845 (1988)].
[3] K. Seeger, *Semiconductor Physics*, Springer-Verlag, Berlin, 1974 [Russian translation, Moscow, 1977].
[4] E. O. Kane, J. Phys. Chem. Solids **1**, 249 (1957).

Translated by M. E. Alferieff

Radiation emitted by quantum-well InGaAs structures I. Spontaneous emission spectra

P. G. Eliseev

Temporary Address: Center of High-Technology Materials University of New Mexico, Albuquerque, USA

I. V. Akimova[a)]

P. N. Lebedev Physical Institute, 117924 Moscow, Russian
(Submitted October 13, 1997; accepted for publication November 10, 1997)
Fiz. Tekh. Poluprovodn. **32**, 472–477 (April 1998)

The spontaneous emission spectra of strained InGaAs quantum wells in the spectral range 1.2–1.5 eV were studied experimentally at 4.2–286 K with pump currents up to $\sim 9.2\ \mathrm{kA\cdot cm^{-2}}$. An interpretation of the observed bands is given. The transition $1e-1hh$ (TE polarization) dominates the spectrum. The position of the peak for this transition is virtually current-independent. No indications of a ''red'' shift, expected at a high carrier density, were observed. The weak forbidden transitions $(1e-2hh)$ were identified. The long-wavelength edge of the band drops off exponentially, by analogy with the well-known Urbach rule for the absorption edge. © *1998 American Institute of Physics.* [S1063-7826(98)01604-4]

1. INTRODUCTION

The spontaneous emission spectra of laser structures contain important information about the physical features of the radiation process occurring in the active medium of the laser. However, the interpretation of a spectra as an integral characteristic is usually complicated by the fact that many factors are superimposed simultaneously. Moreover, only subthreshold spectra are easily accessible experimentally, since once superluminescence and lasing start the shape of the spectrum changes substantially, because parts of it are amplified and because above threshold the carrier density is stabilized — under lasing conditions this density remains nearly the same, despite a substantial increase in the pump current. At the same time, modern laser structures have reached a high degree of perfection, specifically, in the crystallographic respect. This can make it easier to analyze the spectra and to understand their characteristic features. In a previous study[1] we studied ''modeless'' structures with quantum wells based on strained ultrathin InGaAs layers, which had no guided modes because the antiwaveguide effect. We were able to observe in such structures the evolution of the spontaneous emission with a pumping level 100–200 times higher than the inversion threshold at 300 and 77 K. We found that the dominant spectral peak is not shifted for very high degree of filling of the quantum well (onset of filling of the barrier layers). This is in conflict with the expected ''red'' shift due to the well-known renormalization of the band gap with increasing carrier density.[2,3] It was noted that the shape of the spectrum cannot be described on the basis of the conventional approach employing a Lorentzian form factor for homogeneous broadening. To check these data and to analyze them in detail these investigations were extended to lower temperatures.

In the present paper we report the results of a study of the spontaneous emission spectra of a quantum well at 4.2 K. The effect of induced transitions is ruled out right up to currents $6-9\ \mathrm{kA\cdot cm^{-2}}$, when the well is actually filled with excess carriers (whose two-dimensional density exceeds $4\times10^{12}\ \mathrm{cm^{-2}}$). It is shown that there is no red shift, though the spectrum undergoes a change attesting to strong carrier degeneracy. The characteristics of the shape of the spectral band are obtained, specifically, the Urbach rule for the radiation is formulated and the form factor for homogeneous broadening is studied (an analysis of the form factor is given in a companion — the next article in this issue).

2. EXPERIMENT

2.1. InGaAs quantum wells

Structures with quantum wells (QWs) are the basic variant of structures for modern lasers and LEDs. Among such structures InGaAs/GaAs-QW apparently have the highest degree of perfection and the lowest threshold. Because of the strained state of the active layer, the valence-band top is strongly split, and actually only one of the bands (heavy-hole) participates in radiative transitions. When the thickness is small, the number of subbands of quantum-well states can be reduced to one for electrons and one or two for holes. This greatly simplifies the interpretation of the spectrum. For a sufficiently high degree of structural perfection the effect of the nonuniformity, specifically, on the spectral broadening can be reduced to a negligible level.

A sample with the same structure as in previous works[1,4] was investigated: It contained a 6 ± 1-nm-thick $In_{0.2}Ga_{0.8}As$ layer between 15-nm-thick GaAs layers, playing the role of barriers (see Ref. 4 for a description of other, secondary details of the structure). The laser characteristics were quite high: The inversion threshold at 300 K was $80-100\ \mathrm{A\cdot cm^{-2}}$, threshold current density was $190\ \mathrm{A\cdot cm^{-2}}$, and internal optical loss factor was equal to about $5\ \mathrm{cm^{-1}}$.

 © 1998 American Institute of Physics

2.2. Suppression of stimulated emission

Under ordinary conditions the investigation of spontaneous emission in laser structures is limited to the observation of lasing, for which the Fermi quasilevels stay at their threshold value. For this reason, the range of carrier densities is small. This range is larger in structures for optical amplifiers, but the emission spectrum is affected by superluminescence and, moreover, the intensified spontaneous emission also keeps the Fermi quasilevels form rising. In this respect, ''modeless'' structures in which the configuration of the active region corresponds to an antiwaveguide in which there are no guided modes are more convenient for studying spontaneous emission. In this case, as shown in Ref. 1, even the influence of superluminescence is eliminated. Termination of lasing and the negative differential gain effect in such structures due to the antiwaveguide contribution of the carriers are described in Refs. 4 and 5. In our study the sample possessed a narrow ($\sim 2\ \mu$m) strip geometry in which lasing at 300 K does not arise. To quench lasing at low temperatures the cavity was removed by etching the backside mirror. This made it possible to trace the emission spectra with the pumping level exceeding the standard lasing threshold by approximately a factor of 10–15 at room temperature and by more than a factor of 50 at 4.2 K.

In such a structure there is a difficulty in determining the current density and carrier density because of the possibility of lateral spreading. Judging from the size of the emitting spot, this spreading is small, but it still gives an uncertainty of the order of 30–50%. The cross-sectional area of the diode along the base of the mesa strip was equal to 2.175×10^{-5} cm^2. This was used in the calculations of the current density (nominal value). Because of the spreading, the actual current density across the junction can be 50% of this nominal value.

2.3. Measurements and results

The measurements were performed with a direct pump current (up to 20 mA) and with a pulsed current (up to 200 mA) in a 1-μs regime with a pulse period-to-pulse duration ratio of 10^3. For this value of the ratio the average heating was small so that thermal influences were ruled out. In contrast, appreciable heating by the current occurred when the direct current exceeded 5 mA. This was determined by comparing the position of the spectral band with the direct current and pulsed pumping. It was shown that there is virtually no shift of the spectral peak with current if the measurements are performed at a constant temperature.

To estimate the density N^{2D} reached in the experiment we shall employ the formula $N^{2D} \sim (Jd/eB)^{1/2}$, where J is the current density, and B is the radiative recombination coefficient. There are about 4×10^{12} cm^{-2} states in the first electron and hole subbands (taking spin degeneracy into account). If it is assumed that the actual current density equals half the nominal value (because of lateral spreading), while $B = 5\times 10^{-10}$ cm$^3\cdot$s^{-1} at 300 K, then a current $I \simeq 100$ mA will correspond to filling of the subbands. This agrees with our observation that the radiation of secondary subbands becomes visible at high currents. Thus, as the current varies from 1 to 100 mA, N^{2D} hypothetically increases from $\sim 4\times 10^{11}$ up to 4×10^{12} cm^{-2}. The appearance of a second spectral peak 105 meV above the edge emission of the quantum well attests to the achievement of a high degree of filling of the states.

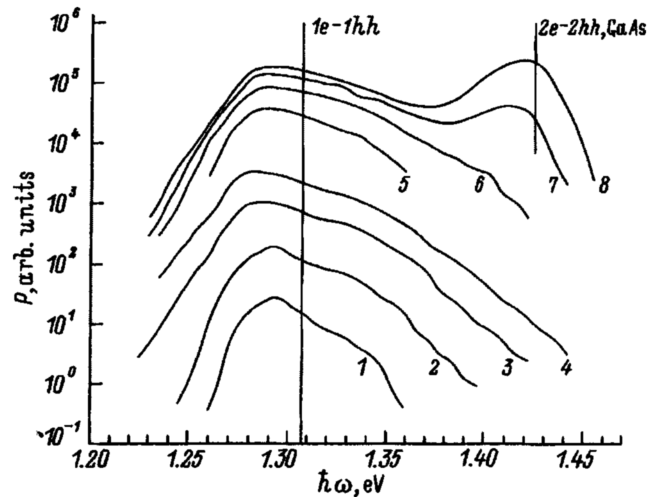


FIG. 1. Relative spectral density P of the radiation of an InGaAs quantum well at 286 K. Pump current I, mA: *1* — 1, *2* — 2, *3* — 5, *4* — 10, *5* — 50, *6* — 100, *7* — 150, *8* — 200. The spectra 3 and 4 were obtained in the direct-current regime.

The emission spectra close to room temperature are shown in Fig. 1. The main peak at 1.29 eV is for all currents; it corresponds to the $1e-1hh$ transitions between the lower states in the electron and heavy-hole subbands (TE polarization predominates). There is no spectral narrowing. In contrast, the spectrum broadens from 30 meV at 1 mA to 60 meV at 200 mA. Its position does not change provided that the 5 and 10 mA curves, which were measured with direct current, are excluded. On these curves the peak shifts by about 1.5 meV in the long-wavelength direction, apparently because of heating by the current (this shift is not confirmed with a further increase in current, but now it is confirmed in the pulsed regime). The shape of the band changes mainly on account of the development of the short-wavelength side, where a peak at ~ 1.42 eV appears with currents of 150 and 200 mA. It could be due to emission on the allowed $2e-2hh$ transitions between the secondary subbands, which is found to be close in photon energy to the emission from GaAs barriers. There are no spectral structures of comparable intensity between the first and second peak. This signifies that the contribution of $1e-1lh$ transitions (TM polarization) and forbidden $1e-2hh$ and $2e-1hh$ transitions is quite small. Thus, in the current range up to 150 mA we are dealing with filling of the lowest subbands. For 150 mA and higher currents appreciable filling of the second subbands and possibly GaAs barriers starts. This indicates that the Fermi quasilevels are located at a high energy — much higher than the energy of the $1e-1hh$ edge.

Figure 2 shows the evolution of the emission spectra at 4.2 K. The main peak has shifted to ~ 1.375 eV. This corresponds approximately to the temperature dependence of the band gap. Its position, just as at 286 K, is current-

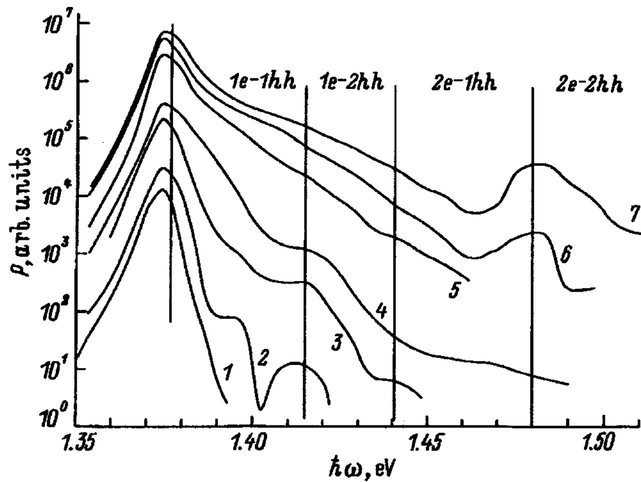


FIG. 2. Relative spectral density P of the radiation of an InGaAs quantum well at 4.2 K. Pump current I, mA: *1* — 1, *2* — 2, *3* — 10, *4* — 20, *5* — 100, *6* — 150, *7* — 200. The spectra 1 and 4 were obtained in the direct current regime.

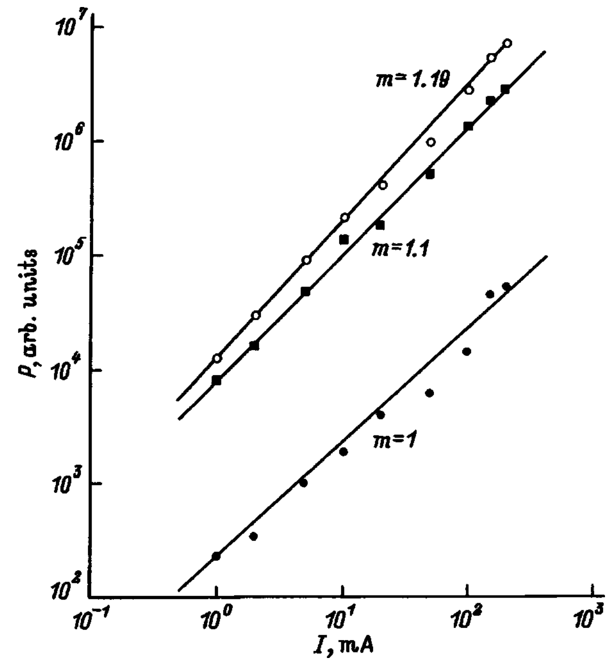


FIG. 3. Relative spectral density P of radiation versus pump current I at 4.2 K. Radiation wavelength λ, nm: *1* — 902, *2* — 904, *3* — 912. m corresponds to the relation $P \sim I^m$.

independent in the entire interval 1–200 mA. The main maximum broadens with current from 7.6 to 8.7 meV. Just as at 286 K, the spectrum develops in the short-wavelength direction, and for 150 mA a second peak becomes visible at ~1.48 eV. In contrast to 286 K, at low temperature and with current weak we see structures at 1.395, 1.415, and 1.44 eV. These structures virtually vanish on the short-wavelength shoulder at high current.

The evolution of the short-wavelength wing reflects the advancement of the Fermi quasilevels into the subbands, and at 150 mA their difference should reach the energy corresponding to the second peak, i.e., 1.48 eV. Otherwise, it is impossible to explain the filling of such high states when kT is low. It turns out that the carriers filling the lower subbands recombine mainly through the states at the $1e-1hh$ edge, where the spectral density of the radiation is much higher than in the entire interval 1.40–1.50 eV.

The shape of the long-wavelength wing is virtually current-independent. The spectral density of the radiation in different sections of the spectrum is shown as a function of current in Fig. 3. It should be noted that growth of the spectral density at the band maximum and in the long-wavelength wing is almost linear (the ''superlinearity'' exponent equals 1.19). Together with the spectral broadening, this confirms the absence of superluminescence. We also note the virtual absence of saturation, as one would expect in the case of degenerate filling of low-lying states.

2.4. Urbach rule for emission spectra

The Urbach rule states that at the intrinsic absorption edge the absorption coefficient a is an exponential function of the photon energy.[6] It has been confirmed in wide-gap materials (single crystals and polycrystals, both direct- and indirect-gap). It is natural to attribute this phenomenon to the form factor determining the broadening of the edge. This is taken into account in the theory: The theory developed in Refs. 7 and 8 explains the Urbach rule by exciton-phonon collisional interaction, which broadens the excitonic line at the intrinsic absorption edge. It is obvious that to obtain an exponential absorption ''tail'' the form factor must have an exponential asymptote (on the long-wavelength side at least). The spectral dependence of the absorption is described by the expression

$$\alpha(h\nu)=\alpha_0 \exp[\sigma(h\nu-E_0)/kT], \tag{1}$$

where α_0 and E_0 are parameters of the material, and σ is the ''slope'' factor. The parameter of the exponential (the Urbach energy) $\varepsilon=(d \ln \alpha/dE)^{-1}=kT/\sigma$ is temperature-dependent because of the fact that optical phonons are frozen out. The conventional formula for the temperature dependence (see, for example, the review in Ref. 9) can be reduced to the form

$$\varepsilon=(\hbar\Omega/2\sigma)\coth(\hbar\Omega/2kT), \tag{2}$$

where $\hbar\Omega$ is the energy of the phonon which predominates in the collisional processes, and σ is the slope parameter (constant). At sufficiently low temperature we simply have $\varepsilon=\hbar\Omega/2\sigma$. The relation (2) has also been confirmed for many wide-gap semiconductors.[9]

An important circumstance is that the long-wavelength wing of the emission spectrum simply drops off exponentially by 2–2.5 orders of magnitude, similar to the Urbach rule for absorption. Thus, the emission of a quantum well also satisfies this rule. The slope parameter $\varepsilon=(d \ln \rho/dE)^{-1}$, where ρ is the spectral density of the radiation, depends only slightly on the current and temperature (see Table I).

Another important fact is the weak saturation of the long-wavelength side right up to 150 mA. For low-temperature conditions ($kT=0.36$ meV) this behavior does not agree with the model of filling of the states according to

TABLE I. Parameter ε of the long-wavelength dropoff of the spectral density of the radiation.

	ε, meV	
T, K	$J=46$ A$\cdot$cm^{-2}	$J=9200$ A$\cdot$cm^{-2}
4.2	3.2	3.3
286	4.4	8.5

which a ''blue'' shift of the spectral peak with saturation of the long-wavelength wing should occur as the current increases and the Fermi quasilevels rise. This corresponds to degenerate filling of low-lying levels, if they exist. The nonsaturating long-wavelength wing is more likely to be due to homogeneous broadening of the main line located at the $1e-1hh$ edge than to the ''tail'' of the density of states. Thus, even in this case the exponential dropoff is due to collisional-type homogeneous broadening.

By analogy with the Urbach rule for absorption (usually observed in the low-intensity limit) the same rule operates for emission, and in a wide range of emission intensities. The variation of the slope parameter ε can be used to study the effect of the density of the electron-hole plasma on intraband relaxation. The temperature decrease of the width on cooling could be due to freezing out of optical phonons. An estimate using Eq. (2) gives $\hbar\Omega \simeq 30\pm 10$ meV, which is consistent in order of magnitude with the energy of optical phonons.

3. DISCUSSION

It follows from the spectral data presented that the scheme of transitions in a quantum well is quite simple (see Fig. 4). The possible position of the levels corresponds to a single $1e$ subband for electrons (the $2e$-subband edge apparently lies in the continuum). There are two subbands for holes, $1hh$ and $2hh$, separated by approximately 40 meV. The subband $1lh$also apparently merges with the continuum. In the calculation of the scheme it was assumed that the band gap in InGaAs equals 1.303 eV (at 4.2 K) and that the ratio $\Delta E_c/\Delta E_v$ of the offsets of the conduction- and valence-band edges equals 57:43.

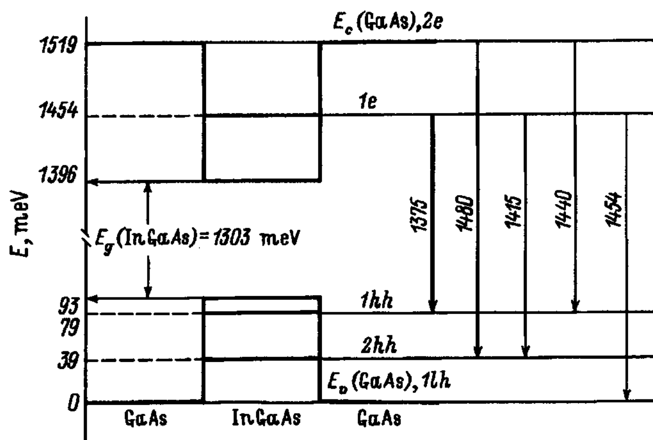


FIG. 4. Scheme of radiative transitions in an InGaAs quantum well at 4.2 K. The transition energies are indicated in meV.

In Ref. 1 we performed polarization investigations of similar quantum wells and found that the $1e-1hh$ band possesses predominantly TE polarization, while a weaker wide band with a maximum 40–50 meV higher in photon energy than the TE peak is observed in TM polarization. It is formed not so much by $1e-1lh$ emission as by the TM component $1e-1hh$, which is also shifted in the direction of high energy, since the corresponding transition matrix element equals zero at the $1e-1hh$ edge but increases with photon energy. We can attribute the spectral shoulder at 1.454 eV to $1e-1lh$ transitions (see Fig. 1). This transition is allowed but weakened because of the low population of the high $1lh$ subband. The short-wavelength band (1.42 eV at 286 K) also is largely due to the TE component. For this reason it is attributed to the transition $2e-2hh$, although states in the GaAs barrier layer also contribute to it. In our study we did not investigate the polarization, since it was noted that in samples with at least one etched face the radiation can be strongly depolarized. The $1e-2hh$ (1.415 eV) and $2e-1hh$ (1.44 eV) bands are weak because they are forbidden. They are seen only at low currents, while as current increases, they are masked by the growing short-wavelength wing of the $1e-1hh$ band.

Only one electronic subband ($1e$) and one hole subband ($1hh$) actually contribute to the radiation in a wide range of currents (over at least a 100-fold range). The carrier densities in these subbands reach a limit (of the order of $5\times 10^{12} = 10^{13}$ cm^{-2}), which corresponds to a degeneracy energy of more than 100 meV. Striking fact under these conditions is that the position of the peak is completely independent of the current: There are no indications of a ''red'' shift due to the well-known narrowing of the band gap with a high electron-hole plasma density.[2] Thus, the $1e-1hh$ edge does not depend on the plasma density nearly up to the maximum filling of the subbands. It is interesting that the same result was obtained earlier in an investigation of the radiation of a GaAs quantum well under the action of powerful ultrashort (0.6 ps) optical pulses.[10] The absence of a displacement of the long-wavelength edge of the band accompanying a large change in the pump level led to the conclusion that band-gap renormalization is not observed, while the calculations show that there should be a shift of at least 30 meV, with density varying from 5×10^{11} to 5×10^{12} cm^{-2} (Ref. 10; see also Ref. 3). In Ref. 1 we pointed out a similar inconsistency in InGaAs also. Here we can state that the direct observation of the $1e-1hh$ edge in a wide current interval (200-fold increase at 4.2 and 286 K) does not give any indications of a red shift, and the question of what compensates for the narrowing of the band gap in this case remains open.

The continuous growth of the main peak (no saturation) can be explained by the fact that radiative recombination occurs mainly at the $1e-1hh$ edge, while the higher states supply carriers on account of intraband relaxation to this edge. The strong drop in the spectral density on the short-wavelength wing agrees with this assumption. It is well known that the matrix element of optical transitions for TE polarization decreases with increasing photon energy. It is possible that the reduced density of states in the quantum well has a maximum on the $1e-1hh$ edge (because of the

flat character of the dispersion curve in the upper valence band and (or) because of Coulomb attraction of the carriers). The absence of saturation in the entire long-wavelength wing attests to the fact that this wing was formed as a result of homogeneous broadening of the $1e-1hh$ edge.

4. CONCLUSIONS

The modeless samples of quantum wells made it possible to trace the evolution of spontaneous emission under conditions which are unique with respect to the high pumping level at low temperatures (nominal current up to $9\ \mathrm{kA\cdot cm^{-2}}$). The analysis of the structure of the spectra appears comparatively simple because of the fact that in the case of a thin active layer we are actually dealing with two allowed transitions, $1e-1hh$ and $2e-2hh$, while other transitions can be seen as weak substructures (at low temperature) or are not seen at all (at room temperature). Almost all observed substructures fit into a consistent scheme, shown in Fig. 4. From analysis of the main spectral peak we conclude the following:

1) Its position is virtually current-independent. There are no indications of a red shift corresponding to the band renormalization theory, although the large range of filling of the quantum well objectively follows from the appearance of transitions 100 meV above the $1e-1hh$ edge.

2) The steep $1e-1hh$ edge is subject to primarily to homogeneous (collisional) broadening. This explains the absence of saturation in the long-wavelength wing with a 200-fold increase in current. This behavior cannot be reconciled with the well-known band-filling model, which deals with states in inhomogeneously broadened band edges;

3) Similarly to the absorption edge, which satisfies the Urbach rule, the long-wavelength spontaneous emission wing is characterized by an exponential dropoff of the spectral density (apparently this is valid as long as the dropoff is due to collisional broadening and is not masked by substructures and other inhomogeneities).

We thank Dr. G. Beister (von Braun Institute, Berlin) for supplying the samples and A. E. Drakin (Physics Institute of the Academy of Sciences) for assisting in the work.

A portion of this work was performed and supported as part of a project of the ''Physics of solid state nanostructures'' program.

[a]Fax: 007(095)1326649; e-mail: drakin@ares.lpi.msk.su

[1] I. V. Akimova and P. G. Eliseev, Proc. SPIE **2693**, 640 (1996).
[2] W. W. Chow, S. W. Koch, M. Sargent III *et al.*, *Semiconductor Laser Physics*, Springer-Verlag, Berlin, 1994, p. 94.
[3] C.-F. Hsu, P. S. Zory, C.-H. Wu, and M. A. Emanuel, IEEE J. Sel. Top. Quantum Electron. **3**, 158 (1997).
[4] P. G. Eliseev, G. Beister, A. E. Drakin, I. V. Akimova, G. Erbert, J. Maege, and J. Sebastian, Kvant. Élektron. **22**, 309 (1995) [J. Quan. Electron. **25**, 291 (1995)].
[5] P. G. Eliseev and A. E. Drakin, Proc. SPIE **2399**, 302 (1995).
[6] F. Urbach, Phys. Rev. **92**, 1324 (1953).
[7] Y. Toyozawa, Prog. Theor. Phys. **20**, 53 (1958).
[8] M. Schreiber and T. Toyozawa, J. Phys. Soc. Jpn. **51**, 1528 (1982).
[9] M. V. Kurik, Phys. Status Solidi A **8**, 9 (1971).
[10] B. Deveaud, F. Clerot, K. Fujivara, and K. Mitsunaga, Appl. Phys. Lett. **58**, 485 (1991).

Translated by M. E. Alferieff

Radiation emitted by InGaAs quantum-well structures. II. Homogeneous lineshape function

P. G. Eliseev

Temporary address: Center for High Technology Materials, University of New Mexico, Albuquerque, USA

I. V. Akimova[a)]

P. N. Lebedev Physical Institute, 117924 Moscow, Russian

(Submitted October 13, 1997; accepted for publication November 10, 1997)
Fiz. Tekh. Poluprovodn. **32**, 478–483 (April 1998)

The shape of the spontaneous-emission spectra of quantum-well InGaAs structures is analyzed in a wide current interval at temperatures 4.2–286 K. Using a modified homogeneous lineshape function, an interpretation of the band shape for transitions between the lowest subbands is given and the nature of its non-Lorentzian contour is discussed. The properties of a family of functions exhibiting a continuous transition from a Lorentzian to a Gaussian lineshape are discussed. They could be suitable for modeling spectra in the case of non-Markovian collisional relaxation. © *1998 American Institute of Physics.* [S1063-7826(98)01704-9]

1. INTRODUCTION

In the description of radiative processes in semiconductor lasers and LEDs, comparatively little attention is devoted to experimental analysis of the homogeneous lineshape function. This is explained by the fact that broadening is ordinarily masked to some extent by inhomogeneous effects and information about the lineshape function must be extracted indirectly, for example, by comparing the computed shape of the spectrum with the experimental shape in the presence of an uncertainty with respect to the inhomogeneous component. In Refs. 1 and 2 we obtained spectra of this kind in a wide range of pump currents. In a previous work[1] we used a non-Lorentzian modified lineshape function (with exponential asymptotes) that describes spectral band shapes including the wings.

In the literature there is no unanimous agreement about the lineshape function for semiconductor laser radiation. Different lineshape functions are used — from Lorentzian to Gaussian — in current computational work. In Ref. 3 it was noted that a Lorentzian contour overestimates the broadening because of the slow dropoff of the intensity in the wings. A contour of the type $1/\cosh[(E_0-E)/\gamma]$, where E is the energy, E_0 is the line center, and γ is the width parameter of the contour, gives more realistic results. Indications of a non-Lorentzian lineshape function have appeared in different investigations,[4,5] and a theoretical explanation of this phenomenon was given by taking into account non-Markovian relaxation processes[6–9] (this is examined in greater detail in Sec. 2). Experimental confirmations of this are a topical problem, since they should shed light on the nature of the fundamental physical processes occurring in semiconductors and also because increasingly more accurate simulation of laser characteristics become necessary as laser technology is perfected.

The lineshape function must actually be taken into account in calculations of the optical gain (we have in mind the peak value), as well as the complete gain, spontaneous emission, and absorption spectra outside the gain band and in the simulation of the mechanisms of spectral restructuring and nonlinear phenomena.

Experimental investigations require perfect objects so as to reduce the influence of spatial inhomogeneities to a minimum. Thermal effects can be taken into account by varying the temperature, including low values, for which the thermal energy is less than the width of the lineshape function. A suitable object is a quantum-well structure with three-dimensional carrier confinement (quantum dots). Such a structure possesses a discrete spectrum similar to atomic spectra. However, in real structures of this type inhomogeneous broadening apparently predominates because of variations in the size, shape, and composition of the quantum cells. Another suitable object is a conventional quantum well (one-dimensional confinement), whose energy spectrum, ideally, possesses steps at the subband edges. Broadening of a step can be interpreted on the basis of the notions of homogeneous and inhomogeneous components.

In this study and in Ref. 2 (continuing the investigations presented in Ref. 1) spectral investigations of the spontaneous emission of an InGaAs quantum well in the temperature interval 4.2–286 K were performed in a wide range of pumping levels, including values exceeding the inversion threshold by a factor of 100. As a result, data on the homogeneous lineshape function (long-wavelength side), the temperature dependence of the lineshape function, and the substantial deviation from the Lorentzian shape were obtained. The nature of the non-Lorentzian contour is discussed and a lineshape function is proposed in an analytical form which is convenient for use in numerical simulations of semiconductor lasers.

 © 1998 American Institute of Physics

2. ON THE HOMOGENEOUS LINESHAPE FUNCTION

2.1. Mechanisms of broadening of optical spectra

Spectral broadening is divided somewhat arbitrarily into inhomogeneous and homogeneous types. In a semiconductor crystal the interband emission band is initially inhomogeneously broadened because of the law of conservation of quasimomentum. The width of the spectrum is related to the population of the working levels and is determined in order of magnitude either by the thermal energy kT or the total Fermi energy for carriers of both signs, depending on the degeneracy. Effects due to spatial inhomogeneity, which broaden the band edges and engender ''tails'' in the density of states in the band gap, can be added. Ultimately, the total inhomogeneous broadening determines the ''static'' structure of the emission spectrum. Its theory was given in Refs. 10–12. Inhomogeneous broadening in a quantum well as a result of thickness fluctuations is studied in Ref. 13.

Homogeneous broadening reflects the dynamics of the quantum states of the carriers, i.e., the effect of all relaxation processes on the damping of these states and, as an unavoidable consequence, on the energy width of the levels according to the uncertainty relation. In this case the full width at half-maximum (FWHM) is 2γ, where $\gamma=\hbar/\tau$, and τ is the decay time of the states. The relaxation processes include recombination (which gives the ''natural'' width) and all types of intraband relaxation, including quasimomentum exchange during emission. If the static structure of the spectrum is given by the function $R_0(h\nu)$, then the homogeneous broadening is usually taken into account by a convolution integral

$$R(h\nu)=\int_0^\infty R_0(E)F(E-h\nu,\gamma)dE, \tag{1}$$

where $F(E-h\nu,\gamma)$ is the homogeneous lineshape function. The homogeneous width is added to the broadening of the function $R_0(h\nu)$. We note that the adding of the partial widths depends, in general, on the form of the corresponding contours. But this addition is always closer to geometric than arithmetic, so that the contribution of the lesser component to the total width is of second order. The effect of dynamic processes on the emission spectrum has been discussed in Refs. 14, 15, and 6–9.

Broadening due to the finite lifetime of the states (lifetime broadening) is ordinarily described by a Lorentzian contour[14,16,17]

$$F(E-h\nu,\gamma)=(\gamma/\pi)/[\gamma^2+(E-h\nu)^2], \tag{2}$$

which clearly limits the nature of the processes which are taken into account. It consists of the assumption that there exists a single relaxation time, i.e., it does not change in the course of the process. In solids collisional-type relaxation, for which a lineshape function in the form (2) is inadequate, plays the dominant role. Nonetheless, this form is very often used in calculations, for example, of the optical gain at the peak of the spectral band, and gives satisfactory results. This can be explained in part as follows: In such calculations the inadequate accounting of the shape is easily compensated for by the fact that the contour width itself (or the intraband relaxation time) is most often an adjustable parameter and is not taken from independent measurements. However, when the question concerns the shape of the gain spectrum, some ''unphysical'' consequences of the contour (2) appear.

Equation (1) is also used in application to the absorption spectrum. In addition, because of the asymmetry of the absorption at the edge of the intrinsic band, a ''creeping'' of this edge occurs and an absorption ''tail'' appears in the transmission region. The Lorentzian contour (2) gives a large overestimation of absorption in the tail (the Lorentzian form is characterized by poor convergence). This was noted in calculations of the gain spectrum, when the gain band is actually superimposed on the absorption tail from noninverted transitions deep in the absorption band. Ultimately, the transmission point is shifted appreciably from the point where the photon energy equals the difference ΔE_F of the Fermi quasilevels. This is a breakdown of detailed balance, initially assumed as the condition of quasiequilibrium in the presence of pumping. Moreover, there appears an apparently spurious long-wavelength ''transmission point'' (where the computed absorption in the ''tail'' starts to predominate over amplification). Since these results are at variance with experiment, they served as a motivation for reexamining the lineshape function[6–9] for both bulk emitting material and quantum-well structures. The crux of the reexamination consists of taking into account the non-Markovian character of collisional relaxation, in the presence of which, as a result of the finite correlation of events, the damping of the states is not described by a single relaxation time (i.e., the dropoffs in time are not exponential) and the width of the contour is determined by a weighted value characterizing the dynamics of the relaxational processes.

2.2. Lineshape function for non-Markovian processes

Let us examine the method employed in Refs. 6 and 7. The kinetics of the radiative process can be represented as an oscillatory packet, referring to the dipole moment $P(t)$ of an electron-hole pair,

$$P(t)=P_0\exp[-L(t)]\exp(i\omega_0 t), \tag{3}$$

where P_0 is the amplitude of the packet at the point $t=0$, $L(t)$ is a damping factor, and ω_0 is the angular frequency of the oscillations. The lineshape function corresponds to the normalized Fourier transform

$$F(\delta)\propto\int_0^\infty P(t)\exp(\pm i\omega t)dt, \tag{4}$$

where δ is the frequency detuning relative to ω_0. The packet function $P(t)$ in Eq. (4) can represent a real wave packet, as in Eq. (3), but it can also represent the autocorrelation function of the oscillations of the dipole moment which are perturbed in amplitude and phase as a result of collisional relaxation. For noncorrelated interruptions the damping factor grows as the time, $L(t)=t/\tau$, where τ is the relaxation time. This gives a packet envelope $\exp(-|t/\tau|)$ and a Lorentzian lineshape function. The packet envelope is a ''sharp'' bell, i.e., it has a cusp at $t=0$, where it is nondifferentiable. This cusp is the reason for the poor convergence of the Lorentzian

shape: It (i.e., the strong change at short times) makes a large contribution to the wings of the lineshape function for large detunings. In the case of non-Markovian processes the damping factor is a nonlinear function of time. This is manifested primarily at short times (for short correlation time τ_{cor}). The calculations performed in Refs. 7 and 8 concerning the function $L(t)$ can serve as an illustration. If τ_{cor} is much longer than the process time, then $L(t)\sim t^2$ and both contours — the packet and lineshape functions — are Gaussian.

Unfortunately, direct theoretical calculations do not give the analytical form of the lineshape function in intermediate cases, and for this reason the Lorentzian lineshape function is still used in studies involving the simulation of amplification in semiconductors. Numerical calculations of the lineshape function lead to a non-Lorentzian, in general, asymmetric form. The non-Lorentzian nature is manifested primarily in the contour wings, where the Lorentzian asymptotes are clipped. In the experimental study[1] it was pointed out that a lineshape function of the form $1/\cosh\delta$ with exponential asymptotes can be used to fit the spontaneous emission spectra. A more general (asymmetric) lineshape function can be introduced

$$F(\delta)=2/[\exp(\delta/\varepsilon_1)+\exp(-\delta/\varepsilon_2)], \quad (5)$$

which has asymmetric exponential asymptotes, which are determined by the constants ε_1 and ε_2. This approach is developed in the present paper. We shall show below that it is convenient to use for the envelope of the packet function $P(t)$ an approximation in the form of a family of functions $1/\cosh^n(t/\tau n)$, which passes continuously from $\exp(-t/|\tau|)$ in the limit $n\rightarrow 0$ to a Gaussian in the limit $n\rightarrow T$. Thus, the parameter n of the family can be compared with the dimensionless parameter $m=\gamma\tau_{cor}$, which characterizes the non-Markovian behavior. This parameter varies in the same interval: from 0 for Markovian processes and a Lorentzian lineshape function up to T in the case of complete correlation of the processes and a Gaussian lineshape function. A lineshape function in the form (5) is proposed in the intermediate case $n=1$.

The order of magnitude of the correlation time is determined from the uncertainty relation, where the energy uncertainty corresponds to the energy transferred in a collisional process. For carriers near band edges this energy is of the same order of magnitude as the homogeneous width. Therefore the likely value of m is of the order of 1.

3. EXPERIMENTAL RESULTS

In a companion paper[2] we presented the results of a spectral investigation of the spontaneous emission of InGaAs quantum wells in a wide range of pump currents (up to $J\simeq 9$ kA$\cdot$cm^{-2}, including experiments at temperature $T=4.2$ K). Here we shall employ the same results, and for this reason we do not describe the structures and the details of the measurements. The following data are of greatest interest from the standpoint of determining the lineshape function.

1. *Shape of the emission spectrum on the transitions $1e-1hh$ in the limit of weak pumping.* In the ideal case the inhomogeneous structure of the spectrum at low temperature can be reduced to a peak of width less than 1 meV. If the lineshape function has a large width, then the emission spectrum will simply repeat the contour of the lineshape function.

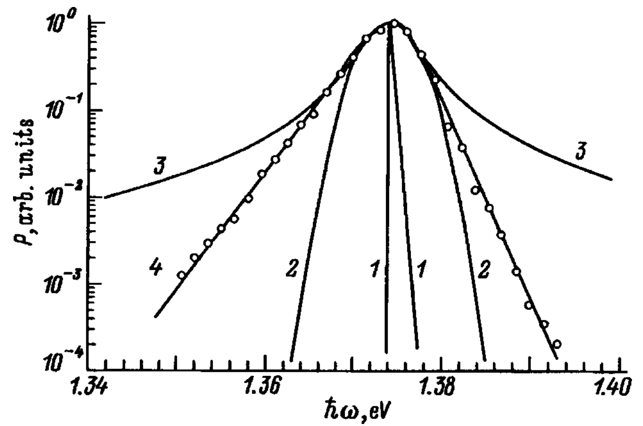


FIG. 1. Normalized spectral density P of spontaneous emission of an InGaAs quantum well at $T=4.2$ K with current $I=1$ mA.[2] The computed curves correspond to the cases: *1* — no homogeneous broadening, *2* — Gaussian broadening, *3* — Lorentzian broadening, *4* — lineshape function in the form (5) (see text).

2. *Form of the long-wavelength dropoff of the emission band on the transitions $1e-1hh$.* If this dropoff is due to homogeneous broadening, then the decay parameter will be the same as for the lineshape function, since here the inhomogeneous structure of the spectrum possesses a sharp step. The variation of the decay parameter with the electron-hole plasma density can be obtained by varying the current.

An example of an analysis of the shape of a peak with an $I=1$ mA pump and temperature $T=4.2$ K is presented in Fig. 1. The dots show the experimental data. The curve *1* is the computed inhomogeneous shape (jump on the long-wavelength side and Fermi dropoff on the short-wavelength side). The width of this peak is 0.25 meV. Curves *2* and *3* correspond to the standard lineshape functions, which are suitable only for obtaining an approximate description of the band top (from 100 to 25% in terms of the spectral density), while neither the Gaussian 2 nor the Lorentzian 3 is suitable for describing the band wings. Curve *4* corresponds to the function (5), where δ is the frequency detuning. Thus it can be shown that in the case where the spectrum of the emission band should be close to the lineshape function it has an exponential dropoff ($\varepsilon_1=3.2$ meV, $\varepsilon_2=1.9$ meV). The width equals 7.5 meV, which gives an estimate of about 175 fs for the effective relaxation time. It is obvious that the lineshape function refers to a non-Markovian relaxation process. Hypothetically, the process is characterized by comparable times for duration and correlation ($n\simeq 1$). The short-wavelength wing can in principle be distorted because of carrier heating. The slope corresponds formally to the maximum temperature ~ 22 K, which gives an upper limit for the possible carrier heating.

Figure 2 shows a similar approximation for the spectrum at $T=77$ K and current $I=1$ mA. Here a lineshape function with exponential asymptotes is likewise required ($\varepsilon_1=6.9$ meV, $\varepsilon_2=6.4$ meV). The solid curve in the form (5) gives a satisfactory approximation. The short-wavelength wing actu-

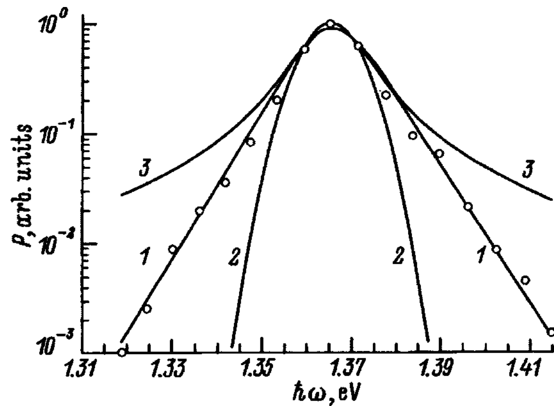


FIG. 2. Approximation for the spectral density of spontaneous emission of an InGaAs quantum well at $T=77$ K with current $I=1$ mA. *1* — Function (5), *2* — Gaussian broadening, *3* — Lorentzian broadening.

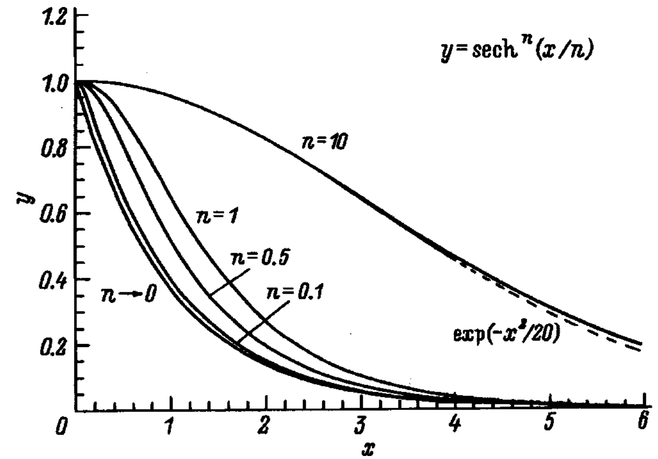


FIG. 3. Functions (A.1) in the limit $n\to 0$ (*1*) and from $n=0$ to $n=10$; the dashed line shows a Gaussian curve close to the curve $n=10$.

ally corresponds to a Boltzmann dropoff. In both cases the contours differ substantially from the standard lineshape functions; the real contours are slightly asymmetric. The exponentials extend over 2.5 orders of magnitude.

The long-wavelength wing of the emission band can be identified with the long-wavelength wing of the lineshape function, since it results from broadening of the sharp $1e-1hh$ edge. The dropoff of the spectral density by 2–2.5 orders of magnitude is almost exponential (in the region 1.35–1.37 eV) and deviates from an exponential in the region of the above-mentioned secondary substructure. In Ref. 2 we identified this property as an Urbach rule for emission. The slope parameter equals 3.2 meV with a weak current $I=1$ mA and 3.3 meV with $I=200$ mA, i.e., it remains nearly constant. At 286 K the variation of the slope is larger — from 4.4 meV (1 mA) to 8.5 meV (200 mA).

The following conclusions can be drawn: a) The long-wavelength dropoff of the lineshape function is virtually exponential over 2–2.5 orders of magnitude (Urbach rule); b) the slope parameter increases slightly with increasing current at room temperature and changes very little at 4.2 K; and c) the slope parameter increases with temperature, which apparently corresponds to acceleration of the collisional relaxation as a result of optical phonons.

4. CONCLUSIONS

Spectral investigations were performed for the purpose of determining the homogeneous lineshape function in an InGaAs quantum well. A sample with an easily interpreted spectrum, sufficiently homogeneous to reduce inhomogeneous broadening to a minimum, was studied in the temperature range 4.2–286 K. To obtain a large range of carrier densities and eliminate superluminescence, a narrow strip structure without a cavity and with a strong antiwaveguide effect (''modeless'' structure) was used. The results which we obtained lead to the following conclusions.

1. The lineshape function has exponential dropoffs. This was most reliably shown for the long-wavelength dropoff (its slope at 4.2 K is 3.2 meV). This corresponds to a non-Markovian collisional relaxation.

2. An Urbach rule was formulated for emission: The long-wavelength dropoff satisfies an exponential law. This holds so long as the dropoff is due to homogeneous broadening and is not masked by phonon repetitions or other spectral substructures.

3. It was shown that a weakly asymmetric or symmetric (non-Lorentzian) contour of the lineshape function can be described analytically. This simplifies the computational procedure in the simulation of laser and LED spectra.

4. When justifying the modification of the lineshape function, we indicated a family of analytical functions $1/\cosh^n(x/n)$ that contains a continuous limiting transition from a sharp bell $\exp(-|x|)$ to a Gaussian. The Fourier transform of these functions gives a continuous transition from a Lorentzian to a Gaussian lineshape function, in accordance with its qualitative evolution with increasing correlation time — gradual smoothing and flattening of the sharp bell $\exp(-|x|)$, characteristic for an uncorrelated process. In the intermediate case (the correlation time is comparable to the duration of the process) the lineshape function can be approximately described as $1/\cosh(x)$.

APPENDIX

Properties of the family of functions Ep(*x*,*n*)

The bell-shaped contours (5) can be united into a family of functions $\mathrm{Ep}(x,n)$, where x is either the time or the frequency detuning, while n is a parameter of the family. Thus, this family can pertain to the packet function in Eq. (3) and to the lineshape function. We shall examine here symmetric contours of the type

$$\mathrm{Ep}(x,n)=1/\cosh^n(x/n), \tag{A.1}$$

which are interesting in that they represent a continuous transition from a sharp bell $\exp(-|x|)$ to a Gaussian. Therefore, their Fourier transform gives a continuous transition from a Lorentzian to a Gaussian contour. From the standpoint of the evolution of the lineshape function this corresponds to $\gamma\tau_{\mathrm{cor}}$ which varies from zero to infinity, i.e., in the same limits as the parameter n. This family can be used to obtain the lineshape function in the transitional region, and it is found to be

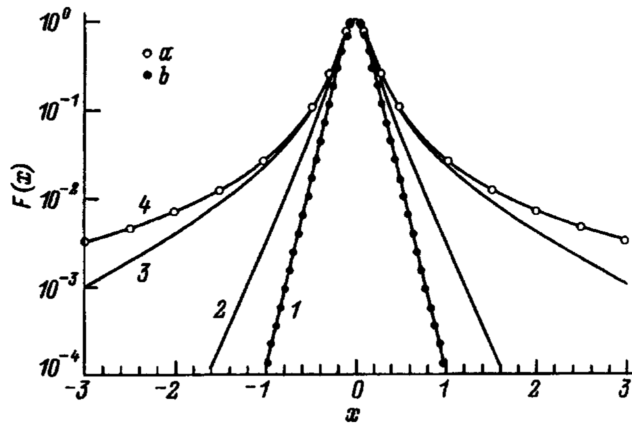


FIG. 4. Amplitude F of the Fourier transform versus the relative frequency detuning x of the functions (A.1) for $n=1$ (*1*), 0.5 (*2*), 0.1 (*3*) and in the limit $n\rightarrow 0$ (*4*); a — Lorentzian contour close to curve *4*, b — function of form (A.1).

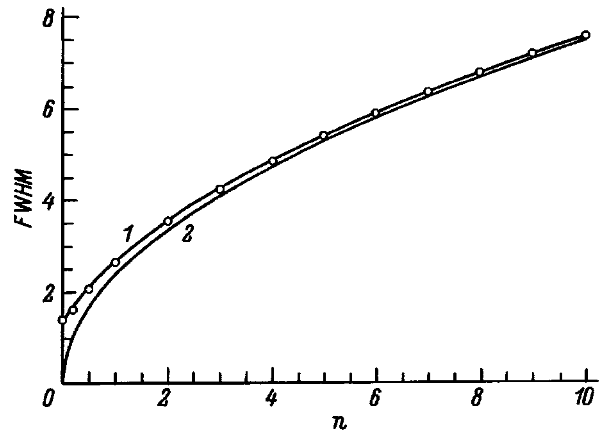


FIG. 5. FWHM of the contour (A.1) in units of ε versus the parameter n (*1*) and width of the Gaussian with $\sigma^2=n$ (*2*).

helpful as a replacement for a Lorentzian contour, which gives inadequate results, in the numerical simulation of spontaneous emission and optical gain in semiconductor media.

Figure 3 shows contours of the function (A.1) as $n\rightarrow 0$ and up to $n=10$. The dashed line shows the Gaussian exp $(-x^2/2\sigma^2)$ with $\sigma^2=n=10$, which in the region $-4.4<x<4.4$ is virtually identical to the function (A.1) for $n=10$. In the indicated interval the difference does not exceed 3%, but outside this interval it increases rapidly, since the asymptotes of the function (A.1) are always exponential. As n increases, the regions of divergence of the asymptotes move farther toward the periphery. By expanding the function (A.1) in a Taylor series and comparing the first three terms of the expansion with the corresponding series for a Gaussian it can be shown that the variance σ^2 of the limiting Gaussian equals n.

The contour 1 in Fig. 3 has a cusp at $x=0$, which renders ''unphysical'' the hypothesis that the processes for large frequency detunings are uncorrelated: In natural processes one can always find reasons why a mathematical cusp is not realized. We recall that if one is talking about the envelope of oscillations, then cusps are physically absent in a region of the order of the period of the oscillations around zero. This corresponds to the inadequacy of the Lorentzian lineshape function for detunings comparable to the frequency of the oscillatory process. Slower relaxation mechanisms smooth the cusp near the zero point — the longer the relaxation time, the more extended the cusp. As a result, the Lorentzian contour deforms: Its far wings are clipped, hyperbolas are replaced by exponentials. The evolution of the Fourier-transform contour from $n=1$ to $n\rightarrow 0$ is shown in Fig. 4 (on a semilogarithmic scale). The curves were obtained by the FFT method. The dots a give a Lorentzian contour ($\gamma=0.167$), which, as expected, is identical to the curve for $n\rightarrow 0$, while the curve with $n=1$ is identical to the dots b for the function $1/\cosh(9.84x)$ which belong to the family (A.1). It is well known that the Fourier transform of bell-shaped contours is also represented by bell-shaped contours; here a Gaussian transforms into a Gaussian. The cosine Fourier transform of the contour $1/\cosh(x/\varepsilon)$ (see, for example, Ref. 18) is

$$F(\delta)=(\pi\varepsilon/2)/\cosh(\pi\delta\varepsilon/2). \quad \text{(A.2)}$$

Thus, for $n=1$ the lineshape function belongs to the family (A.1). Table I gives the comparative characteristics of the standard and modified ($n=0.5$ and 1) spectral contours. It is interesting that for $n=0.5$ the contour (A.1) is a good substitution for a Lorentzian contour: In the interval $-2<x<2$ it does not deviate from a Lorentzian by more than 3%, while outside this interval it passes rapidly to the exponential asymptotes.

A comparison in Fig. 3 of the contours of the function (A.1) for different values of n shows that for the same dropoff parameter ε the parameter n characterizes the relation between the width of the contour and the dropoff parameter. The dependence of the full width on n is shown in Fig. 5 (curve *1*) in units of ε. This is a slowly varying function, but in the limit $n\rightarrow\infty$ the relative width approaches infinity, since the contour approaches a Gaussian contour. Curve *2* shows the width of a Gaussian with variance equal to n; this curve approach curve *1* as n increases.

This work was carried out in part and with the support of a project of the ''Physics of solid-state nanostructures'' program.

TABLE I. Characteristics of spectral lineshape functions (normalized to unit area); $x=E-E_0$.

Lineshape function	Formula	Value at $x=0$	FWHM
Lorentzian	$(y/\pi)(x^2+\gamma^2)$	$0.31831/\gamma$	2γ
Ep, $n=0.5$	$(0.38139/\varepsilon)[\cosh(2x/\varepsilon)]^{1/2}$	$0.38139/\varepsilon$	2.063ε
Ep, $n=1$	$(1/\pi\varepsilon)/\cosh(x/\varepsilon)$	$0.31831/\varepsilon$	2.6339ε
Gaussian	$[(1/2\pi)^{1/2}/\sigma]\exp(-x^2/2\sigma^2)$	$0.39844/\sigma$	2.3548σ

[a]Fax: 007(095) 1326649; e-mail: drakin@ares.lpi.msk.su

[1] I. V. Akimova and P. G. Eliseev, Proc. SPIE **2693**, 640 (1996).

[2] P. G. Eliseev and I. V. Akimova, Fiz. Tekh. Poluprovodn. **32**, 472 (1998) [*sic*].

[3] W. W. Chow, S. W. Koch, M. Sargent III *et al.*, *Semiconductor Laser Physics*, Springer-Verlag, Berlin, 1994, p. 94.

[4] M. Kessler and C. Harder, IEEE J. Quantum Electron. **27**, 1812 (1991).
[5] B. Deveaud, F. Clerot, K. Fujivara, and K. Mitsunaga, Appl. Phys. Lett. **58**, 1485 (1991).
[6] M. Yamanishi and Y. Lee, IEEE J. Quantum Electron. **23**, 367 (1987).
[7] T. Ohtoshi and M. Yamanishi, IEEE J. Quantum Electron. **27**, 46 (1991).
[8] A. Tomita and A. Suzuki, IEEE J. Quantum Electron. **27**, 1630 (1991).
[9] D. Ahn, IEEE J. Sel. Top. Quantum Electron. **1**, 301 (1995).
[10] G. Lasher and F. Stern, Phys. Rev. A **133**, 553 (1964).
[11] F. Stern, in *Semicond. and Semimetals*, edited by R. K. Willardson and A. C. Beer, 1966, Vol. 2, p. 371.
[12] A. G. Aleksanyan, I. A. Poluéktov, and Yu. M. Popov, Kvant. Élektron. **1**, 62 (1974) [Sov. J. Quantum Electron. **4**(1), 32 (1974)].
[13] J. Singh, K. K. Bajaj, and S. Chaudhuri, Appl. Phys. Lett. **44**, 805 (1984).
[14] M. Asada and Y. Suematsu, IEEE J. Quantum Electron. **21**, 434 (1985).
[15] M. Asada, IEEE J. Quantum Electron. **25**, 2019 (1989).
[16] P. T. Landsberg and D. J. Robbins, Solid-State Electron. **28**, 137 (1985).
[17] C. Weisbuch and B. Winter, *Quantum Semiconductor Structures*, Academic Press, N. Y., 1991, p. 168.
[18] H. Bateman, *Table of Integral Transforms*, McGraw-Hill Book Co, Inc., N. Y., 1954, Vol. 1.

Translated by M. E. Alferieff

AMORPHOUS, GLASSY AND POROUS SEMICONDUCTORS

Staebler–Wronski effect as a function of the Fermi level position and structure of nondoped, amorphous, hydrated silicon

O. A. Golikova, M. M. Kazanin, and V. Kh. Kudoyarova

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia
(Submitted July 7, 1997; accepted for publication July 31, 1997)
Fiz. Tekh. Poluprovodn. **32**, 484–489 (April 1998)

The photoconductivity degradation rates γ ($\sigma_{\rm ph}\sim t^{-\gamma}$) of nondoped, amorphous, hydrated silicon films deposited at $T_s=300-400$ °C and subjected to illumination for 5 h at 300 K (light source 100 mW/cm^2, $\lambda<0.9$ μm) were investigated. It was shown that the degradation rate γ depends on the preillumination position of the Fermi level $\varepsilon_c-\varepsilon_F$ and often is not directly related to the hydrogen content in the film. It was found that there are correlations between the value of γ and the bonds in the silicon–hydrogen subsystem [isolated SiH and SiH_2 complexes, clusters $(SiH)_n$, and chains $(SiH_2)_n$]. © *1998 American Institute of Physics.*
[S1063-7826(98)01804-3]

1. INTRODUCTION

The Staebler–Wronski (SW) effect[1] continues to attract attention from a large number of investigators. It is well known that this effect consists of an increase in the defect density N_D in films of amorphous hydrated silicon a-Si:H under the action of prolonged illumination together with a simultaneous drop in photoconductivity $\sigma_{\rm ph}$ and dark conductivity σ_d; the latter satisfies a shift of the Fermi level ε_F toward the center of the mobility gap E_g. In other words, a-Si:H becomes intrinsic silicon ($\varepsilon_c-\varepsilon_F=0.70$ eV) but with a high density of defects — dangling Si–Si bonds which, as ESR investigations showed, are in a neutral state D^0. Thus, degradation of the material occurs. It is obvious that stability is a key problem for device applications of a-Si:H.

Since the SW effect is observed just for hydrated amorphous Si, it is of interest to clarify the effect of hydrogen (both its amount C_H in the a-Si:H film and the type of Si–H bonds) on the density of the light-induced defects. However, the data concerning both C_H and monohydride (SiH) or dihydride (SiH_2) bonds are contradictory: Sometimes directly opposite conclusions are drawn concerning the role of these factors for the SW effect (see, for example, Refs. 2–5). Nonetheless, the most popular model of the formation of light-induced defects remains the model proposed in Ref. 6: The energy released as a result of the recombination of nonequilibrium charge carriers leads to a conversion of weak Si–Si bonds into dangling bonds with mandatory participation of hydrogen diffusion, which exists in a-Si:H at all finite temperatures and which increases with increasing C_H at $T=$const. It has recently been shown[7] that light additionally intensifies diffusion. However, as already mentioned, the effect of hydrogen on the stability of the parameters of a-Si:H films under the action of light is still not understood. This is also true of the other parameters of the a-Si:H structure, despite considerable effort by many investigators to determine the role of each of them for the SW effect.

In our view, the present situation is attributable to the fact that the parameters of the structure of a grown nondoped a-Si:H film (prior to illumination) are functions of the Fermi level position $\varepsilon_c-\varepsilon_F$ and therefore are interrelated.[8] On this basis, groups of nondoped a-Si:H films, which were deposited at $T_s=300-400$ °C and for which

(I) $T_s=$const, $\varepsilon_c-\varepsilon_F=$var, $C_H=$var,

(II) $T_s=$const, $\varepsilon_c-\varepsilon_F=$const, but $C_H=$var,

(III) $T_s=$var, $\varepsilon_c-\varepsilon_F=$const, $C_H=$var

were chosen by us as the objects for investigating the SW effect.

The films in these three groups were characterized by the microstructural parameter $R\leqslant0.3$, i.e., monohydride complexes SiH predominated in them. Moreover, another group of films (IV) was also investigated. For this group, variations of the microstructural parameter $R=0-0.65$ and variations of the type of dihydride complexes SiH_2 or $(SiH_2)_n$ are also possible at the same values of T_s. Our objective was to determine the effect of the Fermi level position, the values of C_H and R, and some other structural characteristics on the SW effect for nondoped a-Si:H.

2. EXPERIMENT

We investigated films deposited at high temperatures $T_s=300-400$ °C. These are the films that stirred greatest interest recently from the standpoint of increasing the stability of a-Si:H.[9] The deposition methods were: glow discharge (PECVD) in a triode reactor[8,9] and decomposition of silane in a magnetron chamber (dc–MASD).[10] Both the technological parameters of film deposition and the results of their characterization are described in detail in Refs. 8–10. Moreover, films which were sent to us from the École Polytechnique (Palaiseau, France) were investigated. One of these

 © 1998 American Institute of Physics

films was deposited in a diode reactor at $T_s = 300\ ^\circ C$ using 100% SiH_4, while the other was deposited at $T_s = 350\ ^\circ C$ using strongly helium-diluted SiH_4 (98% He).[11] It should be noted that to analyze our experimental data we used extensively the data on the deposition of a-Si:H films from SiH_4–He mixtures (with dilution of silane with helium — He-diluted technique), previously obtained at the École Polytechnique (Ref. 4 and other studies, as well as private communications).

The dark conductivity and photoconductivity, defect density and Urbach parameter, hydrogen content, and the microstructure parameter were determined prior to illumination of the films. In a number of cases the Raman frequency ω_{TO} of the TO phonons and the half-width $\Delta\omega_{TO}$ of the corresponding peak were determined. All investigations were performed at room temperature; the methods are described in Refs. 8–10. Nonetheless, here we shall briefly discuss additional information that can be obtained by analyzing the IR and Raman spectra.

As is well known, IR spectra at 1900–2100 cm^{-1} and also near 875 cm^{-1} are analyzed to determine the different types of Si–H bonds. Analysis of absorption at 1900–2100 cm^{-1} makes it possible to find the contributions of isolated complexes SiH and SiH_2. The existence of such complexes is confirmed by the absorption band observed at 875 cm^{-1}. The microstructural parameter is $R = I_{2090}/(I_{2000} + I_{2090})$ (I_i are the intensities of the corresponding bands); if it is comparatively small, the absorption maximum occurs at 2000 cm^{-1}. As shown in Ref. 8, however, the values of R may not be large, but the absorption maximum shifts toward 2100 cm^{-1}. This shows the existence of clustered monohydride complexes $(SiH)_n$. It is believed that such complexes occur in the form of islands[12] or voids[13] on the surface, which produce a porous, so-called ''granular,'' structure. For this reason, in this case one talks about cluster hydrogen located at the grain boundaries, but then it is necessary to determine whether or not such a structure is indeed amorphous: It is known that the IR absorption band at 2019 cm^{-1} is attributed to SiH complexes located at the grain boundaries of microcrystals.[14] For this reason, the TO Raman band near 520 cm^{-1} was analyzed, since a band at this frequency is characteristic of the microcrystalline phase.[14]

We note that to identify dihydride complexes of the type $(SiH_2)_n$, i.e., a chain structure, it is necessary to observe a doublet in the IR absorption spectrum at 840 and 890 cm^{-1}. The authors of Ref. 15 are of the opinion that such complexes usually are located on the surfaces of comparatively small voids which are isolated from one another.

The a-Si:H films investigated by the method enumerated above were exposed for 5 h at $T = 300$ K to light from a source with $W = 100\ mW/cm^2$ and $\lambda < 0.9\ \mu m$. At this, comparatively short, duration of illumination the photoconductivity can be approximated by power-law function $\sigma_{ph} \sim t^{-\gamma}$, where t is the exposure time, and γ is a parameter characterizing the degradation rate.[16] Moreover, the stability of some films (accelerated stability test) was investigated at Philips University (Marburg, Germany). In contrast to the conventional experiments, the a-Si:H film was illuminated at $T = 300$ K, both on the side of its own surface and on the substrate side by two pulsed xenon lamps (pulse duration 2 μs, repetition frequency 300 Hz, spectral range 400–670 nm, peak power 120 mW/cm^2, average power 70 mW/cm^2). Under such conditions the density of induced defects reached the saturation value N_{sat} after 15 h of illumination, while under the conditions of the conventional experiments (source $W = 100\ mW/cm^2$, $T = 300$ K, illumination of the film surface), as is well known, hundreds of hours are required to reach saturation.

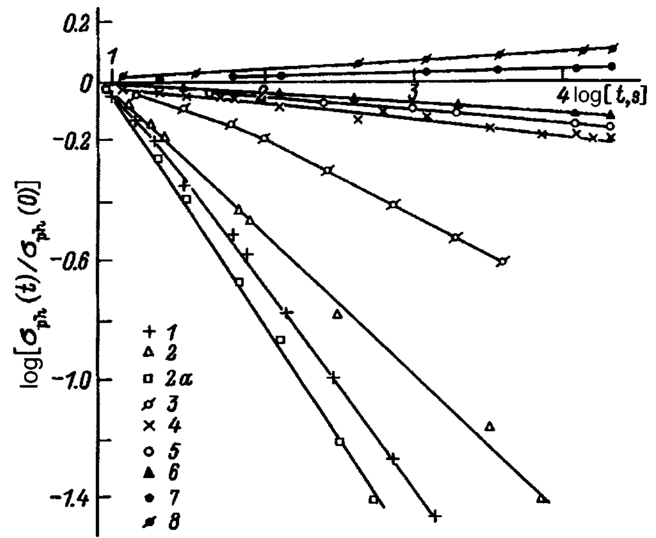


FIG. 1. Relative variation of the photoconductivity σ_{ph} of a-Si:H films deposited at $T_s = 300\ ^\circ C$ versus the illumination time t. The numbers in the figure correspond to the numbers of the films in Table I.

3. DISCUSSION

Let us consider first how the data on the photoconductivity degradation rate, which is characterized by the parameter γ, correlate with the N_{sat} data obtained by means of the accelerated stability test. Figure 1 shows the relative changes in the photoconductivity $\sigma_{ph}/\sigma_{ph}(0)$ as a function of illumination time for a number of films deposited by the PECVD method at $T_s = 300\ ^\circ C$. The characteristics of the films are given in Table I. For two of them, which possess sharply different values of γ (0.66 and 0.05), the absorption coefficients were investigated by the constant photocurrent method (CPM) in states A (preillumination) and B (after illumination

TABLE I. Parameters of films for which data are presented in Fig. 1.

Film No.	$\varepsilon_c - \varepsilon_F$, eV	C_H, at. %	R	γ	Si–H complex types
1	0.45	12	0.30	0.60	SiH, SiH_2
2a	0.67	16	0.27	0.66	SiH, $(SiH)_n$
2	0.65	8	—	0.46	—
3	0.69	9	0	0.2	SiH
4	0.77	—	—	0.05	—
5	0.72	9	0	0.04	SiH
6	0.76	12	—	0.03	—
7	0.82	14	—	−0.02	—
8	0.85	16	0.30	−0.03	SiH, SiH_2

Note: Film 4 was obtained from École Polytechnique (Palaiseau, France).

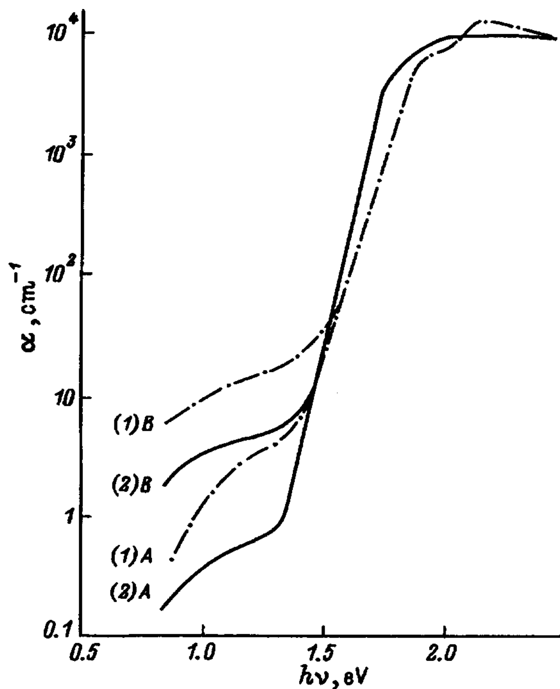


FIG. 2. Absorption spectra (constant current method) for film 2a (curve *1*) and film 4 (curve *2*) (see Fig. 1 and Table I) in the states *A* (preillumination) and *B* (postillumination). Film parameters: No. 2a — $N_D^A=4\times10^{16}$ cm^{-3}, $N_D^B=2\times10^{17}$ cm^{-3}, $E_u=54$ meV; No. 4 — $N_D^A=7\times10^{15}$ cm^{-3}, $N_D^B=5\times10^{16}$ cm^{-3}, $E_u=49$ meV.

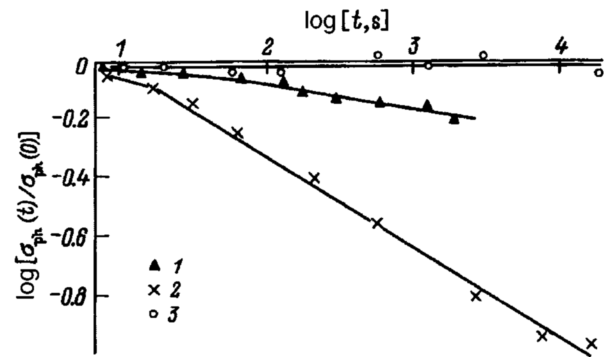


FIG. 3. Relative variation of the photoconductivity σ_{ph} of intrinsic *a*-Si:H films versus the illumination time *t*. The numbers in the figure correspond to the numbers of the films in Table II.

for 15 h) (see Fig. 2). The values of N_{sat} were equal to, respectively, 2×10^{17} cm^{-3} and 5×10^{16} cm^{-3}. Thus, a correlation is observed between γ and N_{sat}. For this reason, in speaking about stability of the material we shall consider the quantity γ.

It follows from Fig. 1 and Table I that for films deposited at T_s=const γ decreases rapidly with increasing gap $\varepsilon_c-\varepsilon_F$ (curves *1–6*). Curve *3* was taken from Ref. 17; it was obtained for close to intrinsic *a*-Si:H. Considering now the preillumination structural parameters of these *a*-Si:H films,[8] the point $\varepsilon_c-\varepsilon_F$=0.70 eV is ''special'' for them. Near this point the defect density, the Urbach parameter E_u, the hydrogen content C_H, and $\Delta\omega_{TO}$ are all at a minimum, the microstructural parameter $R=0$, and ω_{TO} is at a maximum. This material can be regarded as completely homogeneous. In the opinion of the authors of Ref. 8 the boundary between the regions I (0.45 eV $<\varepsilon_c-\varepsilon_F<$0.70 eV) and II (0.70 eV $<\varepsilon_c-\varepsilon_F<$0.85 eV), where defects are in different charge states, is located here. As the position of the Fermi level in nondoped *a*-Si:H changes, the state of the defects changes from predominantly negative D^- (region I) to neutral D^0 (intrinsic *a*-Si:H) and then to predominantly positive D^+ (region II).

On the basis of this concept, in Ref. 18 the rate of both the increase in defect density N_D and decrease of photoconductivity σ_{ph} of the illuminated films were analyzed as a function of $\varepsilon_c-\varepsilon_F$; i.e., the values of γ and β in the expressions $\sigma_{\text{ph}}\sim t^{-\gamma}$ and $N_D\sim t^{\beta}$ were compared. In contrast to Ref. 19, in Ref. 18 it was shown that the rate of increase of the defect density is much lower than the rate of decrease of the photoconductivity only for region-I films, while both rates are comparable for intrinsic *a*-SiH and region-II films. This can be understood if account is taken of the fact that the light-induced defects are in the state D^0. Then even a comparatively small number of them (β for all films does not exceed 0.1) should strongly alter the electron trapping cross section, if prior to illumination all defects were in the state D^- (region I). For this reason γ increases — all the more strongly with greater difference between the material and the intrinsic material. It is obvious that this should not occur for region-II films: For them[18] $\gamma\simeq\beta$.

However, one should note a characteristic feature of region-II films with maximum $\varepsilon_c-\varepsilon_F$: For them σ_{ph} for the given illumination time (5 h) should increase (see Fig. 1 and Table I, films 7 and 8). The same thing was also observed for *a*-Si:H films with high activation energies σ_d, deposited by PECVD method using dilute SiH_4–He mixtures.[4]

On the whole, the region-II films are more stable, despite the fact that, just as in region I, C_H increases as the Fermi level shifts away from the point[8] $\varepsilon_c-\varepsilon_F=0.70$ eV. Hence, it can be concluded that in this case the SW effect is not directly related to C_H.

The role of hydrogen for γ can nonetheless clearly be seen by comparing the data for region-I *a*-Si:H films with $\varepsilon_c-\varepsilon_F$=const (Fig. 1, films 2 and 2a). The second film contains two times more hydrogen and, correspondingly, for it γ is much higher. The same situation is also illustrated in Fig. 3, which shows data on the ratio $\sigma_{\text{ph}}/\sigma_{\text{ph}}$ (0) for two close to intrinsic PECVD films deposited at T_s=390 °C but which nonetheless has different hydrogen content. The film characteristics are given in Table II. We see that γ decreases with C_H. For this reason, one would think that for T_s= const and $\varepsilon_c-\varepsilon_F$=const the stability is determined by the hydrogen content. This conclusion, however, is incorrect, since besides

TABLE II. Parameters of films for which data are presented in Fig. 3.

Film No.	T_s, °C	$\varepsilon_c-\varepsilon_F$, eV	C_H, at.%	R	γ	Si–H complex types
1	390	0.71	4	0	0.1	SiH
2	390	0.71	13	0	0.3	SiH, $(SiH)_n$
3	390	0.71	10	0.65	0	SiH, $(SiH_2)_n$

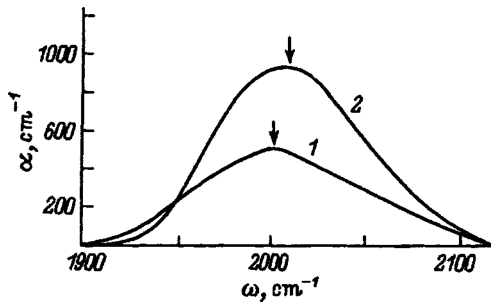


FIG. 4. Infrared absorption spectra $\alpha(\omega)$ for intrinsic a-Si:H film 5 (curve *1*) and film 2a (curve *2*) (see Table I).

different C_H the films also have a different microstructure. This follows from an analysis of their IR spectra (Fig. 4): The shift of the absorption maximum toward 2010 cm^{-1} attests to the presence of SiH cluster complexes located at the boundaries of ''grains'' of the amorphous film. The absence of microcrystalline grains in such films follows from inspection of the Raman spectra: Near the TO-phonon band the characteristic peak located near 520 cm^{-1} is absent. This is shown convincingly in Fig. 5, where the Raman spectra for film 2a and for the region-II film 4 — one of the most stable films — are shown. We see that these spectra are almost completely identical.

Thus, the low stability of films such as film 2a, which have high values of C_H with T_s=const and $\varepsilon_c-\varepsilon_F$=const, is due to their porous inhomogeneous structure and, correspondingly, low density, which promotes hydrogen diffusion, which has a negative effect on stability.[3,20] It is obvious that in this case (T_s=const, $\varepsilon_c-\varepsilon_F$=const, C_H=var) the SW effect cannot be directly associated with the hydrogen content in the films.

A direct relation between the SW effect and the value of C_H was established in investigations of several nearly intrinsic PECVD films deposited at $T_s=300-390$ °C. In the presence of a homogeneous structure, the hydrogen content in the films varied from 9 to 4 at.% and, correspondingly, γ decreased.[17]

We shall now examine the results of the investigations of group–IV a-Si:H films deposited at $T_s=300$ and 350 °C by the MASD method. Together with data for intrinsic PECVD films which was deposited at $T_s=390$ °C Fig. 3 shows data for an intrinsic MASD film deposited at $T_s=300$ °C and which contains 10 at.% hydrogen and the complexes $(SiH_2)_n$ and which has a parameter $R=0.65$. One can see that for this film $\gamma=0$, i.e., this is the most stable film. Figure 6 shows data on the variation of the ratio σ_{ph}/σ_{ph} (0) with increasing illumination time for films deposited at 350 °C. The film characteristics are given in Table III. Two of the films, which are close to intrinsic, were deposited by MASD and PECVD (with strong dilution of silane by helium); they have R values close to zero and differ in hydrogen content C_H. The two other films deposited by PECVD and MASD methods had the same values of $\varepsilon_c-\varepsilon_F=0.64$ eV, but the first film contained $(SiH)_n$ cluster complexes and the second one contained SiH_2 complexes. We see that this is the most stable film among those deposited at the given temperature T_s. Comparing the data for the two MASD films with the same $R=0.65$ but different types of dihydrated complexes (Figs. 3 and 6), one can conclude that they are comparable with respect to degradation rate.

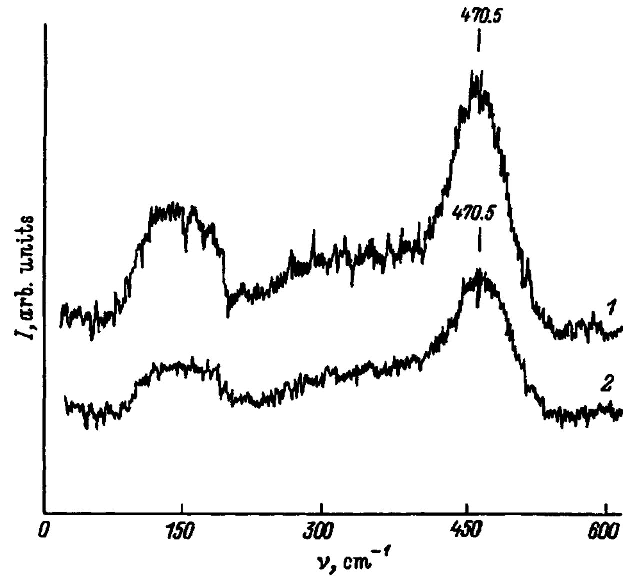


FIG. 5. Raman spectra $I(\nu)$ for a-Si:H film (curve *1*) and film 2a (curve *2*) (see Table I).

According to Ref. 15, the complexes SiH_2 and $(SiH_2)_n$ are located on the surfaces of microvoids, isolated from one another, and for this reason films with such a structure show high thermal stability on annealing: The hydrogen mobility decreases in this case, and therefore hydrogen diffusion occurs with higher annealing temperatures than for standard PECVD films containing predominantly isolated SiH complexes. Apparently, the comparatively low hydrogen mobility is also responsible for the high stability in the case where a-Si:H films with this structure are illuminated. Thus, one would think that to improve stability a-Si:H films containing microvoids, which are isolated from one another and which possess maximum R and minimum C_H, should be prepared. However, then there may not be enough hydrogen to passivate the dangling bonds in the entire volume of the film, and this will have a negative effect on σ_{ph} prior to illumination. Indeed, in Ref. 10 it was noted that σ_{ph} is sharply lower in films where $R\simeq 1$ and $C_H=5-6$ at%. For this reason, to

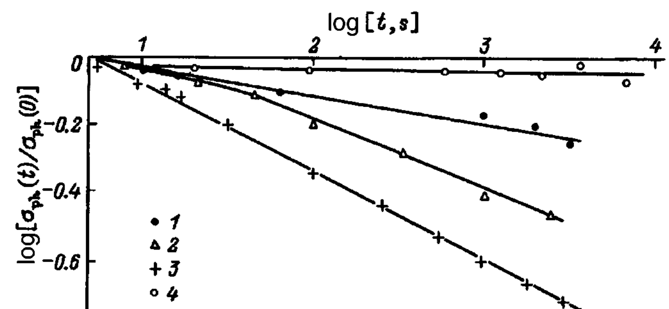


FIG. 6. Relative variation of the photoconductivity σ_{ph} of a-Si:H films deposited at T_s=350 °C versus the illumination time t. The numbers in the figure correspond to the numbers of the films in Table III.

TABLE III. Parameters of films for which data are presented in Fig. 6.

Film No.	$\varepsilon_c-\varepsilon_F$, eV	C_H, at.%	R	γ	Complex types
1	0.71	5	0.0	0.09	SiH
2	—	10	0.1	0.20	SiH, SiH_2
3	0.64	20	0.3	0.3	SiH, $(SiH)_n$
4	0.64	6	0.65	0	SiH, SiH_2

improve stability of a material without degrading its electronic quality an optimal combination of the indicated structural parameters is required.

4. CONCLUSIONS

The following conclusions were drawn as a result of the investigations of the Staebler–Wronski effect for nondoped a-Si:H deposited at high temperatures $T_s=300-400$ °C.

1. The photoconductivity degradation rate of a-Si:H films deposited at T_s=const, which contain isolated SiH and SiH_2 complexes and which possess a microstructural parameter in the range $R=0-0.3$, is determined by the preillumination position of the Fermi level. The degradation rate decreases rapidly with increasing gap $\varepsilon_c-\varepsilon_F$ and is not directly related with the hydrogen content C_H in the films.

2. The large decrease in the photoconductivity degradation rate of a-Si:H films with high C_H and T_s=const and $\varepsilon_c-\varepsilon_F$=const is due to their structural inhomogeneity as a result of the formation of cluster complexes $(SiH)_n$. For this reason it is not possible to establish a direct relation between C_H and γ.

3. Close to intrinsic films deposited at T_s=const are highly stable. The quantity γ for such films decreases systematically with increasing T_s, i.e., with decreasing hydrogen content C_H. In this case a direct relation can be established between C_H and the photoconductivity degradation rate σ_{ph}.

4. The photoconductivity degradation rates of a-Si:H films, which contain the complexes SiH_2 or $(SiH_2)_n$ and which have $R\simeq 0.65$, are close to zero; i.e., other conditions being equal ($\varepsilon_c-\varepsilon_F$=const, T_s=const, C_H=const), these films are highly stable.

This work was supported by INTAS Grant No. 931916.

We thank our colleagues from the École Polytechnique (Palaiseau, France) and Phillips University (Marburg, Germany) for a fruitful collaboration.

[1] D. I. Stabler and C. R. Wronski, Appl. Phys. Lett. **31**, 292 (1977).
[2] H. Fritzsche, MRS Symp. Proc. **336**, 227 (1994).
[3] Y. Liang, N. Maley, and J. R. Abelson, J. Appl. Phys. **75**, 3704 (1994).
[4] R. Meaudre, M. Meaudre, S. Vignoli, P. Poca i Cabarrocas, Y. Bouizem, and M. L. Theye, Philos. Mag. B **67**, 497 (1993).
[5] D. Caputo, G. de Cesare, F. Irrera, F. Palma, M. C. Rossi, G. Conte, G. Nobile, and G. Fameli, J. Non-Cryst. Solids **170**, 278 (1994).
[6] M. Stutzmann, W. B. Jackson, and C. C. Tsai, Phys. Rev. B **32**, 23 (1995).
[7] P. Hari, P. C. Tailor, and R. A. Street, MRS Symp. Proc. **336**, 329 (1994).
[8] O. A. Golikova and V. Kh. Kudoyarova, Fiz. Tekh. Poluprovodn. **29**, 1128 (1995) [Semiconductors **29**, 584 (1995)].
[9] O. A. Golikova, M. M. Kazanin, O. I. Kon'kov, V. Kh. Kudoyarova, and E. I. Terukov, Fiz. Tekh. Poluprovodn. **30**, 405 (1996) [Semiconductors **30**, 336 (1996)].
[10] O. A. Golikova, A. N. Kuznetsov, V. Kh. Kudoyarova, and M. M. Kazanin, Fiz. Tekh. Poluprovodn. **30**, 1879 (1996) [Semiconductors **30**, 983 (1996)].
[11] P. Morin and P. Roca i Cabarrocas, MRS Symp. Proc. **336**, 281 (1994).
[12] W. Paul, in *Amorphous Silicon and Related Materrials*, edited by H. Fritzsche, World Scientific, Singapore, 1988, Vol. 1A, p. 63.
[13] H. Wagner and W. Beyer, Solid State Commun. **48**, 585 (1983).
[14] Z. Igbal and S. Veprec, J. Phys. C: Sol. St. Phys. **15**, 377 (1982).
[15] K. Zellama, L. Chahed, P. Sladec, M. L. Theye, J. H. von Bardeleben, and P. Roca i Cabarrocas, Phys. Rev. B **53**, 3804 (1995).
[16] E. Sauvain, P. Pipoz, A. Shan, and J. Hubin, J. Appl. Phys. **75**, 1722 (1994).
[17] I. A. Kurova, N. N. Ormont, O. A. Golikova, and V. Kh. Kudoyarova, Fiz. Tekh. Poluprovodn. **31**, 536 (1997) [Semiconductors **31**, 452 (1997)].
[18] O. A. Golikova, Fiz. Tekh. Poluprovodn. **31**, 281 (1997) [Semiconductors **31**, 173 (1997)].
[19] P. Tzanetakis, N. Kopidakis, M. Androulidaki, C. Kalpouzos, P. Stradins, and H. Fritzsche, MRS Symp. Proc. **377**, 245 (1995).
[20] S. Bauer, T. Haage, B. Schroder, and H. Oechsner, in *Abstracts of ICAS-16*, Kobe, Japan, 1995, p. 176.

Translated by M. E. Alferieff

Study of optical absorption in thin films of a-As_2Se_3 by photocapacitance spectroscopy

I. A. Vasil'ev and S. D. Shutov

Institute of Applied Physics, Academy of Sciences of Moldova, 2028 Kishinev, Moldova
(Submitted March 6, 1997; accepted for publication October 2, 1997)
Fiz. Tekh. Poluprovodn. **32**, 490–493 (April 1998)

Photocapacitive spectroscopy is used for the first time to study optical absorption in thin films of a-As_2Se_3 for optical photons with energies below the gap width. Photocapacitance spectra are obtained over energies of 0.83–1.94 eV and used to characterize the density of localized states in the tail of the valence band and in deep states of charged defects. Aging and illumination of the film are found to affect the level of optical absorption by deep centers. © *1998 American Institute of Physics.* [S1063-7826(98)01904-8]

1. INTRODUCTION

Optical absorption in thermally deposited films of the chalcogenide glasses at photon energies $h\nu$ below the optical width of the band gap E_{g0} have not been studied adequately. At the same time, for energies $h\nu<E_{g0}$, optical absorption determines the optoelectronic properties of a film which depend on the degree of disorder, the amount of defects, and the presence of impurities. Optical transmission measurements provide information within a limited region near the absorption edge. At lower energies, photoelectron spectroscopy is usually employed, but it is not sensitive enough to reveal and study the structure of weak absorption spectra. To solve this problem it is appropriate to use photocapacitive spectroscopy, which has been used successfully in studies of thin film a-Si:H samples.[1]

Glassy As_2Se_3 is a well-known representative of the chalcogenide glasses and data for it reflect a contradictory picture of the spectrum of localized states involved in optical transitions. Thus, methods based on measurements of the transport of nonequilibrium charge carriers suggest a broad, quasicontinuous distribution of localized states, while methods which control optical transitions (photoexcitation, luminescence) suggest the existence of quasidiscrete energy levels associated with intrinsic defect or impurity centers. These contradictions are fundamental, so that further experimental data are needed in order to choose a model for the defect system.[2,3]

In this paper we use photocapacitive spectroscopy for the first time to study optical absorption at energies $h\nu<E_{g0}$ and determine the parameters of deep centers in a-As_2Se_3 films. Over a wide range of energies (0.83–1.94 eV, i.e., roughly from $E_{g0}/2$ to E_{g0}), we obtain a photocapacitance spectrum which, like the optical absorption spectrum, is determined by electronic transitions between localized states within the mobility gap and delocalized states in the conduction bands, and makes it possible to distinguish the range of energies associated with defect absorption. In the long wavelength portion of the spectrum, this method can be used to observe changes owing to structural relaxation and illumination of the film.

2. Experimental technique

Photocapacitance spectroscopy of deep states is based on measuring the relaxation of the barrier capacitance of a diode with time as a result of illumination by photons with energies $h\nu<E_{g0}$. The optical depopulation of filled states in the depleted region of the diode is accompanied by a rise in the barrier capacitance. In a linear approximation, in the initial stage of relaxation the photocapacitance signal varies with time as[4]

$$\Delta(C^2)=C_\phi^2-C_d^2=\alpha^c B\phi t. \qquad (1)$$

Here C_ϕ and C_d are the capacitance at time t under illumination and in the dark, ϕ is the flux of incident photons, and $B=\varepsilon\varepsilon_0A^2q/[2(V+V_{bi})]$ is a constant determined by the barrier parameters ($\varepsilon\varepsilon_0$ is the dielectric permittivity, A is the contact area, q is the electronic charge, V is the applied bias voltage, and V_{bi} is the band inflection). The proportionality constant α^c has the physical significance and units of an absorption coefficient. Its magnitude is determined by electronic optical transitions, which lead to a change in the charge within the depleted region of the diode, and it disregards the nonphotoactive optical losses.

The photocapacitance relaxation measurements were done at room temperature on Al/a-As_2Se_3 surface-barrier diodes made by successive vacuum thermal deposition on a glass substrate of a lower electrode (antimony), a 1.2-μm-thick As_2Se_3 film, and a semitransparent aluminum film with a contact area of 0.3 cm^2. Here the antimony created a resistive contact, while a potential barrier of height $\simeq 1.16$ eV was formed at the Al/a-As_2Se_3 interface.

The relaxation of the capacitance was measured in the dark and under illumination using a quasistatic $C-V$ method, described previously,[5] in which one detects the bias current that develops when triangular voltage pulses are applied to the reverse biased diode. The duration τ of the driver pulses was chosen subject to the condition $\tau>\tau_D$, where τ_D is the Maxwellian relaxation time ($\tau_D=2.4$ s for As_2Se_3 at 300 K). Before the start of the measurements the diodes were held in the dark with zero bias until a steady-state capacitance was reached. In order to avoid inducing metastable states, the measurements were made by scanning the spec-

 © 1998 American Institute of Physics

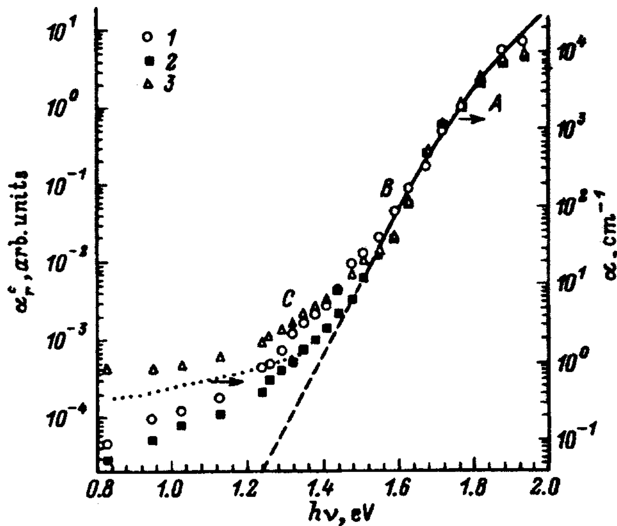


FIG. 1. Spectral variations in the absorption coefficient α_r^c for a fresh a-As_2Se_3 film (*1*), for a sample held more than one year under ordinary conditions (*2*), and for a sample initially illuminated by light that is strongly absorbed (*3*). Published data on the variation in the absorption α are shown by a smooth curve[6] and a dotted curve.[7] The dashed curve is an extrapolation of the exponential edge absorption tail. The values of α_r^c have been calculated from photocapacitance measurements with a scan period of 50 s and a reverse bias of 0.8 V. $T=300$ K.

trum from lower to higher photon energies. Optical excitation from an SPM-2 monochromator ($\phi=10^{14}-10^{16}$ cm^{-2}s^{-1}) was turned on simultaneously with a negative voltage. Values of the absorption coefficient α^c was calculated according to Eq. (1) from the initial rate of relaxation for a series of photon energies within the interval $h\nu=0.83-1.94$ eV and used to construct the resulting absorption spectrum $\alpha^c(h\nu)$. Because of certain differences in the actual values of ϕ (which are attributable primarily to loss of light at the aluminum electrode), the $\alpha^c(h\nu)$ spectra for different samples were displaced relative to one another along the α^c axis. Thus, in order to compare the resulting spectra, they were normalized to the value of the absorption coefficient α^c at $h\nu=1.76$ eV, which corresponds to the room temperature value of E_{g0}: $\alpha_r^c=\alpha^c(h\nu)/\alpha^c(1.76\text{ eV})$. At this energy the values of α^c for the different samples should not be different, since they are predominantly determined by interband optical transitions.

3. Experimental results and discussion

The spectral dependences of the absorption coefficient α_r^c measured by photocapacitive spectroscopy are plotted in Fig. 1 for a fresh film (curve *1*), for a sample kept for more than a year under ordinary conditions (curve *2*), and for the same sample initially irradiated by light that is strongly absorbed (intensity $\simeq 0.1$ mW/cm^2, $h\nu=1.8$ eV, 30 min, curve *3*). For comparison, the figure shows the following: the spectral variation in the absorption coefficient of a-As_2Se_3 films at energies of 1.6-1.8 eV obtained by optical transmission measurements[6] (the smooth curve in the figure) and the spectral variation in the induced absorption of a-As_2Se_3 films in the range 0.85-1.4 eV at $T=300$ K (dotted curve).[7]

The optical absorption coefficient α_r^c varies over a wide interval and corresponds roughly to optical transitions with energies from $E_{g0}/2$ to E_{g0} for a-As_2Se_3. Three characteristic regions, A, B, and C, can be seen in the spectra.[8] Region A corresponds to edge absorption and gives $E_{g0}=1.76$ eV. In region B, α_r^c varies exponentially with the photon energy $h\nu$ as $\alpha_r^c \propto \exp(h\nu/E_c)$, with a characteristic energy $E_c \simeq 0.047$ eV (the dashed curve in the figure). Clearly, within this region (also known as the Urbach edge), the spectral variations $\alpha_r^c(h\nu)$ agree well among themselves and with the data of Ref. 6. Absorption in this region is usually attributed to optical transitions between localized states in the tails of the conduction bands, whose densities fall off exponentially with energy. We may assume that in a-As_2Se_3 the main contribution to the Urbach edge is from the wider tail of the valence band, since its characteristic decay energy, as determined from measurements of transient[9] and modulated[10] light fluxes, is the same as the value of E_c found above.[6]

The data in region C of Fig. 1 are of greatest interest. Absorption in this region is caused by optical transitions involving deep localized states formed by intrinsic defects and impurities.[8] The absorption coefficient in region C is considerably higher than the level to be expected from an extrapolation of the exponential part B (dashed line), decreases as the film is aged (curve *2*), and increases under the influence of preliminary illumination (curve *3*), which may indicate the existence of metastable states.

Published data on the optical transition energies in this part of the spectrum and on the energies of the levels involved in these transitions are quite contradictory.[2] The optical transition at 1.4 eV appears as a shoulder on the photocurrent spectra of unannealed bulk samples, but is absent in films (even thick ones, up to 30 μm); absorption in the region of the shoulder increases when oxygen is present in the glass,[11] but it shifts to higher energies (from 1.4 to 1.6 eV) with deviations from stoichiometry toward higher arsenic contents (up to 60 at. %).[12] The spectral position of the shoulder at 1.4 eV is confirmed by spectra of the optical quenching for the induced absorption (threshold at 1.42 eV)[11] and of the photocurrent quantum efficiency.[13] Another optical transition shows up in the photocurrent spectra in the form of a broad band with a maximum at 1.15 eV;[14] the magnitude of the photocurrent near this peak changes during illumination-annealing cycles and this is evidence of the metastable character of the defects that participate in the absorption. The 1.15-eV level appears as a donor-type recombination center in the temperature dependence of the steady-state photoconductivity.[2] Both of these bands, at 1.4 and 1.1 eV, have been isolated using a high-sensitivity technique and graphical differentiation of photocurrent spectra of As_2Se_3 films.[15] In addition to these transitions, an optical transition has been noticed[2] at 1.6 eV as a weak feature in photocurrent spectra.

Figure 2 shows a plot in this energy range of the absorption coefficient α_{ir}^c owing to optical transitions involving deep levels. It was isolated by subtracting the almost constant absorption level below 1.13 eV and the extrapolated values in the exponential tail (dashed line) from the curves of

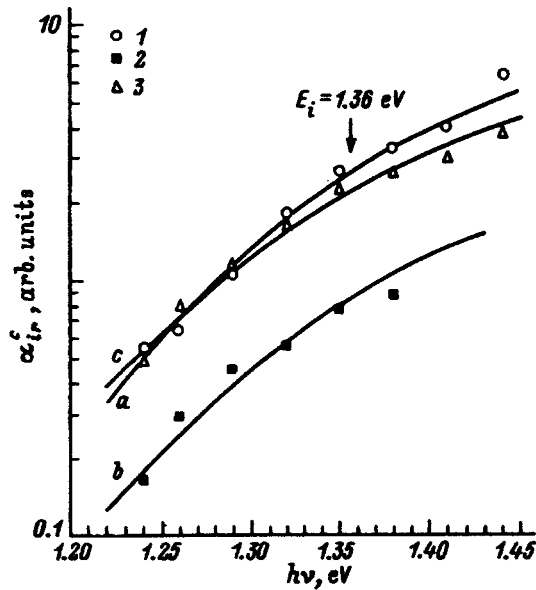


FIG. 2. Spectral variations in the absorption coefficient α^c_{ir} obtained by subtracting the constant absorption level below 1.13 eV and the extrapolated values in the exponential tail from the curves of Fig. 1. The smooth curves a, b, and c are calculated using Eq. (2) for $E_i = 1.36$ eV and Γ (eV) = 0.098 (a), 0.104 (b), and 0.11 (c).

Fig. 1. The smooth curves of Fig. 2 show model calculations of the absorption coefficient α^c_{ir} associated with optical transitions between levels of a defect with a density of states in the form of a Gaussian with half-width Γ for energy E_i and a parabolic edge of the conducting band:[16]

$$\alpha^c_{ir} \propto (h\nu)^{-1} \int_0^\infty \sqrt{E} \exp\left[-\frac{(E_1 + E - h\nu)^2}{2\Gamma^2}\right] dE. \tag{2}$$

The best agreement between the calculated curves and experiment is achieved for $E_i = 1.36$ eV, the same for all the curves, and $\Gamma \simeq 0.1$ eV. Thus, in the absorption spectrum obtained from the photocapacitance measurements on films, we can isolate an energy region corresponding to optical transitions between the conduction band and a 1.36-eV level, in agreement with the spectroscopic data described above. Since the depleted region of the test diodes is formed by a negative space charge, an increase in the capacitance can be related to optical excitation of holes from a defect level in the upper half of the band gap (D^+) into the valence band. The nature of the D^+ defect is related to the supercoordinated As (P_4^+) atom,[17] in accordance with the observed differences in the spectra: in samples subjected to prolonged aging, the defect concentration is lower owing to chemical ordering $As_4Se_4 \rightarrow As_2Se_3$, while in the samples that were initially illuminated, it is higher due to photostructural transformations.

Note that studies of the relaxation of the dark capacitance in Al/aAs$_2$Se$_3$ diodes,[5] as well as experiments on injection currents,[18] indicate the existence of deep levels with thermal ionization energies of $\simeq 1.0-1.05$ eV in the mobility gap of the semiconductor. It may be assumed, as in Ref. 15, that these are also D^+ defect levels, while the observed difference in the thermal and optical ionization energies ($\simeq 0.31$ eV) is caused by the strong electron-phonon interaction typical of the chalcogenide glasses.

A weak, almost constant absorption at low energies ($h\nu \leqslant 1.13$ eV) in the spectra of a-As$_2$Se$_3$ is usually observed at low temperatures ($\simeq 6$ K)[19] and (or) during intense optical illumination,[7] i.e., is associated with the excitation of metastable defect states. Since this kind of induced absorption is always accompanied by the appearance of an electron spin resonance (ESR) signal, the singly occupied neutral defect state D^0 is attributed to a metastable state.[17] The weak absorption spectra in Fig. 1 are consistent with this interpretation, as well as with the above analysis of absorption at charge defect states: the curve *3* obtained after illumination has a $\alpha^c_r(h\nu)$ dependence that is close to the data of Ref. 7 (dotted curve). At the same time, curves *1* and *2* (Fig. 1), which were obtained under near equilibrium conditions without preliminary irradiation of the sample, indicate the existence of a contribution to the absorption that is unrelated to metastable defects. The weak optical absorption near the middle of the mobility gap may be caused by excitation of thermally equilibrium carriers from states in the tails of the conduction bands. This sort of absorption has been analyzed[20] in connection with an effort to estimate the minimum optical losses in chalcogenide glasses at temperatures of 240–400 K. For the present case of an exponential energy spectrum for the traps, a calculation shows that the absorption coefficient emerges in a plateau for photon energies of roughly half the band gap width.

4. Conclusions

Photocapacitance spectroscopy can be used to study the optical absorption spectra of amorphous As_2Se_3 films at photon energies below the optical width of the band gap. The resulting photocapacitance spectra and their features are typical of the optical absorption spectra of the chalcogenide glasses that have become known as a result of a combination of other spectroscopic techniques (photocurrent, ESR, luminescence, etc.). Model calculations have been used to estimate the optical ionization energy $E_i = 1.36$ eV of positively charged D^+ defects. The difference between the energies of thermal[5] and optical ionization ($\simeq 0.31$ eV) presumably characterizes a lattice relaxation defect. It has been shown that optical absorption in the long-wavelength region of the spectrum depends on the history of the sample and this is indicative of a change in the density or degree of filling of deep localized states.

[1] J. D. Cohen, J. Non-Cryst. Solids **114**, 381 (1989).
[2] G. J. Adriaenssens, Philos. Mag. B **62**, 79 (1990).
[3] R. P. Barclay, J. Phys.: Condens. Matter **6**, 9721 (1994).
[4] N. M. Johnson and D. K. Biegelsen, Phys. Rev. B **31**, 4066 (1985).
[5] S. D. Shutov and A. A. Simashkevich, J. Non-Cryst. Solids **176**, 253 (1994).
[6] S. Hosokawa, Y. Sakaguchi, H. Hissa, and K. Tamura, J. Phys.: Condens. Matter **3**, 6673 (1991).
[7] D. A. Vasil'ev, B. T. Kolomiets, and S. K. Pavlov, Pis'ma Zh. Éksp. Teor. Fiz. **2**, 702 (1976) [*sic*].
[8] D. L. Wood and J. Tauc, Phys. Rev. B **5**, 3144 (1972).
[9] J. Orenstein, M. Kastner, and V. Vaninov, Philos. Mag. B **46**, 23 (1982).

[10] P. Brüggeman, C. Main, J. Berkin, and S. Reinolds, Philos. Mag. B **62**, 29 (1990).
[11] M. Hammam, G. J. Adriaenssens, J. Dauwen, G. Seynhaeve, and W. Grevendonk, J. Non-Cryst. Solids **119**, 89 (1990).
[12] M. Hammam, G. J. Adriaenssens, and W. Grevendonk, J. Phys. C: Sol. St. Phys. **18**, 2151 (1985).
[13] D. Monroe and M. Kastner, J. Non-Cryst. Solids **59/60**, 995 (1983).
[14] R. P. Barclay, M. Sarr, and J. L. Brebner, J. Non-Cryst. Solids **97/98**, 687 (1987).
[15] A. Yu. Popov, L. E. Stys, and L. V. Tsybeskov, Fiz. Tekh. Poluprovodn. **21**, 2225 (1987) [Sov. Phys. Semicond. **21**, 1349 (1987)].
[16] W. Grevendonk, M. Verluyten, J. Dauwen, G. J. Adriaenssens, and J. Bezemer, Philos. Mag. B **61**, 393 (1990).
[17] J. Robertson, Adv. Phys. **32**, 361 (1983).
[18] E. S. Vakarov, L. E. Stys, M. G. Foĭgel', and L. V. Tsybeskov, Fiz. Tekh. Poluprovodn. **19**, 931 (1985) [Sov. Phys. Semicond. **19**, 573 (1985)].
[19] S. G. Bishop, U. Strom, and P. C. Taylor, Phys. Rev. B **15**, 2278 (1977).
[20] V. I. Arkhipov and E. V. Emel'yanova, Vysokochist. Veshchestva, No. 6, 198 (1989).

Translated by D. H. McNeill

Effect of thermal annealing and chemical treatment on the photoluminescence of porous silicon

E. A. Shelonin, M. V. Naĭdenkova, A. M. Khort, A. G. Yakovenko, A. A. Gvelesiani, and I. E. Maronchuk

M. V. Lomonosov Moscow State Academy of Precision Chemical Technology, 117571 Moscow, Russia
(Submitted March 17, 1997; accepted for publication November 17, 1997)
Fiz. Tekh. Poluprovodn. **32**, 494–496 (April 1998)

Changes in the intensity of photoluminescence and in the IR absorption spectra of porous silicon samples are studied during chemical annealing and chemical treatment. It is found that processing porous silicon in HCl+Zn increases the photoluminescence intensity by more than a factor of 2 and affects the way the intensity is reduced by annealing. A comparison of the observed changes with IR spectra suggests that $Si-H_x$ bonds play a key role in efficient photoluminescence in porous silicon. © *1998 American Institute of Physics.* [S1063-7826(98)02004-3]

Many papers have been published thus far on photoluminescence in porous silicon.[1–3] The available experimental data, however, are extremely contradictory and lead to ambiguous interpretations of the mechanisms for radiative and nonradiative recombination in porous silicon.

Most of these studies indicate that photoluminescence in porous silicon is related to radiative recombination from quantum-well levels localized in silicon crystallites. As for the quenching (degradation) of photoluminescence in porous silicon, most researchers relate this process to the formation of deep nonradiative recombination centers on the surface of Si nanocrystals.

In this paper we report the results of an experimental study of the thermal degradation of photoluminescence in porous silicon. We also examine the possibility of passivating the nonradiative recombination centers by chemical processing of a porous silicon surface.

As a source material we used type KÉF-10 and KDB-10 silicon slabs with a (100) orientation. Anode etching was carried out in an $H_2O:HF:C_2H_5OH=1:2:2$ electrolyte at a current density of 5–10 mA/cm^2 for 10–20 min. After etching, the samples of porous silicon were cut into parts for taking the photoluminescence spectra and annealing. Prior to annealing, the samples were initially subjected to aging in air at room temperature for 2–3 days in order to eliminate fast processes. The photoluminescence was measured on an SDL-1 spectrometer with an LGI-21 nitrogen laser, operating in a quasicontinuous mode at a wavelength of 334 nm, as a source. Absorption spectra were taken in the infrared (IR) using an FTIR-1710 Fourier spectrometer. Isochronous annealing (20 min) was performed in quartz cells in a helium atmosphere. After each stage of annealing, photoluminescence spectra of the annealed and control samples were taken. The relative change in the photoluminescence intensity was determined from the changes in the intensity at the peak and in the area under the spectral curve. Both methods of measurement yielded the same relative intensity change.

The relative change in the photoluminescence intensity (I/I_0) of typical porous silicon samples obtained under identical etching conditions for p- and n-type substrates is plotted as a function of annealing temperature in Fig. 1. The change in the photoluminescence could be approximated as a linear function $I/I_0=f(T)$ with some deviation from linearity at high temperatures. The photoluminescence vanished almost completely at 400 °C. The photoluminescence fell off faster in the p-type samples than in the n-type. When an attempt was made to fit the intensity variation to an Arrhenius law dependence in $\ln(I/I_0)-(1/T)$ coordinates, the measurement results could not be approximated by a straight lines but could be fit roughly only by a more complex function of the form $y=a\cdot\ln(x)+b$. Thus, the effective activation energy of the process, calculated as the derivative of the approximation curve, varied with temperature from <0.1 eV to >1 eV. This indicated that, in our case, the process of photoluminescence degradation takes place in more than one stage. This is consistent with Ref. 3, in which the changes in photoluminescence due to laser irradiation during heating are attributed to several processes with different activation energies. Evidently, one can distinguish an initial stage, in which there is a sudden change in the photoluminescence, an intermediate stage, in which the photoluminescence varies roughly as $(I/I_0)=a-bT$, and a slow photoluminescence quenching stage at temperatures above 250-300 °C.

The wavelength of the photoluminescence maximum was essentially independent of the annealing temperature. In the IR spectra of the samples taken after annealing, the absorption bands, which appeared as a result of vibrations of $Si-H_x$ bonds,[1,2] vanished. However, the photoluminescence intensity as a function of temperature no longer had the ''step'' form observed in Ref. 1.

When samples with fully quenched photoluminescence were held at room temperature in air for several weeks, they recovered up to 45–50% of their original photoluminescence and the above-mentioned IR absorption bands were partially restored. Repeated annealing of the samples led to repeated loss of the photoluminescence in porous silicon.

The major obstacle to the use of porous silicon in opto-

 © 1998 American Institute of Physics

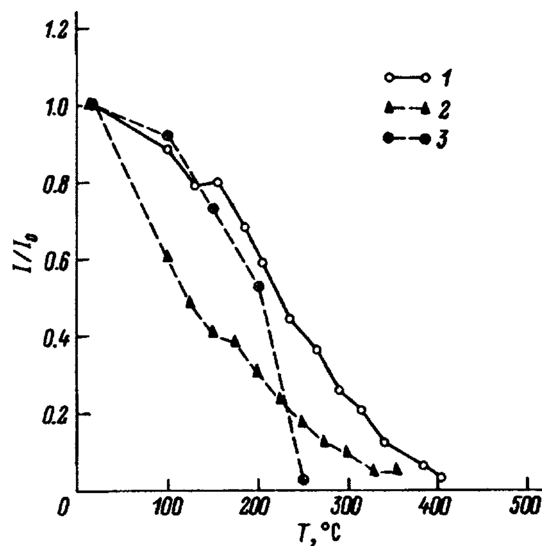


FIG. 1. Relative change in the photoluminescence intensity (I/I_0) for isochronous thermal annealing of porous silicon samples: (*1*) n-type sample, (*2-3*) p-type samples (sample 3 was processed in HCl+Zn).

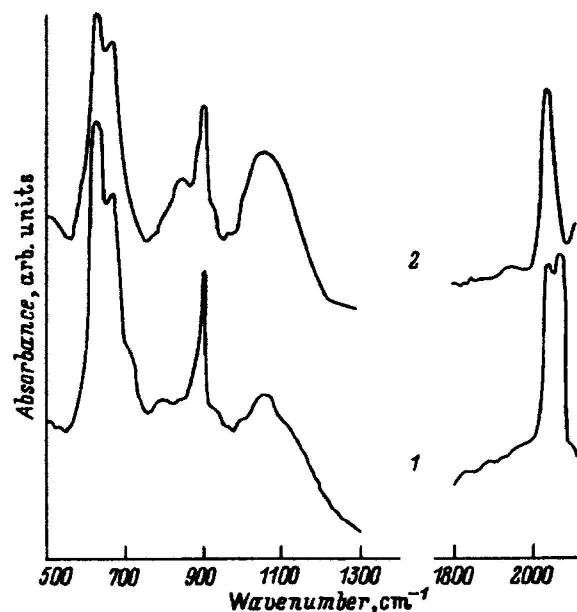


FIG. 2. The change in IR absorption spectra owing to processing in HCl+Zn: (*1*) before processing, (*2*) after processing.

electronics is the instability of its luminescence properties and the degradation of the luminescence as it ages in air or during thermal processing. The instability and degradation of the photoluminescence are mainly attributed to the formation of nonradiative recombination centers on the surface of the nanocrystallites. After anode etching, the porous silicon surface contains a large number of nonradiative recombination centers which are not passivated by the hydrogen and oxygen contained in the electrolyte; nevertheless, the hydrogen concentration in the voids is high. When the residue of electrolyte is removed from the porous silicon surface, passivation by oxygen takes place in the air as the silicon crystallites of the porous silicon come to be surrounded by an amorphous oxide layer.

We have tried to passivate the nonradiative recombination centers in porous silicon by processing the porous silicon in Zn+HCl. Passivation by the hydrogen released in the course of the reaction can be regarded as an alternative to the process of oxidation in air. The porous silicon samples were processed immediately after fabrication by keeping them in hydrochloric acid with zinc for 30 min. The porous silicon was obtained under the same conditions as those used for the studies of the effect of annealing on photoluminescence.

The initial photoluminescence intensity of the processed porous silicon samples was roughly twice that of the control samples. At the same time, the photoluminescence spectra of the porous silicon samples processed in HCl and $ZnCl_2$ were essentially the same as those of the unprocessed samples. Significant differences from the typical spectra of the unprocessed samples were observed in the IR spectra of the processed porous silicon samples (Fig. 2). In the 2120-cm^{-1} doublet corresponding to valence vibrations of SiH and SiH_2 bonds, the relationship of the bands changes as the contribution of the SiH_2 bonds varies. In addition, small absorption bands appear at 890 and 860 cm^{-1}, which may be associated with SiH and SiH_2. The relationship between the bands that form the doublet at 630–670 cm^{-1} also changes. This absorption, which is partially masked by an absorption band owing to Si-Si vibrations at 616 cm^{-1} (in this region, although the substrate absorption spectrum has been subtracted, there may be some enhancement in the measurement error) is also attributed to SiH and SiH_2 vibrations,[1] and, probably, in this case also there is a change in the relationship between the amounts of these bonds. It should also be noted that the magnitude of the absorption on asymmetric Si-O-Si valence vibrations observed at 1100 cm^{-1} was also different, even when the samples were obtained under the same conditions. Furthermore, in a number of cases we observed a region of very high absorption at wave numbers >1100 cm^{-1}, which may be caused by surface oxides. Here the absorption spectrum at wave numbers <1100 cm^{-1} did not change. In these samples the photoluminescence intensity was at roughly the same level as in the samples which did not have a region of this sort in the IR spectrum.

The reduction in the photoluminescence with isochronous annealing of the porous silicon samples processed in HCl+Zn was more rapid. In addition, after annealing at temperatures above 250 °C, there was almost no photoluminescence (Fig. 1, curve *3*); that is, in contrast with the unprocessed samples, we observed no slow reduction in the photoluminescence at temperatures >250 °C.

In summary, we may obviously conclude that the presence of SiH_x bonds on the surface plays an important, if not key, role in the quenching of photoluminescence during thermal processing and in the recovery of the photoluminescence intensity during chemical processing. These studies suggest that, at least in our case, the adsorption of chemical substances on the surface of porous silicon has a significant effect on photoluminescence in porous silicon. Chemical processing of porous silicon can be used to change its photoluminescence characteristics. Thus, the observed changes in the photoluminescence properties of porous silicon during

processing in HCl+Zn can be explained by passivation of nonradiative recombination channels.

[1] C. Tsai, K.-H. Li, J. Sarathy, S. Shih, and J. C. Campbell, Appl. Phys. Lett. **59**, 2814 (1991).
[2] S. M. Prokes, J. Appl. Phys. **73**, 407 (1993).
[3] V. M. Kostishko, A. M. Orlov, and T. G. Emel'yanov, Pis'ma Zh. Tekh. Fiz. **22**, 689 (1996) [Tech. Phys. Lett. **22**, 690 (1996)].

Translated by D. H. McNeill

PHYSICS OF SEMICONDUCTOR DEVICES

Subthreshold characteristics of electrostatically switched transistors and thyristors. II. Deep in-plane gate

A. S. Kyuregyan

V. I. Lenin All-Russian Electrotechnical Institute, 111250 Moscow, Russia
(Submitted April 15, 1997; accepted for publication, June 9, 1997)
Fiz. Tekh. Poluprovodn. **32**, 497–503 (April 1998)

A rigorous analytic theory is developed for the blocking state of electrostatically switched thyristors and transistors. The problems of the potential distribution in the base, the barrier height for electrons, and the current-voltage characteristic in the subthreshold region are solved in quadratures by conformal mapping for electrostatically switched transistors and thyristors with an arbitrary configuration of an in-plane gate in contact with the source. It is shown that the subthreshold current-voltage characteristics of electrostatically switched transistors and thyristors have a universal form independent of the gate configuration. As an example, a study is made of a quasielliptical in-plane gate that corresponds to the configurations of most real devices. Simple formulas are obtained in limiting cases for the blocking coefficient and the parameters of the current-voltage characteristic as functions of the geometric parameters of the electrostatically switched transistor or thyristor, the base doping, and the gate potential. © *1998 American Institute of Physics.* [S1063-7826(98)02104-8]

1. INTRODUCTION

In a previous paper,[1] the author and S. N. Yurkov constructed a rigorous analytic theory for the subthreshold characteristics of electrostatically switched transistors and thyristors (EST) with shallow in-plane source and gate. The major disadvantage of this structure is that a high blocking coefficient g can only be obtained when the source width $2l$ is unacceptably small. Thus, in practice[2] a gate is usually fabricated with a depth $H \geqslant l$ (See Fig. 1), which makes it relatively easy to obtain high g. For quantitative calculations of the parameters of these devices, however, it is evident that the model of EST in the form of a plane capacitor, employed in Ref. 1 and before,[3] is unsuitable.

The purpose of this paper is to generalize the results of Ref. 1 to the case of an in-plane or mesaplanar gate of arbitrary shape and depth. It will be shown that the main results of Ref. 1 remain valid, regardless of the shape of the gate, if the thickness of the space charge region w and the half-width l of the source are replaced in the final formulas by their effective values w' and l', which are calculated, in general, by a method which reduces to determining two parameters of a conformal mapping of the space charge region onto a strip.

2. GENERAL PROPERTIES OF THE CONFORMAL MAPPING OF THE SPACE CHARGE REGION ONTO A STRIP

In this paper we use the same assumptions as in Ref. 1, except for one: the shape of the gate is assumed for now to be arbitrary and, as a model for the space charge region, we use the ''ribbed'' translationally symmetric capacitor whose cross section is shown in Fig. 2 in the complex $z=x+iy$ plane. The plane segments of the lower electrode are sources and the ''ribs,'' which are described by a piecewise continuous periodic function $y=y_g(x)$, are gates, while the upper electrode is the boundary of the space charge region which is in contact with the drain. Let the function $T(z)$ execute a conformal mapping of the capacitor cell (region D) onto the upper half-plane of the $t=\theta+i\tau$ plane in a way such that the boundary of D transforms to the real axis $\tau=0$ and the points $z=\pm l$, $z=\pm L+iH$, and $z=\pm L+iw$ map to the points $t=\pm\beta$, $t=\pm y$, and $t=\pm 1$, respectively (Fig. 3). A gate of arbitrary shape can be approximated to a specified accuracy by a $2m$-polygon with vertices at the points $z=z_n=x_n+iy_g(x_n)$. Then, if the boundary $z=iw$ of the space charge region can be regarded as flat (this assumption is justified later), the region D will form a closed $(2m+6)$-gon that is symmetric with respect to the imaginary axis, while the inverse function $Z(t)$ of $T(z)$ can be written in the form of a Schwarz-Christoffel integral.[4] Since the arcs of the gates intersect the lines $y=0$ and $x=\pm L$ at right angles,

$$Z(t)=C\int_0^t \prod_{n=1}^{m} (t^2-\lambda_n^2)^{(\omega_n-1)} \frac{dt}{\sqrt{(t^2-1)(t^2-\gamma^2)(t^2-\beta^2)}}, \tag{1}$$

where $\pi\omega_n$ are the external angles of an m-gon (half gate), and the λ_n are real numbers, with

$$\sum_{n=1}^{m} (\omega_n-1)=\frac{1}{2}, \quad \beta<\lambda_n<\lambda_{n+1}<\gamma, \tag{2}$$

while for the matching of points specified above it is necessary, in particular, that

 © 1998 American Institute of Physics

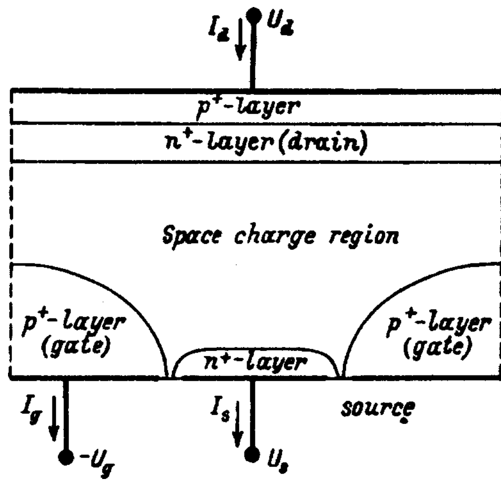


FIG. 1. A schematic illustration of the cross section of a transistor with electrostatic switching; in a transistor there is no p^+-layer above the drain.

$$l=C\int_0^{\beta}\prod_{n=1}^{m}(\lambda_n^2-\theta^2)^{(\omega_n-1)}\frac{d\theta}{\sqrt{(1-\theta^2)(\gamma^2-\theta^2)(\beta^2-\theta^2)}},\tag{3}$$

$$L=C\int_1^{\infty}\prod_{n=1}^{m}(\theta^2-\lambda_n^2)^{(\omega_n-1)}\frac{d\theta}{\sqrt{(\theta^2-1)(\theta^2-\gamma^2)(\theta^2-\beta^2)}},\tag{4}$$

and

$$w=H+C\int_{\gamma}^{1}\frac{\Pi(\theta^2-\lambda_n^2)^{(\omega_n-1)}d\theta}{\sqrt{(1-\theta^2)(\theta^2-\gamma^2)(\theta^2-\beta^2)}}$$
$$=C\int_0^{\infty}\frac{\Pi(\tau^2+\lambda_n^2)^{(\omega_n-1)}d\theta}{\sqrt{(\tau^2+1)(\tau^2+\gamma^2)(\tau^2+\beta^2)}}.\tag{5}$$

In turn, the half-plane $\tau\geqslant 0$ is mapped into the rectangle $|x'|\leqslant L$, $0\leqslant y'\leqslant w'$ in the $z'=x'+iy'$ plane (Fig. 4) by the function

$$Z'(t)=C'\int_0^{t}\frac{dt}{\sqrt{(t^2-1)(t^2-\gamma^2)}},\tag{6}$$

where a correspondence between the points $t=\pm\gamma$ and $t=\pm 1$ and the points $z'=\pm L$ and $z'=\pm L+iw'$ is reached for

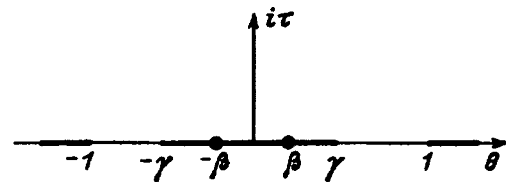


FIG. 3. A stage of an electrostatically switched transistor in the $t=\theta+i\tau=T(z)$ plane.

$$L=C'\int_1^{\infty}\frac{dt}{\sqrt{(\theta^2-1)(\theta^2-\gamma^2)}},\tag{7}$$

and

$$w'=C'\int_{\gamma}^{1}\frac{d\theta}{\sqrt{(1-\theta^2)(\theta^2-\gamma^2)}}$$
$$=C'\int_0^{\infty}\frac{d\tau}{\sqrt{(\tau^2+1)(\tau^2+\gamma^2)}}.\tag{8}$$

Because of the symmetry principle,[4] the function $Z'[T(z)]$ maps the cross section of the infinite "ribbed" capacitor onto the strip $0\leqslant y'\leqslant w'$ in the z' plane, while the sources, gates, and sink transform to the segments $[2mL-l'<x'<2mL+l', y=0]$ and $[2mL+l'<x'<2m(L+1)-l', y'=0]$ and the straight line $y'=w'$, respectively, where $m=0,\pm 1,\pm 2,\ldots,$

$$l'=\frac{L}{K(\gamma)}F(\beta/\gamma,\gamma),\tag{9}$$

and $F(x,k)$ is the elliptic integral of the first kind in the normal Legendre form. The function $Z'[T(z)]$ defined in this way has several important general properties. First of all, we always have

$$0\leqslant\Delta=w-w'\leqslant H.\tag{10}$$

The validity of the first inequality can be confirmed simply by substituting the second equalities from Eqs. (5) and (8) in Eq. (10) and noting that Eqs. (4) and (7) imply that $C\geqslant C'$. To prove the second inequality, we note [see Eqs. (4) and (7) and the first equalities from Eqs. (5) and (8)] that

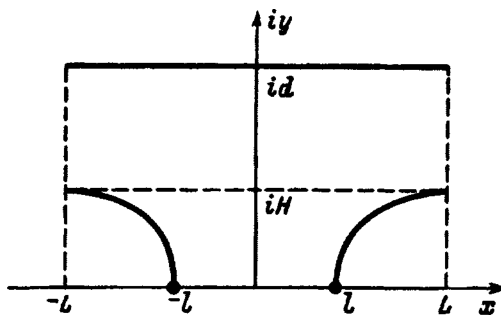


FIG. 2. The model of a stage of an electrostatically switched transistor used to calculate the potential distribution.

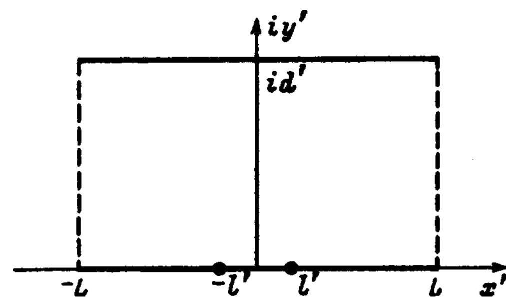


FIG. 4. A stage of an electrostatically switched transistor in the $z'=x'+iy'=Z'(t)$ plane.

$$H-\Delta \propto \int_\gamma^1 \frac{d\theta}{\sqrt{(1-\theta^2)(\theta^2-\gamma^2)}} \times \int_1^\infty \frac{\Pi(\theta^2-\lambda_n^2)^{(\omega_n-1)}d\theta}{\sqrt{(\theta^2-1)(\theta^2-\gamma^2)(\theta^2-\beta^2)}} - \int_\gamma^1 \frac{\Pi(\theta^2-\lambda_n^2)^{(\omega_n-1)}d\theta}{\sqrt{(1-\theta^2)(\theta^2-\gamma^2)(\theta^2-\beta^2)}} \times \int_1^\infty \frac{d\theta}{\sqrt{(\theta^2-1)(\theta^2-\gamma^2)}}. \quad (11)$$

If we replace all the λ_n by β on the right hand side of Eq. (11), then it becomes equal to zero, but then the first term increases by less than a factor of Q, while the second term increases by more than a factor of Q, where $Q=\sqrt{(1-\beta^2)}\Pi_{n=1}^m(1-\lambda_n^2)^{(1-\omega_n)}>1$. Thus, before the λ_n were replaced by β, the first term in Eq. (11) was greater than the second, which demonstrates the validity of Eq. (10). Secondly, Eqs. (7) and (8) imply that

$$w'=LK(\sqrt{1-\gamma^2})/K(\gamma),$$

where $K(\gamma)$ is the complete elliptic integral of the first kind. Using the known series expansions of $K(\gamma)$ in terms of γ, it is easy to show that, for $\gamma \ll 1$,

$$\gamma=4\exp\left(-\frac{\pi w'}{2L}\right). \quad (12)$$

The error in this formula is on the order of $\gamma^2/2$ and is less than 1.5% provided [see Eq. (10)]

$$w-H \geqslant 2L. \quad (13)$$

In the following we shall assume that Eq. (13) is valid everywhere, since it is usually satisfied with a large reserve and makes the calculations much simpler. In particular, Eqs. (4) and (7) yield $C=C'=2L/\pi$, Eq. (9) simplifies to

$$l'=\frac{2}{\pi}L \arcsin\frac{\beta}{\gamma}, \quad (14)$$

and Eqs. (1) and (6) imply that

$$1-\left.\frac{dy}{dy'}\right|_{y'=w'} = 1-\left.\frac{dz}{dz'}\right|_{z'=iw'} = 1-\frac{C}{C'}\left(1-\frac{\beta^2}{\theta^2}\right)^{-1/2} \times \prod_{n=1}^m \left(1-\frac{\lambda^2}{\theta^2}\right)^{(\omega_n-1)} < \frac{\gamma^2-\beta^2}{2} \ll 1.$$

The last, strong inequality means that, near the drain, the function $Z'[T(z)]$ is essentially the same as a shift by Δ along the imaginary axis and, therefore, a gate of arbitrary shape distorts the field just as weakly under condition (13) as a shallow in-plane gate with $w \geqslant 2L$.[1] Thus, the boundary of the space charge region can be regarded as flat with an accuracy of order $\gamma^2/2$, even if it does not fill the entire base region. Furthermore, the magnitude of Δ, as such, is determined only by the gate configuration and is essentially independent of the thickness of the space charge layer. In fact, using the second equalities of Eqs. (5) and (8), it is easy to confirm that

$$\Delta=\frac{2}{\pi}L\int_0^\infty \frac{\Pi(\theta_1^2+\lambda_n^2/\gamma^2)^{(\omega_n-1)}-\sqrt{\theta_1^2+\beta^2/\gamma^2}}{\sqrt{(\theta_1^2+1)(\theta_1^2+\beta^2/\gamma^2)}}\,d\theta_1 + 0\left(\frac{\gamma^2-\beta^2}{4}\ln \gamma_1^{-2}\right), \quad (15)$$

where γ_1 is an arbitrary quantity which satisfies the inequality $\gamma \ll \gamma_1 \ll 1$; for example, we can set $\gamma_1=\sqrt{\gamma}$. Formulas with similar structures can be obtained from the $(2m+1)$ conditions for matching the points $z=l$, $z=z_n$ to the points $t=\beta$, $t=\lambda_n$. For example,

$$l=\frac{2}{\pi}L\int_0^{\beta/\gamma} \frac{\Pi(\lambda_n^2/\gamma^2-\theta_1^2)^{(\omega_n-1)}d\theta_1}{\sqrt{(1-\theta_1^2)(\beta^2/\gamma^2-\theta_1^2)}} + 0(\beta^2). \quad (16)$$

The integrals in Eq. (16) and in the remaining $2m$ equations do not depend on the absolute values of β, λ_n, and γ, but only on the ratios β/γ and λ_n/γ, which to within the exponentially small final terms are determined only by the gate parameters l, L, H, and z_n, but do not depend on w. Thus [see Eq. (15)], with a slightly greater, but still exponentially small error, we can assume that Δ is also independent of w. On the other hand,

$$b=\left.\frac{dy}{dy'}\right|_{y'=x'=0} = \frac{1}{\beta}\prod_{n=1}^m \lambda^{2(\omega_n-1)}, \quad (17)$$

from which we find $1<b<\gamma/\beta$; that is, near the source the shape of the gate can have a strong influence on the field distribution for $\gamma \gg \beta$.

3. POTENTIAL DISTRIBUTION AND BARRIER PARAMETERS IN THE GENERAL CASE

Assuming that the base is uniformly doped (see Ref. 1), we shall seek the potential distribution $\varphi(x,y)$ in the form

$$\varphi(x,y)=\varphi_0(x,y)-V_w y^2/w^2, \quad (18)$$

where $\varphi_0(x,y)$ is the solution of the Laplace equation in a space charge region with the boundary conditions

$$\varphi_0(x,0)=0 \quad \text{for } 2mL-l<x<2mL+l, \quad (19)$$

$$\varphi_0(x,y)=-U_g+V_H y_g^2(x)/H^2$$

$$\text{for } 2mL+l<x<2(m+1)L-l \quad \text{and } y=y_g^2(x), \quad (20)$$

and

$$\varphi_0(x,w)=U_d+V_w, \quad (21)$$

$$V_w=qNw^2/2\varepsilon\varepsilon_0, \quad V_H=qNH^2/2\varepsilon\varepsilon_0,$$

while remaining notation is defined in Ref. 1. Calculating and analyzing the potential distribution as a function of the variables x' and y', which differs from that given in Ref. 1 only in some minor details related to the appearance of a second term on the right-hand side of Eq. (20), we obtain the following results. If the drain potential

$$U_d > U_{d\,\min} = V_d\left(1-2\frac{\Delta}{d}\right) - U_g\left(1-\frac{l'}{L}\right),$$

then the space charge region fills the entire base and $w=d$, $w'=d'$. For $U_d<U_{d\,\min}$, the thickness of the space charge layer is less than d, but the real boundary of the space charge region on the drain side deviates from the plane

$$y=w=\Delta+\sqrt{\Delta^2+\frac{2\varepsilon\varepsilon_0}{qN_D}\left[U_d+U_g\left(1-\frac{l'}{L}\right)\right]} \tag{22}$$

by a negligible amount on the order of $w\gamma^2/2$, while condition Eq. (13) is satisfied for

$$U_d > U_{w\,\min} = 4V_L - V_\Delta - U_g\left(1-\frac{l'}{L}\right),$$

where $V_\Delta = qN_D\Delta^2/2\varepsilon\varepsilon_0$, and $V_L = qN_DL^2/2\varepsilon\varepsilon_0$. The potential distribution along the channel axis has the form

$$\varphi(0,y') = U_g\left\{M\frac{y'}{L} - \frac{2}{\pi}\arctan\left[M_0\tanh\left(\frac{\pi y'}{2L}\right)\right]\right\} - V_H\left\{\frac{y^2(0,y')}{H^2} - \Phi_0(M_0)\frac{y'}{L} - \int_{l'}^{L}\frac{y_g^2(x)\sinh(\pi y'/L)}{\cosh(\pi y'/L)-\cos(\pi x'/L)}\frac{dx'}{LH^2}\right\}, \tag{23}$$

where

$$M = \left[\frac{U_d+V_w}{U_g} - \frac{U_H}{U_g}\Psi(M_0)\right]\frac{L}{w-\Delta} + \frac{L-l'}{w-\Delta} + \frac{U_H}{U_g}\Phi_0(M_0), \quad M_0 = \cot\left(\frac{\pi l'}{2L}\right),$$

$$\Phi_k(a) = f^{-2}(\sqrt{1+a^2})\int_0^a \chi^k f^2\left(\sqrt{\frac{1+a^2}{1+\chi^2}}\right)d\chi,$$

$$\Psi(a) = \frac{2}{\pi}f^{-2}(\sqrt{1+a^2})\int_0^a f^2\left(\sqrt{\frac{1+a^2}{1+\chi^2}}\right)\frac{d\chi}{1+\chi^2},$$

$$f(\theta/\beta) = \mathrm{Im}[Z(\theta)]f(\gamma/\beta)/H.$$

In an electrostatically switched transistor or thyristor with an undoped base, the potential at the saddle point located at

$$y' = \tilde{y}' = \frac{2}{\pi}L\arctan\sqrt{\frac{1-M/M_0}{1+MM_0}} \tag{24}$$

is given by

$$\tilde{\varphi} = \frac{2}{\pi}U_g\left[M\arctan\sqrt{\frac{1-M/M_0}{1+MM_0}} - \arctan\left(M_0\sqrt{\frac{1-M/M_0}{1+MM_0}}\right)\right], \tag{25}$$

where $M=(U_dL+U_g(L-l')/U_g(d-\Delta))$. Near the threshold, when

$$\sqrt{1-\frac{M}{M_0}} \ll \sqrt{1+M_0^{-2}}, \tag{26}$$

we obtain the following expression from Eq. (25):

$$\tilde{\varphi} = -\frac{4U_g}{3\pi}\sqrt{\frac{M_0^2}{M_0^2+1}}\left(1-\frac{M}{M_0}\right)^{3/2}. \tag{27}$$

The blocking coefficient is given by

$$g_0 = \frac{d-\Delta}{L}M_0 - 1 + \frac{l'}{L}, \tag{28}$$

and the parameter $\xi=1^{1,S}$. If $N\neq 0$, then the coordinate $\tilde{y}'$ of the saddle point can be calculated analytically only near the threshold, when $M_0\tilde{y}'\ll L$ and the right-hand side of Eq. (24) can be expanded in a series in y' limited to terms proportional to $(y')^3$. As a result, given Eq. (17), we obtain

$$\tilde{y}' = \frac{1}{2}\left(y_0' + \sqrt{(y_0')^2 + 2y_0'L\frac{U_g}{V_L}\frac{M_0-M}{b^2}}\right), \tag{29}$$

$$\tilde{\varphi} = \varphi_0\left(\frac{\tilde{y}'}{y_0'}\right)^2\left(4\frac{\tilde{y}'}{y_0'} - 3\right), \tag{30}$$

where

$$\varphi_0 = -V_Lb^2(y_0')^2/3L^2,$$

$$y_0' = \frac{8}{\pi^2}b^2L\frac{V_L}{U_g}\times\left\{M_0(1+M_0^2) - \frac{V_H}{U_g}[\Phi_0(M_0)+3\Phi_2(M_0)]\right\}^{-1}. \tag{31}$$

If $M_0|y_0'|\ll L$, then along with the condition $M_0\tilde{y}'\ll L$ for applicability of Eqs. (29) and (30), the following inequality will also be satisfied:

$$\sqrt{|y_0'|} \ll \sqrt{2L\frac{U_g}{V_L}\frac{M_0-M}{b^2}}. \tag{32}$$

In this case Eq. (30) yields a formula for $\tilde{\varphi}$ that differs from Eq. (27) only in having an additional factor $\{1-V_H[\Phi_0(M_0)+3\Phi_2(M_0)]/U_gM_0(1+M_0^2)\}^{-1/2}$ on the right. It follows from Eq. (30) that, as in the case of a shallow gate,[1] the barrier vanishes as $M\to M_0$ only in electrostatically switched transistors or thyristors with an acceptor doped base ($N=N_A\leqslant 0$), for which the blocking coefficient is given by

$$g = g_0 - \frac{V_d}{U_g}\left\{1+\frac{H^2}{d^2}\left[\frac{d-\Delta}{L}\Phi_0(M_0) - \Psi(M_0)\right]\right\}. \tag{33}$$

If, on the other hand, $N=N_D\geqslant 0$, then as $M\to M_0$, the barrier does not vanish, since $\tilde{y}'\to y_0'$ and $\tilde{\varphi}\to\varphi_0$. When $M>M_0$, a potential well, which fills up with electrons, develops near the source, and our approximation of a completely depleted base is no longer valid. As shown in Ref. 1, however, Eq. (33) provides entirely adequate accuracy for

practical purposes, even in the case $N=N_D\geqslant 0$. In particular, for $U_d<U_{d\min}$, on replacing d and d' by w and w' in Eq. (33) and using Eq. (22), we obtain

$$g=\frac{U_g}{4V_L}\left[M_0-\frac{V_H}{U_g}\Phi_0(M_0)\right]^2-\frac{\Delta}{L}\left[M_0-\frac{V_H}{U_g}\Phi_0(M_0)\right]-1+\frac{l'}{L}+\frac{V_H}{U_g}\Psi(M_0). \tag{34}$$

Finally, for $N\neq 0$, the parameter

$$\xi=\sqrt{1-\frac{y_0'}{2\bar{y}'}}. \tag{35}$$

Therefore, arbitrary distortions in the shape of the gate do not lead to qualitative changes in the subthreshold characteristics; the form of the current-voltage characteristics obtained in Ref. 1 for a shallow gate has turned out to be universal. The most important quantitative changes involve the replacement of w, l by w', l' and the appearance of a factor b^2 when $N\neq 0$. As will be shown in the next section, the remaining differences in the formulas for y_0' and g from the analogous formulas from Ref. 1 are negligible in the cases of greatest interest, when $l'\ll L$ and the blocking coefficient is large.

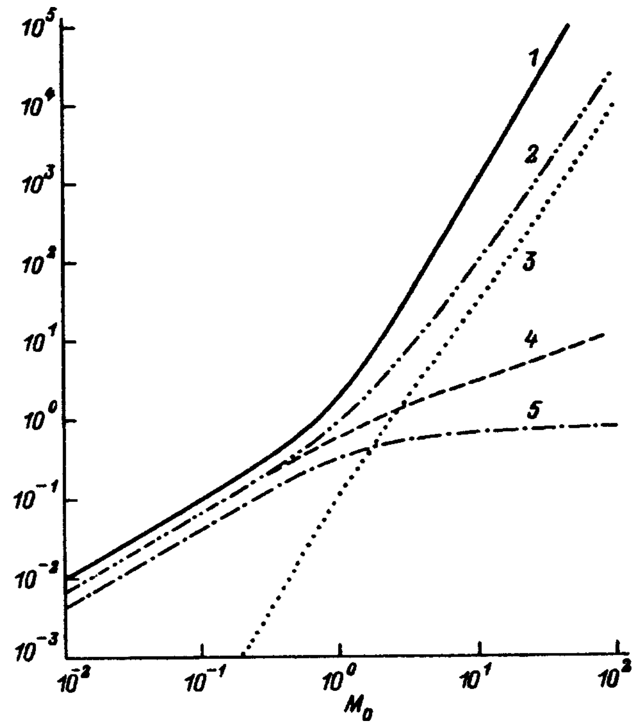


FIG. 5. Calculated variations in the functions (*1*) $M_0(M_0^2+1)$, (*2*) $\Phi_0(M_0)+3\Phi_3(M_0)$, (*3*) $\Phi_2(M_0)$, (*4*) $\Phi_0(M_0)$, (*5*) $\Psi_0(M_0)$ for an oval gate.

4. OVAL-SHAPED IN-PLANE GATE

As a specific example of applying the general theory, here we shall consider the simple, but typical and rather general, case where

$$Z(t)=\int_0^t\left\{\frac{L-l}{\sqrt{t^2-\gamma^2}\cos^{-1}(\beta/\gamma)}+\frac{H}{\sqrt{t^2-\beta^2}\cosh^{-1}(\gamma/\beta)}\right\}\frac{dt}{\sqrt{(t^2-1)}}, \tag{36}$$

where β and γ are solutions of the system of equations

$$\frac{L-l}{\cos^{-1}(\beta/\gamma)}+\frac{H}{\cosh^{-1}(\gamma/\beta)}=\frac{2}{\pi}L, \tag{37}$$

$$\frac{\ln(4/\gamma)}{\cos^{-1}(\beta/\gamma)}+\frac{\ln(4/\beta)}{\cosh^{-1}(\gamma/\beta)}=\frac{w}{H}. \tag{38}$$

The function (36) was obtained as a result of modifying the integral (1) for the values $m=1$ and $\omega_1=3/2$, which correspond to a rectangular gate, by rounding the obtuse angles.[4] The cross section of the space charge region, onto which this function maps the upper half-plane, is illustrated schematically in Fig. 2. The gate has the form of an oval with semiaxes H and $(L-l)$, which is close to an ellipse for $\beta>0.1\gamma$, and is described by the parametric equations

$$x(\theta)=L-(L-l)\frac{\cos^{-1}(\theta/\gamma)}{\cos^{-1}(\beta/\gamma)}, \tag{39}$$

and

$$y(\theta)=H\frac{\cosh^{-1}(\theta/\beta)}{\cosh^{-1}(\gamma/\beta)}. \tag{40}$$

Equation (40) yields the formula $f(\theta/\beta)=\cosh^{-1}(\theta/\beta)$, which makes it possible to calculate the functions $\Phi_k(M_0)$ and $\Psi(M_0)$ by numerical integration, where results are shown in Fig. 5. Equations (12), (14), and (36)–(38) yield the equation

$$\frac{L-l}{H}=\left(\frac{2L}{\pi H}-\frac{1}{\cosh^{-1}M_0}\right)\tan^{-1}M_0, \tag{41}$$

of which some numerical solutions are shown in Fig. 6, as well as the formulas

$$\Delta=H\frac{\ln(M_0^2+1)}{2\sinh^{-1}M_0}, \tag{42}$$

and

$$b=1+\frac{\pi H}{2L}\frac{\sqrt{M_0^2+1}-1}{\sinh^{-1}M_0}, \tag{43}$$

which determine the quantities M_0, Δ, and b to be substituted in the general formulas of the preceding section.

The material presented above can be used to calculate all the subthreshold characteristics of electrostatically switched transistors and thyristors for arbitrary relationships between the source and gate dimensions. The results of Ref. 1 for a shallow gate are obtained when $H\ll L$. The variant of a slit gate, formed by diffusion into a narrow, deep slot, is described by the formulas

$$M_0\approx\sinh\left(\frac{\pi H}{2L}\right),\quad \Delta\approx\frac{2}{\pi}L\ln\left[\cosh\left(\frac{\pi H}{2L}\right)\right],$$

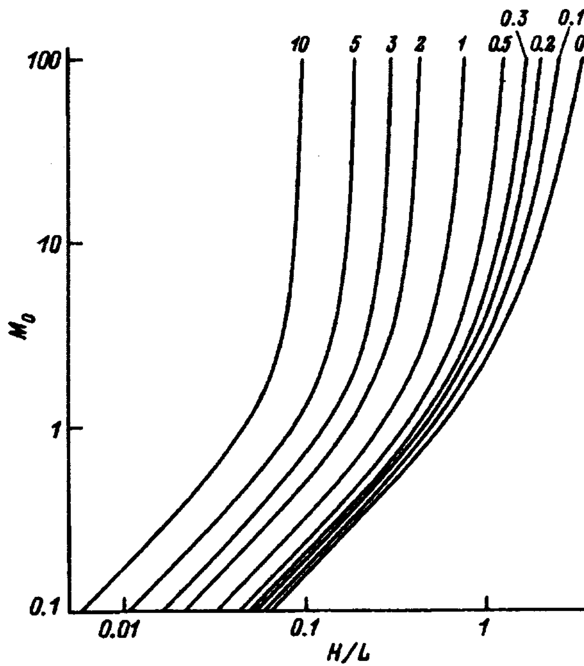


FIG. 6. Calculated dependence of M_0 on the ratio H/L for an oval gate. The numbers on the curves denote values of the parameter $(L-l)/H$.

$$b \approx \cosh\left(\frac{\pi H}{2L}\right), \tag{44}$$

which hold for $(L-l) \ll L$.

The case of greatest practical importance, with large blocking coefficients, is realized for $(L-l) \ll l \exp(\pi H/2l)$, when

$$M_0 \approx \frac{1}{2} \exp\left[\frac{\pi H}{2l}\right], \quad \Delta \approx H - l^2 \frac{2}{\pi} \ln 2, \quad b \approx M_0, \tag{45}$$

with $M_0 \gg 1$, so that, as can be seen from Fig. 6, $\Psi(M_0) \ll \Phi_0(M_0) \ll M_0$ and $\Phi_0(M_0) + 3\Phi_2(M_0) \ll M_0(1+M_0^2)$. With these inequalities it is possible to write down much simpler formulas for the parameters of electrostatically switched transistors or thyristors with doped bases and for the conditions under which they apply. First, given that usually $U_g \gg V_H$, we can use simplified expressions for M and y_0':

$$M \approx \frac{U_d + V_w}{U_g} \frac{L}{w - \Delta}, \quad y_0' \approx \frac{8V_L}{\pi^2 U_g} \frac{L}{M_0}.$$

Secondly, Eqs. (29) and (30), are valid for electrostatically switched transistors or thyristors with a p-type base when $\sqrt{1 - M/M_0} \ll 1$, while electrostatically switched transistors or thyristors with an n-type base require additionally that $U_g \gg V_L$. If $V_L/U_g \ll \sqrt{1 - M/M_0} \ll 1$, then

$$\tilde{\varphi} \approx -\frac{4U_g}{3\pi}\left(1 - \frac{M}{M_0}\right)^{3/2}. \tag{46}$$

Finally, the blocking coefficient is given by

$$g \approx \frac{d-\Delta}{L} M_0 - \frac{V_d}{U_g} \tag{47}$$

for $U_d > U_{d\,\min}$ and

$$g \approx \frac{U_g}{V_L} \frac{M_0^2}{4}. \tag{48}$$

for $U_{w\,\min} < U_d < U_{d\,\min}$.

The author thanks T. T. Mnatsakanov and S. N. Yurkov for numerous useful discussions of the questions touched on in this paper.

This work was supported by the Russian Fund for Fundamental Research (Project No. 95-02-05767) and the INTAS fund (Project No. 94-0417).

[1] A. S. Kyuregyan and S. N. Yurkov, Fiz. Tekh. Poluprovodn. **32**, 249 (1998) [Semiconductors **32**, 225 (1998)].
[2] B. J. Baliga, *Modern Power Devices*, Singapore (1987), p. 132.
[3] A. V. Gorbatyuk and I. V. Grekhov, Fiz. Tekh. Poluprovodn. **15**, 1353 (1981) [Sov. Phys. Semicond. **15**, 781 (1981)].
[4] M. A. Lavrent'ev and B. V. Shabat, *Methods in the Theory of Functions of a Complex Variable* [in Russian], Nauka, Moscow (1987).
[5] A. V. Gorbatyuk and A. S. Kyuregyan, Mikroélektronika **20**, 254 (1991).

Translated by D. H. McNeill

Amplification of electromagnetic waves at odd harmonics of the cyclotron resonance of heavy holes with negative effective masses in semiconducting diamond

V. A. Chuenkov

P. N. Lebedev Physics Institute, Russian Academy of Sciences, 117924 Moscow, Russia

(Submitted April 11, 1997; accepted for publication November 12, 1997)
Fiz. Tekh. Poluprovodn. **32**, 504–508 (April 1998)

It is shown theoretically that the absorption coefficient for circularly polarized electromagnetic waves at the cyclotron resonance of heavy holes with negative effective masses in diamond in parallel electric and magnetic fields oriented along the [001] crystal axis takes negative values at the frequency of any of the $n+1$ harmonics ($n=0,4,8$, etc.) for the right (electron) polarization and at the frequency of any of the $n-1$ harmonics ($n=4,8,12$, etc.) for the left (hole) polarization. In an electric field $E\approx 10^4$ V/cm and magnetic fields $H=30-80$ kOe, at lattice temperatures of 77–100 K, and for a hole concentration of $(3-5)\times 10^{15}$ cm^{-3}, the absorption coefficient for an electromagnetic wave at the third harmonic $\omega_3=3\omega=2.5\times 10^{12}$ s^{-1} (wavelength $\lambda_3=0.92$ mm) is as high as $\eta_3=(-7)-(-30)$ cm^{-1}. © *1998 American Institute of Physics.* [S1063-7826(98)02204-2]

The idea of resonant amplification of electromagnetic waves in p-type germanium at holes with negative effective masses rotating in a magnetic field $\mathbf{H}\|\mathbf{E}\|\mathbf{k}\|[\mathbf{001}]$ (where $\mathbf{E}$ is a constant electric field and $\mathbf{k}$ is the wave vector of a circularly polarized electromagnetic wave) was proposed by Dousmanis.[1] A correct theoretical justification of this idea has been provided by Andronov *et al.*[2] The experimental generation of stimulated electromagnetic radiation with a wavelength $\lambda=2$ mm in relatively pure Ge (total hole concentration $P_0\leqslant 10^{14}$ cm^{-3}) at liquid-helium temperatures with the above-mentioned orientation of the electric and magnetic fields has been observed by Andronov *et al.*[3]

There is theoretical and practical interest in the problem of generating electromagnetic waves based on this principle at higher temperatures and at shorter wavelengths, especially using other materials. One of the most promising materials in this regard is p-type semiconducting diamond. The acceptor in p-type semiconducting diamond is the boron atom (ionization energy 0.37 eV).

Solving this problem reduces to finding the dispersion relation and distribution function of the heavy holes in diamond for the above orientation of the electric and magnetic fields and calculating the high-frequency differential conductivity σ_ω along the electric field of an electromagnetic wave.

The dependence of the heavy hole energy ε on the momentum $\mathbf{p}$ (dispersion relation) which we have calculated using Kane's model,[4] has a complicated, highly anisotropic form. A detailed analysis showed, however, that the dynamics of heavy holes in diamond, both those with negative and those with positive effective cyclotron masses, can be described reasonably correctly by using the following comparatively simple dispersion relation:

$$\varepsilon(\mathbf{p})=\frac{1}{2m}\left[p^2-\xi(\mathbf{p})(p_\perp^2p_z^2+p_\perp^4\sin^2\varphi\cos^2\varphi)^{1/2}\right],$$

$$\xi(\mathbf{p})=\alpha_0+\frac{\beta_0}{p^4}(p_\perp^2p_z^2+p_\perp^4\sin^2\varphi\cos^2\varphi)-\frac{\gamma_0}{p^6}p_\perp^4p_z^2\sin^2\varphi\cos^2\varphi, \tag{1}$$

$$\alpha_0=1.40,\quad \beta_0=1.24,\quad \gamma_0=18.48.$$

Here $p^2=p_\perp^2+p_z^2$; $p_\perp$ and p_z are the components of the hole momentum in the direction of the [001] axis and in a direction perpendicular to this axis, respectively; φ is the azimuthal angle (the polar axis is taken to coincide with the [001] axis); $m=0.28m_0$, where m_0 is the free electron mass; and α_0, β_0, and γ_0 are dimensionless parameters. Expanding $\varepsilon(\mathbf{p})$ in a Fourier series in φ, we obtain

$$\varepsilon(\mathbf{p})=\sum_n \varepsilon_n(p_\perp,p_z)\exp(in\varphi), \tag{2}$$

where the expansion coefficients $\varepsilon_n(p_\perp,p_z)$ are real; $n=0$, $\pm 4,\pm 8,\ldots$; and i is the square root of -1. The expansion coefficients $\varepsilon_n(p_\perp,p_z)$ become smaller as n increases, so that the series (2) converges rather rapidly.

Here we write only the expressions for the first terms of the expansion, which we shall need below:

$$\varepsilon_0(p_\perp,p_z)=\frac{1}{2m}\left[p^2-\alpha_0\left(p_\perp^2p_z^2+\frac{1}{8}p_\perp^4\right)^{1/2}\right], \tag{3}$$

and

$$\varepsilon_4(p_\perp,p_z)=\frac{p_\perp^2}{64m}\left[\alpha_0\frac{p_\perp}{p_z}\left(1-\frac{1}{16}\frac{p_\perp^2}{p_z^2}\right)+3\beta_0\frac{p_\perp^3p_z}{p^4}\times\left(1+\frac{1}{16}\frac{p_\perp^2}{p_z^2}\right)-2\gamma_0\frac{p_\perp^3p_z^3}{p^6}\left(1+\frac{1}{8}\frac{p_\perp^2}{p_z^2}\right)\right] \tag{4}$$

Equation (3) is valid for any value of the polar angle θ be-

 © 1998 American Institute of Physics

tween the momentum **p** and the [001] axis and Eq. (4), is valid for $\theta \leqslant 45°$. The values of $\varepsilon_4(p_\perp, p_z)$ for $\theta > 45°$ are of no importance whatever in the theory.

According to Eqs. (2) and (3), the cyclotron effective mass and cyclotron frequency, both averaged over φ, can be calculated using the formulas

$$\frac{1}{m_c(\theta)} = \frac{1}{m}\left[1 - \frac{\alpha_0}{2|\tan\,\theta|}\,\frac{1+0.25\,\tan^2\,\theta}{(1+0.125\,\tan^2\,\theta)^{1/2}}\right]$$

$$\omega_c = \bar{\omega}_c\,\frac{m}{m_c(\theta)}, \quad \bar{\omega}_c = \frac{eH}{mc}, \tag{5}$$

where e is the electronic charge, c is the speed of light, and θ is the angle between the momentum **p** and the [001] axis.

It is easy to confirm that, inside the cone $\theta = \theta_m \approx 36°$ (θ_m is defined from the condition $\omega_c = 0$) the cyclotron effective mass and cyclotron frequency of the heavy holes take negative values. This implies that the high frequency differential conductivity of the heavy holes will be negative and, therefore, that amplification of electromagnetic waves is possible if a large number of holes lies within the cone $\theta = \theta_m$, i.e., moves within relatively small angles $\theta < \theta_m$. The hole distribution function in this case is highly extended along the direction of the electric and magnetic fields. This can all be attained when the following conditions are satisfied:

$$\hbar\omega_0 \gg k_0 T, \quad \tau_0 < \tau_E < \tau_p, \quad E \gg E_\omega, \tag{6}$$

where ω_0 is the optical phonon frequency; k_0 is the Boltzmann constant; T is the lattice temperature; $\tau_E = \sqrt{2m\hbar\omega_0}/eE$ is the time over which a hole, moving without collisions along the electric field **E**, acquires energy $\hbar\omega_0$; τ_p is the mean free time for holes with $\varepsilon < \hbar\omega_0$ (passive region), determined by scattering on impurities and acoustic phonons, as well as by scattering with absorption of optical phonons; τ_0 is the mean free time for $\varepsilon > \hbar\omega_0$ (active region), determined by spontaneous emission of optical phonons; and, E_ω and ω are the amplitude and frequency, respectively, of the electromagnetic field.

One consequence of the second inequality of Eq. (6) is the inequality $E_0 \gg E$, where $E_0 = \hbar\omega_0/el_0$ is a characteristic electric field and $l_0 = \tau_0\sqrt{2\hbar\omega_0/m}$ is the characteristic mean free path of the holes in the active region. We shall also assume that the following inequalities are satisfied:

$$\frac{eH}{mc}\,\tau_0 \ll 1, \quad \omega_c \tau_E \gg 1. \tag{7}$$

According to the first inequality (7), the magnetic field has no effect on the form of the hole distribution function in the active region. The second inequality (7) is a necessary condition for observation of a cyclotron resonance.

Under conditions (6) and (7), the hole distribution function in the active region is determined by the motion of the holes in the electric field and by spontaneous emission of optical phonons, while in the passive region it is determined by the motion of the holes in the electric and magnetic fields and by the arrival of holes from the active region owing to the emission of optical phonons. The solution of the kinetic equation and the calculation of the hole distribution function under these conditions have been described in detail elsewhere.[5,6] The following expression is obtained for the high frequency resonance differential conductivity at the frequency of any of the $|n+1|$ harmonics ($n = 0, \pm 4, \pm 6, \ldots$):

$$\frac{\sigma_{\omega r|n+1|}}{\sigma_0} = \frac{m_c^2(\theta_r)}{|n+1|}\left(\frac{1}{p_\perp}\frac{\partial\varepsilon_n}{\partial p_\perp} - n\,\frac{\varepsilon_n}{p_\perp^2}\right)^2_{\theta=\theta_r}\left(\frac{\sigma_{\omega r1}}{\sigma_0}\right)_{\theta=\theta_r}, \tag{8}$$

where the cyclotron mass $m_c(\theta_r)$ is given by Eq. (5); θ_r is the resonance angle determined by the resonance condition $\omega = (n+1)\omega_c$; the $\varepsilon_n(p_\perp, p_z)$ are the expansion coefficients in Eq. (2);

$$\sigma_0 = e^2 P_0/m\bar{\omega}_c = ecP_0/H; \tag{9}$$

and $\sigma_{\omega r1}$ is the high-frequency resonance differential conductivity at the frequency of the first harmonic (fundamental), with

$$\left(\frac{\sigma_{\omega r1}}{\sigma_0}\right)_{\theta=\theta_r} = \frac{16\pi}{3\alpha_0}\,\frac{m}{m_c(\theta_r)}\left(\frac{E_0}{E}\right)^2\frac{\sin^2\,\theta_r}{F(\theta_r)}$$
$$\times\frac{(1+0.125\,\tan^2\,\theta_r)^{3/2}}{L+R(E/E_0)^{1/3}}\int dx x^{5/2}$$
$$\times\exp\left(-\frac{2}{3}\,\frac{E_0}{E}\,x^{3/2}\right)\int d\theta\,\frac{\sin^2\,\theta}{F^2(\theta)}$$
$$\times\left[1-\frac{\alpha_0}{2\,\tan\,\theta}\,\frac{1+0.25\,\tan^2\,\theta}{(1+0.125\,\tan^2\,\theta)^{1/2}}\right]. \tag{10}$$

In Eq. (10) we have used the following notation: L and R are constants, $x = \varepsilon/\hbar\omega_0$, the angle θ is less than $\pi/2$ (the region $\pi/2 \leqslant \theta \leqslant \pi$ is taken into account by doubling the integral), and

$$F(\theta_r) = 1 - \alpha_0(\sin^2\,\theta_r\,\cos^2\,\theta_r + 0.125\,\sin^4\,\theta_r)^{1/2}.$$

The integral over θ in Eq. (10) is calculated exactly. The limits of integration with respect to x and θ are chosen so that the point with coordinates (ε, θ) lies within the cylinder

$$p_\perp = p_{\perp r} = \frac{\sqrt{2m\hbar\omega_0}\,\sin\,\theta_r}{\sqrt{F(\theta_r)}},$$

because there are no holes outside this cylinder that are at resonance with the electromagnetic field. L, R, and $(\sigma_{\omega r1}/\sigma_0)_{\theta=\theta_r}$ have been calculated in detail elsewhere.[5]

The absorption coefficient $\eta_{|n+1|}$ at the frequency of the $|n+1|$ harmonic is given by

$$\eta_{|n+1|} = \frac{2\pi\sigma_{\omega r|n+1|}}{c\sqrt{\varepsilon_0}}, \tag{11}$$

where ε_0 is the dielectric constant.

An analysis of Eqs. (10) and (11) and some detailed calculations based on it yielded a number of interesting results. It turns out that the absorption coefficient for electromagnetic waves takes negative values at the frequency of any of $n+1$ harmonics ($n = 0,4,8$, etc.) for the right (electron) polarization and at the frequency of any of $n-1$ har-

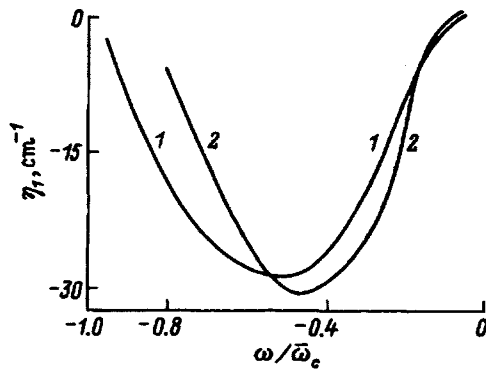


FIG. 1. The absorption η_1 at the first harmonic (right polarized electromagnetic wave) as a function of the ratio $\omega/\bar{\omega}_c$ at $T=100$ K. Computational parameters for the curves: (*1*) $\omega=2\times10^{12}$ s^{-1} ($\lambda=0.92$ mm), $P_0=3\times10^{15}$ cm^{-3}, $E=5\times10^3$ V/cm, $H=35-550$ kOe; (*2*) $\omega=4\times10^{12}$ s^{-1} ($\lambda=0.46$ mm), $P_0=10^{16}$ cm^{-3}, $E=10^4$ V/cm, $H=110-1110$ kOe.

monics ($n=4,8,12$, etc.) for the left (hole) polarization. Thus, amplification of electromagnetic waves is possible for the appropriate polarization at the frequency of any of the odd harmonics.

Figure 1 shows the absorption coefficient η_1 at the first harmonic (right polarized electromagnetic wave) as a function of $\omega/\bar{\omega}_c$ at $T=100$ K.

The curves in Fig. 1 show that conditions exist under which the absorption coefficient takes negative values and, therefore, amplification and generation of electromagnetic radiation can occur in p-type semiconducting diamond. These conditions, which are determined by the inequalities (6) and (7) and the optimum values $\sigma_{\omega r1}\approx-0.4\sigma_0$ and $\omega/\bar{\omega}_c\approx-0.5$ (see Fig. 1), reduce to the following:

1. The concentration of holes and charged impurities must not exceed 10^{16} cm^{-3} and the concentration of neutral impurities is less than 10^{20} cm^{-3}.

2. The lattice temperature and electric and magnetic fields are limited to the intervals $70\leqslant T<300$ K, $5\times10^3\leqslant E\leqslant4\times10^4$ V/cm, and $4\times10^4\leqslant H\leqslant4\times10^5$ Oe.

3. In order to avoid overheating of the sample, the duration of the electric field pulse must not exceed $10^{-6}-10^{-7}$ s.

For the values of T, P_0, E, and H given above, the wavelength of the generated radiation varies over $0.15\leqslant\lambda\leqslant1.5$ mm, the absorption coefficient varies over $\eta_1=(-15)-(-30)$ cm^{-1}, and the maximum power emitted from the crystal surface is

$$W_s=\frac{c\sqrt{\varepsilon_0}E_\omega^2}{4\pi}=(2.5\times10^4\text{-}1.6\times10^6)\,\text{W/cm}^2$$

for $E_\omega=0.4E=2\times10^3-2\times10^4$ V/cm.

The most surprising result from this theory is the rather high gain coefficient at the third harmonic. Figures 2 and 3 show the dependence of the gain coefficient η_3 at the third harmonic (left polarized electromagnetic wave) on $\omega/3\bar{\omega}_c$ for $T=100$ K. In Fig. 2 the values of the frequency ω are the same as in Fig. 1, while the magnetic field strength is reduced by a factor of 3. In Fig. 3 the values of the magnetic field strength H are the same as in Fig. 1, but the frequencies

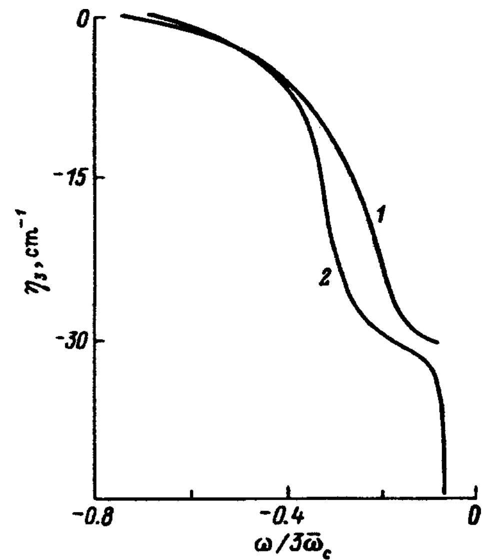


FIG. 2. The absorption η_3 at the third harmonic (left polarized electromagnetic wave) as a function of $\omega/3\bar{\omega}_c$ at $T=100$ K. The parameters used for calculating these curves are the same as in Fig. 1, except for the magnetic field H (kOe): (*1*) 12–180; (*2*) 37–367.

ω are higher by a factor of 3. It follows from Fig. 2 that at a wavelength $\lambda=0.92$ mm, the gain coefficient $\eta_3=(-6.6)-(-22)$ cm^{-1} for $H=2.7-5.5\times10^4$ Oe (curve 1), while at $\lambda=0.46$ mm, $\eta_3(-7.3)-(-30)$ cm^{-1} for $H=5.5-11\times10^4$ Oe. Figure 3 implies that at a wavelength $\lambda=0.31$ mm, the gain coefficient $\eta_3=(-2.2)-(-10)$ cm^{-1} for $H=8.2-23.7\times10^4$ Oe (curve 1), while at $\lambda=0.15$ mm, it equals $(-0.43)-(-10)$ cm^{-1} for $H=1.1-3.3\times10^5$ Oe. In the numerical calculations, the parameters of diamond (τ_0, $\hbar\omega_0$, E_0, hole mobility, etc.) were taken from Ref. 5.

We have written the values of η_1 and η_3 for the case in which the electric and magnetic fields are not too large and the conditions (6) and (7) for validity of the theory are satisfied with a wide safety factor. (See Tables I and II.) For $T=100$ K, $P_0=3\times10^{15}$ cm^{-3}, and $E=3\times10^3$ V/cm, the absorption coefficients η_1 and η_3 take their maximum absolute values. (See the first three lines of Table I.)

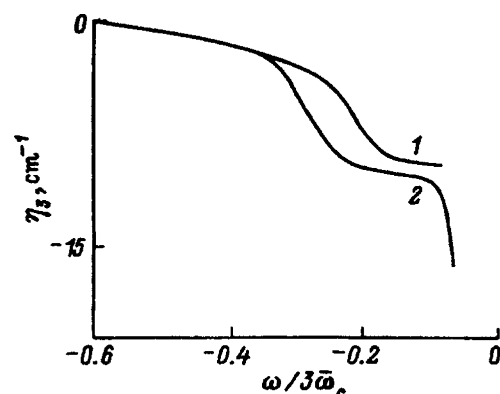


FIG. 3. The absorption η_3 at the third harmonic (left polarized electromagnetic wave) as a function of $\omega/3\bar{\omega}_c$ at $T=100$ K. The parameters used for calculating these curves are the same as in Fig. 1, except for the frequency ω (s^{-1}): (*1*) 6×10^{12} ($\lambda=0.31$ mm); (*2*) 1.2×10^{13} ($\lambda=0.15$ mm).

TABLE I.

$T=100$ K, $P_0=3\times10^{15}$ cm^{-3}				
E, 10^3 V/cm	H, kOe	λ, mm	η_1, cm^{-1}	η_3, cm^{-1}
3	37	1.4	−80	–
3	18-53	1.4	–	(−18)-(−80)
3	55-160	0.46	–	(−6)-(−27)
5	55	0.92	−27	–
5	27-79	0.92	–	(−6.6)-(−30)
5	82-237	0.31	–	(−2)-(−10)

The absorption coefficient also takes negative values (for the appropriate polarization of the electromagnetic wave) at the 5th, 7th, etc., harmonics. However, its absolute magnitude at these harmonics is 1-2, etc., orders of magnitude smaller than that of the absorption coefficients η_1 and η_3.

Our estimates[5] showed that the above mechanism for resonance amplification of electromagnetic waves at $0.15 \leqslant \lambda \leqslant 1.5$ mm overcomes all the mechanisms for absorption in diamond (lattice absorption, intraband nonresonant absorption by holes of all types, absorption associated with interband hole transitions). Including these absorption mechanisms does not affect our conclusions.

The problem of creating the required hole concentration in diamond ($P_0=10^{15}-10^{16}$ cm^{-3}) requires a separate discussion. One possible way of creating such a concentration of holes is to illuminate the diamond with 0.37-eV photons (photoionization). The number n_q of photons required to create a hole concentration P_0 can be estimated using the formula

$$n_q=\frac{\sigma_{\rm rec}\bar{\nu}P_0^2}{\sigma_{\rm ph}N_a},$$

where $\sigma_{rec}=1.5\times10^{-17}$ cm^2 is the experimental value of the cross section for hole recombination at boron acceptor centers in diamond,[7] $\sigma_{ph}=1.56\times10^{-16}$ cm^2 is the experimental value of the cross section for photoionization of a boron atom in diamond,[7] N_a is the acceptor concentration, and $\bar{\nu}$ is the average velocity of the recombining holes. For $P_0=10^{15}-10^{16}$ cm^{-3}, $N_a=10^{17}-10^{19}$ cm^{-3}, and $\bar{\nu}=10^7$ cm/s (a clearly excessive value, since the slow holes recombine), we obtain $n_q=10^{17}-10^{19}$ cm^{-2}s^{-1}. The energy flux incident on 1 cm^2 per second is $W=\hbar\omega n_q=0.004-0.4$ W ($\hbar\omega=0.37$ eV). It may also be possible to obtain the required hole concentration by injection.

TABLE II.

$T=70\div200$ K, $P_0=10^{16}$ cm^{-3}				
E, 10^3 V/cm	H, kOe	λ, mm	η_1, cm^{-1}	η_3, cm^{-1}
10	83-164	0.46	(−15)-(−30)	–
10	55-109	0.46	–	(−7)-(−30)
10	200	0.15	–	−3

Thus, amplification and generation of electromagnetic waves at the cyclotron resonance of heavy holes with negative masses may be achieved in diamond at higher temperatures and at shorter wavelengths than in germanium. Here the gain coefficient is 10–50 times higher, while the maximum power emitted from the sample surface is several orders of magnitude higher than for Ge.

The author thanks V. S. Vavilov, Yu. A. Mityagin, V. N. Murzin, and N. A. Penin for a discussion of the results of this study and for useful advice.

This work was supported in part by the Russian Fund for Fundamental Research (Project 96-02-18593a).

[1] G. C. Dousmanis, Phys. Rev. Lett. **1**, 55 (1958).
[2] A. A. Andronov, E. P. Dodin, and Z. F. Krasil'nik, Fiz. Tekh. Poluprovodn. **16**, 212 (1982) [Sov. Phys. Semicond. **16**, 133 (1982)].
[3] A. A. Andronov, A. M. Belyantsev, V. I. Gavrilenko, E. P. Dodin, Z. F. Krasil'nik, V. V. Komarov, S. A. Pavlov, and M. M. Shvarts, Zh. Éksp. Teor. Fiz. **90**, 367 (1986) [Sov. Phys. JETP **63**, 211 (1986)].
[4] E. O. Kane, J. Phys. Chem. Solids **1**, 82 (1956).
[5] V. A. Chuenkov, Tr. FIAN Im. P. N. Lebedev **224**, 213 (1993).
[6] V. A. Chuenkov, Izv. RAN, Ser. Fiz. **58**, 51 (1994).
[7] V. S. Vavilov and E. A. Konorova, Usp. Fiz. Nauk **118**, 611 (1976).

Translated by D. H. McNeill

PERSONALIA

Yaroslav Evgen'evich Pokrovskiĭ (On his 70th birthday)

Fiz. Tekh. Poluprovodn. **32**, 509 (April 1998)
[S1063-7826(98)02304-7]

In January, 1998, Professor Yaroslav Evgen'evich Pokrovskiĭ, a member of the Editorial Board of *Fizika i Tekhnika poluprovodnikov*, an outstanding physicist, a corresponding member of the Russian Academy of Sciences, had his 70th birthday. His scientific activity has all been inseparably associated with the physics of semiconductors and it has determined the modern state of the field to a great extent.

Yaroslav Evgen'evich's major scientific discoveries include the discovery of a new phase transition– exciton condensation in electron-hole fluids. For this discovery he, along with other scientists, was awarded a prize by the European Physical Society in 1975, and his major article on this topic has been recognized by the Philadelphia Institute of Scientific Information as a ''citation classic.''

Another important work of his was the discovery of exciton-impurity complexes— a new quantum mechanical object inherent only to semiconductors. Studies of the luminescence of these complexes have made it possible to determine their unusual electronic structure. This work was recognized as a discovery, and a USSR State Prize was awarded for it in 1988. Continued research in this area has made it possible to develop a sensitive technique for analyzing the impurity content of semiconductors.

The fundamental studies conducted by Ya. E. Pokrovskiĭ have always been closely related to practical problems– technologies for obtaining pure materials and the development of semiconductor devices. At the very dawn of the development of semiconductor physics in this country, in 1956, he proposed and developed a silane method for the production of pure silicon, which is still the main method used in industry. For his great contribution to the creation and development of infrared photodetectors, he was awarded the ''Znak Pocheta'' (Medal of Distinction) in 1977.

Ya. E. Pokrovskiĭ, a talented experimental physicist and leader of a prominent scientific school, has a well deserved status among the world scientific community.

We send cordial greetings to Yaroslav Evgen'evich on his birthday and wish him a long, productive life and success in realizing his plans.

His colleagues, friends, and the Editorial Board

Translated by D. H. McNeill

© 1998 American Institute of Physics